

HAND- BOOKS IN ECONOMICS

Benjamin M. Friedman
Michael Woodford

Monetary Economics

VOLUME 3A

NORTH-HOLLAND

HANDBOOK OF MONETARY ECONOMICS

VOLUME

3A

INTRODUCTION TO THE SERIES

The aim of the *Handbooks in Economics* series is to produce Handbooks for various branches of economics, each of which is a definitive source, reference, and teaching supplement for use by professional researchers and advanced graduate students. Each Handbook provides self-contained surveys of the current state of a branch of economics in the form of chapters prepared by leading specialists on various aspects of this branch of economics. These surveys summarize not only received results but also newer developments, from recent journal articles and discussion papers. Some original material is also included, but the main goal is to provide comprehensive and accessible surveys.

The Handbooks are intended to provide not only useful reference volumes for professional collections but also possible supplementary readings for advanced courses for graduate students in economics.

KENNETH J. ARROW and **MICHAEL D. INTRILIGATOR**

HANDBOOK OF MONETARY ECONOMICS

VOLUME

3A

Edited by

BENJAMIN M. FRIEDMAN

MICHAEL WOODFORD



Amsterdam • Boston • Heidelberg • London • New York • Oxford
Paris • San Diego • San Francisco • Singapore • Sydney • Tokyo
North-Holland is an imprint of Elsevier



North-Holland in an imprint of Elsevier
525 B Street, Suite 1800, San Diego, CA 92101-4495, USA
Radarweg 29, 1000 AE Amsterdam, The Netherlands

First edition 2011

Copyright © 2011 Elsevier B.V. All rights reserved

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means electronic, mechanical, photocopying, recording or otherwise without the prior written permission of the publisher

Permissions may be sought directly from Elsevier's Science & Technology Rights Department in Oxford, UK: phone (+44) (0) 1865 843830; fax (+44) (0) 1865 853333; email: permissions@elsevier.com. Alternatively you can submit your request online by visiting the Elsevier web site at <http://elsevier.com/locate/permissions>, and selecting *Obtaining permission to use Elsevier material*

Notice

No responsibility is assumed by the publisher for any injury and/or damage to persons or property as a matter of products liability, negligence or otherwise, or from any use or operation of any methods, products, instructions or ideas contained in the material herein. Because of rapid advances in the medical sciences, in particular, independent verification of diagnoses and drug dosages should be made

Library of Congress Cataloging-in-Publication Data

A catalog record for this book is available from the Library of Congress

British Library Cataloguing in Publication Data

A catalogue record for this book is available from the British Library

ISBN Vol 3A: 978-0-444-53238-1

ISBN Vol 3B: 978-0-444-53454-5

SET ISBN: 978-0-444-53470-5

For information on all North-Holland publications
visit our website at elsevierdirect.com

Printed and bound in the USA

11 12 13 10 9 8 7 6 5 4 3 2 1

Working together to grow
libraries in developing countries

www.elsevier.com | www.bookaid.org | www.sabre.org

ELSEVIER

BOOK AID
International

Sabre Foundation

CONTENTS-VOLUME 3A

<i>Contributors</i>	xv
<i>Preface</i>	xvii

Part One: Foundations: The Role of Money in the Economy

1. The Mechanism-Design Approach to Monetary Theory	3
Neil Wallace	
1. Introduction	4
2. Some Frictions	5
3. An Illustrative Model with Perfect Recognizability	8
4. Imperfect Recognizability and Uniform Currency	14
5. Optima Under a Uniform Outside Currency	16
6. Extensions of the Illustrative Model	18
7. Concluding Remarks	22
References	23
2. New Monetarist Economics: Models	25
Stephen Williamson and Randall Wright	
1. Introduction	26
2. Basic Monetary Theory	31
3. A Benchmark Model	38
4. New Models of Old Ideas	57
5. Money, Payments, and Banking	71
6. Finance	79
7. Conclusion	89
References	90
3. Money and Inflation: Some Critical Issues	97
Bennett T. McCallum and Edward Nelson	
1. Introduction	98
2. The Quantity Theory of Money	99
3. Related Concepts	102
4. Historical Behavior of Monetary Aggregates	104
5. Flawed Evidence on Money Growth-Inflation Relations	108

6. Money Growth and Inflation in Time Series Data	112
7. Implications of a Diminishing Role for Money	134
8. Money Versus Interest Rates in Price Level Analysis	136
9. Conclusions	146
References	148

Part Two: Foundations: Information and Adjustment

4. Rational Inattention and Monetary Economics 155

Christopher A. Sims

1. Motivation	156
2. Information Theory	157
3. Information Theory and Economic Behavior	160
4. Implications for Macroeconomic Modeling	171
5. Implications for Monetary Policy	174
6. Directions for Progress	176
7. Conclusion	178
References	180

5. Imperfect Information and Aggregate Supply 183

N. Gregory Mankiw and Ricardo Reis

1. Introduction	184
2. The Baseline Model of Aggregate Supply	185
3. Foundations of Imperfect-Information and Aggregate-Supply Models	191
4. Partial and Delayed Information Models: Common Predictions	196
5. Partial and Delayed Information Models: Novel Predictions	207
6. Microfoundations of Incomplete Information	213
7. The Research Frontier	217
8. Conclusion	222
References	223

6. Microeconomic Evidence on Price-Setting 231

Peter J. Klenow and Benjamin A. Malin

1. Introduction	232
2. Data Sources	234
3. Frequency of Price Changes	238
4. Size of Price Changes	257
5. Dynamic Features of Price Changes	258

6. Ten Facts and Implications for Macro Models	271
7. Conclusion	278
References	279

Part Three: Models of the Monetary Transmission Mechanism

7. DSGE Models for Monetary Policy Analysis 285

Lawrence J. Christiano, Mathias Trabandt, and Karl Walentin

1. Introduction	286
2. Simple Model	289
3. Simple Model: Some Implications for Monetary Policy	302
4. Medium-Sized DSGE Model	331
5. Estimation Strategy	345
6. Medium-Sized DSGE Model: Results	351
7. Conclusion	362
References	364

8. How Has the Monetary Transmission Mechanism Evolved Over Time? 369

Jean Boivin, Michael T. Kiley, and Frederic S. Mishkin

1. Introduction	370
2. The Channels of Monetary Transmission	374
3. Why the Monetary Transmission Mechanism may have Changed	385
4. Has the Effect of Monetary Policy on the Economy Changed? Aggregate Evidence	388
5. What Caused the Monetary Transmission Mechanism to Evolve?	396
6. Implications for the Future Conduct of Monetary Policy	415
References	418

9. Inflation Persistence 423

Jeffrey C. Fuhrer

1. Introduction	424
2. Defining and Measuring Reduced-Form Inflation Persistence	431
3. Structural Sources of Persistence	449
4. Inference about Persistence in Small Samples: "Anchored Expectations" and their Implications for Inflation Persistence	473
5. Microeconomic Evidence on Persistence	478
6. Conclusions	482
References	483

10. Monetary Policy and Unemployment	487
Jordi Galí	
1. Introduction	488
2. Evidence on the Cyclical Behavior of Labor Market Variables and Inflation	491
3. A Model with Nominal Rigidities and Labor Market Frictions	495
4. Equilibrium Dynamics: The Effects of Monetary Policy and Technology Shocks	515
5. Labor Market Frictions, Nominal Rigidities and Monetary Policy Design	528
6. Possible Extensions	535
7. Conclusions	537
References	543
11. Financial Intermediation and Credit Policy in Business Cycle Analysis	547
Mark Gertler and Nobuhiro Kiyotaki	
1. Introduction	548
2. A Canonical Model of Financial Intermediation and Business Fluctuations	551
3. Credit Policies	566
4. Crisis Simulations and Policy Experiments	574
5. Issues and Extensions	581
6. Concluding Remarks	589
References	597
12. Financial Intermediaries and Monetary Economics	601
Tobias Adrian and Hyun Song Shin	
1. Introduction	602
2. Financial Intermediaries and the Price of Risk	606
3. Changing Nature of Financial Intermediation	615
4. Empirical Relevance of Financial Intermediary Balance Sheets	623
5. Central Bank as Lender of Last Resort	631
6. Role of Short-Term Interest Rates	636
7. Concluding Remarks	646
References	648
Index-Volume 3A	11
Index-Volume 3B	139

CONTENTS-VOLUME 3B

<i>Contributors</i>	xv
<i>Preface</i>	xvii

Part Four: Optimal Monetary Policy

13. The Optimal Rate of Inflation	653
Stephanie Schmitt-Grohé and Martín Uribe	
1. Introduction	654
2. Money Demand and the Optimal Rate of Inflation	658
3. Money Demand, Fiscal Policy and the Optimal Rate of Inflation	664
4. Failure of the Friedman Rule Due to Untaxed Income: Three Examples	667
5. A Foreign Demand For Domestic Currency and the Optimal Rate of Inflation	675
6. Sticky Prices and the Optimal Rate of Inflation	684
7. The Friedman Rule Versus Price-Stability Trade-Off	695
8. Does the Zero Bound Provide a Rationale for Positive Inflation Targets?	701
9. Downward Nominal Rigidity	704
10. Quality Bias and the Optimal Rate of Inflation	706
11. Conclusion	715
References	720
14. Optimal Monetary Stabilization Policy	723
Michael Woodford	
1. Introduction	724
2. Optimal Policy in a Canonical New Keynesian Model	726
3. Stabilization and Welfare	759
4. Generalizations of the Basic Model	790
5. Research Agenda	818
References	826
15. Simple and Robust Rules for Monetary Policy	829
John B. Taylor and John C. Williams	
1. Introduction	830
2. Historical Background	830
3. Using Models to Evaluate Simple Policy Rules	833

4. Robustness of Policy Rules	844
5. Optimal Policy Versus Simple Rules	850
6. Learning from Experience Before, During and after the Great Moderation	852
7. Conclusion	855
References	856
16. Optimal Monetary Policy in Open Economies	861
Giancarlo Corsetti, Luca Dedola, and Sylvain Leduc	
1. Introduction and Overview	862
2. Part I: Optimal Stabilization Policy and International Relative Prices with Frictionless Asset Markets	869
3. A Baseline Monetary Model of Macroeconomic Interdependence	870
4. The Classical View: Divine Coincidence in Open Economies	886
5. Skepticism on the Classical View: Local Currency Price Stability of Imports	894
6. Deviations from Policy Cooperation and Concerns with "Competitive Devaluations"	909
7. Part II: Currency Misalignments and Cross-Country Demand Imbalances	915
8. Macroeconomic Interdependence Under Asset Market Imperfections	915
9. Conclusions	928
References	929
Part Five: Constraints on Monetary Policy	
17. The Interaction Between Monetary and Fiscal Policy	935
Matthew Canzoneri, Robert Cumby, and Behzad Diba	
1. Introduction	936
2. Positive Theory of Price Stability	937
3. Normative Theory of Price Stability: Is Price Stability Optimal?	973
References	995
18. The Politics of Monetary Policy	1001
Alberto Alesina and Andrea Stella	
1. Introduction	1002
2. Rules Versus Discretion	1003
3. Central Bank Independence	1013
4. Political Business Cycles	1027
5. Currency Unions	1034
6. The Euro	1041
7. Conclusion	1046
References	1050

19. Inflation Expectations, Adaptive Learning and Optimal Monetary Policy 1055

Vitor Gaspar, Frank Smets, and David Vestin

1. Introduction	1056
2. Recent Developments in Private-Sector Inflation Expectations	1059
3. A Simple New Keynesian Model of Inflation Dynamics Under Rational Expectations	1061
4. Monetary Policy Rules And Stability Under Adaptive Learning	1065
5. Optimal Monetary Policy Under Adaptive Learning	1071
6. Some Further Reflections	1089
7. Conclusions	1091
References	1092

20. Wanting Robustness in Macroeconomics 1097

Lars Peter Hansen and Thomas J. Sargent

1. Introduction	1098
2. Knight, Savage, Ellsberg, Gilboa-Schmeidler, and Friedman	1100
3. Formalizing a Taste for Robustness	1104
4. Calibrating a Taste for Robustness	1109
5. Learning	1117
6. Robustness in Action	1133
7. Concluding Remarks	1148
References	1155

Part Six: Monetary Policy in Practice

21. Monetary Policy Regimes and Economic Performance: The Historical Record, 1979–2008 1159

Luca Benati and Charles Goodhart

1. Introduction	1160
2. Monetary Targetry, 1979–1982	1168
3. Inflation Targets	1183
4. The “Nice Years,” 1993–2006	1185
5. Europe and the Transition to the Euro	1204
6. Japan	1209
7. Financial Stability and Monetary Policy During the Financial Crisis	1216
8. Conclusions and Implications for Future Central Bank Policies	1221
References	1231

22. Inflation Targeting	1237
Lars E.O. Svensson	
1. Introduction	1238
2. History and Macroeconomic Effects	1242
3. Theory	1250
4. Practice	1275
5. Future	1286
References	1295
23. The Performance of Alternative Monetary Regimes	1303
Laurence Ball	
1. Introduction	1304
2. Some Simple Evidence	1306
3. Previous Work on Inflation Targeting	1313
4. The Euro	1318
5. The Role of Monetary Aggregates	1325
6. Hard Currency Pegs	1328
7. Conclusion	1332
References	1341
24. Implementation of Monetary Policy: How Do Central Banks Set Interest Rates?	1345
Benjamin M. Friedman and Kenneth N. Kuttner	
1. Introduction	1346
2. Fundamental Issues in the Mode of Wickseil	1353
3. The Traditional Understanding of "How they do that"	1360
4. Observed Relationships Between Reserves and the Policy Interest Rate	1375
5. How, Then, Do Central Banks Set Interest Rates?	1385
6. Empirical Evidence on Reserve Demand and Supply within the Maintenance Period	1399
7. New Possibilities Following the 2007–2009 Crisis	1414
8. Conclusion	1432
References	1433
25. Monetary Policy in Emerging Markets	1439
Jeffrey Frankel	
1. Introduction	1441
2. Why Do We Need Different Models for Emerging Markets?	1443
3. Goods Markets, Pricing, and Devaluation	1445

4. Inflation	1453
5. Nominal Targets for Monetary Policy	1456
6. Exchange Rate Regimes	1461
7. Procyclicality	1465
8. Capital Flows	1472
9. Crises in Emerging Markets	1481
10. Summary of Conclusions	1498
References	1499
Index-Volume 3B	11
Index-Volume 3A	141

This page intentionally left blank

CONTRIBUTORS

Tobias Adrian

Federal Reserve Bank of New York

Jean Boivin

Bank of Canada

Lawrence J. Christiano

Northwestern University

Jeffrey C. Fuhrer

Federal Reserve Bank of Boston

Jordi Galí

Universitat Pompeu Fabra

Mark Gertler

New York University

Michael T. Kiley

Board of Governors of the Federal Reserve System

Nobuhiro Kiyotaki

Princeton University

Peter J. Klenow

Stanford University

Benjamin A. Malin

Federal Reserve Board

N. Gregory Mankiw

Harvard University and Columbia University

Bennett T. McCallum

Carnegie Mellon University

Frederic S. Mishkin

Columbia University

Edward Nelson

Federal Reserve Board

Ricardo Reis

Harvard University and Columbia University

Christopher A. Sims

Princeton University

Hyun Song Shin

Princeton University

Mathias Trabandt

European Central Bank

Karl Walentin

Sveriges Riksbank

Neil Wallace

Pennsylvania State University

Stephen Williamson

Washington University

Randall Wright

University of Wisconsin – Madison

PREFACE

These new volumes supplement and bring up to date the original *Handbook of Monetary Economics* (Volumes I and II of this series), edited by Benjamin Friedman with Frank Hahn. It is now twenty years since the publication of those earlier volumes, so a reconsideration of the field is timely if not overdue. Some of the topics covered in the previous volumes of *Handbook of Monetary Economics* were updated in the *Handbook of Macroeconomics*, edited by Michael Woodford with John Taylor, but it is now ten years since the publication of those volumes as well. Further, that publication, with its broader focus on macroeconomics, could not fully substitute for a new edition of the *Handbook of Monetary Economics*. The subject here is macroeconomics, to be sure, but it is *monetary* macroeconomics.

Publication of a “handbook” in some area of intellectual inquiry usually means that researchers in the field have made substantial progress that is worth not only reviewing but also adding, in summary form, to the canonical presentation of work made conveniently available to students and other interested scholars. As the 25 chapters included in these new volumes make clear, this has certainly been the case in monetary macroeconomics. While many chapters of both the 1990 *Handbook of Monetary Economics* and the 2000 *Handbook of Macroeconomics* will remain valuable resources, the pace of recent progress has been such that a summary from even as recently as a decade ago is incomplete in many important respects. These new volumes are intended to fill that gap.

Publication of a handbook also often means that a field has reached a sufficient stage of maturity so that it is safe to take stock without concern that new ideas, or the press of external events, will soon result in significant new directions. Today, however, the opposite is likely to be true in monetary macroeconomics. The extraordinary economic and financial events of 2007–2010 seem highly likely to prod researchers to consider new lines of thinking, and to evaluate old ones against new bodies of evidence that in many key respects differ sharply from prior experience. It is obviously too early for us to anticipate what the full consequences of such reconsideration would be. We believe, however, that it is valuable to take stock of the state of the field “before the deluge.” Further, a number of the chapters included here present early attempts to pursue lines of inquiry suggested by the 2007–2010 experience.

Developments in the world economy since the publication of the earlier volumes of this *Handbook* provided much new ground for economic thinking, even prior to the recent crisis, and these had already spurred significant developments in monetary macroeconomics as well. Among the notable monetary experiments of the past two decades, we should mention two in particular. The creation of a monetary union in

Europe has not only introduced a new major world currency and a new central bank, but has revived interest in the theory of monetary unions and “optimal currency areas” and raised novel questions about the degree to which it is possible to separate monetary policy from fiscal policy and from financial supervision (the latter issues are handled at a completely different level of government in the Euro Zone). And the spread of inflation targeting as an approach to the conduct of monetary policy — first adopted mainly by members of the OECD, now increasingly popular among emerging market economies as well, but still resisted by a number of highly visible central banks (including, most clearly, the U.S. Federal Reserve System) — has brought not only a stronger degree of emphasis on inflation stabilization as a policy goal but also greater explicitness about central banks’ policy targets and a more integrated role for quantitative modeling in policy deliberations. It has also changed central banks’ communications with the public about those deliberations. Both of these developments have been the subject of extensive scholarly analysis, both theoretical and empirical, and they are treated in detail in several chapters of these new volumes.

The past two decades have witnessed important methodological advances in monetary macroeconomics as well. One of the more notable of these has been the development of empirical dynamic stochastic general equilibrium (DSGE) models that incorporate serious (although also seriously incomplete) efforts to capture the monetary policy transmission mechanism. While these models are doubtless still at a fairly early stage of development, and the adequacy of current-generation DSGE models for practical policy analysis remains a topic of lively debate, for at least the past decade they have been an important focus of research efforts, particularly in central banks around the world and in other policy institutions. Quite a few of the chapters included here rely on these models, while several others examine these models’ structure and the methods used to estimate and evaluate them, with particular emphasis on the account that they give of the transmission mechanism for monetary policy.

There have also been important changes in the methods used to assess the empirical realism of particular models. One important development has been the increasing use of structural vector autoregression methodology to estimate the effects of monetary policy shocks under relatively weak theoretical assumptions. The chapter on this topic in the *Handbook of Macroeconomics* (Chapter 7; Christiano, Eichenbaum, and Evans, 1999) provides a sufficient exposition of this method; but several of the chapters included in these volumes illustrate how this method is now routinely used in applied work. Another notable development in empirical methodology has been increasing use by macroeconomists of individual or firm-level data sets, and not simply aggregate time series, as sources of evidence about aspects of behavior that are central to macroeconomic models. Some of the work surveyed in these new volumes illustrates this importation of micro-level data into monetary macroeconomics.

Finally, there have been important methodological innovations in monetary policy analysis as well. Research on monetary policy rules has exploded over this period, having received considerable impetus from the celebrated proposal of the “Taylor rule” (Taylor, 1993), which not only suggested the possibility that some fairly simple rules might have desirable properties, but also indicated that some aspects of the behavior of actual central banks might be usefully characterized in terms of simple rules. Among other notable developments, an active literature over the past decade has assessed proposed rules for the conduct of monetary policy in terms of their implications for welfare as measured by the private objectives (household utility) that underlie the behavioral relations in microfounded models of the monetary transmission mechanism — essentially applying to monetary policy the method that had already become standard in the theory of public finance. Many of the chapters in these new *Handbook* volumes address these issues, and others related to them as well.

The events of the years immediately preceding publication of these new *Handbook* volumes have presented further challenges and opportunities for research in much of economics, but in monetary macroeconomics in particular. The 2007–2010 financial crisis and economic downturn constituted one of the most significant sequences of economic dislocations since World War II. In many countries the real economic costs — costs in terms of reduced production, lost jobs, shrunken investment, and foregone incomes and profits — exceeded those of any prior post-war decline. It was in the financial sector, however, that this latest episode primarily stood out. The collapse of major financial firms, the decline in asset values and consequent destruction of paper wealth, the interruption of credit flows, the loss of confidence both in firms and in credit market instruments, the fear of default by counterparties, and above all the intervention by central banks and other governmental institutions, were extraordinary.

Large-scale and unusual events often present occasions for introspection and learning, especially when they bring unwanted consequences. David Hume (1987), residing in Edinburgh during the Scottish banking crisis of 1772, wrote of that distressing sequence of events to his close friend Adam Smith. After recounting the bank failures, spreading unemployment, and “Suspicion” surrounding yet other industrial firms as well as banks, including even the Bank of England, Hume asked his friend, “Do these Events any-wise affect your Theory?” They certainly did. In *The Wealth of Nations*, published just four years later, Smith took the 1772 crisis into account in describing the interrelation of banking and nonfinancial economic activity and recommended a set of policy interventions that he thought would preclude or at least soften such disastrous episodes in the future.

The field of monetary macroeconomics has always been especially subject to just this kind of influence stemming from events in the world of which researchers are attempting to gain an understanding. Even the very origins of the field reflect the influence of real-world events. For all practical purposes it was the depression of the 1930s

that created monetary macroeconomics as a recognizable component within the broader discipline, placing the obvious fact of limited price flexibility, and its consequences, at the center of the field's attention, and introducing new intellectual constructs like aggregate demand. In the 1970s, as high inflation rates became both widespread and chronic across most industrialized economies, further new constructs such as dynamic inconsistency, again together with its consequences, profoundly influenced the field's approach to issues of monetary policy. In the 1980s, the experience of disinflation led the field to change its direction and focus once again, as the costs associated with disinflation in many countries contradicted key lines of thinking spawned during the prior decade, and it was difficult to identify first-order differences in the disinflation experiences of countries that had pursued different policy paths and under different policy institutions.

There is no reason to expect the events of 2007–2010 to have any lesser impact. One influence that is already evident in new work in the field, and reflected in several of the chapters included in these new *Handbook* volumes, is an enhanced focus on credit; that is, the liability side of the balance sheets of households and firms and, conversely, the asset side (as opposed to the deposit, or “money” side) of the balance sheets of banks and other financial institutions. The reason is plain enough. In most economies that experienced severe crises and economic downturns in 2007–2010, the quantity of money did not decline and there was no evident scarcity of reserves supplied to the banking system by the central bank. Instead, what mattered, both in the origins of the crisis and for its consequences for nonfinancial economic activity, was the volume and price and availability of credit.

Another aspect of the crisis that has inspired new lines of research, also reflected in some of the chapters included in these new volumes, is the role of nonbank financial institutions. Traditional monetary economics, with its emphasis on the presumed central role of households' and firms' holdings of deposits as assets, naturally focused on deposit-issuing institutions. In some economies in recent decades, nonbank institutions began to issue deposit-like instruments, and therefore they too became of interest; but the volumes involved were normally small, and as an intellectual matter it was easy enough to consider these firms merely as a different form of “bank.” By contrast, once the emphasis shifts to the credit side of financial activity, the path is open for entertaining a key role for institutions that are very unlike banks and that may issue no deposit-like liabilities at all. At the same time, it becomes all the more important to understand the role played by prevailing *institutions*, including matters of financial regulation as well as more general aspects of business organization and practice (limited liability and the consequent distortion of incentives, broadly dispersed stockownership and the consequent principal-agent conflicts, and the like). Several of the chapters included here summarize the most recent research, or present entirely new research, along just these lines.

Yet further lines of inquiry motivated by the 2007–2010 experience remain sufficiently new, or as yet untried in a satisfactorily fleshed-out way, or even fundamentally uncertain, that it is still too early for these new *Handbook* volumes to reflect them. Will the experience of pricing of some credit market instruments — most obviously, claims against U.S. residential mortgages, but many others besides — lead to a broader questioning of what have until now been standard presumptions about rationality of asset markets? Will new theoretical advances make it possible to render the degree of market rationality, in this and other contexts, endogenous with respect to either economic outcomes or economic policy arrangements? Will the surprising (to many economists) use of discretionary anti-cyclical fiscal policy in many countries, or the sharp and seemingly sudden deterioration in governments' fiscal positions, lead to renewed interest in fiscal-monetary connections, possibly with new normative implications? Most generally of all, will the experience of the deepest and longest lasting economic downturn in six decades lead to new thinking about the business cycle itself, including its origins as well as potential policy remediation?

As of 2010, the answer in each case is that no one knows. All that seems certain, given past experience, is that monetary macroeconomics will continue to evolve — and, we trust, to progress. In another decade, or two, there will be room for yet a further *Handbook* to supplement these new volumes. But for now, the 25 chapters published for the first time here speak to the status of a field that has been and will continue to be central to the discipline of economics. We hope students of the field, both new and experienced, will learn from them.

Our foremost debt in presenting these new *Handbook* volumes is to the authors who have contributed their work to be published here. Their own research and their review of the research of others is ample testimony to the effort they have put into this project, and we are grateful to every one of them for it.

We are also grateful to many others who have also added their efforts to this endeavor. Each of the chapters published here was presented, in early draft form, at one of two conferences held in the fall of 2009: one hosted by the Board of Governors of the Federal Reserve System and the other by the European Central Bank. We thank the Board and the ECB for their support of this project and for their generous hospitality. We are also grateful to the economists at these two institutions who took the lead in organizing these two events: at the Federal Reserve Board, Christopher Erceg, Michael Kiley, and Andrew Levin; and at the ECB, Frank Smets and Oreste Tristani. The planning of these conferences required an enormous amount of personal effort on their part, and we certainly appreciate it. We also thank Sue Williams at the Federal Reserve Board and Iris Bettenhauser at the ECB for the efficient and friendly staff support that they rendered.

The presentation of each draft chapter, at one or the other of these two conferences, involved a prepared response by a designated discussant. We are especially

grateful to the over two dozen fellow economists who devoted their efforts to offering extremely thoughtful discussions that in most cases turned out to be both highly constructive and helpful. Their commentaries are not explicitly included in these volumes, but the ideas that they suggested are well reflected in the revised chapters published here. With few exceptions, these chapters are better — better thought out, better organized, better written, and more comprehensive in surveying the relevant research in their assigned areas — because of the comments that the authors received at the conferences.

Finally, we are grateful to Kenneth Arrow and Michael Intriligator, the long-time general editors of this *Handbook* series, for urging us to undertake these new volumes of the *Handbook of Monetary Economics*. We would not have done so without their encouragement.

Benjamin M. Friedman

Harvard University

Michael Woodford

Columbia University

May, 2010

REFERENCES

- Christiano, L.J., Eichenbaum, M., Evans, C.L., 1999. Monetary policy shocks: What have we learned and to what end? In: Taylor, J.B., Woodford, M. (Eds.), *Handbook of macroeconomics*, vol. 1A. Elsevier, Amsterdam.
- Hume, D., 1987. Letter to Adam Smith, 3 September 1772. In: Mossner, E.C., Ross, I.S. (Eds.), *Correspondence of Adam Smith*. Oxford University Press, Oxford, UK, p. 131.
- Taylor, J.B., 1993. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series in Public Policy* 39, 195–214.

PART *One*

Foundations: The Role of Money in the Economy

This page intentionally left blank

The Mechanism-Design Approach to Monetary Theory[☆]

Neil Wallace

The Pennsylvania State University, Department of Economics

Contents

1. Introduction	4
2. Some Frictions	5
2.1 Imperfect monitoring	6
2.2 Costly connections among people	7
2.3 Imperfect recognizability	8
3. An Illustrative Model with Perfect Recognizability	8
3.1 The model	8
3.2 A class of allocations	9
3.3 Incentive-feasible allocations	10
3.4 Results	11
4. Imperfect Recognizability and Uniform Currency	14
5. Optima Under a Uniform Outside Currency	16
6. Extensions of the Illustrative Model	18
6.1 Capital	18
6.2 Endogenous monitored status	19
6.3 Other information structures and other financial instruments	20
6.4 Production and consumption at the centralized stage	21
7. Concluding Remarks	22
References	23

Abstract

The mechanism-design approach to monetary theory is the search for fruitful settings in which money is necessary for the achievement of some desirable allocations. Fruitfulness means that the settings provide insights about puzzling observations and policy questions. Settings with three frictions are considered: imperfect monitoring, costly connections among people, and imperfect recognizability of assets. An illustrative model with those frictions is used to explain as an optimum the following features of actual economies: currency is a uniform object, currency is (usually) dominated in rate of return, some transactions are accomplished using currency and others are accomplished in other ways.

JEL classification: E4, E5

[☆] I am indebted to Ed Green and the editors for helpful comments on an earlier draft.

Keywords

Money
Frictions
Inside-money
Mechanism-design
Monetary and Fiscal policy
Outside-money

1. INTRODUCTION

The mechanism-design approach to monetary theory is the search for fruitful settings or environments in which something that resembles monetary trade actually accomplishes something — or, in Hahn’s (1973) terminology, settings in which money is essential. Fruitfulness means that the settings provide new insights about puzzling observations and policy questions.

The search for settings in which money is essential is hardly new. Suggestions about absence-of-double-coincidence difficulties go back at least to the first millennium (see Monroe, 1966). However, despite being repeated over and over again ever since, those statements are incomplete. After all, if they were regarded as satisfactory, then the search would long ago have been regarded as over. If it were over, then the problem of integrating price theory and monetary theory would not have been one of the big unsolved problems in economics throughout the twentieth century.¹

Monetary trade accomplishes something if monetary trade is necessary for the achievement of some desirable allocations. To establish such necessity, it must be shown there is no other way to achieve those allocations. That, in turn, requires that all other ways be considered. Mechanism design is the tool that can be used to consider all other ways.

Is essentiality in the above sense a reasonable goal? I think so. Monetary trade has been a pervasive phenomenon. While it is conceivable that its appearance is accidental in the sense that it is one of many equivalent ways of achieving desirable allocations, I find that far-fetched—in part, because the settings described below in which monetary trade is essential are intrinsically attractive.

So what kinds of settings lend themselves to a mechanism-design analysis of monetary trade and are fruitful? Needless to say, models with cash-in-advance constraints — or, more generally, models with asset-specific transaction costs — and models with real balances as arguments of utility or production functions are not among the candidates for such settings. The former are ruled out because their structure does not permit us to ask about other ways of achieving allocations and the latter are ruled out because they are at best implicit versions of the former. My general suggestion is that we study environments with three types of frictions: imperfect monitoring, costly connections among people, and imperfect recognizability of assets.

¹ For example, Banerjee and Maskin (1996) allude to that problem when they begin their 1996 paper on money by saying: “Money is something of an embarrassment to economic theory.”

One of the biggest payoffs from doing mechanism-design analysis against the background of such frictions is that it allows us to bypass the distinction between monetary and fiscal policy, and, more generally, to bypass the need to make assumptions about what policies are feasible. The frictions dictate what policies are feasible. Ignoring the frictions and their implications for feasible policies leads to extreme results. For example in Correia, Nicolini, and Teles (2008), an optimal allocation can be achieved in a variety of ways — including by command. Hence, in particular, money is not essential. Is it surprising, then, that there are policies that achieve an optimal allocation? Frictions are also ignored in getting the equivalence (Modigliani–Miller) results in Wallace (1981) and Sargent and Smith (1987). In those models, people can commit to future actions and there is no private information. It is doubtful that such results, and related results like the equivalence between open-market operations and money creation achieved by way of lump-sum transfers, would hold in the presence of frictions that make money essential.

The best that can be said about approaches that ignore the frictions that give monetary trade a role and the implied connections to feasible policies is that they rest on the view that the unmodeled features that give monetary trade a role have no implications for feasible policies. Such a view seems inconsistent with the kinds of frictions previously listed that have been shown to give monetary trade a role. It also seems inconsistent with pervasive observations. Consider currency. Despite claims to the contrary, it is the best analog of money in most existing models because currency is *the* outside asset that does not bear explicit interest. We know that currency is widely used in what we label the underground economy. Underground activities are those that are difficult to monitor and, therefore, difficult to tax. Hence, there seems to be a close connection between frictions that give currency a role and feasible taxes.²

I begin this chapter by briefly discussing the three frictions: imperfect monitoring, costly connections among people, and imperfect recognizability of assets. Then, I turn to a specific illustrative model and use it to consider how close we can get to explaining as an optimum the following features of actual economies: currency is a uniform object, currency is (usually) dominated in rate of return, some transactions are accomplished using currency and others are accomplished in other ways.

2. SOME FRICTIONS

If money is to be essential, then we need to stay away from the Arrow–Debreu model and its second welfare theorem. That is easy enough: competitive trade is not a mechanism and the Arrow–Debreu model assumes that people can commit to future actions. I assume that trade is accomplished through a mechanism and that people cannot commit to future actions. We also need to stay away from folk-theorem results. This is accomplished by assuming sufficient discounting, a sufficiently large number of agents, and imperfect monitoring.

² Despite that, no applied work on issues like the welfare cost of inflation takes into account the connection between currency usage and activities that are difficult to monitor and tax.

2.1 Imperfect monitoring

The ancient absence-of-double-coincidence suggestion is incomplete in at least one important sense. Does it apply if the two people being described are part of a small isolated community such as a small kibbutz, a small Amish community, or a family? It seems as if the two people are meant to be strangers. One of the first discussions of the sense in which they are meant to be strangers is by Ostroy (1973). He suggests that money is a substitute for knowledge of previous actions. The modern term for describing what is known about previous actions is monitoring: perfect monitoring means common knowledge of all previous actions; imperfect monitoring means anything else. Townsend (1989) use imperfect monitoring to motivate the use of money in an explicit intertemporal model, and Kocherlakota (1998) combines it with no commitment. Given no commitment, which I maintain throughout, the crucial proposition implicit in this work is that *imperfect monitoring* is necessary for money to be essential.

A proof of such necessity would proceed by contradiction. Suppose there is perfect monitoring and that there is an implementable allocation that makes use of fiat money, an intrinsically useless object. Perfect monitoring means that previous actions are common knowledge. So suppose that some initial condition, which includes the distribution of money holdings, and previous actions determine the evolution of actions and holdings of money. In other words, there is a composite mapping from previous actions to current actions, composite in the sense that an intermediate stage involves money holdings and transfers of money among people. Now, consider the implied direct mapping from previous actions to current actions without the use of money. The claim is that *implementability* of the actions implied by the composite mapping implies implementability of the same actions using the direct mapping. Hence, money is not essential.

The above sketch of a proof uses fiat money rather than commodity money. Fiat money is convenient because the alternative mechanism that uses the direct mapping can simply ignore the fiat money — can treat it as worthless. This could not be done with commodity money. And, with commodity money, it is not easy to distinguish between monetary trade and non-monetary trade. Indeed, the advantage of using fiat money in the argument is similar to the advantage of using it in the quantity theory of money and its neutrality proposition; something that was done by Hume (1752) and others even when actual money was a commodity.

The necessity claim is supposed to apply to any model and, in particular, to models with private information about types. And, there is no assumption about discounting. No commitment and discounting can help determine the conditions for implementability, which can always be stated in terms of actions that do not involve fiat money.

Why might money help if there is imperfect monitoring? If the people that a person will meet in the future do not directly observe what is done today, then it may help for the person to collect some evidence that can subsequently be shown. That is, acquiring money today can *weaken* the person's future truth-telling constraints about today's actions. If we

think of fiat money as a physical and durable object like currency, then, counterfeiting aside, it can serve that role. Others can say “show me” if the person tries to overstate holdings of it.

The necessity claim implies that one route to a cashless economy is better and better monitoring. But better monitoring is not the only route to a cashless economy. More generally, while the claim asserts that imperfect monitoring is necessary for monetary trade to be essential, it says nothing about sufficient conditions. It does suggest that no monitoring at all — each person’s previous actions are private information to the person — offers the best shot at making money essential. However, if we want a setting in which some form of credit exists, then no monitoring is too extreme.

Credit of any sort requires some monitoring in the sense that someone has to observe that a person has borrowed. Therefore, if we want both monetary trade and credit in the same model, we need something between perfect monitoring and no monitoring. As in other areas of economics — for example, transport costs in international-trade theory — extreme versions are both easy to describe and easy to analyze. The challenge is to specify and analyze intermediate situations.

2.2 Costly connections among people

Absence-of-double-coincidence has almost always been described in terms of meetings between two people. This description has led to a large literature in which it is assumed that people meet in pairs. Any such model should be interpreted as one in which connections among people are costly. Models of pairwise meetings in discrete time assume that one pairwise meeting per period is free and that all others are infinitely costly. Models with pairwise meetings at random, one of which I will use below, assume that the free meeting is determined randomly. Any such model is very different from having everyone together or at least connected as in the Arrow-Debreu model.

It is evident that pairwise meetings were originally invoked as a way to limit the role of quid pro quo or spot trade in commodities. However, pairwise meetings are not necessary for there to be a role for intertemporal trade. All we need is a potential role for credit and frictions that inhibit credit (see, for example, Levine, 1990).

So why bother with models of pairwise meetings? One reason for studying these models is that such meetings can provide a rationale for imperfect monitoring. In a large economy, if people meet in pairs and, therefore, know only what they have experienced or what they have been told by people they meet, then imperfect monitoring emerges as an implication. This point of view is explored in Kocherlakota (1998) and Araujo (2004). Also, models of pairwise meetings are attractive settings for exploring issues like counterfeiting (see Nosal & Wallace, 2007), imperfect divisibility of money (see Lee, Wallace, & Zhu, 2005), and float (see Wallace & Zhu, 2007).

However, models of pairwise meetings come with complications. One is the wide range of equilibrium concepts used to answer the old question: What do a pair who meet to trade do? One approach taken in the literature is descriptive; for example,

the buyer and seller make alternating offers, buyers make take-it-or-leave-it offers, or sellers commit to posted prices. Another approach explores all implementable outcomes subject either to individual defection or such defection and cooperative defection by the pair in the meeting. In keeping with the spirit of mechanism-design analysis, I will, for the most part, adopt the second approach.

2.3 Imperfect recognizability

Recognizability has often appeared as one among a list of desirable properties of a medium of exchange. Settings with imperfect recognizability are usually modeled by supposing that the current holder of an object knows more about its qualities than a potential acquirer of it. I will suggest that such asymmetric information is one explanation for our seeming preference for uniform currency. However, because my discussion of imperfect recognizability is far from complete, I start by assuming perfect recognizability.

3. AN ILLUSTRATIVE MODEL WITH PERFECT RECOGNIZABILITY

Central banks in the UK, the United States, and several other countries emerged as legally mandated monopoly issuers of banknotes from systems in which there were many private banks issuing banknotes. In an attempt to model and compare the latter (which I call an inside- or private-money system) to the former (which I call an outside-money system), Cavalcanti and Wallace (1999) use a model with an extreme form of imperfect monitoring: an exogenous fraction of people are perfectly monitored (the potential issuers of private money) and the rest are not monitored at all. Indeed, the rest are assumed to be anonymous. In the next section I set out that model more generally than has been previously done and prove some simple results about implementable allocations in it.

3.1 The model

The background environment is an elaboration of that seen in Shi (1995) and Trejos and Wright (1995). Time is discrete. There is a nonatomic and unit measure of infinitely lived people. Preferences are additively separable over dates, and each person maximizes expected discounted utility with discount factor $\delta \in (0, 1)$. Period utility is $u(x) - c(y)$, where $x \in \mathbb{R}_+$ is consumption and $y \in \mathbb{R}_+$ is production. The functions u and c are strictly increasing and differentiable with u strictly concave, c convex, $c(0) = u(0) = 0$, and are such that there exists $\tilde{y} > 0$ that satisfies $c(\tilde{y}) = u(\tilde{y})$. In addition, there are no intertemporal technologies (production is perishable).

The set of people is partitioned initially and permanently into two sets: the fraction α are monitored (m people) and the fraction $1 - \alpha$ are not (n people), where α should be interpreted as the economy's exogenous monitoring capacity. The history of each m person is common knowledge, while that of each n person is private to the person. (It is as if each m person wears a computer chip that transmits everything about the

person to everyone else.) The only thing known about an n person is the person's producer–consumer status in a meeting and that the person is not an m person.

To allow a discussion of inside money, each person has a printing press capable of turning out identical, divisible, and durable objects. Those turned out by the printing press of any one person are, however, distinguishable from those turned out by other peoples' printing presses. This is the perfect recognizability assumption.

There are two stages at each date. Stage 1 has pairwise meetings at random: a person is a producer (seller) at each date with probability θ , a consumer (buyer) with probability θ , and is neither (meets no one) with probability $1 - 2\theta$, where $\theta \leq 1/2$. Any production and consumption necessarily occurs at stage 1 and no one ever both consumes and produces at the same date.³ Stage 2 has a centralized meeting that can be used for transfers of money among agents. It is intended to be the model's analog of a clearing house, a federal funds market, or a commercial paper market. Because there are no goods at stage 2, there are no separate stage 2 preferences.

One benchmark allocation in the previous model is production (and consumption) equal to $\arg \max_x [u(x) - c(x)]$, denoted x^* , in every (single-coincidence) meeting. According to a representative-agent welfare criterion that views people as identical before being assigned type, m or n , initial money holdings, and histories, that allocation is the first-best allocation — first best in the sense of best subject only to the pairwise structure.

One convenient feature of the this setting is the simple description of the first best. Notice that first-best actions are the same at every date: produce x^* whenever you are a producer in a meeting and consume x^* whenever you are a consumer in a meeting. As might be expected, the difficulty is getting the producer to produce x^* . One possible difficulty arises solely from discounting and is present even if everyone is an m person. But, as noted above, money cannot help if everyone is an m person. The presence of n people gives money a role. However, as we will see, money is necessarily accompanied by history-dependent actions, and, hence, a departure from the first best.

3.2 A class of allocations

Although richer classes of allocations could be considered, I limit allocations to those in which all monies issued by m people who have not defected (and any initial money) and money issued by the planner are treated as perfect substitutes and all monies issued by n people are worthless. (Hence, I simply assume that n people do not issue money.) Therefore, a person's state at the beginning of date t prior to pairwise meetings is an element in the set $\mathbb{S}_t = (\mathbb{I} \times H^{t-1} \times \mathbb{R}_+)$, where $\mathbb{I} = \{m, n\}$ is the person's type, and H^{t-1} is the set of possible histories starting from the initial date, $t = 0$, up through

³ As is well-known, one underlying setting is a K -good, K -type setting in which there is an equal measure of each type. A type- k person consumes only type- k good and produces only type- $(k + 1)$ good for $k \in \{1, 2, \dots, K\}$, where addition is modulo K . If $K > 2$, then we get the model in the text with $\theta = 1/K$.

date $t - 1$. A history for a person describes who was met in the past in pairwise meetings and includes the state of each meeting partner. A generic element of $\mathbb{S}_t$ is denoted $s_t = (i, h^{t-1}, z)$, where z is holdings of money issued by others (other monitored people or the planner). If $i = m$, then s_t is common knowledge. If not, then (h^{t-1}, z) is private information. In particular, an n person can hide money. The post-meeting state of a person is the same kind of object except that it includes what happened at stage 1.

Given a starting distribution of people over states, an allocation is a sequence that describes what happens in meetings at stage 1 and at stage 2 as a function of the states of people. The state of a date- t pairwise meeting is $(s_t^p, s_t^c) \in \mathbb{S}_t^2$, where the first component describes the producer and the second the consumer as they enter the meeting. In a pairwise meeting, the actions are some amount of production and consumption and state transitions for the two people. At stage 2, the only action is a state transition. At both stages, it is convenient to allow for randomization so that there can be a distribution of actions at stage 1 at a given date for the same kind of meeting.

An allocation describes what happens in the economy in the following sense. Given the initial distribution over $\mathbb{S}_0$ and the assumption that meetings occur at random, the date-0 actions imply a date-1 distribution over $\mathbb{S}_1$, and so on.

3.3 Incentive-feasible allocations

There are two kinds of constraints on allocations: physical feasibility restrictions and incentive constraints (IC). One physical constraint is that consumption in a meeting is bounded above by production in the meeting. Also, in a meeting between two n people, people who by assumption do not issue money, total end-of-trade money holdings cannot exceed total pre-trade money holdings. The transitions at stage 2 permit transfers of money to and from the planner.

Regarding ICs, I can allow either of two kinds of Nash implementation: one requires that the allocation be immune to individual defection and the other requires that it be immune to both individual and cooperative pairwise defection of those in a pairwise meeting.⁴ Nash means that each person or pair takes a given no-defection by everyone else. Common to both notions are the following assumed punishments. Defection by an n person has no future consequences for the person except those implied by the current trade to which the person defects. Defection by an m person is common knowledge and is assumed to be punished by permanent expulsion from the set of m people to the set of n people starting at the next stage.

⁴ Throughout I use weak implementability of allocations in the sense that I require that an allocation be the outcome of some equilibrium. In particular, lurking in the background of what I do is always an equilibrium in which all money is ignored. I do not deal with ruling out such equilibria (see Aiyagari & Wallace, 1997 and Wallace & Zhu, 2004 for attempts to do that). I also leave implicit the game that supports the outcomes. See Zhu (2008) for an explicit definition of implementability that can be used to support the outcomes associated with either the individual or cooperative defection versions and Hu, Kennan, and Wallace (2009) for an application of the cooperative defection game.

Such exclusion may seem like a weak punishment. One alternative would be economy-wide reversion to autarky as a response to a defection. That would not be best if there were a small probability of errors in actions. And, even without such errors, it would not be time-consistent for the society. If economy-wide or even positive-measure punishments are not imposed, then the assumed punishment can be justified by assuming that there is free exit at any time from the set of m people to the set of n people. Even if that were not assumed, it would be delicate to impose stronger individual punishments. Even if an m person is a known defector, n people would generally want to trade with that person.

Given the structure of the model, an individual defection is always to no trade at the current stage: in a pairwise meeting, it is zero production and consumption and an unchanged holding of money; at stage 2, it is no transfer of money. If the defector is an n person, then there are no further consequences. If the defector is an m person, then that person begins the next stage as an n person with the money held and with a useless printing press – useless because the defection is assumed to make that person's money worthless.

Regarding cooperative defections in pairwise meetings, there are three kinds of meetings. In a meeting between two m people, there is no private information. Any cooperative defection has both people becoming n people at the next date with both monies issued by those people worthless. Hence, in any defection their total money holdings are limited by the money holdings they bring into the meeting and two m people cannot make each other rich by issuing money to each other. The restriction implied by the possibility of cooperative defection is that their payoffs (the profile of the current utility payoff plus the discounted continuation value for the producer and the consumer) must be weakly outside the payoff frontier of a meeting between two n people with the same profile of money holdings. In a meeting between an n person and an m person, there is one-sided asymmetric information. Again, any cooperative defection has the m person becoming an n person at the next date. Both for those meetings and for meetings between two n people, a full analysis requires that some notion of the core under asymmetric information must be adopted. (When two n people meet, there is two-sided asymmetric information if only because both the producer and the consumer can hide money.)

The results presented next, which are not existence results, do not depend on which notion is adopted. In particular, the arguments take as given the trades and payoffs of n people.

3.4 Results

There are three simple results about the set of IC allocations. The first is that more monitoring is better.

Claim 1 In terms of production and consumption, the set of IC allocations is weakly increasing in the fraction who are monitored.

Proof. Consider two economies, economy 1 and economy 2, that are identical except for α : let $\alpha_2 > \alpha_1$. If the allocation A_1 is IC for economy 1, then there exists

A_2 that is IC for economy 2 and has the same production and consumption. The allocation A_2 is constructed by having the additional monitored people behave exactly as do the non-monitored people under A_1 that they “replace.” In other words, in economy 2, select at random a fraction $(\alpha_2 - \alpha_1)/\alpha_2$ of the m people and give them a special starting history, a label, and have them behave exactly as n people do in A_1 . Have everyone else behave as they do in A_1 . Then because A_1 is IC in economy 1, it follows that A_2 is IC in economy 2. In other words, having an m person behave like an n person is always IC because defection of any sort is always to n status. ■

The next claim says that allocations can be limited to those in which m people enter stage 1 without money —with only their printing presses. In general, m people acquire money in pairwise meetings when they produce for n people. Therefore, such an allocation calls for them to immediately destroy any money received or, equivalently, turn it in to the planner at the next stage 2.⁵ A consequence is that any spending by an m person in a meeting involves the issue of that person’s money. This result uses the restriction that the only allocations I consider are those in which all monies issued by monitored nondefectors are perfect substitutes.

Claim 2 If an allocation is IC, then there is another IC allocation with the same production and consumption in which monitored people enter stage 1 without money.

Proof. Consider an arbitrary IC allocation in which some m person enters a pairwise meeting with some money at some date. Consider an alternative that is identical except that this person has turned in that money at the previous stage 2, but keeps spending unchanged by issuing the person’s own money instead of spending the money issued by others. Because all monies are perfect substitutes, trading partners are not affected, and, therefore, no-defection payoffs are not affected. What about defection payoffs? A consequence of the ability of n people to hide money is that the discounted utility of an n person is weakly increasing in money holdings. That implies that the defection payoffs implied by the alternative are no larger than those of the given arbitrary allocation. Hence, the alternative is also IC. ■

Notice that the converse of this claim does not hold. Start with an allocation in which m people hold no money and consider an alternative that differs only because at some date an m person has not turned in the money received earlier. Does willingness to turn in the money imply that the alternative is IC? It does not. The money is turned in prior to the next meeting (before the next stage 1 meeting realization occurs). It is based on an expected value over such realizations and the defection realizations. But that implied inequality does not imply no defection in each subsequent stage 1 meeting realization.

Why have money transferred to an m person in a pairwise meeting if the person will simply turn it in? If the person making the transfer is an n person, then the transfer

⁵ If money were costly to produce, then it would be wasteful to destroy it. In that case, stage 2 could be used as a kind of clearing stage during which m people turn in other issuers’ monies and receive their own money for subsequent use.

provides an additional inducement for that person to have acquired money in the past. Also, if an allocation is to have m people issue money when they are consumers in meetings with n people, then unless they collect money from n people when they are producers in meetings with n people, holdings of money by n people would be growing. That would necessarily produce the model's analog of inflation.

If the person making the transfer is another m person, then the transfer plays no role and can be eliminated. If it is eliminated, then an outside observer would see production and consumption occur without any transfer of money. That is the model's version of a credit transaction. It follows, in accord with the necessity of imperfect monitoring, that if everyone is monitored, then money is not needed. It also follows that any production by an m person — whether for another m person or for an n person — is supported entirely by threatened expulsion from the set of m people.

Claim 3 If not everyone is monitored, then the first-best is not IC.

Proof. Suppose it is IC and consider two mutually exhaustive possibilities. Either the support of the distribution of money holdings across n people at some date prior to stage 1 contains two different holdings, $m_1 < m_2$, and money is valuable in the sense that the discounted value of the holding m_2 exceeds that of the holding m_1 , or not. The former — a nondegenerate distribution and valuable money — contradicts the first-best because the first-best implies that everyone has the same discounted utility prior to pairwise meetings at every date. If the latter, then discounted utilities are degenerate at every date either because there is a degenerate distribution of money holdings or because all holdings in the support have equal discounted value. However, if this holds at some date, then the first-best actions imply that it does not hold at the next date. In particular, those n people who are supposed to produce x^* in a pairwise meeting must see a future reward from doing so or they defect to no trade. But, for them, that future reward can only take the form of higher money holdings prior to stage 1 at the next date — to which is attached higher discounted utility. Hence, the degeneracy and first-best actions cannot hold at every date. ■

Before I go on to discuss the consequences of imperfect recognizability, several comments about the previous model should be made. First, the assumption that n people do not issue money seems innocuous because I permit the planner to make positive transfers of money at stage 2 to n people. Such transfers — perhaps, sprinkled in a random way among n people — would seem to be a good substitute for making the money issued by a subset of the n people acceptable.

Second, no mention has been made of a commonly studied intervention in models of money — the use of taxes to finance the payment of interest on money — either explicitly or through deflation. A special case is real interest that exactly offsets discounting, which is called the Friedman rule. Such schemes do not have to be considered separately because they are included in the above class of allocations. For example, a deflation can be produced by having money in the hands of n people decline over time. Although that cannot be achieved by an explicit tax on n people, it can be

achieved in other ways. One way is by having m people issue less money when they are consumers in meetings with n people than they collect and destroy when they are producers in meetings with n people. Another way is having m people consume less per unit of money transferred in meetings with n people than they produce per unit of money received in meetings with n people.

Whether such schemes are IC and optimal cannot be addressed without imposing additional structure on the model. However, even at this level of generality, any such analysis seems very different from the analysis of deflation or paying interest on money in representative-agent models. The financing of any such scheme has to come from taxes on m people. Such taxes are scarce because m people can defect and because good allocations have m people giving gifts to n people — gifts that are not reciprocated. In addition, the dependence of an n person's current ability to spend on recent realizations gives rise to a risk-sharing role for transfers to n people, even if those transfers cannot be contingent on their wealth, which is private information (see Deviatov, 2006; Green & Zhou, 2005; Levine, 1990).

Finally, although the model was introduced to contrast inside and outside money, so far nothing has been said about that. Outside money is the special case in which no one but the planner issues money. The restriction that no one issue money is IC because if money-issue is a defection, then that money becomes worthless at the next date and, therefore, is worthless when issued. However, because outside money is a special case with additional restrictions, imposing it in the above setting cannot help.

Does the restriction hurt? Without imposing additional structure, I cannot demonstrate that it hurts. But I can describe why it might hurt. Under outside money, the spending of m people seems to be tied to their individual histories (as it necessarily is for n people). However, the introduction of stage 2 goes some way toward removing that dependence. In particular, stage 2 can be used for transfers among the m people (something like borrowing and lending among themselves or, more precisely, insurance among them) and there can be transfers to and from the planner — all of which are subject to defection constraints. However, those defection constraints tend to be tighter under outside money because the result in claim 2 is lost; namely, that m people enter pairwise meetings without money. According to the model, that is why imposing outside money might hurt.

4. IMPERFECT RECOGNIZABILITY AND UNIFORM CURRENCY

Despite the benefit of private currencies just identified, we almost always observe uniform currencies. There are many possible reasons. One that potentially fits within our mechanism-design framework is recognizability problems with many distinct currencies. Such problems could take a variety of forms. Here, I consider the threat of counterfeiting.

In the context of the earlier model, suppose some n people have a costly counterfeiting technology. At any stage 2, they can produce counterfeits subject to a positive

fixed cost and a constant marginal cost. In stage 1 meetings, suppose producers cannot distinguish between genuine currency and counterfeits until after they acquire the currency. Then they learn whether they have acquired genuine currency or counterfeits.

There are two conceivable kinds of allocations in these circumstances. In one, counterfeits are produced and known counterfeits and genuine currency are perfect substitutes. Even if this kind of allocation is implementable, it has obvious welfare shortcomings. Aside from the costs of counterfeiting, it is identical to one without counterfeiting, but in which the genuine currencies of the potential counterfeiters are treated as perfect substitutes with other currencies. In such an allocation, those n people never produce and they issue currency period after period, imposing costs on others.

The other kind of allocation is one in which known counterfeits are less valuable than genuine currency. Here, there is asymmetric information between the producer and the consumer in a pairwise meeting: the consumer knows whether he or she has genuine or counterfeit currency and the producer does not. Any such allocation is either a pooling allocation or a separating allocation. A separating allocation in which counterfeiting actually occurs hardly fits our notion of counterfeiting, because producers end up accepting known counterfeits. Hence, most analyses focus on pooling allocations.

However, because there is no standard notion of the core under asymmetric information, all existing analyses adopt a particular game form in these situations. The most common is a signaling-game framework in which buyers make take-it-or-leave-it offers. In the context of such a game, Nosal and Wallace (2007) showed that imposition of the Cho-Kreps intuitive criterion rules out pooling with counterfeiting. The deviating offer that destroys a pooling equilibrium has the consumer with genuine currency offering a smaller trade—less currency for less output—and has the producer inferring from this offer that the consumer has genuine currency.

Given that no counterfeiting occurs in equilibrium, what are the possibilities for equilibria? That depends on other aspects of beliefs about out-of-equilibrium actions. Nosal and Wallace (2007) implicitly assumed if there is no counterfeiting in equilibrium, then any offer of currency at stage 1 is an offer of genuine currency. They, therefore, conclude that an equilibrium in which genuine currency is valuable and no counterfeiting occurs exists only if counterfeiting is more costly than the value of currency in the absence of a counterfeiting threat. Otherwise, the only equilibrium is autarky. However, as pointed out by Li and Rocheteau (2009), another out-of-equilibrium belief is possible. They consider an equilibrium in which the value of genuine currency in trades is small enough to make counterfeiting unprofitable. That equilibrium is supported by the belief that an out-of-equilibrium offer of any larger amount of currency comes from a counterfeiter.

Both analyses conclude that counterfeiting is a serious threat. Hence, an implication of their analyses is that there should be sufficient enforcement to make counterfeiting a very costly activity. That implication takes us in the direction of a single uniform currency if we assume, as seems plausible, that the prevention of counterfeiting is much easier with a single uniform currency than with many distinct private currencies.⁶

5. OPTIMA UNDER A UNIFORM OUTSIDE CURRENCY

As suggested above, under outside money, stage 2 in the illustrative model takes on added importance. First, it could be desirable for there to be transfers of currency among m people — transfers that accomplish insurance. In particular, it may be desirable and incentive feasible to have those m people who recently earned outside money make transfers to those who recently spent such money. Second, the planner who controls outside money might participate in the transfer scheme. To get a sense about what optima might look like under outside money, Deviatov and Wallace (2009) study an example of the illustrative model in which there is an exogenous two-date periodic productivity process, a deterministic seasonal.

In their example, the discount factor is 0.95, the utility of consuming, $u(x)$, is $2x^{1/2}$, and the disutility of producing, $c(x)$, is $x/0.8$ when t is odd (low productivity dates) and $x/(1.2)$ when t is even (high productivity dates). Therefore, $(0.8)^2$ is the first-best output at low productivity dates and $(1.2)^2$ at high productivity dates. Each person is a producer with probability $1/3$ and is a consumer with probability $1/3$, and monitored people are one-quarter of the population. Aside from the discount factor being sufficiently high in a sense to be described, this specification is arbitrary.

For this example, Deviatov and Wallace (2009) compute the maximum of ex ante representative-agent utility, prior to the assignment into types, monitored or non-monitored, and prior to the assignment of initial currency holdings — the distribution of which is treated as among the choice variables of the planner. The constraints are individual defection in both stages, and, in addition, pairwise cooperative defection at stage 1 pairwise meetings. They simplify the problem in three important respects. First, they search only over two-date periodic allocations. Second, currency is indivisible and holdings prior to stage 1 are restricted to be in $\{0, 1\}$. Third, while lotteries are allowed (and play a role), randomization is not allowed. They study and compare two versions of this problem: in one, called no-intervention, the quantity of currency is constrained to be constant; in the other, it, like everything else, is permitted to be two-date periodic.

⁶ The implication that counterfeiting does not occur may seem inconsistent with observed counterfeiting. However, some judgment has to be used. In the United States, it is estimated that one in ten thousand dollars are counterfeits (see Judson & Porter, 2003). That is so close to zero that in a pooling equilibrium with that proportion of counterfeits, it would not be worthwhile to think about a deviating offer.

Constraining currency holdings to $\{0, 1\}$ is an extreme assumption, but is not misleading. As noted above, the economic problem in this model is to free current actions from previous realizations. The restriction to $\{0, 1\}$ money holdings exacerbates this problem, but does not change its nature. Regarding the discount factor, it is high in two senses, each related to providing simple benchmark allocations. First, it is high enough so that if everyone were monitored, then the first-best allocation, one with first-best outputs in every meeting, would be implementable. Second, the discount factor is high enough so that the best allocation subject to treating everyone like an n person, which is an implementable allocation, has half the agents with a unit of currency and has first-best outputs in one-quarter of the meetings. Moreover, in that allocation, the intervention and nonintervention versions are identical. Ex ante utility for that allocation is equal to one-quarter of the first best, one quarter because trade in a pairwise meeting requires that the potential consumer have money and that the potential producer not have money.

Some features of the optimum are common whether or not there is intervention. First, the discounted utility of an m person is roughly $2/3$ of the first best, while that of an n person prior to the assignment of money is roughly $1/3$ of it. (Thus, everyone benefits relative to an allocation in which all people are treated as n people or relative to the same economy with no m people.) Second, there are no transfers of currency to n people.⁷ Third, all m people enter stage 1 with a unit of currency and the constraint that they not defect when called on to produce is binding in many meetings. It follows that production by m people is supported entirely by threatened expulsion from the set of m people, as happens generally under inside money. As noted previously, the threatened punishment under inside money would be greater, and that would allow better allocations to be achieved under perfect recognizability.

Intervention in the example looks like the granting of zero interest loans to the aggregate of m people at stage 2 following pairwise meetings at the high-productivity date, with repayment at stage 2 following the next low productivity meetings. In terms of outputs, intervention shows up mainly in meetings between m and n people. Intervention frees the distribution of money holdings between m and n people from the constraint that net currency trades between the two groups are zero at every date. A consequence is that intervention has smoother outputs in those meetings, smoother relative to the respective first-best outputs. Perhaps, the simplest way to describe the role of intervention is that loans at stage 2 after the high-productivity date restore the currency holdings of m people, permitting higher spending by them in the pairwise meetings with n people at the high-productivity date. Obviously, the distinction between m and n people and the fact that they interact are crucial features of the example.

⁷ This finding may depend on $\{0, 1\}$ money holdings. The possibility that transfers to n people may change the distribution of money holdings in a favorable way (see Deviatov, 2006) requires a richer set of individual holdings.

6. EXTENSIONS OF THE ILLUSTRATIVE MODEL

The illustrative model is extreme in many respects. In this section, I comment on some directions in which it could be generalized.

6.1 Capital

An unrealistic and, therefore, seemingly troublesome feature of the illustrative model is the absence of forms of wealth other than currency. Indeed, in the inside-money version, if money is treated as a liability of m people, then net wealth is zero. Here I describe a way to remedy that unrealistic feature by introducing putty-clay capital into the model.

In the previous version, the production technology in pairwise meetings has a single input, labor or effort. As has been recognized by others, that technology could be amended so that output produced in a meeting is a function of the producer's labor and *capital*, say $f(k, l)$, where l is the person's effort, k is the person's capital at a date, and where, f could be assumed to be standard; for example, linearly homogeneous and strictly quasi-concave. As in the model without capital, the period utility for a person is $u(y) - c(l)$, where y denotes consumption.

The crucial assumptions concern the law of motion for a person's capital. One specification is that a person's gross investment good and consumption good are the same object. If so, then in a meeting with a producer at a date, the usual putty restriction on non-negative consumption and non-negative gross investment would hold: the sum of the two is bounded by the output acquired from the producer in the meeting. The clay aspect of capital is that existing capital cannot be transferred to another person.

This generalization of the model does not change what is traded at either stage. If there were only n people, then the model would be unchanged except that the state space would be richer. Now each person would be characterized by a portfolio, money and capital, although capital could not be traded. If there were only m people, then the implication that money is not needed is unaffected. Even with a mixture of m and n people, the only difference is the richer state space. Whether under inside money or outside money, there is, perhaps, a greater potential role for insurance. But n people are no better able to participate, while m people have essentially the same defection constraints on their participation as they have in the version without capital. In particular, the existence of capital of the above sort would not seem to enlarge the risk-sharing possibilities for either type of person.

One plausible conjecture about this version is that an optimum would display less dispersion in capital holdings across m people than among n people. In other words, capital would be more efficiently distributed among m people than among n people.

6.2 Endogenous monitored status

The illustrative model has an exogenous fraction who are perfectly monitored (m people). That exogeneity assumption can be reconciled with one-time free entry into m status if each person makes such a choice subject to a one-time, additively separable utility cost that is distributed in a very special way across people. Let $F: R_+ \rightarrow [0, 1]$, where $F(k)$ is the fraction of people who can become permanently monitored with a one-time additively separable utility cost no greater than k . For the illustrative model, F takes the special form,

$$F_x(k) = \begin{cases} \alpha & \text{if } k < K \\ 1 & \text{if } k \geq K \end{cases}, \quad (1)$$

where K is so large that a person with cost K would never choose to be an m person for any incentive-feasible allocation. (Notice that for the fraction α of the population, there is a zero cost of becoming monitored, where zero is the payoff from autarky.)

To allow for initial free entry into m status, the sequence of actions at the initial date is as follows. People are ex ante identical. Then the planner announces an allocation, including initial distributions of money dependent on m or n status. Then, in accord with the distribution function F , each person privately learns the cost of becoming an m person and chooses whether to become an m person. Then, initial money holdings are distributed. If the allocation satisfies the incentive constraints, which include the restriction that those with low enough utility costs of becoming an m person choose m status and that the rest do not, then trades are undertaken in accord with the planner's suggestion.

Let $v_t^j(x)$ be the discounted expected utility of a type $j \in \{m, n\}$ at the beginning of date t of someone who holds x amount of money, the only asset. The assumption that there is free-exit from m -status implies that one of the constraints on the planner is $v_t^m(x) \geq v_t^n(x)$ for all (t, x) . This, in turn, implies $v_0^m \geq v_0^n$, where $v_0^j = E_x v_0^j(x)$ and E_x denotes expectation taken over the relevant initial distribution of money. For the previous model in which F is given by Eq. (1), the restriction that people choose m status as intended by the planner is nothing more than $K \geq v_0^m \geq v_0^n \geq 0$; that is, nothing more than is implied by free exit from m status at any time.

However, initial free-entry imposes additional constraints if F is continuous and strictly increasing, and if a person's realization from F is private information. (For the F in Eq. 1, private information about the realization plays no role.) Under such assumptions, initial money holdings can be assigned based only on whether the person chooses to become monitored; and $v_0^m - v_0^n = \hat{k}$, where $F(\hat{k})$ is the fraction who choose m -status.

In such a version, the shape of F will play a role in determining the optimum allocation. That is, the planner will, in effect, be concerned about how the allocation affects the fraction who choose to become monitored. I suspect that a smooth F contributes to making the taxation of m people scarce.

6.3 Other information structures and other financial instruments

When compared to the actual economy, the illustrative model is deficient in terms of the limited financial instruments that appear in it. Indeed, whether under inside money or outside money, the only financial instrument seems to be something like a uniform currency. Mainly because of the extreme monitoring assumptions, instruments like checking accounts, debit cards, credit cards, cell phone money, and bills of exchange are either not feasible or not needed.

All such instruments are supported by an informational network. By assumption, n people cannot be part of any such network. As for m people, they are part of a perfect and costless network that reveals their actions to everyone. Hence, other financial instruments that potentially convey information about m people are not needed. There is, however, a caveat. I have imposed that currency, whether inside or outside, is treated as a uniform object in equilibrium. Conceivably, there might be a role for distinct objects to be held by n people.

The set of implementable allocations could conceivably be enriched by having different assets with different rates of return available to n people. In Bryant and Wallace (1984), people are faced with nonlinear and increasing returns on saving among which they self-select. Such nonlinearity can enrich the set of implementable allocations. In Kocherlakota (2003), facing people with some assets that can be traded for goods and others with higher returns that can only be saved enlarges the set of implementable allocations in a beneficial way. But those analyses are silent about what allows nonlinear returns to be implementable. Thus, for example, Kocherlakota (2003) simply assume that his higher return assets, called bonds, cannot be traded for goods directly. In Bryant and Wallace (1984), the bonds are explicitly indivisible and large, but nothing is said about why they cannot be shared or intermediated.

The illustrative model can help in that regard. To fix ideas, suppose the planner at stage 2 sells high-denomination, one-period, payable-to-the-bearer bonds intended to be bought by some n people. Their high denomination limits their use in trade in pairwise meetings among n people. However, if the bonds dominate currency in rate of return across stage 2 at adjoining dates, then they give rise to a profitable arbitrage opportunity: hold the bonds as assets and issue small-denomination, one-period payable-to-the-bearer claims. But who could engage in such intermediation? It involves promises, which, in the illustrative model, can only be made by m people. But the actions of m people are public. Hence, if the planner wants to prevent such intermediation, then it can be prevented. Put somewhat differently, a legal restriction against such intermediation is easy to enforce in the illustrative model. Indeed, for closely related reasons, it would seem to be easy to enforce in the actual economy.⁸

⁸ A different route to coexistence of currency and higher return assets is pursued in Zhu and Wallace (2007). In a model with only n people and with observed portfolios in meetings, they show that an extraneous property like color can be used to produce implementable allocations that are immune to cooperative defection and are consistent with different rates of return.

Thus, the illustrative model is one example in which assumptions about the information structure have implications for the kinds of financial instruments that might exist and that might play a beneficial role.

6.4 Production and consumption at the centralized stage

There is a substantial literature that resembles the illustrative model but has production and consumption at stage 2 when everyone is together. Most of it follows Lagos and Wright (2005) in assuming that everyone has identical and quasi-linear preferences at stage 2. A general version with production and consumption at stage 2 would qualitatively resemble the illustrative model, but the version with identical and quasi-linear preferences at stage 2 does not.

The essential features of the Lagos-Wright model are most easily seen in a version with identical and linear preferences at stage 2: there is a stage 2 perishable good for which preferences are identical, additively separable, and linear. Positive consumption of this good is interpreted as consumption and negative consumption is interpreted as production of it. Trade at stage 2 is modeled as competitive price-taking trade. (In a mechanism-design version, if group defection is permitted, then such trade is without loss of generality because it is equivalent to the static core at stage 2.) The insight of Lagos and Wright is that the outcome of such trade absorbs all wealth differences among people entering stage 2 through differences in consumption of that good. It follows that the distribution of wealth entering stage 2 is not a state of the economy going forward from the end of stage 2. A similar role is played by the *large family* in Shi (1997).⁹ In each case, the economy starts anew with an essentially exogenous distribution of money at each date.¹⁰

While there is a huge gain in terms of tractability from the Lagos and Wright (2005) specification or Shi's (1995) large family model, insufficient attention has been paid to what is lost from those specifications. Put differently, insufficient attention has been paid to studying the robustness of conclusions to what are very special assumptions. One thing lost is the result that the first best is not implementable if there are n people (see Hu, Kennan, & Wallace, 2009). Another thing lost is the risk-sharing role of monetary transfers among m people at stage 2 and the potential risk-sharing role of positive monetary transfers to n people. In Lagos and Wright (2005), such risk-sharing is accomplished through trade in the linear good, while in Shi's (1995) large family model it is accomplished within the large family.

⁹ On the modeling of the large family model, see Zhu (2008).

¹⁰ Often, this distribution is degenerate, but not in every model. (See Galenianos & Kircher, 2008 for a model in which the crucial starting-anew property holds, but in which there is a nondegenerate distribution at the end of stage 2.)

7. CONCLUDING REMARKS

I began by defining the mechanism-design approach to monetary theory and then turned immediately to a discussion of a particular model. The model builds on the pioneering work on matching models of Kiyotaki and Wright (1989), Trejos and Wright (1995), and Shi (1995) who were the first to formulate coherent intertemporal models of trade in which people meet in pairs and use money. It also builds on the ideas about the connection between imperfect monitoring and monetary trade of Ostroy (1973), Townsend (1989), and Kocherlakota (1998).

The goal of the discussion is to explain as an optimum three features of most actual economies: currency is a uniform object; currency is (usually) dominated in rate of return; some transactions are accomplished using currency and others are accomplished in other ways. Toward that end, I first described how private money would work and its advantages under the assumption of perfect recognizability. Then, I invoked imperfect recognizability in the form of a counterfeiting threat as a disadvantage of many distinct private monies. Finally, to explain why currency is dominated in return, I invoked a connection between the main feature of the model that gives currency a role, the imperfect monitoring, and feasible forms of taxation. I suggested that the implied restrictions on taxation will in many settings imply that the optimum does not have currency earning the Friedman-rule rate of return.

The models I have described seem both special and complicated. That may be inevitable. First, models with imperfect monitoring, costly connections among people, and imperfect recognizability are unlikely to be simple. Second, monetary trade is a descriptive or positive feature of an economy. It will not be an implication of every environment that we can imagine.

Despite that, some progress has been made. First, by its very nature, the mechanism-design approach accomplishes the longstanding goal of integrating monetary economics with the rest of economics. However, it is not the integration that seemed to be the goal a century or more ago; namely, integration with the then current version of the Arrow-Debreu model. That form of integration is not the right goal. Instead, the goal is integration with the rest of economics that deals with frictions. Second, there have been new insights about puzzling observations and policy questions. Among the issues addressed in recent work are private versus government currency, the issue focused on earlier, the long-standing puzzle concerning profitability of private currency systems in the nineteenth century (see Wallace & Zhu, 2007), the denomination structure of currency (see Lee, Wallace, & Zhu, 2005), and the analysis of counterfeiting. These and other contributions illustrate the fruitfulness of studying issues in monetary economics against the background of models in which something that resembles monetary trade is the best way to achieve good outcomes.

REFERENCES

- Aiyagari, S.R., Wallace, N., 1997. Government transaction policy, the medium of exchange, and welfare. *J. Econ. Theory* 74, 1–18.
- Araujo, L., 2004. Social norms and money. *J. Monetary Econ.* 51, 241–256.
- Banerjee, A., Maskin, E., 1996. A Walrasian theory of money and barter. *Quarterly Journal of Economics* 111, 955–1005.
- Bryant, J., Wallace, N., 1984. A price discrimination analysis of monetary policy. *Review of Economic Studies* 51, 279–288.
- Cavalcanti, R., Wallace, N., 1999. Inside and outside money as alternative media of exchange. *J. Money Bank Credit* 31 (part 2), 443–457.
- Correia, I., Nicolini, J., Teles, P., 2008. Optimal fiscal and monetary policy: equivalence results. *J. Polit. Econ.* 168, 141–170.
- Deviatov, A., 2006. Money creation in a random matching model. *Topics in Macroeconomics* 6 (3), Article 5.
- Deviatov, A., Wallace, N., 2009. A model in which monetary policy is about money. *J. Monetary Econ.* 56, 283–288.
- Galenianos, M., Kircher, P., 2008. A model of money with multilateral matching. *J. Monetary Econ.* 55, 1054–1066.
- Green, E.J., Zhou, R., 2005. Money as a mechanism in a Bewley economy. *Int. Econ. Rev.* 46, 351–371.
- Hahn, F., 1973. On the foundations of monetary theory. In: Parkin, M., Nobay, A.R. (Eds.), *Essays in Modern Economics*. Barnes and Noble, New York (Chapter 13).
- Hu, T.W., Kennan, J., Wallace, N., 2009. Coalition-proof trade and the Friedman rule in the Lagos–Wright model. *J. Polit. Econ.* 117, 116–137.
- Hume, D., 1752. On money. Reprinted 1970. In: Eugene, R. (Ed.), *Writings on Economics*. University of Wisconsin Press, Madison, pp. 33–46.
- Judson, R.A., Porter, R.D., 2003. Estimating the worldwide volume of counterfeit U.S. currency: Data and extrapolation. *Federal Reserve Board Finance and Economics Discussion Series* 2003–52.
- Kiyotaki, N., Wright, R., 1989. On money as a medium of exchange. *J. Polit. Econ.* 97, 927–954.
- Kocherlakota, N., 1998. Money is memory. *J. Econ. Theory* 81, 232–251.
- Kocherlakota, N., 2003. Societal benefit of illiquid bonds. *J. Econ. Theory* 108, 179–193.
- Lagos, R., Wright, R., 2005. A unified framework for monetary theory and policy analysis. *J. Polit. Econ.* 113, 463–484.
- Lee, M., Wallace, N., Zhu, T., 2005. Modeling denomination structures. *Econometrica* 73, 949–960.
- Levine, D., 1990. Asset trading mechanisms and expansionary policy. *J. Econ. Theory* 54, 148–164.
- Li, Y., Rocheteau, G., 2009. Liquidity constraints. Manuscript. <http://www.grocheteau.com/wp.html>.
- Monroe, A.E., 1966. *Monetary Theory Before Adam Smith*. Kelley, New York.
- Nosal, E., Wallace, N., 2007. A model of (the threat of) counterfeiting. *J. Monetary Econ.* 54, 994–1001.
- Ostroy, J., 1973. The informational efficiency of monetary exchange. *Am. Econ. Rev.* 63, 597–610.
- Sargent, T., Smith, B., 1987. Irrelevance of open market operations in some economies with government currency being dominated in rate of return. *Am. Econ. Rev.* 77, 78–92.
- Shi, S., 1995. Money and prices: a model of search and bargaining. *J. Econ. Theory* 67, 467–498.
- Shi, S., 1997. A divisible search model of money. *Econometrica* 65, 75–102.
- Townsend, R.M., 1989. Currency and credit in a private information economy. *J. Polit. Econ.* 97, 1323–1344.
- Trejos, A., Wright, R., 1995. Search, bargaining, money and prices. *J. Polit. Econ.* 103, 118–141.
- Wallace, N., 1981. A Modigliani–Miller theorem for open-market operations. *Am. Econ. Rev.* 71, 267–274.
- Wallace, N., Zhu, T., 2004. A commodity money refinement in matching models. *J. Econ. Theory* 117, 246–258.
- Wallace, N., Zhu, T., 2007. Float on a note. *J. Monetary Econ.* 54, 229–246.
- Zhu, T., 2008. Equilibrium concepts in the large household model. *Theor. Econ.* 3, 257–281.
- Zhu, T., Wallace, N., 2007. Pairwise trade and coexistence of money and higher return assets. *J. Econ. Theory* 133, 524–535.

This page intentionally left blank

New Monetarist Economics: Models[☆]

Stephen Williamson* and Randall Wright**

*Washington University in St. Louis and Federal Reserve Banks of Richmond and St. Louis

**University of Wisconsin — Madison and Federal Reserve Bank of Minneapolis

Contents

1. Introduction	26
2. Basic Monetary Theory	31
2.1 The simplest model	31
2.2 Prices	35
2.3 Distributions	37
3. A Benchmark Model	38
3.1 The environment	39
3.2 Results	44
3.3 Unanticipated inflation	48
3.4 Money and capital	49
3.5 The long-run Phillips curve	53
3.6 Benchmark summary	55
4. New Models of Old Ideas	57
4.1 The Old Monetarist Phillips curve	57
4.2 New Keynesian sticky prices	61
4.3 New Monetarist sticky prices	66
5. Money, Payments, and Banking	71
5.1 A payments model	71
5.2 Banking	75
6. Finance	79
6.1 Asset trading and pricing	80
6.2 Capital markets	83
7. Conclusion	89
References	90

Abstract

The purpose of this paper is to discuss some of the models used in New Monetarist Economics, which is our label for a body of recent work on money, banking, payments systems, asset markets, and related topics. A key principle in New Monetarism is that solid microfoundations

[☆] This essay was written as a chapter for the new Handbook of Monetary Economics, which is being edited by Benjamin Friedman and Michael Woodford. We thank the editors, as well as Boragan Aruoba, Guillaume Rocheteau, Robert Shimer, Jiang Shi, Liang Wang and Lucy Liu for useful discussions and comments. We thank the NSF for financial support. Wright also thanks the Ray Zemon Chair in Liquid Assets at the Wisconsin School of Business.

are critical for understanding monetary issues. We survey recent papers on monetary theory, showing how they build on common foundations. We then lay out a tractable benchmark version of the model that allows us to address a variety of issues. We use it to analyze some classic economic topics, like the welfare effects of inflation, the relationship between money and capital accumulation, and the Phillips curve. We also extend the benchmark model in new ways, and show how it can be used to generate new insights in the study of payments, banking, and asset markets.

JEL classification: E0, E1, E4, E5

Keywords

Monetary Theory

Monetary Policy

New Monetarism

1. INTRODUCTION

Our goal is to present some models in current use, plus work in progress, in a distinct school of thought in monetary economics. Any school needs a name, and we call ours *New Monetarist Economics*. A key principle in New Monetarism is that we need solid microfoundations for institutions that facilitate the process of exchange — institutions like money, banks, financial intermediaries more generally, and so on — if we are to make progress in monetary economics. That this view is not universally accepted is clear from the fact that many currently popular models used for monetary policy analysis either have no money (or banks or related institutions), or if they do, they slip it in with ad hoc approaches by assuming a cash-in-advance constraint or by putting money in utility or production functions (some even resort to putting government bonds and commercial bank reserves in utility or production functions). We do not go far into methodology or history of thought here, but we will say this by way of explaining our name. New Monetarists find much that is appealing in Old Monetarism, epitomized by the writings of Friedman and his followers, although we also disagree with them in several important ways. And New Monetarists have little in common with Old or New Keynesians, although this may have as much to do with the way they approach monetary economics and microfoundations generally as with sticky prices. An extended discussion of these issues has been relegated to a companion paper.¹

¹ In “New Monetarist Economics: Methods” (Williamson & Wright, 2010) we lay out what we think are the unifying principles of New Monetarism, and indicate where and why it differs from Old Monetarism, and New or Old Keynesianism. We also argue that the New Keynesian consensus that some people think characterizes the current state of affairs, at least among more policy-oriented monetary and macro economists, is not healthy. The Old Keynesians had Old Monetarists continuously engaging them in debate over issues and models. We think the current situation would be healthier if there was more discussion and appreciation of alternatives to textbook New Keynesianism. This is one of the reasons we were interested in writing this essay. A discussion along these lines was originally meant to be included here, but to keep the *Handbook* chapter focused, on the advice of the editors, we moved that material to the companion paper.

New Monetarism encompasses a body of research on monetary theory and policy, and on banking, financial intermediation, payments, and asset markets, that has occurred over the last few decades. In monetary economics, this includes the seminal work using overlapping generations models by Lucas (1972) and some of the contributors to the *Models of Monetary Economics* volume edited by Kareken and Wallace (1980), although antecedents exist, including Samuelson (1958). More recently, much monetary theory has adopted the search and matching approach, early examples of which are Kiyotaki and Wright (1989, 1993), although there are also antecedents for this, including Jones (1976) and P. Diamond (1982, 1984). In the economics of banking, intermediation, and payments, which builds on advances in information theory that occurred mainly in the 1970s, examples of what we have in mind include Diamond and Dybvig (1983), D. Diamond (1984), Williamson (1986, 1987), Bernanke and Gertler (1989), and Freeman (1996). Much of this research is abstract and theoretical in nature, but the literature has turned more recently to empirical and policy issues.

A key principle, laid out first in the introduction to Kareken and Wallace (1980), and elaborated in Wallace (1998), is that progress can be made in monetary theory and policy analysis only by modeling monetary arrangements explicitly. In line with the arguments of Lucas (1976), to conduct a policy experiment in an economic model, the model must be structurally invariant to the experiment under consideration. One interpretation is the following: if we are considering experiments involving the operating characteristics of the economy under different monetary policy rules, we need a model in which economic agents hold money not because it enters utility or production functions, in a reduced-form fashion, but because money ameliorates some fundamental frictions. Of course the view that monetary theory should “look frictions in the face” goes back to Hicks (1935). Notice that here we are talking about explicit descriptions of frictions *in the exchange process*, as opposed to frictions *in the price setting process*, like the nominal rigidities in Keynesian theory, where money does not help (it is really the cause of the problem).

We now know that there are various ways to explicitly model frictions. There are many important frictions to consider in monetary and financial economics, including private information, limited commitment, and spatial separation, and this potentially makes the modeling difficult. There is an element of art and skill in capturing key frictions while allowing for tractability. Overlapping generations models can be simple, although one can also complicate them as one likes. Much research in monetary theory in the last 20 years, as mentioned above, has been conducted using matching models, building on ideas in search and game theory.² Matching models are very tractable for many questions in monetary economics, although a key insight that eventually arose from this literature is that spatial separation per se is not the critical friction making money essential. As

² Individual contributions to the search and matching literature will be discussed in detail below. The previous *Handbook of Monetary Economics* has a survey by Ostroy and Starr (1990) of earlier attempts at building microfoundations for money using mainly general equilibrium theory, as well as a survey of overlapping generations models by Brock (1990).

emphasized by Kocherlakota (1998), with credit due to earlier work by Ostroy (see Ostroy & Starr, 1990) and Townsend (1987, 1989), money is essential because it overcomes a double coincidence of wants problem in the context of limited commitment and imperfect record keeping. Perfect record keeping would imply that efficient allocations can be supported through insurance and credit markets, or various other institutions, without money. Random bilateral matching among a large number of agents is a convenient way to generate a double coincidence problem, and to motivate incomplete record keeping, but it is not the only way, as we discuss.

While it is important to understand the above issues, New Monetarism is not just about the role of currency in the exchange process. It also attempts to study a host of related institutions. An important departure from Old Monetarism is to take seriously the role of financial intermediaries and their interactions with the central bank. Developments in intermediation and payment theories over the last 25 years are critical to our understanding of credit and banking arrangements. By way of example, a difference between Old and New Monetarists regarding the role of intermediation is reflected in their respective evaluations of Friedman's (1960) proposal for 100% reserve requirements on transactions deposits. His argument was based on the premise that tight control of the money supply by the central bank was key to controlling the price level. Since transactions deposits at banks are part of what he means by money, and the money multiplier is subject to randomness, even if we could perfectly control the stock of outside money, inside money would move around unless we impose 100% reserves. Old Monetarists therefore viewed 100% reserves as desirable. What this ignores is that banks perform a socially beneficial function in transforming illiquid assets into liquid liabilities, and 100% reserve requirements inefficiently preclude this activity.

The 1980s saw important developments in the theory of banking and financial intermediation. One influential contribution was the model of Diamond and Dybvig (1983), which we now understand to be a useful approach to studying banking as liquidity transformation and insurance (it does however require some auxiliary assumptions to produce anything resembling a banking panic or run; see Ennis and Keister, 2008). Other work involved well-diversified intermediaries economizing on monitoring costs, including D. Diamond (1984) and Williamson (1986). In these models, financial intermediation is an endogenous phenomenon. The resulting intermediaries are well-diversified, process information in some manner, and transform assets in terms of liquidity, maturity, or other characteristics. The theory of financial intermediation has also been useful in helping us understand the potential for instability in banking and the financial system (Ennis & Keister, 2009a, 2009b, 2010), and how the structure of intermediation and financial contracting can affect aggregate shocks (Bernanke & Gertler, 1989; Williamson, 1987).

A relatively new sub-branch of this theory studies the economics of payments. This involves the study of payments systems, particularly among financial institutions, such as Fedwire in the United States, where central banks can play an important role. Freeman (1996) is an early contribution, and Nosal and Rocheteau (2011) provide a

recent survey. The key insights from this literature are related to the role played by outside money and central bank credit in the clearing and settlement of debt, and the potential for systemic risk as a result of intraday credit. Even while payment systems are working well, this area is important, since the cost of failure is potentially big, given the volume of payments processed through such systems each day. New Monetarist economics not only has something to say about these issues, it is almost by definition the only approach that does. How can one hope to understand payments and settlement without explicitly modeling the exchange process?

Our objective is to explain the kinds of models people are using to study these issues. As an overview, what we do is this. First we survey the papers on monetary theory with microfoundations building on matching theory, showing how several models that are apparently different actually build on common foundations. Indeed, they can all be considered special cases of a general specification. We then lay out a benchmark version of the model that is very tractable, but still allows us to address a variety of important issues. We show how it can be used to analyze classic economic topics, like the welfare effects of inflation, the relationship between money and capital accumulation, and the short- and long-run Phillips curve. We then extend the benchmark model in some new ways, and show through a series of applications how it can be used to generate new insights in the study of payments, banking, and asset markets.

To go into more detail, in Section 2 we start with models of monetary economies that are very simple because of the assumption that money, and sometimes also goods, are indivisible. We try to say why the models are interesting, and why they were constructed as they were — what lies behind the abstractions and simplifications. In Section 3 we move to more recent models, with divisible money. These models are better suited to address many empirical and policy issues, but are still tractable enough to deliver sharp analytic results. We lay out a benchmark New Monetarist model, based on Lagos and Wright (2005), and show how it can be used to address various issues. Again, we explain what lies behind the assumptions, and we discuss some of its basic properties (e.g., money is neutral but not superneutral, the Friedman rule is typically optimal, but may not yield the first best, etc.). We also show how this benchmark can be extended to incorporate capital accumulation, unemployment, and other phenomena. As one example, we generate a traditional Phillips curve — a negative relation between inflation and unemployment — that is structurally stable in the long run. In this example, anticipated policy *can* exploit this trade-off, but it *ought not*: the Friedman rule is still optimal. This illustrates the value of being explicit about micro details.

While much of the material in Sections 2 and 3 is already in the literature, Section 4 presents novel applications. First, we show how the benchmark model can be used to formalize Friedman's (1968) view about the short-run Phillips curve, using a signal extraction problem as in Lucas (1972). This yields some conclusions that are similar to those of Friedman and Lucas, but also some that are different. We then use the model to illustrate New Keynesian ideas by introducing sticky prices. This generates policy conclusions similar to those in Clarida, Gali, and Gertler (1999) or Woodford (2003), but there are also differences, again

illustrating how details matter. In addition, we present a New Monetarist model of *endogenously* sticky prices, with some very different policy implications. Although some of the applications in this Section re-derive known results, in a different context, they also serve to make it clear that other approaches are not inconsistent with our model. One should not shy away from New Monetarism even if one believes sticky prices, imperfect information, and related ingredients are critical, since these are relatively easily incorporated into micro-based theories of the exchange process.³

In Section 5, we discuss applications related to banking and payments. These extensions contain more novel modeling choices and results, although the substantive issues have been raised in earlier work. One example incorporates ideas from payments economics similar in spirit to Freeman (1996), but the analysis looks different through the lens of the New Monetarist approach. Another example incorporates existing ideas in the theory of banking emulating from Diamond and Dybvig (1983), but again the details look different. In particular, we have genuinely monetary versions of these models, which seems relevant, or at least realistic, since money plays a big role in actual banking and payments systems (previous attempts to build monetary versions of Diamond-Dybvig include Freeman, 1988 and Champ, Smith, & Williamson, 1996). In Section 6, we present another application, exploring a New Monetarist approach to asset markets. This approach emphasizes liquidity, and studies markets where asset trade can be complicated by various frictions.

We think these applications illustrate the power and flexibility of the New Monetarist approach. As we hope readers will appreciate, the various models may differ with respect to details, but they share many features and build upon common principles. This is true for the simplest models of monetary exchange, as well as the extensions that integrate banking, credit arrangements, payments mechanisms, and asset markets. We think that this is not only interesting in terms of economic theory, but that there are also lessons to be learned for understanding the current economic situation and shaping future policy. To the extent that the recent crisis has at its roots problems related to banking, mortgage markets, and other credit arrangements, or information problems in asset markets, one cannot address the issues without models that take seriously the exchange process. We do not claim New Monetarist economics provides all of the answers for all of the recent economic problems; we do believe it has a great deal to contribute to the discussion.

³ Since part of our mandate from the editors was to illustrate how standard results in other literatures can be recast in the context of modern monetary theory, we thought it would be good to discuss topics such as the relationship between money and capital, the long- and short-run Phillips curve, signal extraction, and sticky prices. But our New Keynesian application should not be read as condonation of the practice of assuming nominal rigidities in an ad hoc fashion. It is rather meant to show that even if one cannot live without such assumptions, this does not mean one cannot think seriously about money, banking, and so forth. Also, our examples are meant to be simple, but one can elaborate as one wishes. Craig and Rocheteau (2008), for example, have a version of our benchmark model with sticky prices as in Benabou (1988) and Diamond (1993), while Aruoba and Schorfheide (2010) have a version on par with a typical New Keynesian model that they estimate. Similarly, Faig and Li (2009) have a more involved version with signal extraction that they take to data. The goal here is mainly to illustrate basic qualitative effects, although in various places we discuss aspects of calibration and report some quantitative results.

2. BASIC MONETARY THEORY

An elementary model in the spirit of New Monetarist Economics is a version of the first-generation monetary search theory, along the lines of Kiyotaki and Wright (1993), which is a stripped-down version of Kiyotaki and Wright (1989, 1991), and uses methods from equilibrium search theory (e.g., Diamond 1982). This model makes some strong assumptions, which will be relaxed later, but even with these assumptions in place it captures something of the essence of money as an institution that facilitates exchange. What makes exchange difficult in the first place is a double-coincidence problem, generated by specialization and random matching, combined with limited commitment and imperfect memory. Frictions like this, or at least informal descriptions thereof, have been discussed in economics for a long time, and certainly versions of the double-coincidence problem can be found in Adam Smith, and much further back, if one looks. The goal of recent theory is to formalize these ideas, to see which are valid under what assumptions, and hopefully to develop new insights along the way.

Before proceeding, since we start with search-based models, it is perhaps worth saying why. Clearly, random matching is an extreme assumption, but it captures well the notion that people trade with each other and not only against budget constraints; and yet it is all too easy to criticize. As Howitt (2005) said:

In contrast to what happens in search models, exchanges in actual market economies are organized by specialist traders, who mitigate search costs by providing facilities that are easy to locate. Thus when people wish to buy shoes they go to a shoe store; when hungry they go to a grocer; when desiring to sell their labor services they go to firms known to offer employment. Few people would think of planning their economic lives on the basis of random encounters.

Based in part on such criticism, much of the theory, including the models in this section, has been redone using directed rather than random search (Corbae, Temzelides, & Wright, 2003; Julien, Kennes, & King, 2008). While some results change, the basic theory remains intact. Hence we start with random matching, hoping readers understand that the theory also works with directed search. Later, search is replaced by preference and technology shocks.

2.1 The simplest model

Time is discrete and continues forever. There is a $[0, 1]$ continuum of infinite-lived agents. To make exchange interesting, these agents specialize in production and consumption of differentiated commodities, and trade bilaterally. It is an old idea that specialization is intimately related to monetary exchange, so we want this in the environment. Although there are many ways to set it up, here we assume the following: There is a set of goods, that for now are indivisible and nonstorable. Each agent produces, at cost $C \geq 0$, goods in some subset, and derives utility $U > C$ from

consuming goods in a different subset. It is formally equivalent, but for some applications it helps the discussion, to consider a pure exchange scenario. Thus, if each agent is endowed with a good each period that he can consume to yield utility C , but he may meet someone with another good that gives him utility U , the analysis is basically the same, except C is interpreted as an opportunity cost rather than a production cost.

Let α be the probability of meeting someone each period. There are different types of potential trade meetings. Let σ be the probability that you like what your partner can produce but not vice versa — a single coincidence meeting — and δ the probability that you like what he can produce and vice versa — a double-coincidence meeting.⁴ The environment is symmetric, and for the representative agent, the efficient allocation clearly involves producing whenever someone in a meeting likes what his partner can produce. Let V^C be the payoff from this cooperative allocation, described recursively by

$$\begin{aligned} V^C &= \alpha\sigma(U + \beta V^C) + \alpha\sigma(-C + \beta V^C) + \alpha\delta(U - C + \beta V^C) \\ &\quad + (1 - 2\alpha\sigma - \alpha\delta)\beta V^C \\ &= \beta V^C + \alpha(\sigma + \delta)(U - C). \end{aligned}$$

If agents could commit, ex ante, they would all agree to execute the efficient allocation. If they cannot commit, we have to worry about ex post incentive conditions.

The binding condition is this: to get agents to produce in single-coincidence meetings we require $-C + \beta V^C \geq \beta V^D$, where V^D is the deviation payoff, depending on what punishments we have at our disposal. Suppose we can punish a deviator by allowing him in the future to only trade in double-coincidence meetings. It is interesting to consider other punishments, but this one has a nice interpretation in terms of what a mechanism designer can see and do. We might like to trigger to autarky — no trade at all — after a deviation, but it is not so obvious we can enforce this in double-coincidence meetings. Having trade only in double-coincidence meetings — a pure barter system — is self-enforcing, and implies payoff $V^B = \alpha\delta(U - C)/(1 - \beta)$. If we take the deviation payoff to be continuing with pure barter, $V^D = V^B$, the relevant incentive condition can be reduced to

$$[1 - \beta(1 - \alpha\sigma)]C \leq \beta\alpha\sigma U. \quad (1)$$

If every potential trade meeting involves a double-coincidence; that is, if $\sigma = 0$, then pure barter suffices to achieve efficiency and there is no incentive problem. But with $\sigma > 0$, given imperfect commitment, Eq. (1) tells us that we can achieve efficiency iff production is not too expensive (C is small), search and specialization

⁴ Many extensions and variations are possible. In Kiyotaki and Wright (1991), for example, agents derive utility from all goods, but prefer some over others, and the set of goods they accept is determined endogenously. In Kiyotaki and Wright (1989) or Aiyagari and Wallace (1991, 1992) there are N goods and N types of agents, where type n consumes good n and produces good $n + 1 \pmod{N}$. In this case, $N = 2$ implies $\sigma = 0$ and $\delta = 1/2$, while $N \geq 3$ implies $\sigma = 1/N$ and $\delta = 0$. The case $N = 3$ has been used to good effect by Wicksell (1967) and Jevons (1875).

frictions are not too severe (α and σ are big), and so forth.⁵ If Eq. (1) holds, one can interpret exchange as a credit system, as in Sanches and Williamson (2009), but there is no role for money. A fundamental result in Kocherlakota (1998) is that money is not essential; that is, it does nothing to expand the set of incentive-feasible allocations when we can use trigger strategies as previously described. Obviously this requires that deviations can be observed and recalled. Lack of perfect monitoring or record keeping, often referred to as incomplete memory, is necessary for money to be essential.

There are several way to formalize this. Given a large number of agents that match randomly, suppose that they observe what happens in their own but not in other meetings. Then, if an agent deviates, the probability someone he meets later will know it is 0. This is often described by saying agents are anonymous. In addition to Kocherlakota (1998), see Araujo (2004); Araujo, Camargo, Minetti, and Puzello (2010); Aliprantis, Camera, and Puzello (2006, 2007a,b); Kocherlakota and Wallace (1998); and Wallace (2001) for more discussion. Also note that we only need *some* meetings to be anonymous; in applications below we assume that with a given probability meetings are monitored, and credit may be used in those meetings. But for now, we assume all meetings are anonymous, so there is no credit, and hence no one ever produces in single-coincidence meetings. In this case, absent money, we are left with only direct barter.

Therefore we want to introduce money. Although we soon generalize this, for now, there are $M \in (0, 1)$ units of some object that agents can store in units $m \in \{0, 1\}$. This object is worthless in consumption and does not aid in production, and so if it is used as a medium of exchange it is, by definition, fiat money (Wallace, 1980). One could also assume the object gives off a flow utility $\gamma > 0$ — say a dividend yield — and interpret it as commodity money. Alternatively, if $\gamma < 0$, we can interpret it as a storage cost. To ease the presentation we set $\gamma = 0$ for now (but see Section 6). While m may not have all the properties that undergraduate textbooks say money tends to or ought to have, and in particular it lacks divisibility, it does have other desirable properties, like storability, portability, and recognizability. We assume it is initially distributed randomly across agents, and from then on the matching process is such that, conditional on a meeting, your partner has $m = 1$ with probability M and $m = 0$ with probability $1 - M$.

Let V_m be the payoff to an agent with money holdings $m \in \{0, 1\}$. Then the value function of an agent with $m = 0$ is given by

$$V_0 = \beta V_0 + \alpha \delta (U - C) + \alpha \sigma M \max_{\xi} \xi [-C + \beta (V_1 - V_0)], \quad (2)$$

since he can still barter in double-coincidence meetings, and now has another option: if he meets someone with money who likes his good but cannot produce anything he likes, he could trade for cash, and ξ is the probability he agrees to do so. Similarly, the value function of an agent with $m = 1$ is

⁵ Do not get confused by the fact that $\sigma = 0$ implies (1) fails. It is true that if there were no single-coincidence meetings then we could not sustain cooperative trade in single-coincidence meetings, but it does not matter.

$$V_1 = \beta V_1 + \alpha \delta(U - C) + \alpha \sigma(1 - M) \Xi[U + \beta(V_0 - V_1)], \quad (3)$$

because he can still barter, and now he also can make a cash offer in single-coincidence meetings, which is accepted with some probability Ξ that he takes as given.⁶

The best response condition gives the maximizing choice of ξ taking Ξ as given: $\xi = 1$ or 0 or $[0, 1]$ as $-C + \beta(V_1 - V_0)$ is positive or negative or 0 , where V_1 and V_0 are functions of Ξ obtained by solving Eq. (2)–Eq. (3). An equilibrium is a list $\{\xi, V_0, V_1\}$ satisfying Eq. (2)–Eq. (3) and the best response condition. Obviously $\xi = 0$ always constitutes an equilibrium, and $\xi = 1$ does as well iff

$$[1 - \beta + \beta \alpha \sigma(1 - M)]C \leq \beta \alpha \sigma(1 - M)U$$

(there are also mixed strategy equilibria, but one can argue they are not robust, as in Shevchenko & Wright, 2004). Hence, there is a monetary equilibrium $\xi = 1$ iff C is below an upper bound. This bound is less than the one we had for credit equilibrium when triggers were available. Moreover, even if we can support $\xi = 1$, payoffs are lower with money than with triggers. So when monitoring or memory is bad, money may allow us to do better than barter, but not as well as perfect credit. In other words, money may be a substitute, but it is not a perfect substitute, for credit.

This model is crude, with its indivisibilities, but without a doubt it captures the notion that money is a beneficial institution that facilitates exchange. This contrasts with cash-in-advance models, where money is a hindrance, or sticky-price models, where money plays a purely detrimental role when it is assumed agents must quote prices in dollars and are not allowed to change them easily. Also note that, contrary to standard asset-pricing theory, in monetary equilibria an intrinsically worthless object has positive value. Naturally, it is valued as a medium of exchange, or for its liquidity. Monetary equilibria have good welfare properties relative to barter, even if they do not achieve first best. The fact that $\xi = 0$ is always an equilibrium points to the tenuousness of fiat money. Yet it is also robust, in the sense that the equilibrium with $\xi = 1$ survives even if we endow the fiat object with some bad characteristics, like a transaction or storage cost, or if we tax it, as long as the costs or taxes are not too big. So, while it may be crude, the model makes many predictions that ring true.⁷

⁶ The presentation here is slightly different from the original search models, which usually assumed agents with money could not produce. The version here is arguably more natural, and for some issues simpler. See Rupert et al. (2001) for an extended discussion and references.

⁷ Other applications of these first-generation models include the following: Aiyagari and Wallace (1991); Kehoe, Kiyotaki, and Wright (1993); Kiyotaki and Wright (1989); and Wright (1995), among others, allow goods to be storable and discuss commodity money. Kiyotaki and Wright (1991, 1993); Camera, Reed, and Waller (2003); and Shi (1997a), endogenized specialization in production and consumption. Kiyotaki et al. (1993) and Zhou (1997) pursued issues in international monetary economics. Kim (1996), Li (1995), and Williamson and Wright (1994), introduced private information to show how money can ameliorate lemons problems. Li (1994, 1995) discussed the optimal taxation of money in the presence of search externalities. Ritter (1995) asked how fiat currency might first get introduced. Green and Weber (1996) discussed counterfeiting. Cavalcanti, Erosa, and Temzelides (1999); He, Huang, and Wright (2005); and Lester (2009) studied banking and payments issues.

2.2 Prices

Up to now prices were fixed, since every trade involves a one-for-one swap. Beginning the second generation of papers in this literature, Shi (1995) and Trejos and Wright (1995) endogenized prices by keeping $m \in \{0, 1\}$ but allowing divisible goods. Although we relax $m \in \{0, 1\}$ soon enough, the advantage of this approach is that one can talk about prices while maintaining a simple fixed distribution of money holdings across agents: it is still the case that at any point in time M agents each hold $m = 1$ and $1 - M$ agents each hold 0. When a producer gives output x to a consumer, their instantaneous utilities are $U = u(x)$ and $C = c(x)$, where $u' > 0$, $c' > 0$, $u'' < 0$, $c'' \geq 0$, and $u(0) = c(0) = 0$. Letting x^* solve $u'(x^*) = c'(x^*)$, it is easy to show that the efficient outcome is for agents to produce x^* in every meeting where their partner likes their output. A credit system with perfect memory could support this if β is big enough. We instead want to talk about monetary equilibria, so we assume imperfect memory, as previously discussed.

We focus on the case where money is accepted with probability $\xi = 1$, and to ease the presentation, we start with $\delta = 0$ so there is no direct barter. Now, to determine x in a monetary exchange, we use the generalized Nash bargaining solution.⁸ One virtue of this is simplicity; another is the well-known result that Nash bargaining can be interpreted as a natural limit of a simple noncooperative bargaining game (see e.g., Binmore, Osborne, & Rubinstein, 1992). Letting the bargaining power of the consumer be θ and letting threat points be given by continuation values, x then solves

$$\max [u(x) + \beta V_0 - \beta V_1]^\theta [-c(x) + \beta V_1 - \beta V_0]^{1-\theta}. \quad (4)$$

For now we consider the notion of a stationary equilibrium, or steady state, which is a list $\{x, V_0, V_1\}$ such that: given V_0 and V_1 , x solves Eq. (4); and given x , V_0 and V_1 solve Eqs. (2) and (3).

For the sake of illustration, consider the case $\theta = 1$, which means that buyers get to make take-it-or-leave-it offers, so that $c(x) = \beta(V_1 - V_0)$. Solving for $V_1 - V_0$ from Eqs. (2)–(3), this reduces to

$$c(x) = \frac{\beta \alpha \sigma (1 - M) u(x)}{1 - \beta + \beta \alpha \sigma (1 - M)}. \quad (5)$$

This condition holds at $x = 0$, which is a nonmonetary equilibrium, and at a unique monetary equilibrium $x > 0$, where it is easy to check $\partial x / \partial M < 0$, so the price level $p = 1/x$ increases with the number of buyers. When we relax $\delta = 0$, there are generically either multiple monetary equilibria or no monetary equilibria. This generalization is straightforward, although note that $\delta > 0$ means one has to also solve for x in a barter

⁸ Other solution concepts can also be used: Curtis and Wright (2004) used price posting; Julien et al. (2008) used auctions in a version with some multilateral meetings; and Wallace and Zhou (2007a,b) used mechanism design.

exchange, which generally differs from the x in a monetary trade (for general results with $\delta > 0$, as well as any bargaining power θ and alternative specifications for the threat points, see Rupert, Schindler, & Wright, 2001).

In the symmetric case $\theta = 1/2$ and $M = 1/2$, which is the one used in Shi (1995) and Trejos and Wright (1995), it can be shown that $x < x^*$ in any equilibrium. Hence, monetary exchange does not achieve the efficient allocation. However, it is easy to verify that $x \rightarrow x^*$ as $\beta \rightarrow 1$. To understand this, consider an Arrow-Debreu version of this environment, which means the same preferences and technology but no frictions. In such an economy, since given agents can turn their production into instantaneous consumption through the market, they choose $x = x^*$. But in our economy, they must first turn production into cash, which can then only be used in the future. Therefore, as long as $\beta < 1$, agents are willing to produce less than they would in a frictionless model. Now, one can get x to increase, say by raising θ , and for big enough θ we may have $x > x^*$, but the model still illustrates a basic tendency for $x < x^*$, other things being symmetric.⁹

Before moving on, we briefly mention nonstationary equilibria in this simple setup. For illustration, assume $\delta = 0$, and add a flow utility γ of holding $m = 1$; as discussed above, if $\gamma > 0$ then m is commodity money, and if $\gamma < 0$ then m has a storage cost. Also, purely for convenience, we move to continuous time by letting the length of a period (in both the search and bargaining processes) vanish, implying

$$\begin{aligned} rV_0 &= \alpha\sigma M[-c(x) + V_1 - V_0] + \dot{V}_0 \\ rV_1 &= \gamma + \alpha\sigma(1 - M)[u(x) + V_0 - V_1] + \dot{V}_1. \end{aligned}$$

Subtracting yields a differential equation in the difference

$$\dot{V}_1 - \dot{V}_0 = -\gamma - \alpha\sigma(1 - M)u(x) - \alpha\sigma M c(x) + (r + \alpha\sigma)(V_1 - V_0). \quad (6)$$

To reduce notation, without loss in generality, set $\alpha\sigma = 1$, and let $c(x) = x$. Also, assume for simplicity $\theta = 1$. Then we get $V_1 - V_0 = x$, $\dot{V}_1 - \dot{V}_0 = \dot{x}$, and

$$\dot{x} = -\gamma + (r + 1 - M)x - (1 - M)u(x),$$

Define $F(x)$ by the RHS of Eq. (6). Then equilibrium can be defined as a non-negative time path for x satisfying $\dot{x} = F(x)$, plus a side condition that says buyers want to trade, $u(x) + V_0 - V_1 \geq 0$ (the seller wants to trade by construction when $\theta = 1$). This side condition holds if and only if $x \leq \bar{x}$, where $u(\bar{x}) = \bar{x}$, and tells us that an equilibrium path for x cannot leave $[0, \bar{x}]$. By plotting $F(x)$ versus x it is now easy to see the following:

⁹ One can argue that $x > x^*$ is an artifact of indivisible money here as follows: if we allow lotteries, which are useful with $m \in \{0, 1\}$, and in a sense approximate divisible m , it can be shown that x can never exceed x^* (see Berentsen, Molico, & Wright 2002; Berentsen & Rocheteau, 2002). Soon enough we can check this in models that have divisible money.

1. When $\gamma = 0$, which means fiat money, there are two steady states, $x = 0$ and $x = x^0 \in (0, \bar{x})$, plus a continuum of dynamic equilibria starting from any $x \in (0, x^0)$ and converging to 0.
2. When $\gamma > 0$, which means commodity money, the $F(x)$ curve shifts down. As long as γ is not too big the unique equilibrium is a steady state with $x = x^\gamma \in (x^0, \bar{x})$, since no other path satisfying $\dot{x} = F(x)$ remains in $[0, \bar{x}]$. This illustrates the venerable idea that commodity money can eliminate indeterminacies associated with fiat money. If γ gets too big, however, then $x^\gamma > \bar{x}$, which means an agent with $m = 1$ prefers to hoard rather than spend it, and is reminiscent of Gresham's Law (or at least it would be if we introduce a second money, which is easy enough to do).
3. When $\gamma < 0$, there is always a steady state equilibrium with $x = 0$, where agents freely dispose of money, and if $|\gamma|$ is big then this is the only equilibrium. If $|\gamma|$ is not too big then there are two steady states in $(0, x^0)$, say x^1 and x^2 , plus a continuum of dynamic equilibria starting at any $x \in (0, x^2)$ and converging to x^1 .

These results illustrate some interesting properties of fiat and commodity money systems, and show how different types of interesting dynamic equilibria may emerge (as is true in most monetary theories, of course). There are many other applications of this simple model, but without further ado, we now move to relax the inventory restriction $m \in \{0, 1\}$.¹⁰

2.3 Distributions

Although there are various alternatives, consider the approach in Molico (2006), who allowed $m \in [0, \infty]$.¹¹ This means that we have to deal with the endogenous distribution of money across agents, $F(m)$, while previously this was trivial. Now, in a single-coincidence meeting where the consumer has m and the producer has $\tilde{m}$, let $x(m, \tilde{m})$ be the amount of output and $d(m, \tilde{m})$ the amount of money traded. Again setting $\delta = 0$, for expositional purposes, the generalization of Eqs. (2)–(3) is

¹⁰ A few applications include the following: Shi (1996) introduced bilateral borrowing and lending to study the relation between money and credit. Aiyagari, Wallace, and Wright (1996) studied the interaction between money and bonds. Coles and Wright (1998), Ennis (2001) and Shi (1995) further studied nonstationary equilibria. Katzman, Kennan, and Wallace (2003) and Wallace (1997) studied the inflation–output relation. Wallace and Zhou (1997) studied currency shortages. Ales, Carapella, Maziero, and Weber (2008); Burdett, Trejos and Wright (2001); Redish and Weber (2010); and Velde, Weber, and Wright (1999) used the model to analyze various issues in monetary history. Lee, Wallace, and Zhu (2005) studied denomination structures. Williamson (1999) considered private money. Cavalcanti and Wallace (1999a,b) introduced banks. Trejos (1999) studied private information. Johri and Leach (2002), Li (1999), and Schevchenko (2004) studied middlemen. Nosal and Wallace (2007) analyzed counterfeiting.

¹¹ Other approaches to relaxing $m \in \{0, 1\}$ include Berentsen (2002), Camera and Corbae (1999), Deviatov and Wallace (2001), and Zhu (2003, 2005). There is also a series of papers following up on Green and Zhou (1998); rather than list them all here, see the references in Jean, Stanislav, and Wright (2010). Some of these models assume $m \in \{0, 1, \dots, \bar{m}\}$, where the upper bound $\bar{m}$ may or may not be finite. The value function in Eq. (7) is still valid in such cases, including the case $\bar{m} = 1$.

$$V(m) = \beta V(m) + \alpha \sigma \int \{u[x(m, \tilde{m})] + \beta V[m - d(m, \tilde{m})] - \beta V(m)\} dF(\tilde{m}) \\ + \alpha \sigma \int \{-c[x(\tilde{m}, m)] + \beta V[m + d(\tilde{m}, m)] - \beta V(m)\} dF(\tilde{m}). \quad (7)$$

The first term is the expected value of buying from a producer with $\tilde{m}$ dollars, and the second the expected value of selling to a consumer with $\tilde{m}$ dollars (notice how the roles of m and $\tilde{m}$ are reversed in the two integrals).

In this model, we can easily add injections of new currency, say by lump sum or proportional transfers, which was not so easy with $m \in \{0, 1\}$. With lump-sum transfers, we simply change m on the RHS to $m + \mu M$, where M is the aggregate money supply, governed by $M_{t+1} = (1 + \mu)M_t$. This greatly extends the class of policies that can be analyzed. However, to illustrate the basic idea, for now we keep $M = \int m dF(m)$ fixed. Then a stationary equilibrium is a list of functions $\{V(\cdot), x(\cdot), d(\cdot), F(\cdot)\}$ such that: given $x(m, \tilde{m})$, $d(m, \tilde{m})$ and $F(m)$, $V(m)$ solves Eq. (7); given $V(m)$, $x(m, \tilde{m})$ and $d(m, \tilde{m})$ are determined by some bargaining solution, such as

$$\max [u(x) + \beta V(m - d) - \beta V(m)]^\theta [-c(x) + \beta V(\tilde{m} + d) - \beta V(\tilde{m} + d) - \beta V(\tilde{m})]^{1-\theta} \quad (8)$$

where the maximization is s.t. $d \leq m$; and given $x(m, \tilde{m})$ and $d(m, \tilde{m})$, $F(m)$ solves a stationarity condition omitted in the interest of space. From this we can calculate other interesting objects, such as the distribution of $p(m, \tilde{m}) = d(m, \tilde{m})/x(m, \tilde{m})$.

This model is complicated, even using numerical methods. Heterogeneous-agent, incomplete-market, macro models of the sort analyzed by Huggett (1993) or Krusell and Smith (1998) also have an endogenous distribution as a state variable, but the agents in those models do not care about this distribution per se, they only care about prices. Prices depend on the distribution, but one can typically characterize accurately prices as functions of a small number of moments. In a search model, agents care about $F(m)$ directly, since they are trading with each other and not merely against their budget equations. Still, Molico (2006) computed equilibria, and the model is used to discuss issues such as the effects of inflation (see also Chiu & Molico, 2006, 2010). An alternative approach used by Dressler (2009, 2010) is to assume competitive pricing, rather than bargaining (see Section 3). This makes computation easier, on a par with Huggett-Krusell-Smith models. But while it is easier, this approach also loses some of the interesting elements from bargaining models, including the endogenous distribution of prices.

3. A BENCHMARK MODEL

Some search models with divisible money use devices that allow one to avoid having to track $F(m)$. There are two main approaches.¹² The first, originating with Shi (1997b), uses the assumption of large households to render the distribution degenerate. Thus,

¹² Recently, Menzio et al. (2009) proposed a new method for dealing with distributions, based on directed search.

each decision making unit consists of many members who search randomly, as in the previous models, but at the end of each trading round they return to the homestead, where they share the money they bring back with their siblings. Loosely speaking, by the law of large numbers, each household starts the next trading round with the same m . The large household is a natural extension for random-matching models of the “worker-shopper pair” discussed in the cash-in-advance literature (Lucas, 1980). A number of interesting papers use this environment; rather than cite them all here, we refer the reader to Shi (2006). We focus instead on a different approach, following Lagos and Wright (2005), and use markets instead of families.

We use the Lagos-Wright model because it allows us to address a variety of other issues, in addition to rendering the distribution of money tractable (although some of the applications could in principle also use Shi’s model). In particular, it serves to reduce the gap between monetary theory with some claim to microfoundations and standard macroeconomics. Whatever one thinks of the models discussed earlier, they are pretty far from mainstream macro. As Azariadis (1993) said:

Capturing the transactions motive for holding money balances in a compact and logically appealing manner has turned out to be an enormously complicated task. Logically coherent models such as those proposed by Diamond (1982) and Kiyotaki and Wright (1989) tend to be so removed from neoclassical growth theory as to seriously hinder the job of integrating rigorous monetary theory with the rest of macroeconomics.

And as Kiyotaki and Moore (2002) put it, “The matching models are without doubt ingenious and beautiful. But it is quite hard to integrate them with the rest of macroeconomic theory —not least because they jettison the basic tool of our trade, competitive markets.”

To pursue the analogy, the setup in Lagos-Wright (2005) allows one to bring competitive markets back on board, in a way that can make monetary theory much closer to standard macro, as we show below. And rather than complicating matters, integrating competitive markets and search markets makes the analysis easier. We also believe this is a realistic way to think about economic activity. In reality, there is some activity in our economic lives that is relatively centralized — it is fairly easy to trade, credit is available, we take prices as given, and so forth — which can be well captured by the notion of a competitive market. But there is also much activity that is relatively decentralized — it is not easy to find trading partners, it can be hard to get credit, and so forth — as captured by search theory. One might imagine that there are various alternative ways to integrate search and competitive markets. Here we present one that we think is useful.

3.1 The environment

We now divide each period into two subperiods. In one, agents interact in a decentralized market (DM) with frictions as in the search models discussed earlier. In the other, they interact in a frictionless centralized market (CM) as in standard general equilibrium theory. Sometimes the setup is described by saying the DM convenes during the day and the

CM at night; this story about day and night is not important for the theory, but we sometimes use it when it helps keep the timing straight.¹³ There is one consumption good x in the DM and another X in the CM, although it is easy to have x come in many varieties, or to interpret X as a vector, as in standard GE theory (Rocheteau, Rupert, Shell, & Wright, 2008). For now x and X are produced one-for-one using labor h and H , but this is relaxed later. The implication is that for now the real wage in the CM is $w = 1$.

Preferences in any period, encompassing one DM and CM, are described by a standard utility function $\mathcal{U}(x, h, X, H)$. What is important for tractability, although not for the theory, in general, is quasi-linearity: $\mathcal{U}$ should be linear in either X or H . To be clear, with general preferences, the model requires numerical methods (see Chiu & Molico, 2007b); with quasi-linearity, we can derive many results analytically. Actually, as discussed below, we can use general utility and still get analytic tractability if we assume indivisible labor. For now, we assume divisible labor and take quasi-linearity as the benchmark. Here we assume $\mathcal{U}$ is linear in H , and in fact for now we assume

$$\mathcal{U} = u(x) - c(h) + U(X) - H;$$

later we consider cases where $\mathcal{U}$ is not necessarily separable in (x, h, X) .

If we shut down the CM, these are the same preferences used in Molico, and the models become equivalent. Since the Molico model collapses to Shi-Trejos-Wright when we impose $m \in \{0, 1\}$, and to Kiyotaki-Wright when we further make x indivisible, these ostensibly different environments can be interpreted as special cases of one framework. Faig (2006, 2008) further argued that the alternating-market model and the large household model in Shi (1997a) can be encompassed in a more general setup. We think this is good, but not because we want one all-purpose vehicle for every issue in monetary economics. Rather, we do not want people to get the impression that New Monetarist economics consists of a huge set of mutually inconsistent models. The models reviewed so far, as well as the extensions seen next to incorporate banking, a payment system, and asset markets, all use similar fundamental building blocks, even if some applications make certain special assumptions.¹⁴

¹³ One can also proceed differently without changing basic results. Williamson (2007), for example, assumed both markets are always open and agents randomly transit between them. For some issues, it is also interesting to have more than one round of trade in the DM between meetings of the CM, as in Berentsen, Camera, and Waller (2005) and Ennis (2008), or more than one period of CM trade between meetings of the DM, as in Telyukova and Wright (2008). Chiu and Molico (2006) allowed agents to transit between markets whenever they liked, at a cost, embedding something like the model of Baumol (1952) and Tobin (1956) into general equilibrium where money is essential, but that requires numerical methods.

¹⁴ An assumption not made explicit in early presentations of the model, but clarified by the work of Aliprantis et al. (2006, 2007a,b) is that in the CM agents observe only prices, and not other agents' actions. If they did observe others' actions there is a potential to use triggers, rendering money inessential. Aliprantis et al. (2007b) also described variations on the environment where triggers cannot be used, and hence money is essential, even if agents' actions can be observed in the CM. This was perhaps less of an issue in models with no CM — or perhaps not — since multilateral trade is neither necessary nor sufficient for public observability or communication. Some of these issues are not yet completely settled. For a recent discussion, see Araujo et al. (2010).

In the DM, the value function $V(\cdot)$ would be described exactly by Eq. (7) in the Section 2.3, except for one thing: wherever $\beta V(\cdot)$ appears on the RHS, replace it with $W(\cdot)$, since before going to the next DM agents now get to visit the CM, and $W(\cdot)$ denotes the CM payoff. In particular,

$$W(m) = \max_{X, H, \hat{m}} \{U(X) - H + \beta V(\hat{m})\} \\ \text{st } X = \phi(m - \hat{m}) + H - T,$$

where ϕ is the value of money, or the inverse of the nominal price level, in the CM, and T is a lump-sum tax. Assuming an interior solution (see Lagos & Wright, 2005 for details), we can eliminate H and write

$$W(m) = \phi m - T + \max_X \{U(X) - X\} + \max_{\hat{m}} \{-\phi \hat{m} + \beta V(\hat{m})\}.$$

From this several results are immediate: $W(m)$ is linear with slope ϕ ; $X = X^*$ where $U'(X^*) = 1$; and $\hat{m}$ is independent of wealth $\phi m - T$.

Based on this last result, we should expect (and we would be right) a degenerate $F(\hat{m})$, where everyone takes the same $\hat{m} = M$ out of the CM, regardless of the m they brought in.¹⁵ Using the fact that $F(\cdot)$ is degenerate and $W'(m) = \phi$, and replacing $\beta V(\cdot)$ with $W(\cdot)$, Eq. (7) simplifies rather dramatically to

$$V(m) = W(m) + \alpha \sigma \{u[x(m, M)] - \phi d(m, M)\} + \alpha \sigma \{-c[x(M, m)] + \phi d(M, m)\}. \quad (9)$$

Effectively, the CM here is a settlement subperiod where agents reset their liquidity positions. Without this feature the analysis is more difficult, and we think it is nice to have a benchmark model that is tractable. By analogy, while models with heterogeneous agents and incomplete markets are obviously interesting, it is nice to have the basic neoclassical growth theory with complete markets and homogeneous agents as a benchmark. Since serious monetary theory with complete markets and homogeneous agents is a nonstarter, we need to find another benchmark, and this is our suggestion.

A degenerate distribution is not all we get in terms of tractability. Replacing $\beta V(\cdot)$ with $W(\cdot)$ and using $W'(m) = \phi$, the bargaining solution Eq. (8) reduces to

$$\max [u(x) - \phi d]^\theta [-c(x) + \phi d]^{1-\theta}$$

st $d \leq m$. In any equilibrium the constraint binds (see Lagos & Wright, 2005). Inserting $d = m$, taking the FOC for x , and rearranging, we get $\phi m = g(x)$, where

¹⁵ The fact that $\hat{m}$ is independent of m does not quite imply that all agents choose the same $\hat{m}$. In a version of the model with some multilateral meetings, and auctions instead of bargaining, Galenianos and Kircher (2008) showed that agents are indifferent over $\hat{m}$ in some set, and equilibrium entails a nondegenerate distribution $F(\hat{m})$. This cannot happen in our baseline model.

$$g(x) \equiv \frac{\theta c(x)u'(x) + (1 - \theta)u(x)c'(x)}{\theta u'(x) + (1 - \theta)c'(x)}. \quad (10)$$

This expression may look complicated but it is easy to use, and simplifies a lot in some special cases; for example, $\theta = 1$ implies $g(x) = c(x)$, and real balances paid to the producer ϕm exactly compensate him for his cost. More generally, it says ϕm is determined by the sharing rule:

$$\phi m = \frac{\theta u'(x)}{\theta x'(x) + (1 - \theta)c'(x)} c(x) + \frac{(1 - \theta)c'(x)}{\theta u'(x) + (1 - \theta)c'(x)} u(x).$$

Notice $\partial x / \partial m = \phi / g'(x) > 0$, so bringing more money increases DM consumption, but in a nonlinear way, unless $\theta = 1$ and $c(x) = x$.

We have established $d(m, \tilde{m}) = m$ and $x(m, \tilde{m})$ depends on m but not $\tilde{m}$. Differentiating Eq. (9), we get

$$V'(m) = (1 - \alpha\sigma)\phi + \alpha\sigma\phi u'(x)/g'(x). \quad (11)$$

The marginal benefit of DM money is the value of carrying it into the next CM with probability $1 - \alpha\sigma$, plus the value of spending it on x with probability $\alpha\sigma$. Updating this one period and combining it with the FOC for $\hat{m}$ from the CM, we arrive at

$$\phi_t = \beta\phi_{t+1}[1 + \ell(x_{t+1})], \quad (12)$$

where

$$\ell(x) \equiv \alpha\sigma \left[\frac{u'(x)}{g'(x)} - 1 \right]. \quad (13)$$

The function defined in Eq. (13) is the *liquidity premium*, giving the marginal value of spending a dollar, as opposed to carrying it forward, times the probability $\alpha\sigma$ of spending it. Using the bargaining solution $\phi m = g(x)$ plus market clearing $m = M$, Eq. (12) becomes

$$\frac{g(x_t)}{M_t} = \beta \frac{g(x_{t+1})}{M_{t+1}} [1 + \ell(x_{t+1})]. \quad (14)$$

Equilibrium can be defined as a list including $V(\cdot)$, $W(\cdot)$, $x(\cdot)$, and so on, satisfying the obvious conditions, but Eq. (14) reduces all this to a simple difference equation determining a path for x , given a path for M . Here we focus on stationary equilibria, where x and ϕM are constant (nonstationary equilibria, including sunspot, cyclic, and chaotic equilibria, are studied in Lagos & Wright, 2003). For this to make sense, we impose $M_{t+1} = (1 + \mu)M_t$ with μ constant. Of course, one has to also consider the consolidated monetary-fiscal budget constraint $G = T + \mu\phi M$, where G is government consumption in the CM. But notice that it does not matter for Eq. (14) whether

changes in M are offset by changing T or G . Individuals would of course prefer lower taxes, other things being equal, but this does not affect their decisions about real balances or consumption in our quasi-linear model. Hence we actually do not have to specify how money transfers are accomplished for the purpose of describing equilibrium x and ϕ . In a stationary equilibrium, or steady state, Eq. (14) simplifies to $1 + \mu = \beta[1 + \ell(x)]$.

Before moving to results, we illustrate one aspect of the framework's flexibility by showing what happens if we replace Nash bargaining with Walrasian pricing in the DM.¹⁶ This can be motivated by interpreting agents as meeting in large groups in the DM, rather than bilaterally, and assuming that whether one is a buyer or seller is determined by preference and technology shocks, rather than by whom one meets. It might help to think about labor search models, like Mortensen-Pissarides (1994), which uses bargaining, and Lucas-Prescott (1974), which uses price taking. A standard interpretation of the latter is that workers and firms meet on islands representing "local labor markets," but on each island there are enough workers and firms that it makes sense to take wages parametrically. The same is true in monetary models. Specialization and anonymity can lead to an essential role for money despite agents meeting in large groups.

Assume for now that the shocks determining if an agent is to be a producer or a consumer in the DM are realized after the CM closes. Then we have

$$V(m) = \gamma V^b(m) + \gamma V^s(m) + (1 - 2\gamma)W(m),$$

where γ is the probability of being a buyer and the probability of being a seller (so that we have the same number of each, but this is easy to relax), while $V^b(m)$ and $V^s(m)$ are the payoffs. These payoffs solve

$$\begin{aligned} V^b(m) &= \max \{u(x) + W(m - \hat{p}x)\} \text{ s.t. } \hat{p}x \leq m \\ V^s(m) &= \max \{-c(x) + W(m + \hat{p}x)\} \end{aligned}$$

where $\hat{p}$ is the DM price of x in terms of dollars, which obviously is different from the CM price $p = 1/\phi$ in general. One can show the constraint for buyers binds, $\hat{p}x = m$, just like in the bargaining model. Then, market clearing in the DM and optimization imply that, to use Walrasian pricing, simply replace $g(x)$ with $c(x)$ and $\alpha\sigma$ with γ . In particular, the same simple condition $l(x) = i$ in (13) determines the unique stationary monetary equilibrium, as long as in the formula for $l(x) = \alpha\sigma \left[\frac{u'(x)}{g'(x)} - 1 \right]$ we replace $\alpha\sigma$ with γ and $g'(x)$ with $c'(x)$. The results are otherwise qualitatively the same.

¹⁶ The use of price taking instead of bargaining in this model follows Rocheteau and Wright (2005). They also considered price posting with directed search, as did Faig and Huangfu (2007) and Dong (2010a) among others. Other mechanisms people consider include the following: Aruoba et al. (2007) used several alternative (to Nash) bargaining solutions. Dutu et al. (2009), and Galeanois and Kircher (2008) used auctions. Dong and Jiang (2009), Ennis (2008), Faig and Jerez (2006), and Sanches and Williamson (2010) studied pricing with private information. Hu, Kennan, and Wallace (2009) used pure mechanism design. And as we show explicitly in Section 4.3, one can also use price posting with random search.

3.2 Results

We have defined monetary equilibrium in the benchmark model, where money has a desirable role, similar to the role it had in the more primitive search-based models in the previous section. We now discuss some of its properties. To facilitate comparison to the literature, we proceed as follows. Suppose one uses standard methods to price real and nominal bonds between any two meetings of the CM, assuming these bonds cannot be traded in the DM (say, maybe because they are merely book entries that cannot be transferred between agents, although we are well aware that this deserves much more discussion). Then the real and nominal interest rates r and i satisfy $1 + r = 1/\beta$ and $1 + i = (1 + \mu)/\beta$, where the latter is a version of the standard Fisher equation. Then we can rewrite the steady state condition $1 + \mu = \beta[1 + \ell(x)]$ derived above as

$$\ell(x) = i. \quad (15)$$

In the Walrasian version of the model, the same condition holds, except in the formula for $\ell(x) = \alpha\sigma[u'(x)/g'(x) - 1]$ we replace $\alpha\sigma$ with γ and $g'(x)$ with $c'(x)$.

Notice Eq. (15) equates the marginal benefit of liquidity to its cost, given by the nominal interest rate, as is standard. In what follows we assume $i > 0$, although we do consider the limit $i \rightarrow 0$ (it is not possible to have $i < 0$ in equilibrium). A stationary monetary equilibrium, or steady state, is *almost* any solution $x > 0$ to Eq. (15). We say *almost* because this condition is really just the FOC for the CM choice of $\hat{m}$, and in principle one needs to check the SOC to be sure we have a maximum, and when there are multiple solutions we have to be sure we pick the global maximum. The existence of a solution to $\ell(x) = i$ is immediate given standard assumptions like $u'(0) = \infty$, and if $\ell(x)$ is monotone then $\ell'(x) < 0$ at the solution, which means it is unique and satisfies the SOC. In this case, there exists a unique stationary monetary equilibrium. Unfortunately, however, $\ell(x)$ is not generally monotone.¹⁷ Still, one can establish, as in Wright (2010), that there is generically a unique stationary monetary equilibrium even if $\ell(x)$ is not monotone. Basically this is because, even if there are multiple local maximizers solving Eq. (15), generically only one of them constitutes a global maximizer for the underlying CM problem.

This establishes the existence and uniqueness of stationary monetary equilibrium. In terms of welfare and policy implications, the first simple observation is that it is equivalent here for policymakers to target either the money growth rate or the inflation rate, since both are equal to μ ; or they can target the nominal interest rate i , which is tied to μ through the Fisher equation. Second, it is clear that the initial stock of money M_0 is irrelevant for the real allocation (money is neutral), but the growth rate μ is not

¹⁷ Under some additional assumptions one can show $\ell(x)$ is monotone. One such assumption is $\theta \approx 1$. Another is that $c(x)$ is linear and $u(x)$ displays decreasing absolute risk aversion. In the version with Walrasian pricing, it is monotone if $c(x)$ is convex and $u(x)$ is concave.

(money is not superneutral). These are properties shared by many monetary models, including typical overlapping-generations, cash-in-advance, and money-in-the-utility-function constructs. Next, since $\ell'(x) < 0$ in equilibrium, (14) implies $\partial x / \partial i < 0$. Hence DM output is unambiguously decreasing in i , because i represents the cost of participating in monetary exchange or, in other words, because inflation is a tax on DM activity. Since CM output $X = X^*$ is independent of i in this basic setup, total output is decreasing in i . However, X is not generally independent of i if we allow nonseparable utility (see Section 3.5).

One can also show that x is increasing in bargaining power θ . And one can show $x < x^*$ for all $i > 0$, and in fact, $x = x^*$ if and only if $i = 0$ and $\theta = 1$.¹⁸ The condition $i = 0$ is the Friedman rule, and is standard, while $\theta = 1$ is a version of the Hosios (1990) condition describing how to efficiently split the surplus. This latter condition is specific to monetary theory with bargaining. To understand it, note that in general there is a holdup problem in money demand analogous to the usual problem with ex ante investments and ex post negotiations. Thus, agents make an investment when they acquire cash in the CM, which pays off in single-coincidence meetings in the DM since it allows them to trade. But if $\theta < 1$ producers capture some of the gains from trade, leading agents to initially underinvest in $\hat{m}$. The Hosios condition tells us that investment is efficient when the payoff to the investor is commensurate with his contribution to the total surplus, which in this case means $\theta = 1$, since it is the money of the buyer (not the seller) that allows the pair to trade.

There is reason to think that this is important in terms of quantitative and policy analysis, and not merely a technical detail. To make the case, first consider the typical quantitative exercise using something like a cash-in-advance model, without other explicit frictions, where one asks about the welfare cost of fully anticipated inflation. If as is standard we measure this cost by asking agents what fraction of consumption they would be willing give up to go from, say, 10% inflation to the Friedman rule, the answer is generally very low. There are many such studies, but we can summarize them accurately by saying that consumers would be willing to give up around 1/2 of 1%, or perhaps slightly more, but rarely above 1%, of their consumption. See Cooley and Hansen (1989) for a representative paper, Lucas (2000) for a somewhat different analysis, or Craig and Rocheteau (2008) for a survey. This has led many economists to conclude that the distortion introduced by inflation is not large.

Why is the distortion implied by those models so small? It seems hard to reconcile with the aversion many politicians and regular people seem to have to inflation. The

¹⁸ The argument is straightforward, if slightly messy. First compute $g'(x)$ and check $u'(x^*) < g'(x^*)$, which by Eq. (13) means $\ell(x^*) < 0$. Hence, $x < x^*$. We can actually say more. One can show $x < \bar{x}$ where $\bar{x}$ solves $u'(x) = g'(x)$, and $\bar{x} < x^*$ unless $\theta = 1$. In fact, $\bar{x}$ is the x that maximizes a buyer's surplus, $u(x) - \phi \hat{m} = u(x) - g(x)$, which we use below.

intuition is actually straightforward. In the standard cash-in-advance or other reduced-form model, at the Friedman rule we get the first best. Hence, by the envelope theorem, the derivative of welfare with respect to i is 0 at the Friedman rule, and a small inflation matters little. This is indeed consistent with what one finds in our benchmark model when we set $\theta = 1$ and calibrate other parameters using standard methods. But if $\theta < 1$ then the envelope theorem does not apply, since while $i = 0$ is still optimal it is a corner solution (remember that $i < 0$ is not feasible). Hence, the derivative of welfare is not 0 at $i = 0$, and a small deviation from $i = 0$ has a first-order effect. The exact magnitude of the effect depends on parameter values, but in calibrated versions of the model it can be an order of magnitude bigger than the cost found in reduced-form models. These results lead New Monetarists to rethink the previously conventional wisdom that anticipated inflation does not matter much.

One should look at the literature for all of the details, but we can sketch the basic method here. Assume $U(X) = \log(X)$, $u(x) = Ax^{1-a}/(1-a)$, and $c(x) = x$. Then calibrate the parameters as follows. First set $\beta = 1/(1+r)$ where r is the average real rate in the data (which data and which real rate are interesting issues). In terms of arrival rates, we can at best identify $\alpha\sigma$, so normalize $\alpha = 1$. In fact, it is not that easy to identify $\alpha\sigma$, so for simplicity set σ to its maximum value of $\sigma = 1/2$, although this is not very important for the results. We need to set bargaining power θ , as discussed below. Then, as in Cooley and Hansen (1989), Lucas (2000), and virtually all other quantitative monetary models, we set the remaining parameters A and a to match the so-called money demand observations, which means the empirical relationship between i and the inverse of velocity, M/PY .

The relationship between M/PY and i is interpreted as money demand by imagining agents setting real balances M/P proportional to income Y , with a factor of proportionality that depends on the opportunity cost i . Here, with $U(X) = \log(X)$, real CM output is $X^* = 1$ (a normalization), and so nominal CM output is $PX = 1/\phi$. Nominal DM output is $\alpha\sigma M$, since in every single-coincidence meeting M dollars change hands. Hence, total nominal output is $PY = 1/\phi + \alpha\sigma M$. Using $\phi M = g(x)$, we get

$$\frac{M}{PY} = \frac{g(x)}{1 + \alpha\sigma g(x)}, \quad (16)$$

and since x is decreasing in i , so is M/PY . This is the money demand curve implied by theory.¹⁹ Given θ , $g(x)$ depends on preferences, and we can pick the parameters a and A of $u(x)$, by various methods, to fit Eq. (16) to the data (assuming, for simplicity, say, that each observation corresponds to a stationary equilibrium of the model, although

¹⁹ In another guise, holding M and P constant and plotting the same relationship in (Y, i) space, it becomes the LM curve from undergraduate Keynesian economics.

one can also do something more sophisticated). Roughly speaking, average M/PY identifies A , and the elasticity wrt i identifies a .

To do this one has to choose an empirical measure of M , which is typically $M1$. People have tried other measures, and it does make a difference (as it would in any model of money, with or without microfoundations). One might think a more natural measure would be M_0 based on a narrow interpretation of the theory, but this may be taking the model too literally. In any case, this empirical research program is ongoing, and some of the modeling approaches used to incorporate financial intermediation and alternative assets into the benchmark model (see Sections 5 and 6) are potentially useful in matching the theory with measurement.

This describes how one can quantify the benchmark model. The only nonstandard parameter is bargaining power θ , which does not show up in theories with price taking, and so we spend some time on it. A natural target for calibrating θ is the markup, price over marginal cost, since it seems intuitive that this should convey information about buyers' bargaining power. One can compute the average markup implied by the model using standard formulae as in Aruoba, Waller, and Wright (2009) and set θ so that this number matches the data. In terms of data, evidence discussed by Faig and Jerez (2005) from the Annual Retail Trade Survey describes markups across retailers as follows. At the low end, in Warehouse Clubs, Superstores, Automotive Dealers, and Gas Stations, markups range between 1.17 and 1.21; and at the high end, in Specialty Foods, Clothing, Footware, and Furniture, they range between 1.42 and 1.44. Aruoba et al. (2009) targeted 1.3, right in the middle of these data. Lagos and Wright (2005), used 1.1, as one might see in other macro applications (e.g., Basu & Fernald 1997). However, in this range, the exact value of θ turns out to not matter too much.

It is now routine to compute the cost of inflation. What is the final answer? It is hard to summarize all the results with one number, since the exact results depend on many factors, such as the sample period, frequency (monthly, quarterly, or annual), whether one includes complications like capital or fiscal policy, and so on. However, it is safe to say that Lagos and Wright (2005) can get agents to willingly give up 5% of consumption to eliminate a 10% inflation, which is an order of magnitude larger than previous findings. In the model with capital presented in Section 3.4, Aruoba et al. (2009) reported findings closer to 3%, which is still quite large. There are many recent studies using variants of the benchmark model that come up with similar numbers (again see Craig & Rocheteau, 2008). Two points to take away from this are the following. First, inflation may well be more costly than most economists used to think. Second, getting into the details of monetary theory, which in this application means thinking about search and bargaining, can make a big difference for quantitative as well as qualitative work.

3.3 Unanticipated inflation

So far we have been concerned only about fully anticipated inflation; we now describe one way to introduce aggregate shocks.²⁰ Suppose the money supply is given by $M_t = \zeta_t M_{t-1}$, where we now include time subscripts explicitly, and $\zeta_t = 1 + \mu_t$ in the earlier notation. Assume ζ_t is i.i.d., drawn from some distribution G . Also, suppose that at the start of the DM at each date t , agents receive a perfect signal about the value of ζ_t to be implemented in the CM later that period, which in general affects ϕ_t . However, when they chose $\hat{m}_t$ in the CM at t they do not know ζ_{t+1} . Then the CM problem is as before, except we replace $\beta V(\hat{m}_t)$ with $\beta E_t V_{t+1}(\hat{m}_t)$. Thus, the relevant FOC becomes

$$\phi_t = \beta E_t V'_{t+1}(\hat{m}_t). \quad (17)$$

In the DM, at $t + 1$, upon observing ζ_{t+1} , buyers are holding $\hat{m}_t$ and cannot increase it, as they might like to do when inflation is higher than expected. Here we must get into a technicality that comes with Nash bargaining. It turns out that the surplus of the buyer $u(x) - \phi \hat{m} = u(x) - g(x)$ is not globally increasing in x ; typically there is some $\bar{x}$ satisfying $u'(\bar{x}) = g'(\bar{x})$ where the surplus is maximized, and $\bar{x} < x^*$ unless $\theta = 1$ (see Aruoba, Rocheteau, & Waller, 2007 for more discussion). Hence, if a buyer has more than required to buy $\bar{x}$ he would rather not bring it all to the bargaining table. This is not a problem in the deterministic case, since agents never choose $\hat{m}$ to purchase more than $\bar{x}$; now, however, it could be that the realized ζ_{t+1} and ϕ_{t+1} are sufficiently low that buyers can afford more than $\bar{x}$. In this case we assume that they leave some of their cash “at home” before going shopping in the DM.²¹ In any case, we assume buyers after seeing ζ_{t+1} decide how much money to take shopping, which is in real terms denoted z . Letting $\bar{z} = g(\bar{x})$, nominal expenditure in the DM is

$$d_{t+1} = \begin{cases} \hat{m}_t & \text{if } \phi_{t+1} \hat{m}_t < \bar{z} \\ \bar{z} / \phi_{t+1} & \text{if } \phi_{t+1} \hat{m}_t \geq \bar{z} \end{cases}$$

Given i.i.d. shocks, it makes sense to look for a stationary equilibrium where real balances are constant: $\phi_t M_t = z \forall t$. This implies $\phi_t / \phi_{t+1} = \zeta_{t+1}$ and

$$d_{t+1} = \begin{cases} \bar{z} \zeta_{t+1} / \phi_t & \text{if } \zeta_{t+1} < \phi_t \hat{m}_t / \bar{z} \\ \hat{m}_t & \text{if } \zeta_{t+1} \geq \phi_t \hat{m}_t / \bar{z} \end{cases}$$

²⁰ Although there are many ways one could apply this extension, we do not do much here other than present it, in the spirit of using the *Handbook* as a teaching tool. As with many of the subsections, one could skip this and move on to more substantive material without much loss in continuity.

²¹ This is not meant to be a big deal, and we could proceed differently, but here we are following earlier models where agents sometimes leave something behind when they go to the DM. See Geromichalos, Licari, and Lledo (2007); Lagos and Rocheteau (2008); and Lester, Postlewaite, and Wright (2009). The issue could be avoided if we set $\theta = 1$, or we use an alternative pricing mechanism, like proportional instead of Nash bargaining, or Walrasian price taking, since in these cases buyers' surplus is globally increasing in m .

Therefore we can write

$$\begin{aligned}
 E_t V_{t+1}(\hat{m}_t) &= \alpha \sigma \int_0^{\phi_t \hat{m}_t / \bar{z}} [u(\bar{x}) + W_{t+1}(\hat{m}_t - \bar{z} \zeta_{t+1} / \phi_t)] dG(\zeta_{t+1}) \\
 &\quad + \alpha \sigma \int_{\phi_t \hat{m}_t / \bar{z}}^{\infty} [u(x_{t+1}) + W_{t+1}(0)] dG(\zeta_{t+1}) \\
 &\quad + \alpha \sigma E_t [-c(x_{t+1}^s) + W_{t+1}(\hat{m}_t + d_{t+1}^s)] + (1 - 2\alpha \sigma) E_t W_{t+1}(\hat{m}_t),
 \end{aligned} \tag{18}$$

where x_{t+1}^s and d_{t+1}^s are the terms of trade when selling, which as above do not depend on the seller's money. Indeed, the bargaining solution is still given by

$$g(x_{t+1}) = \phi_{t+1} \hat{m}_t = z / \zeta_{t+1}.$$

Using this, we can differentiate Eq. (18) and insert $E_t V'_{t+1}(\hat{m}_t)$ into Eq. (17) to get

$$1 + r = \alpha \sigma \int_{z/\bar{z}}^{\infty} \left[\frac{u'(x_{t+1})}{g'(x_{t+1})} - 1 \right] \frac{dG(\zeta_{t+1})}{\zeta_{t+1}} + E_t \left(\frac{1}{\zeta_{t+1}} \right). \tag{19}$$

To find the equilibrium, simply solve Eq. (19) for z . In fact, note that no-arbitrage implies the following version of the Fisher equation for our stochastic economy,

$$1 + i_t = \frac{1 + r}{E_t(1/\zeta_{t+1})}, \tag{20}$$

where $1 + r = 1/\beta$. Given this, Eq. (19) can be rewritten

$$i_{t+1} \zeta_t E_t \left(\frac{1}{\zeta_{t+1}} \right) = \int_{z/\bar{z}}^{\infty} \ell(x_{t+1}) dG(\zeta_{t+1}) \tag{21}$$

where $\ell(x)$ is the marginal benefit of liquidity defined in Eq. (13). In the stochastic version, agents still equate the cost and benefit of liquidity at the margin, but since they need to take expectations Eq. (21) replaces Eq. (15).

Also, in the stochastic economy we need to be a more careful with central bank policy, since setting the nominal rate i is not the same as pinning down a path for M . That is, a given i is consistent with many different stochastic processes for money growth, as long as the average return on cash $E_t(1/\zeta_{t+1})$ satisfies Eq. (20). Nevertheless, it is not hard to verify that the Friedman rule, $i_t = 0$ for all t , is optimal, and that it still achieves the first best iff $\theta = 1$. But there can be many paths for M_t that are consistent with $i_t = 0$ for all t . See Lagos (2009) for an in-depth analysis of these issues. We return now to the effects of fully anticipated inflation.

3.4 Money and capital

Because of worries about the theory being “removed” from mainstream macro, we sketch the extension that includes investment and fiscal policy in Aruoba et al. (2009). For

simplicity, we ignore long-run technical change (see Waller, 2010). Also, in this version, capital K is a factor of production, but it does not compete with M as a medium of exchange. To motivate this, one can assume K is not portable, making it hard to trade directly in the DM, but of course this does not explain why claims to capital cannot circulate. On the one hand, this is no different from the result that agents in the DM cannot trade claims to future income: this is precluded by imperfect commitment and monitoring. On the other hand, if capital trades in the CM, one can imagine certified claims on K that might also circulate in the DM. We think monetary theorists do not yet have a definitive stance on this issue, but one approach is to introduce additional informational frictions. It would suffice, for example, to assume counterfeit claims to K can be costlessly produced, and are not recognizable in the DM, even if they are in the CM. Then agents will not accept claims to K in the DM, and M must serve as the medium of exchange.²²

Assume the CM technology produces output $f(K, H)$ that can be allocated to consumption or investment, while the DM technology is represented by a cost function $c(x, k)$ that gives an agent's disutility of producing x when he has k , where lower (upper) case denotes individual (aggregate) capital. The CM problem is

$$\begin{aligned} W(m, k) = \max_{X, H, \hat{m}, \hat{k}} \{ & U(X) - H + \beta V(\hat{m}, \hat{k}) \} \\ \text{st } x = & \phi(m - \hat{m}) + w(1 - t_h)H + [1 + (\rho - \Delta)(1 - t_k)]k - \hat{k} - T, \end{aligned} \quad (22)$$

where ρ is the rental rate, Δ the depreciation rate, and we incorporate income taxes in the CM. The FOC for $(X, \hat{m}, \hat{k})$ are

$$\begin{aligned} U'(X) &= \frac{1}{(1 - t_h)w} \\ \frac{\phi}{w(1 + t_h)} &= \beta V_1(\hat{m}, \hat{k}) \\ \frac{1}{w(1 - t_h)} &= \beta V_2(\hat{m}, \hat{k}). \end{aligned} \quad (23)$$

Generalizing what we found in the baseline model, $(\hat{m}, \hat{k})$ is independent of (m, k) , and W is linear with $W_1(m, k) = \phi/w(1 - t_h)$ and $W_2(m, k) = [1 + (\rho - \Delta)(1 - t_k)]/w(1 - t_h)$.

In the DM, instead of assuming that agents may be consumers or producers depending on who they meet, we now proceed as follows. After the CM closes, as

²² This line is not especially elegant, but seems logically consistent. Lester et al. (2009, 2010) attempted to take the idea more seriously, following models of money and private information like Williamson-Wright (1994) or Berentsen-Rocheteau (2004), and earlier suggestions by Freeman (1989), but it raises technical challenges. A promising route has been proposed by Rocheteau (2009) (see also Li & Rocheteau, 2009, 2010). Alternatively, Lagos and Rocheteau (2008) allowed K and M to both be used as media of exchange, and show M can still be essential if K is not sufficiently productive or the need for liquidity is great, although in that model K and M must pay the same return in equilibrium.

discussed earlier, we assume agents draw preference and technology shocks determining whether they can consume or produce, with γ denoting the probability of being a consumer and of being a producer. Then the DM opens and consumers and producers are matched bilaterally. This story helps motivate why capital cannot be used for DM payments: one can say that it is fixed in place physically, and consumers have to travel without their capital to producers' locations to trade. Thus, producers can use their capital as an input in the DM but consumers cannot use their capital as payment. With preference and technology shocks, the equations again look exactly the same as when we had random matching and specialization except γ replaces $\alpha\sigma$. Also, it is possible under this interpretation to easily replace Nash bargaining with Walrasian pricing, which allows us to quantify the holdup problems.

Using bargaining for now, one can again show $d = m$, and that the Nash outcome depends on the consumer's m but not the producer's M , and on the producer's K but not the consumer's k . Abusing notation slightly, $x = x(m, K)$ solves $g(x, K) = \phi m / w(1 - t_h)$, where

$$g(x, K) \equiv \frac{\theta c(x, K)u'(x) + (1 - \theta)u(x)c_1(x - K)}{\theta u'(x) + (1 - \theta)c_1(x, K)}$$

generalizes Eq. (10). Then we have the following version of Eq. (9)

$$\begin{aligned} V(m, k) &= W(m, k) + \gamma \left\{ u[x(m, K)] - \frac{\phi m}{w(1 - t_h)} \right\} \\ &\quad + \gamma \left\{ \frac{\phi M}{w(1 - t_h)} - c[x(M, k), k] \right\}. \end{aligned}$$

Differentiating this, then inserting V_1 and V_2 , market clearing $k = K$ and $m = M$, and equilibrium prices $\phi = w(1 - t_h)g(x, K)/M$, $\rho = f_1(K, H)$, and $w = f_2(K, H)$, into the FOC, we have

$$U'(X_t) = \frac{1}{(1 - t_h)f_2(K_t, H_t)} \quad (24)$$

$$\frac{g(x_t, K_t)}{M_t} = \frac{\beta g(x_{t+1}, K_{t+1})}{M_{t+1}} \left[1 - \gamma + \gamma \frac{u'(x_{t+1})}{g_1(x_{t+1}, K_{t+1})} \right] \quad (25)$$

$$\begin{aligned} U'(X) &= \beta U'(X_{t+1}) \{ 1 + [f_1(K_{t+1}, H_{t+1}) - \Delta](1 - t_k) \} \\ &\quad - \beta \gamma \left[c_2(x, K) - c_1(x, K) \frac{g_2(x, K)}{g_1(x, K)} \right]. \end{aligned} \quad (26)$$

And we have the resource constraint

$$X_t + G = f(K_t, H_t) + (1 - \Delta)K_t - K_{t+1}. \quad (27)$$

Equilibrium is defined as (positive, bounded) paths for $\{x, X, K, H\}$ satisfying Eqs. (24)–(27), given monetary and fiscal policy, plus an initial condition K_0 . As a special case, in non-monetary equilibrium we have $x = 0$ while $\{X, H, K\}$ solves the system ignoring Eq. (25) and setting the last term in Eq. (26) to 0. Those conditions are exactly the equilibrium conditions for $\{X, H, K\}$ in the standard nonmonetary growth model described, for example, in Hansen (1985).²³ So we nest standard real business cycle theory as a special case. In monetary equilibria, we get something even more interesting. The last term in Eq. (26) generally captures the idea that if a producer buys an extra unit of capital in the CM, his marginal cost is lower in the DM for a given x , but x increases as an outcome of bargaining. This is a holdup problem on investment, parallel to the one on money demand discussed earlier. With a double holdup problem there is no value of θ that delivers efficiency, which has implications for the model's empirical performance and welfare predictions.

Aruoba et al. (2009) calibrate the model with bargaining and with price taking and compare the quantitative predictions. Interestingly, although the bargaining version generates a somewhat bigger welfare cost of inflation, the price-taking version generates much bigger effects of monetary policy on investment. Intuitively this is because K in the bargaining version is relatively low and unresponsive to what happens in the DM due to the holdup problem. That is, the returns to investing accrue mostly from CM trade, since the seller has to split with the buyer whatever surplus arises from having more K in the DM. This makes K unresponsive to taxing DM trade via inflation. In the price-taking version the effects of inflation on K are big compared to what has been found in earlier work, because with no holdup problem, the returns to investing are affected by taxing DM trade. One can put this model to many other uses, such as quantifying the impact of these holdup problems.

We do not have space to go into all the numerical results, but we do want to emphasize the methodological point that it is not hard to integrate modern monetary theory and mainstream macro. The only quantitative result we mention is this. In case one wonders what fraction of output is produced in the DM, it is easy to see the answer is less than 10%. To verify this, note the following: Since there are γ buyers in the DM each period, and they each spend M , the share of total output produced in the DM is $\gamma M / PY = \gamma / \nu$, where $\nu = PY / M$ is velocity. If M is measured by $M1$ then ν is around 5 in annual data, and since $\gamma \leq 1/2$, we are done. For actual calibrated values of γ , the share is slightly less than this upper bound. Of course if we change

²³ At least, in the deterministic version of Hansen (1985), but at this stage it is not hard to add technology and other shocks, as in Aruoba (2009), Aruoba and Schorfheide (2010), or Telyukova and Visschers (2009).

the frequency (from annual to quarterly, e.g.) PY changes, but so does the calibrated value of γ , keeping the DM share about the same. This would not work in standard cash-in-advance models, where agents always spend all their money each period. This is important because it shows that details, like stochastic trading opportunities, as well as the two-sector structure, matter, even though 90% of output here is produced in a CM that looks exactly like standard neoclassical growth theory.

3.5 The long-run Phillips curve

In the baseline model, without capital, we saw that DM output is decreasing in anticipated inflation, while CM output is independent of anticipated inflation. It is not true that CM output is independent of anticipated inflation in the model with capital in the previous section, because we assumed K enters $c(x, K)$. If this is not the case, and $c_K(x, K) = 0$, then the last term in Eq. (26) vanishes, K drops out of Eq. (25), and the system dichotomizes: we can independently solve Eq. (25) for the DM allocation x and the other three equations for the CM allocation (X, K, H) . In this dichotomous case, monetary policy affects x but not (X, K, H) . This is why we assumed K enters $c(x, K)$. In this section, without capital, we break the dichotomy using nonseparable utility. In fact, here we take the Phillips curve literally, and model the relation between inflation and unemployment. To make this precise, first, we introduce another friction to generate unemployment in the CM, and second, we re-cast the DM as a pure exchange market, so that unemployment is determined exclusively in the CM.

To give some background, a principle explicated in Friedman (1968) is that, while there may exist a Phillips curve trade-off between inflation and unemployment in the short run, there is no trade-off in the long run. The *natural rate* of unemployment is defined as “the level that would be ground out by the Walrasian system of general equilibrium equations, provided there is embedded in them the actual structural characteristics of the labor and product markets” (although, as Lucas, 1980 noted, Friedman was “not able to put such a system down on paper”). Friedman (1968) said monetary policy cannot engineer deviations from the natural rate in the long run. However, he tempered this view in Friedman (1977) where he said

There is a natural rate of unemployment at any time determined by real factors. This natural rate will tend to be attained when expectations are on average realized. The same real situation is consistent with any absolute level of prices or of price change, provided allowance is made for the effect of price change on the real cost of holding money balances.

Here we take this real balance effect seriously.

Of the various ways to model unemployment, in this presentation we adopt the indivisible labor model of Rogerson (1988).²⁴ This has a nice bonus feature: we do

²⁴ The approach follows Dong (2010b) and Rocheteau, Rupert, and Wright (2007). Alternatively, Berentsen et al. (2010) and Liu (2009) used the unemployment theory in Mortensen and Pissarides (1994).

not need quasi-linearity, because in indivisible-labor models agents act *as if* utility were quasi-linear. To make the point, we revert to the case where X is produced one-for-one with H , but now $H \in \{0, 1\}$ for each individual. Also, as we said, to derive cleaner results we use a version where there is no production in the DM. Instead, agents have an endowment $\bar{x}$, and gains from trade arise due to preference shocks. Thus, DM utility is $v^j(x, X, H)$ where j is a shock realized after (X, H) is chosen in the CM. Suppose $j = b$ or s with equal probability, where $\partial v^b(\cdot)/\partial x > \partial v^s(\cdot)/\partial x$, and then in the DM everyone that draws b is matched with someone that draws s . The indices b and s indicate which agents will be buyers and sellers in matches, for obvious reasons. We also assume here that there is discounting between one DM and the next CM, but not between the CM and DM, but this is not important. What is interesting is nonseparability in $v^j(x, X, H)$.

As in any indivisible labor model, agents choose a *lottery* $(\ell, X_1, X_0, \hat{m}_1, \hat{m}_0)$ in the CM where ℓ is the probability of working $H = 1$, while X_H and $\hat{m}_H$ are CM purchases of goods and cash conditional on H (if one does not like lotteries, the equilibrium can also be supported using pure Arrow-Debreu contingent commodity markets, as in Shell & Wright 1993). There is no direct utility generated in the CM; utility is generated by combining (X, H) with x in the DM. Hence,

$$\begin{aligned} W(m) &= \max_{\ell, X_1, X_0, \hat{m}_1, \hat{m}_0} \{ \ell V(\hat{m}_1, X_1, 1) + (1 - \ell) V(\hat{m}_0, X_0, 0) \} \\ \text{st } 0 &\leq \phi m - \ell \phi \hat{m}_1 - (1 - \ell) \phi \hat{m}_0 + w\ell - T - \ell X_1 - (1 - \ell) X_0. \end{aligned} \quad (28)$$

As is well known, X and $\hat{m}$ depend on H , in general, but if V is separable between X and H then $X_0 = X_1$, and if V is separable between $\hat{m}$ and H then $\hat{m}_1 = \hat{m}_0$. But the function V is endogenous. This is another argument for making the role of money explicit, instead of, say, simply sticking it in the utility function: one cannot simply *assume* V is separable (or homothetic or whatever), one has to *derive* its properties, and this imposes discipline on both theory and quantitative work.²⁵

Letting λ be the Lagrangian multiplier for the budget constraint, FOC for an interior solution are

$$0 = V_2(\hat{m}_H, X_H, H) - \lambda, \text{ for } H = 0, 1 \quad (29)$$

$$0 = V_1(\hat{m}_H, X_H, H) - \lambda \phi, \text{ for } H = 0, 1 \quad (30)$$

$$0 = V(\hat{m}_0, X_0, 0) - V(\hat{m}_1, X_1, 1) + \lambda(X_1 - X_0 - 1 + \phi \hat{m}_1 - \phi \hat{m}_0) \quad (31)$$

$$0 = \ell - \ell X_1 - (1 - \ell) X_0 + \phi[m + \gamma M - \ell \hat{m}_1 - (1 - \ell) \hat{m}_0]. \quad (32)$$

One can guarantee $\ell \in (0, 1)$, and show the FOC characterize the unique solution, even though the objective function is not generally quasi-concave (Rocheteau et al.,

²⁵ This point is played up in Aruoba and Chugh (2008), in the context of optimal tax theory, where properties of $V(\cdot)$ can matter a lot for the results.

2007). Given $V(\cdot)$, Eqs. (29)–(31) constitute five equations that can be solved under weak regularity conditions for $(X_1, X_0, \hat{m}_1, \hat{m}_0, \lambda)$, independent of ℓ and m . Then Eq. (32) can be solved for individual labor supply as a function of money holdings at the start of the period, $\ell = \ell(m)$. Notice $\hat{m}_H$ may depend on H , but not m , and hence we get at most a two-point distribution in the DM. Also, $W(m)$ is again linear, with $W'(m) = \lambda\phi$. This is what we meant earlier when we said that agents act as if they had quasi-linear preferences in the model with indivisible labor and lotteries.

In DM meetings, for simplicity we assume take-it-or-leave-it offers by the buyer ($\theta = 1$). Also, although it is important to allow buyers' preferences to be nonseparable, we do not need this for sellers, so we make their preferences separable. Then as in the baseline model, the DM terms of trade do not depend on anything in a meeting except the buyer's m : in equilibrium, he pays $d = m$, and chooses the x that makes the seller just willing to accept, independent of the seller's (X, H) . In general, buyers in the DM who were employed or unemployed in the CM get a different x since they have different m . In any case, we can use the methods discussed above to describe $V(\cdot)$, differentiate it, and insert the results into Eqs. (29)–(31) to get conditions determining $(x_1, x_0, X_1, X_0, \lambda)$. From this we can compute aggregate employment $\bar{\ell} = \ell(M)$.

It is now routine to see how endogenous variables depend on policy. First, it is easy to check $\partial x / \partial i < 0$, since as in any such model the first-order effect of inflation is to reduce DM trade. A calculation then implies that the effect on unemployment depends on the cross derivatives of buyers' utility function as follows:

1. if $v^b(x, X, H)$ is separable between (X, H) and x , then $\partial \bar{\ell} / \partial i = 0$
2. if $v^b(x, X, H)$ is separable between (x, X) and H , then $\partial \bar{\ell} / \partial i > 0$ iff $v_{Xx}^b < 0$
3. if $v^b(x, X, H)$ is separable between (x, H) and X , then $\partial \bar{\ell} / \partial i > 0$ iff $v_{xH}^b < 0$

The economic intuition is simple. Consider case 2. Since inflation reduces x , if x and X are complements then it also reduces X , and hence reduces the $\bar{\ell}$ used to produce X ; but if x and X are substitutes then inflation increases X and $\bar{\ell}$. In other words, when x and X are substitutes, inflation causes agents to move from DM to CM goods, increasing CM production and reducing unemployment. A similar intuition applies in Case 3, depending on whether x is a complement or substitute for leisure. In either case, we can get a downward-sloping Phillips curve under simple and natural conditions, without any complications like imperfect information or nominal rigidities. This relation is exploitable by policy makers in the long run: given the right cross derivatives, it is indeed feasible to achieve permanently lower unemployment by running a higher anticipated inflation, as Keynesians used to (still?) think. But it is not optimal: it is easy to check that the efficient policy is still Friedman's prescription, $i = 0$.

3.6 Benchmark summary

We believe this benchmark delivers a lot of insight. A model with only CM trade could not capture the fundamental role of money, which is why one has to resort to

shortcuts like cash-in-advance or money-in-the-utility-function specifications. The earlier work on microfoundations with only DM trade gets at the salient role of money, but requires harsh restrictions or it becomes analytically intractable. There are other devices, including Shi (1997b) and Menzio, Shi, and Sun (2009), which achieve some similar results, but one reason to like this benchmark model is that, in addition to imparting tractability, it integrates search and competitive markets, and this reduces the gap between the microfoundations literature and mainstream macro. Alternating markets do not yield tractability; we also need something like quasi-linearity or indivisibilities. This does not seem a huge price to pay, especially for anyone who uses the indivisible labor model anyway, but we could also dispense with these assumptions if we were willing to rely on numerical methods.²⁶

Before we move to new results, however, we mention a variation by Rocheteau and Wright (2005), since this is something we use in several applications later. This extension considers an environment with two permanently distinct types, called buyers and sellers, where the former are always consumers in the DM and the latter are always producers in the DM. One could not have permanent buyers or permanent sellers in the DM if there were no CM, since no one would produce in one DM if they cannot spend the proceeds in a subsequent DM. Here sellers may want to produce in every DM, since they can spend the money in the CM, and buyers may want to work in every CM, since they need the money for the DM. Monetary equilibrium no longer entails a degenerate distribution, but all sellers choose $m = 0$, while all buyers choose the same $m > 0$. Notice that with two types the distribution of money holdings is degenerate only conditional on type, as we encountered earlier in Section 3.5, but this is still tractable. Indeed, the key property of the model in terms of tractability is that the choice of $\hat{m}$ is history independent, not that it is the same for all agents.

Having two types is interesting for several reasons, including the fact that one can introduce a generalized matching technology, and one can incorporate a participation decision for either sellers or buyers. By way of analogy, Pissarides (2000) had two types (workers and firms), while Diamond (1982) had only one (traders), which allows the former to consider more general matching and entry. Note also that, in a sense, having

²⁶ There are many applications of this model. A sample includes: Aruoba and Chugh (2008), Gomis-Porqueras and Peralta-Alva (2009), Martin (2009), and Waller (2009) studied optimal monetary and fiscal policy. Banks are introduced by Bencivenga and Camera (2008); Berentsen, Menzio, and Wright (2008); Chiu and Meh (2010); He, Huang, and Wright (2008); and Li (2007). Berentsen and Waller (2009) and Boel and Camera (2006) studied the interaction between money and bonds. Andolfatto (2010a,b); Berentsen and Monnet (2008); Hoerova, Monnet, and Temzelides (2007); and Kahn (2009) discussed details of monetary policy implementation. Guerrieri and Lorenzoni (2009) analyzed the effects of liquidity on business cycles. Lagos and Rocheteau (2005); Lui, Wang, and Wright (2010); and Nosal (2010) studied how velocity (or the time it takes to spend one's money) depends on inflation. These last applications are also relevant for the following reason. One sometimes hears that anything one can do with a search-based theory could be replicated with a cash-in-advance or money-in-the-utility-function specification. That is definitely *not* the case in these papers, which are concerned mainly about the effect of inflation on search behavior (as is true of some papers in the first-generation, including Li 1994, 1995).

two types makes the model similar to the models presented in Section 2 with $m \in \{0, 1\}$. And there are many applications where two types just seems more natural. Actually, for all of this, we do not really need permanently distinct types: it would be equivalent to have types determined each period, as long as the realization occurs before the CM closes — the important distinction concerns whether agents can choose $\hat{m}$ conditional on type. This would be the case, for example, if we took the model at the end of Section 3.1, with preference and technology shocks in the DM replacing random matching, but alternatively assumed the realizations of these shocks were known before agents chose $\hat{m}$.

4. NEW MODELS OF OLD IDEAS

Although one of our goals is to survey existing models, we also want to present new material. In this section we lay out some new models of ideas in earlier Monetarist or Keynesian traditions. This shows how similar results can be derived in our framework, although sometimes with interesting differences. We first introduce additional informational frictions to show how signal extraction problems can lead to a short-run Phillips curve, as in Old Monetarist economics. Then we analyze what happens when prices are sticky, for some unspecified reason, as in Keynesian models. Then we give a New Monetarist spin on sticky prices with some very different implications. As discussed in the introduction, there are some papers in New Monetarist economics that already explore some of these issues, with embellishments that allow one to take the theories to the data. The goal here is to come up with simple models to illustrate basic qualitative properties, although we also discuss a few empirical implications.

4.1 The Old Monetarist Phillips curve

Here we discuss some ideas about the correlations defining the short-run Phillips curve, and the justification for predictable monetary policy, in Old Monetarist economics. Given that we already discussed a model where unemployment appears explicitly in Section 3.5, we now for simplicity take the Phillips curve to mean a positive relation between money growth or inflation, on the one hand, and output, on the other hand. Also, we use the setup where there are two distinct types called buyers and sellers. In particular, there is a unit mass of agents, half buyers and half sellers. Further, during a period CM trade occurs first, followed by DM trade, and we sometimes describe the CM and DM subperiods as the day and night markets to keep track of the timing. Finally, to yield clean results we sometimes use $u(x) = \log x$.²⁷

We already studied a certain type of unanticipated inflation in Section 3.3, but in order to build a model in the spirit of Lucas (1972), we now include both real and

²⁷ Many applications of the general framework assume $u(0) = 0$, for technical reasons; we do not need this because we assume $\theta = 1$ in the bargaining solution below.

monetary shocks. First, some fraction of the population is inactive each period: a fraction ω_t of buyers participates in both markets in period t , while the fraction $1 - \omega_t$ rests. As well, a fraction ω_t of sellers will not participate in the DM of period t and in the CM of period $t + 1$. Assume that ω_t is a random variable, and realizations are not publicly observable. Second, the money growth rate μ_t is random, and realizations are not publicly observable. So that agents have no direct information on the current money injection, only indirect information coming from prices, we add some new actors to the story. We call them government agents, and assume that in the CM in each period t , a new set of such agents appears. They have linear utility $X - H$, and can produce X one-for-one with H . If $\mu_t > 0$, the central bank prints money and gives it to these agents, and they collectively consume $\phi_t M_{t-1} \mu_t$, and if $\mu_t < 0$ they retire money by collectively producing $-\phi_t M_{t-1} \mu_t$. Their role is purely a technical one, designed to make signal extraction interesting.

In the CM, agents learn last period's money stock M_{t-1} and observe the price ϕ_t , but not the current aggregate shocks ω_t and μ_t . For an individual buyer acquiring money in the CM, the current value of money may be high (low), either because the demand for money is high (low) or because money growth is low (high). To ease the presentation, assume take-it-or-leave-it offers by buyers in the DM, $\theta = 1$, and assume that a seller's cost function is $c(h) = h$. This implies

$$x_t = \beta m_t E[\phi_{t+1} | \phi_t]. \quad (33)$$

An active buyer's FOC from the CM reduces by the usual manipulations to

$$-\phi_t + \beta E[\phi_{t+1} | \phi_t] u'(x_t) = 0. \quad (34)$$

Given that the mass of buyers is $1/2$, market clearing implies

$$\omega_t m_t / 2 = (1 + \mu_t) M_{t-1}. \quad (35)$$

If μ_t were a continuous random variable, in principle we could solve for an equilibrium as in Lucas (1972). For illustrative purposes, however, we adopt the approach in Wallace (1992), using a finite state space (see also Wallace, 1980). To make the point, it suffices to assume μ_t and ω_t are independent i.i.d. processes, where μ_t is μ_1 or $\mu_2 < \mu_1$ each with probability $1/2$, and ω_t is ω_1 or $\omega_2 < \omega_1$ each with probability $1/2$. We then assume that

$$\frac{\omega_1}{1 + \mu_1} = \frac{\omega_2}{1 + \mu_2}, \quad (36)$$

so that agents cannot distinguish between high money demand and high money growth, on the one hand, or low money demand and low money growth, on the other.

Using Eqs. (33)–(35) we obtain closed-form solutions for prices and quantities. Let $\phi(i, j)$ and $q(i, j)$ denote the CM price and the DM quantity when $(\mu_t, \omega_t) = (\mu_i, \omega_j)$. Then

$$\phi(i, j) = \frac{\omega_j}{2(1 + \mu_i)M_{t-1}}, \text{ for } i = 1, 2 \quad (37)$$

$$q(i, j) = \frac{\beta(\omega_1 + \omega_2)(2 + \mu_1 + \mu_2)}{4(1 + \mu_1)(1 + \mu_2)\omega_j}, \text{ for } (i, j) = (1, 2), (2, 1) \quad (38)$$

$$q(1, 1) = q(2, 2) = \frac{\beta(\omega_1 + \omega_2)^2(2 + \mu_1 + \mu_2)}{8(1 + \mu_1)(1 + \mu_2)\omega_1\omega_2}. \quad (39)$$

Let total output in the day and night be $Q^d(i, j)$ and $Q^n(i, j)$ in state (μ_i, ω_j) . Given $\mu_1 > \mu_2 \geq 0$, we have

$$Q^d(i, j) = \phi_i M_t = \omega_j / 2, \quad (40)$$

for $i, j = 1, 2$ from Eq. (37). Further, from Eqs. (36), (38), and (39),

$$Q^n(1, 2) = Q^n(2, 1) = \frac{\beta(\omega_1 + \omega_2)(2 + \mu_1 + \mu_2)}{8(1 + \mu_1)(1 + \mu_2)} \quad (41)$$

$$Q^n(1, 1) = \frac{\beta(\omega_1 + \omega_2)^2(\mu_1 + \mu_2)}{16\mu_1\mu_2\omega_2} \quad (42)$$

$$Q^n(2, 2) = \frac{\beta(\omega_1 + \omega_2)^2(2 + \mu_1 + \mu_2)}{16(1 + \mu_1)(1 + \mu_2)\omega_1}. \quad (43)$$

Total real output is $Q(i, j) = Q^d(i, j) + Q^n(i, j)$. From Eq. (40), Q^d depends only on the real shock. That is, when the number of active buyers is high (low), money demand is high (low), and the price of money is high (low). Thus, active buyers collectively produce more (less) in the day to acquire money when the number of active buyers is high (low). And at night, one can show that $Q^n(2, 2) < Q^n(1, 2) = Q^n(2, 1) < Q^n(1, 1)$.

Figure 2.1 displays the scatterplot of aggregate output Q against money growth μ , using time series observations generated by the model. The four dots represent money and output in each of the four states, depicting a clear positive correlation between μ and Q . This results from agents' confusion, since if there were full information about the shocks we would have

$$Q^n(i, j) = \frac{\beta(\omega_1 + \omega_2)(2 + \mu_1 + \mu_2)}{8(1 + \mu_1)(1 + \mu_2)} \text{ for all } (i, j)$$

as in Figure 2.2. Confusion results from the fact that, if money growth and money demand are both high (low), then agents' subjective expectation of ϕ_{t+1} is greater (less) than the objective expectation, so more (less) output is produced in the DM than under full information. Except for technical details, this non-neutrality of money is essentially that in Lucas (1972) and Wallace (1980, 1992).

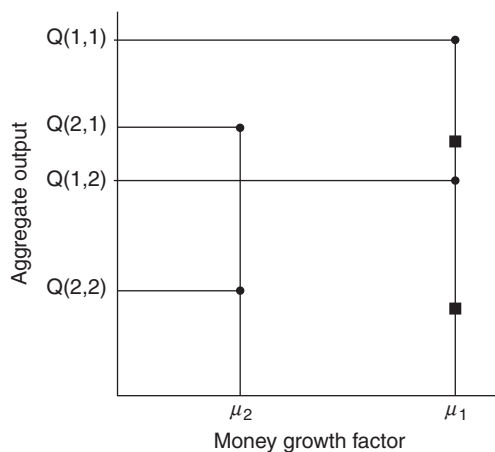


Figure 2.1 Imperfect information.

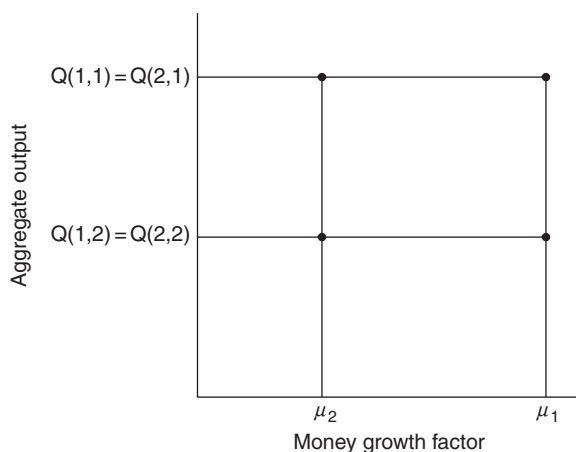


Figure 2.2 Perfect information.

A standard narrative associated with ideas in Friedman (1968) and Lucas (1972, 1976) is that 1960s and 1970s macroeconomic policy erred because policymakers treated the dots in (their empirical version of) Figure 2.1 as capturing a structural relationship between money growth and output. Policymakers took for granted that more output is good and more inflation is bad, and they took the observed correlation as evidence that if the central bank permanently increased money growth this would achieve permanently higher output. Although we saw in Section 3.5 that permanent trade-offs are a theoretical possibility, the point to be emphasized is that observed empirical relations by no means constitute evidence that there is an actual trade-off. What happens in

this example if we permanently set money growth to μ_1 ? The data points we would generate would be the two squares in Figure 2.1, with high (low) output when money demand is high (low). Rather than increasing output, higher inflation lowers output in all states of the world.

What is optimal policy? If we can find a monetary policy rule that achieves $x = x^*$ in all states, it is optimal. From Eq. (34), we require $\phi_t = \beta E[\phi_{t+1}]$, from which we can obtain

$$1 + \mu_{t+1} = \frac{\beta \omega_{t+1}}{\omega_t}. \quad (44)$$

This is the Friedman rule, dictating that the money supply decrease on average at the rate of time preference, with higher (lower) money growth when money demand is high (low) relative to the previous period. It might appear hard for the monetary authority to implement such a rule, because it seems to require that they know the shock ω_t . However, all we need is $\phi_{t+1} = \phi_t/\beta$, so they need not observe the shock, and can attain efficiency simply by engineering a constant rate of deflation. In equilibrium, the price level is predictable, and carries no information about the aggregate state. It is not necessary for the price level to reveal aggregate information, since efficiency requires that buyers acquire the same real balances in the CM and receive the same quantity in the DM, independent of the shocks.

In a sense, these results are consistent with the thrust of Friedman (1968) and Lucas (1972). Monetary policy can confuse price signals, and this can result in a non-neutrality that generates a Phillips curve. However, the policy prescription derived from the model is in line with Friedman (1969) rather than Friedman (1968): the optimal money growth rate is not constant, and should respond to aggregate real disturbances to correct intertemporal distortions. This feature of the model appears consistent with some of the reasons that money growth targeting by central banks failed in practice in the 1970s and 1980s. Of course we do not intend the model in this section to be taken literally. It is meant mainly as an example to illustrate once again, but here in the context of our benchmark framework, the pitfalls of naive policymaking based on empirical correlations that are incorrectly assumed to be structural.²⁸

4.2 New Keynesian sticky prices

We now modify our benchmark model to incorporate sticky prices, capturing ideas in New Keynesian economics along the lines of Woodford (2003) and Clarida et al. (1999). We will first construct a cashless version, like Woodford (2003), where all

²⁸ Faig and Li (2009) have a more general quantitative analysis of signal extraction and the cost of unanticipated inflation. They find that the welfare cost of signal extraction is very low. They also find the cost of anticipated inflation is fairly low, but note that they use Walrasian pricing and not Nash bargaining in their DM.

transactions are carried out using credit, then modify it to include currency transactions. New Keynesian models typically use monopolistic competition, where individual firms set prices, usually according to a Calvo (1983) mechanism. Here, to fit into our benchmark model, we assume that some prices are sticky in the DM. Again we use the version with permanently distinct buyer and seller types, with the mass of each set to $1/2$, and set $c(h) = h$.

In the cashless model, in spite of the fact that money is not held or exchanged, prices are denominated in dollars. Sticky price modelers do not usually attempt any justification for this, other than stating that they observe this. We follow in that tradition in this section. As in the benchmark model, the price of money in the CM, ϕ_t , is flexible. In the DM, each buyer-seller pair conducts a credit transaction where goods are received by the buyer in exchange for a promise to pay in the next CM. To support these credit transactions we assume that there is perfect memory or record keeping in every meeting. That is, if a buyer defaults on an obligation, it is observed and an exogenous legal system imposes a severe punishment. Thus, in equilibrium, all borrowers pay their debts.

In the DM, suppose that in an individual match the terms of trade between a buyer and seller are either flexible with probability $1/2$, or fixed with probability $1/2$. In a flexible match, the buyer makes a take-it-or-leave-it offer. Let $1/\psi_t$ be the number of dollars a buyer offers to pay in the following CM for each unit produced by a flexible-price seller in the DM, and s_t^1 be the quantity of goods produced by the seller. Then the bargaining outcome satisfies $s_t^1 = \beta s_t^1 \phi_{t+1} / \psi_t$, so that $\psi_t = \beta \phi_{t+1}$. Now, assume that in each fixed-price exchange in the DM, the seller is constrained to offering a contract that permits buyers to purchase as much as they like in exchange for $1/\psi_{t-1}$ dollars in the next CM per unit purchased. In a flexible price contract, the buyer chooses $s_t^1 = x^*$. However, in a fixed-price contract, the buyer chooses the quantity s_t^2 to maximize $u(s_t^2) - s_t^2 \phi_{t+1} / \phi_t$, which gives

$$u'(s_t^2) = \phi_{t+1} / \phi_t. \quad (45)$$

So far there is nothing to determine the sequence $\{\phi_t\}_{t=0}^{\infty}$. In Woodford (2003), one solution approach involves first determining the price of a nominal bond. In our model, in the CM of period t the price z_t in units of money of a promise to pay one unit of money in the CM during period $t + 1$ is given by

$$z_t = \beta \phi_{t+1} / \phi_t. \quad (46)$$

Following Woodford one could then argue that z_t can somehow be set by the central bank, perhaps in accordance with a Taylor rule. Then, given determinacy of z_t , we can solve for $\{\phi_t\}_{t=0}^{\infty}$ from Eq. (46). It seems consistent with New Keynesian logic to consider $\{\phi_t\}_{t=0}^{\infty}$ as an exogenous sequence of prices that can be set by policy. In terms of what matters, it is equivalent to say that government sets the path for the inflation rate, $\pi_t = \phi_{t-1} / \phi_t$.

Since $s_t^1 = x^*$, the path for inflation is irrelevant for s_t^1 , but from Eq. (45) s_t^2 is increasing in π_{t+1} . In fixed-price transactions, buyers write a credit contract under which the nominal payment in the CM is determined by the flexible-price contract from the previous period. When inflation increases, the implicit real interest rate on credit in fixed-price contracts falls, and the buyer purchases more. Note that, when the buyer in a fixed-price meeting at t repays the loan in period $t + 1$, he produces $s_t^2/\beta\pi_{t+1}$. Generally, the effect of inflation depends on preferences, but if we set $u(x) = \log x$, then CM production is invariant to the path of π_t , and the only component of aggregate output affected by inflation is production in fixed-price DM meetings. From Eq. (45), $s_t^2 = \pi_{t+1}$, so there is a short- and long-run Phillips curve: a temporarily higher rate of inflation increases output temporarily, and a permanently higher rate increases it permanently. The model predicts that the Phillips curve exists in the data and can be exploited by policy. Should policy exploit this? No. Equilibrium is generally inefficient due to sticky prices, and this shows up in a suboptimal quantity of output in fixed-price contracts. For efficiency, we require that $s_t^2 = x^*$ which implies from Eq. (45) $\phi_t = \phi$, which means 0 inflation. Further, from Eq. (46), the optimal nominal bond price consistent with price stability is $z_t = \beta$, the “Wicksellian natural rate.”

To get money to play a role, assume a fraction α of meetings are non-monitored in the DM, so the seller does not have access to the buyer’s history, and anything that happens in the meeting is private information to the pair.²⁹ Further, assume the same set of sellers engage in non-monitored meetings for all t . The remaining fraction $1 - \alpha$ of DM meetings is monitored, as in the cashless economy: the seller observes the buyer’s history and their interaction is public information. The buyer and seller continue to be matched into the beginning of the next day, before the CM opens, so default is publicly observable, and we continue to impose punishments that preclude default. The CM, where money and goods are traded, opens in the latter part of the day, and here only prices (not individual actions) are observable. As with credit transactions, half of the money transactions have flexible and half have fixed prices. The type of meeting (monitored or non-monitored, flexible-price or fixed-price) is determined at random, but a buyer knows in the CM what type of meeting he will have in the following DM.

As in the cashless model, the quantities of goods traded in flexible-price and fixed-price credit transactions are s_t^1 and s_t^2 , with $s_t^1 = x^*$ and s_t^2 determined by Eq. (45). For flexible-price transactions where there is no monitoring and money is needed, the buyer carries m_t^1 from the CM to the DM and makes a take-it-or-leave-it offer, which involves giving up all the money for

²⁹ This setup has a superficial resemblance to reduced-form models with cash goods and credit goods (Lucas & Stokey, 1987), just like the baseline model has a resemblance to simple cash-in-advance models. This is as it should be, since reduced-form models were designed to be descriptive of reality, but it should be clear that there are ingredients in the models presented here that are not in those models.

$$x_t^1 = \beta \phi_{t+1} m_t^1, \quad (47)$$

so the implicit flexible price of goods in terms of money is $1/\beta \phi_{t+1}$. In a fixed-price money transaction, the seller must charge a price equal to the flexible money price in the previous period. Therefore, a buyer in a fixed-price money transaction carries m_t^2 into the meeting and spends it all to get x_t^2 , where

$$x_t^2 = \beta \phi_t m_t^2. \quad (48)$$

As buyers choose money balances optimally in the daytime, we obtain the following FOC for buyers in monetary flexible-price and fixed-price transactions, respectively:

$$-\phi_t + \beta \phi_{t+1} u'(x_t^1) = 0 \quad (49)$$

$$-\phi_t + \beta \phi_t u'(x_t^2) = 0. \quad (50)$$

Assume that money is injected by the government by lump-sum transfers to sellers during the day, and that M grows at rate μ . In equilibrium, the entire money stock must be held by buyers at the end of the day who will be engaged in monetary transactions at night. Thus, we have the equilibrium condition

$$\frac{\alpha}{2} (m_t^1 + m_t^2) = M_t \quad (51)$$

Now, consider the equilibrium where $1/\phi_t$ grows at the rate μ and all real quantities are constant for all t . From Eq. (45) and Eqs. (47)–(51), equilibrium quantities are

$$\begin{aligned} s_t^1 &= x^* \\ u'(s_t^2) &= 1/(1 + \mu) \\ u'(x_t^1) &= (1 + \mu)/\beta \\ u'(x_t^2) &= 1/\beta. \end{aligned}$$

In equilibrium the money growth rate is equal to the inflation rate, and higher money growth increases output in fixed-price relative to flexible-price transactions.

From a policy perspective, we cannot support the efficient allocation $s_t^i = x_t^i = x^*$ for $i = 1, 2$. However, we can maximize the weighted average welfare criterion

$$W(\mu) = \frac{\alpha}{2} [u(x_t^1) - x_t^1 + u(x_t^2) - x_t^2] + \frac{(1 - \alpha)}{2} [u(s_t^1) - s_t^1 + u(s_t^2) - s_t^2].$$

Then we have

$$W'(\mu) = \frac{\alpha}{2\beta u''(x_t^1)} \left(\frac{1 + \mu}{\beta} - 1 \right) - \frac{(1 + \alpha)}{2(1 + \mu)^2 u''(s_t^2)} \left(\frac{1}{1 + \mu} - 1 \right). \quad (52)$$

From Eq. (52) one can check that the optimal money growth rate is between the Friedman rule and a constant price level. This reflects a trade-off between two

distortions: inflation distorts the relative price of flexible- and fixed-price goods, which is corrected by price stability; and inflation results in the standard intertemporal distortion, in that too little of the flexible-price good is purchased with cash, which is corrected by the Friedman rule. We are not the first to point this out (Aruoba & Schorfheide, 2010 provide references to the literature); we simply recast this trade-off in terms of our New Monetarist model.

What do we learn from this? A central principle of New Monetarism is that it is important to be explicit about the frictions underlying the role for money and related institutions. What do models with explicit frictions tell us that New Keynesian models do not? One line of argument in Woodford (2003) argued that it was sufficient to use a cashless model to analyze monetary policy, and the intertemporal monetary distortions corrected by the Friedman rule are secondary to sticky price considerations. Further, he argues that one can construct monetary economies that behave essentially identically to the cashless economy, so that it is sufficient to analyze the cashless limit. This cashless limit is achieved here if we let $\alpha \rightarrow 0$. In the model, quantities traded in different types of transactions are independent of α , and the only effects of changing α are on the price level and the fraction of credit trades. As well, the optimal money growth rate tends to rise as α decreases, with $\mu^* \rightarrow 0$ as $\alpha \rightarrow 0$.

So while we can construct explicitly a cashless limit in our model, it is apparent to us that confining policy analysis to the cashless economy is not innocuous. A key feature of equilibrium in our model is that the behavior of prices is tied to the aggregate money stock, in line with the quantity theory of money. Thus the model with both cash and credit gives the central bank control over a monetary quantity, not direct control over market interest rates, prices, or inflation. In reality, central banks intervene mainly through exchanges of their liabilities for other assets and by lending to financial institutions. Though central banks may conduct such interventions to target an interest rate, it seems important to model accurately the means by which this is done. How else could one evaluate, for example, whether it is preferable in the short run for the central bank to target a short-term nominal interest rate or the growth rate in the money stock?

Moreover, we have to emphasize that it is important to be open minded, *ex ante*, concerning which frictions are relevant for policy, and recall from Section 3.2 that New Monetarist models predict that quantitatively the cost of inflation can be quite high. Aruoba and Schorfheide (2010) built a full-fledged model incorporating both New Keynesian rigidities and elements of our New Monetarist framework, and estimated it using Bayesian methods to explicitly compare the two channels identified above, what they called the Friedman channel, and the New Keynesian channel (inefficiency generated by sticky prices and monopolistic competition). They estimate their model under four different scenarios, having to do with whether there is Nash bargaining or Walrasian pricing in the DM, and whether they try to fit the short- or long-run

elasticity of money demand. In the version with bargaining designed to fit the short-run elasticity, despite a reasonably sized New Keynesian friction, the Friedman rule turns out to be optimal after all. The other three versions yield optimal inflation rates around -1.5 , -1 , and -0.75% . Even considering parameter uncertainty, they never find an optimal inflation rate very close to 0, and conclude that the two channels are about equally important. Moreover, microfoundations matter for this: in a similar model, except that money demand is generated by putting M in the utility function, 0 inflation is close to optimal.

So while one can build nominal rigidities into our model and examine cashless limits, we are not at all convinced that it is harmless to ignore monetary matters or to sweep all of the frictions other than sticky prices under the carpet. Further, we are generally uncomfortable with sticky-price models even when there are explicit costs to changing prices. The source of these menu costs is typically unexplained, and once one opens the door to such costs of adjustment it seems that one should consider many other similar types of costs in the model if we are to take them seriously. Again, our motivation for presenting a New Keynesian sticky-price model is mainly to show that if one thinks it is desirable to have nominal rigidities in a model, this is not inconsistent with being relatively explicit about the exchange process or the role of money and related institutions.

4.3 New Monetarist sticky prices

Temporarily leaving aside qualms about exactly how one introduces stickiness into the model, we have to admit that it is desirable to do so, for the simple reason that stickiness seems to be a feature of reality. How can New Monetarists — or Old Monetarists or New Classicists or anyone else — ignore this? Indeed, it is apparent to us that this is one of the main driving forces, if not *the* main force that makes Keynesians Keynesian. Consider Ball and Mankiw (1994), who we think are fairly representative. As they put it, “We believe that sticky prices provide the most natural explanation of monetary non-neutrality since so many prices are, in fact, sticky.” Moreover, “based on micro-economic evidence, we believe that sluggish price adjustment is the best explanation for monetary non-neutrality.” And “As a matter of logic, nominal stickiness requires a cost of nominal adjustment.” *Fait accompli*.

But healthy science has to be willing to challenge and confront all aspects of theory, even fundamental canons like those passed down by Ball and Mankiw (1994). To show one way to potentially confront the sticky-price issue, here we sketch the recent analysis by Head, Liu, Menzio, and Wright (2010). What they show is that some natural models generate nominal price stickiness endogenously, as a result, and not an assumption. These models seem consistent not just with the broad observation that prices are, in fact, sticky, but also with some of the more detailed micro evidence discussed next. Yet, as we will soon see, such models have policy implications that are very different

from those of Keynesian economics. That is, these models predict that sticky prices can emerge without Calvo (1983) pricing, Mankiw (1985) costs, or other such devices, and yet these models are consistent with monetary neutrality. And they certainly do not imply that Keynesian monetary policy prescriptions are either feasible or desirable.³⁰

Consider the benchmark New Monetarist model with one change: we swap out the Nash bargaining module for price setting by sellers as in Burdett and Judd (1983). The Burdett-Judd model has every seller posting a price p taking as given the distribution of other prices, say $F(p)$, and then buyers search for prices in the sense of sampling from $F(p)$. What prevents the distribution from collapsing to a single price, as in Diamond (1970), is that buyers generally get to sample more than one draw from $F(p)$. Although there are many ways to set this up, let us assume here that the representative buyer gets to see n prices with probability α_n . Also, assume for simplicity that they each want to buy 1 unit of an indivisible good, and that each seller can satisfy any demand at cost c per unit. What drives Burdett-Judd pricing is this: Suppose all sellers charge $\bar{p}$; then any buyer that samples more than one seller will pick one at random; this gives any individual seller an incentive to shade down to $\bar{p} - \varepsilon$. In the end, equilibrium must have a nondegenerate $F(p)$. Quite naturally, sellers posting high p make more per unit, while sellers posting low p earn less per unit but make it up on the volume, so that in equilibrium their profits are the same.³¹

Taking as given for now the price distribution, the DM value function for a buyer can be written

$$V(m) = W(m) + \sum_n \alpha_n \int_p^m (u - \phi p) dJ_n(p) \quad (53)$$

where J_n is the distribution of the lowest p sampled from $F(\cdot)$ given $n \geq 1$ draws. When a buyer samples $n > 1$ prices, he obviously buys at the lowest one, generating a distribution of transactions prices (those actually paid, as opposed to posted) denoted by $J(p)$, which generally differs from $F(p)$. For ease of presentation, from now on we assume $\alpha_n = 0$ for $n \geq 3$. The distribution of transactions prices in this case is simply

³⁰ This model presented in this section, while based on Head et al. (2010), has antecedents in Head and Kumar (2005) and Head, Kumar, and Lapham (2008). The idea is obviously also related to earlier work by Caplin and Spulber (1987), although their model is really very different, as are some of the implications.

³¹ One reason to work with the Burdett-Judd model is that it can generate price dispersion even without inflation, since the original version is a nonmonetary model. This is consistent with the observation that we see price dispersion in the data even during periods when inflation was very low (see e.g., Campbell & Eden 2007). That observation is a problem for Calvo pricing models, since the only reason for dispersion in the baseline version of that model is inflation: all firms set p in nominal terms and are only allowed to adjust it at random times, so that at any point during an inflation some (who got to adjust recently) will have a price above others (who did not). Without inflation all sellers charge the same price. Of course there are other ways to generate price dispersion. But Burdett-Judd seems reasonable, is certainly tractable, and can be generalized along many interesting dimensions. Additionally, we like that similar search-type frictions are at the heart of what makes money essential and what drives price dispersion.

$$J(p) = \frac{\alpha_1 F(p) + \alpha_2 \{1 - [1 - F(p)]^2\}}{\alpha_1 + \alpha_2}.$$

One can also define the distribution of prices posted in real terms $H(z)$, where $z = \phi p$, as well as the distribution of real transactions prices.

For the same reason trade is monetary in all the models presented earlier, sellers in this model post prices in nominal terms (in dollars), since it is dollars that buyers must trade for goods. So posting nominal prices is natural, although of course they could post in other units, like the number of dollars needed to buy X in the next CM. In any case, profit from posting p is

$$\Pi(p) = (\phi p - c)b\{\alpha_1 + 2\alpha_2[1 + F(p)]\}, \quad (54)$$

where b is the buyer-seller ratio. Notice the number of units sold is the measure of buyers who show up with no other option, $b\alpha_1$, plus the measure who show up with a second option that is not as good, $2b\alpha_2[1 - F(p)]$. This multiplied by $\phi p - c$ is profit in real terms. Let $\mathcal{F}$ be the support of the price distribution. Then profit maximization means:

$$\Pi(p) = \bar{\Pi} \forall p \in \mathcal{F} \text{ and } \Pi(p) \leq \bar{\Pi} \forall p \notin \mathcal{F}. \quad (55)$$

It is standard to show in Burdett-Judd models that the distribution can have no mass points, and $\mathcal{F} = [\underline{p}, \bar{p}]$ is an interval. At the upper bound, profit is

$$\Pi(\bar{p}) = \bar{\Pi} = (\phi \bar{p} - c)b\alpha_1 \quad (56)$$

since the highest price seller only serves customers with no other option. Combining Eqs. (56) and (54), we can immediately solve for the closed form of the price distribution,

$$F(p) = 1 - \frac{\alpha_1}{2\alpha_2} \left(\frac{\phi \bar{p} - \phi p}{\phi p - c} \right). \quad (57)$$

To get the bounds, simply note that $\bar{p} = M$, assuming all buyers choose the same $\hat{m} = M$ in the CM, as in the benchmark model, and solve $F(p) = 0$ for

$$\underline{p} = \frac{\alpha_1 \phi \bar{p} + 2\alpha_2 c}{(\alpha_1 + 2\alpha_2)\phi}.$$

From this one easily gets the real distribution $H(z)$, given the CM price level $1/\phi$.

Consider a stationary equilibrium where all real variables, including distributions, are constant while all nominal variables grow at the same rate as M . We need to satisfy two conditions: given $\phi = z/M$, the distributions are as constructed in the previous equations; and given the distributions, z solves a version of our benchmark CM problem (see below). One can also generalize the model to allow entry by buyers into the DM,

at some participation cost. This determines the buyer-seller ratio b , therefore we can determine the arrival rates α_n endogenously through a standard matching technology, which is of interest for reasons discussed next.

This is textbook Burdett-Judd, except that we are in a monetary economy, which raises a slight complication. There are typically many equilibria in models with fiat money, price posting, and indivisible goods, for reasons related to coordination, and one needs some sort of refinement to make things determinate.³² Since any possible equilibrium is qualitatively the same, for our purposes, and we do not want to get into refinement issues here, we simply select the equilibrium that satisfies

$$i = \alpha_1 H'(\bar{z})(u - \bar{z}). \quad (58)$$

This seems the natural analog to the unique stationary monetary equilibrium in our benchmark model, as Eq. (58) equates the marginal cost of carrying a dollar to the benefit, which is the probability of sampling a price which in real terms is $\bar{z}$, times the surplus $u - \bar{z}$. One can show that an equilibrium of this form exists for any nominal rate below some threshold.

What happens in equilibrium? Although the distribution of real prices $H(z)$ is pinned down, individual sellers do not care where they are in the support of that distribution, since all $p \in \mathcal{F}$ earn equal profit. As we said, it is natural to imagine sellers posting prices in nominal terms, not because a dollar is some abstract unit of account, but because it is a medium of exchange. What happens when M increases? In a stationary equilibrium ϕ decreases, and since the real distribution $H(z)$ is invariant, the nominal distribution $F(p)$ shifts to the right. But for any seller that was at t charging $p_t \in \mathcal{F}_t$, when M_t increases to M_{t+1} and $\mathcal{F}_t$ shifts to $\mathcal{F}_{t+1}$, as long as p_t is still in $\mathcal{F}_{t+1}$ there is no incentive to raise the price. Sure, profit per sale goes down, but he makes it up on the volume. He could change to some other $p_{t+1} \in \mathcal{F}_{t+1}$, and some sellers typically must change, because we need the right number of sellers at each p to keep the same real distribution (see Head et al., 2010 for details). But many sellers with prices posted in nominal terms may not bother to adjust in any period. Thus sticky prices emerge as an equilibrium outcome, even though we let sellers adjust whenever they want, at no cost.

Many sellers not adjusting nominal prices even as the aggregate price level rises is exactly what Ball and Mankiw (1994) correctly claim to observe in the real world (although they were evidently wrong to think this implies we need menu costs in models as a matter of logic). The model is consistent with this, but also with many other observations. Consider this list of facts that people think are noteworthy:³³

³² See Jean et al. (2010). We can of course relax the assumption of indivisible goods, and the results go through, but this increases the algebra and raises other issues, like whether sellers post a price, a price-quantity pair, a price-quantity schedule, and so forth. So here we keep goods indivisible.

³³ Klenow and Malin (2010) in Chapter 6 of this *Handbook* emphasize facts 1, 2, 4, 6, 7, and 8. Nakamura and Steinsson (2008) cover facts 3 and 5. Both also provide many other references.

1. Prices change slowly, with a median frequency of adjustment between 4 and 7 months, or 8 and 10 months, depending on details.
2. The frequency of price changes varies a lot across goods.
3. The size of prices changes varies a lot across goods.
4. All sellers that change prices at a point in time do not all change to the same price.
5. About one-third of price changes are reductions even during general inflation.
6. Hazard rates for price changes are flat or declining, with an eventual spike.
7. Many price changes are quite small.
8. Frequency of price changes is positively related to inflation.

The New Monetarist sticky price model can in principle match all of these observations, although only time will tell just how well. But it is already known that other more popular models do not do so well, including the basic Calvo-pricing and menu-cost models. Some parts of this claim are obvious, like the fact that standard (s , S) models predict all sellers should jump to the same price when they do change, in contradiction of item 3 (although we are aware there are “fixes” one can tack on). Other parts of our claim are perhaps less obvious. Consider item 7, the fact that many p changes are small. As Klenow and Kryvtsov (2008) said, this is “hard to reconcile with the large menu costs needed to rationalize large average price changes.” The model presented here has no problem with this.³⁴ There is clearly more work to be done on taking this kind of model to the data and, again, time will tell. But given the success at matching the data with the labor market version of Burdett-Judd, the well-known Burdett-Mortensen (1998) model, there is reason to think it is worth pursuing.

To return to the issue of monetary neutrality and implications for policy, here, we use the extended version of the model where the measure of buyers in the DM b is determined endogenously by an entry condition. First, as we said above, the distribution of real prices is invariant to the price level along the equilibrium path, although of course there are real effects to changing the inflation rate, as in any New Monetarist model. Moreover, a one-time surprise increase in M will be exactly neutral: $F(p)$ shifts up with the aggregate price level, while all real variables, including $H(z)$, b , and so forth, stay the same. This is *very different* from what happens in a Keynesian version of the model, where prices are sticky for Calvo reasons. In such a model, when the surprise increase in M hits, it is not possible for all nominal prices to adjust (in a menu-cost version, it may be possible, but it is not generally going to happen, and the story is

³⁴ Admittedly, at least in part, the reason the model has no problem with some of these observations is that it has a lot of indeterminacy. Still, our main point is that other models do not do very well. For instance, to be precise, define a small price change as less than 5%. Klenow and Kryvtsov (2008) report around 39% of changes are small in the data, and cannot match this in their model. In the Golosov and Lucas (2005) model, which was designed to generate approximate monetary neutrality, less than 10% of price changes are small. Midrigan (2007) can match the observation in question with some effort, but then he loses approximate neutrality. The model here can match the facts easily and is consistent with exact neutrality.

similar). The distribution of nominal prices will not shift the way our model predicts, and the shape of the real price distribution changes.

Generally, in a Keynesian version of the model, after a surprise M increase, buyers will expect lower real prices — there are some real bargains out there with many sellers stuck at low prices. This increases b , and hence output, since there are more buyer-seller matches. Indeed, in the very short run, when Calvo has not yet allowed any seller to adjust, the increase in M lowers all real prices. This sets off a shopping frenzy, which means a production boom, as sellers are obliged to meet demand at the posted prices. We do not go into whether a central bank would want to engineer such a boom here, or whether they could do so systematically over time. Instead we emphasize the following. Suppose we concede the observation that some prices are sticky. We have demonstrated that this does not imply monetary injections are non-neutral, let alone that particular Keynesian policy prescriptions are feasible or desirable. To be clear, the New Monetarist position is not that non-neutralities do not exist, and this chapter contains many examples where obviously money matters (e.g., Section 4.1). Our position is that the observation that prices appear to be sticky in the data does not logically imply that Keynesian models or policy implications are correct.

5. MONEY, PAYMENTS, AND BANKING

In this section we analyze extensions of the benchmark model that incorporate payments arrangements, along the lines of Freeman (1996), and banks, along the lines of Diamond and Dybvig (1983). The goal is to construct environments where outside money is important not only for accomplishing the exchange of goods but for supporting credit arrangements.

5.1 A payments model

For this application we include two types of buyers and two types of sellers. It is convenient to refer to CM meetings as occurring in the day and DM transactions at night. A fraction α of buyers and a fraction α of sellers are type 1 buyers and sellers, respectively, and they meet in the night in non-monitored matches. When a type 1 buyer meets a type 1 seller, they can trade only if the former has money. As well, there are $1 - \alpha$ type 2 buyers and $1 - \alpha$ type 2 sellers, who are monitored at night and hence can trade on credit, which again is perfectly enforced. During the day, we will have a more elaborate set of meetings among agents, with limited participation in the CM. This is slightly complicated, but we think that it is an improvement over some of the models used in the payments literature, including Freeman (1996).

Thus, in the morning of the day, type 1 sellers and type 2 buyers meet in a Walrasian market where money trades for goods at the price ϕ_t^1 , and type 2 buyers can produce. Type 1 buyers and type 2 sellers do not participate in this market. Then, at

mid-day, bilateral meetings occur between type 2 buyers and type 2 sellers who were matched the previous night. This is essentially another DM, but neither buyers or sellers produce, and this market is only an opportunity for the type 2 buyers to settle their debts. Finally, in the afternoon, type 1 buyers meet in a second Walrasian market with type 2 sellers, with the price of money denoted by ϕ_t^2 . Here, type 1 buyers can produce. Neither type 2 buyers or type 1 sellers participate in this second CM. The government can make lump-sum money transfers in the Walrasian markets during the day, so that there are two opportunities to intervene each period. We assume these interventions are lump-sum transfers in equal quantities to sellers. As in the benchmark model, we must have $\phi_t^i \geq \beta\phi_{t+1}^i$, for $i = 1, 2$.

We are interested in an equilibrium where trade occurs as follows. First, in order to purchase goods during the night, type 1 buyers need money, which they acquire in the afternoon Walrasian market. They trade all this money at night for goods, so that type 1 sellers go into the next day with all the money. In the Walrasian market in the next morning, type 2 buyers produce in exchange for the money held by type 1 sellers. Then, at mid-day, type 2 buyers meet type 2 sellers and use money to settle their debts acquired in the previous night. Then, in the second Walrasian market, during afternoon, type 2 sellers exchange money for the goods produced by type 1 buyers. Finally, at night, meetings between type 1 buyers and sellers involve the exchange of money for goods, while meetings between type 2 buyers and sellers are exchanges of IOU's for goods. For clarity, we show agents' itineraries and patterns of trade in Figure 2.3.

In bilateral meetings at night, buyers make take-it-or-leave-it offers. Letting x_t denote the quantity of goods received by a type 1 buyer at night, his optimal choice of money balances yields the FOC

$$-\phi_t^2 + \beta\phi_{t+1}^1 u'(x_t) = 0. \quad (59)$$

To repay the debt that supported the purchase of s_t units of goods, the type 2 buyer must acquire money in Walrasian market 1 at price ϕ_{t+1}^1 , and give it to a type 2 seller, who then exchanges the money for goods in Walrasian market 2 at the price ϕ_{t+1}^2 . Therefore, s_t satisfies the FOC

$$-\phi_{t+1}^1 + \phi_{t+1}^2 u'(s_t) = 0. \quad (60)$$

Let M_t^i denote the quantity of money (post transfer) supplied in the i^{th} Walrasian market during the day, for $i = 1, 2$. Then market clearing in Walrasian markets 1 and 2 implies

$$(1 - \alpha)s_{t-1} = \beta\phi_t^2 M_t^1, \quad (61)$$

$$\alpha x_t = \beta\phi_{t+1}^1 M_t^2. \quad (62)$$

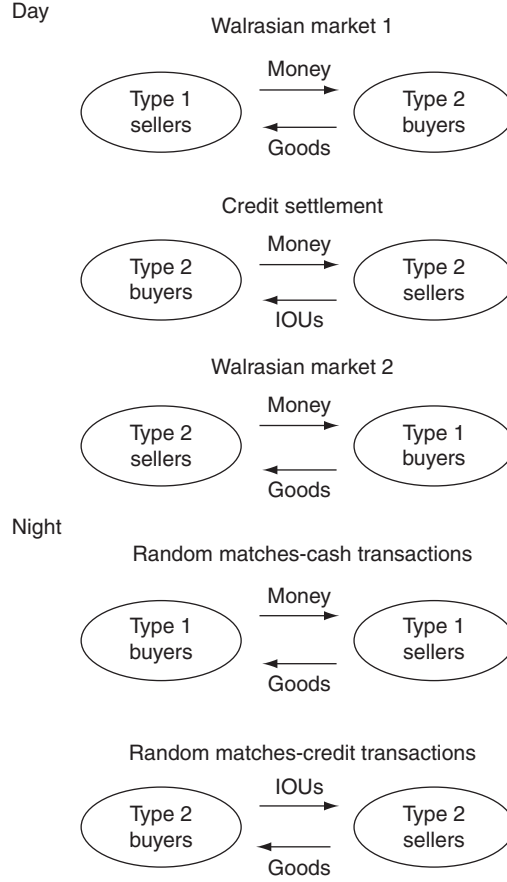


Figure 2.3 Interaction in the payments system model.

To solve for equilibrium, substitute for prices in Eqs. (59) and (60) using Eqs. (61) and (62) to obtain

$$-\frac{\alpha x_t}{M_t^2} + \frac{(1 - \alpha)s_t u'(s_t)}{M_{t+1}^1} = 0 \quad (63)$$

$$-\frac{(1 - \alpha)s_{t-1}}{\beta M_t^1} + \frac{\alpha q_t u'(x_t)}{M_t^2} = 0. \quad (64)$$

Given $\{M_t^1, M_t^2\}_{t=0}^\infty$, we can determine $\{x_t, s_t\}_{t=0}^\infty$ from Eqs. (63) and (64), and then $\{\phi_t^1, \phi_t^2\}_{t=0}^\infty$ can be determined from Eqs. (61) and (62). Note that, in general, intervention in both Walrasian markets matters. For example, suppose that $M_t^1/M_t^2 = 1 + \gamma$ for all t , $M_{t+1}^i/M_t^i = 1 + \mu$, where $\gamma > -1$ and $\mu \geq \beta$ so that the ratio of money in the two

markets is constant for all t and in individual Walrasian markets money grows at a constant rate over time. Further, suppose $u(x) = \ln x$. Then, in an equilibrium where $s_t = s$ for all t and $x_t = x$ for all t , Eqs. (63) and (64) yield

$$x = \frac{(1 - a)}{\alpha(1 + \gamma)(1 + \mu)}$$

$$s = \frac{\alpha\beta(1 + \gamma)}{(1 - \alpha)}$$

A higher money growth rate μ decreases the quantity of goods traded in cash transactions during the night, as is standard. However, a higher γ (relatively more cash in the first Walrasian market) increases the quantity of goods bought on credit and reduces goods bought with cash at night. What is efficient in general? To maximize the total surplus in the two types of trades, we need $x_t = s_t = x^*$. From Eqs. (63) and (64), this gives $\mu = \beta - 1$ and $\gamma = [1 - \alpha(1 + \beta)]/\alpha\beta$. At the optimum, in line with the Friedman rule, money should shrink over time at the rate of time preference, but we also need a monetary injection in the first market that increases with the fraction of credit relative to cash transactions to support the optimal clearing and settlement of debt.

Outside money plays two roles here: it is used as currency in some transactions, and it is used to accommodate credit in other transactions where it is needed to settle debts. This second role is similar to the one played by central bank balances in interbank payments systems, such as Fedwire. In the model, central bank intervention in the morning Walrasian market relative to the afternoon Walrasian market stands in for real-world intervention via daylight overdrafts, while intervention in the afternoon Walrasian market relative to the next Walrasian market is similar to real-world central bank intervention in overnight financial markets. There are two dimensions to monetary policy, and both are important. The optimal policy sets both the intraday nominal interest rate (the nominal interest rate on bonds issued in the morning and paying off in the afternoon Walrasian market) and the overnight rate (the nominal interest rate on bonds issued in the afternoon Walrasian market paying off the next morning) to zero.

It seems clear that it would not be easy to come up with such insights without modeling the details of the exchange process carefully. Although the example is obviously special, it is not contrived. It is meant to capture some of what goes on in actual economies, albeit in an abstract and stylized way. This is a nascent research area, and we think there are many possible applications and extensions of these types of models. Nosal and Rocheteau (2011), Chapman et al. (2008), and the references contained therein provide additional examples and references to other work on payments.

5.2 Banking

We now extend the benchmark model by including banking, in the spirit of Diamond and Dybvig (1983). The original Diamond-Dybvig model appears to have been intended mainly as a model of bank runs and deposit insurance. Subsequent research (see Ennis & Keister, 2009a, 2009b, 2010) has shown that auxiliary assumptions are required to obtain runs, and it is not clear if there is a role for government deposit insurance in this modeling framework. However, what survives is a nice model of financial intermediaries that act to provide insurance against liquidity needs, and they do so by diversifying across liquid and illiquid assets. The model does not capture all important features of banks, such as the fact that they issue liabilities that compete with government currency in transactions. And since it ignores monetary factors, the basic framework cannot be used to address some key features of historical banking panics, like currency shortages and high nominal interest rates (Friedman & Schwartz, 1963).

Champ et al. (1996) provided an attempt to capture these features by integrating Diamond-Dybvig banks into an overlapping generations model of money. But that model is incomplete, and has the unfortunate implication that, at the optimum, the central bank should intermediate all illiquid assets. In this subsection, we build on Champ et al. (1996) in the context of our benchmark model. This is an example of how recent advances in monetary theory allow us to do more than we could in the earlier overlapping generations framework. In the model constructed here, currency and bank liabilities are both used in transactions, and a diversified bank provides risk sharing services that avoid waste. Thus, there is a Diamond-Dybvig risk-sharing role for banks, but banking provides other efficiency gains as well.

We begin with a version of the model with no aggregate uncertainty. Again we refer to the first subperiod with CM exchange as day, and the second with DM exchange as night, and there are α type 1 sellers who engage in non-monitored exchange at night using currency and $1 - \alpha$ type 2 sellers who engage in monitored exchange at night using credit. At night there are α type 1 buyers each matched with a type 1 seller, and $1 - \alpha$ type 2 buyers each matched with a type 2 seller, but a buyer's type is random, revealed at the end of the previous day after production and portfolio decisions are made. There is an intertemporal storage technology that takes goods produced by buyers during the afternoon of the day, and yields R goods per unit invested during the morning of the next day, with $R > 1/\beta$. All buyers and type 1 sellers are together in the Walrasian market that opens during the afternoon of the day, while only type 2 sellers are present during the morning of the day.

First suppose banking is prohibited. To trade with a type 2 seller at night, a buyer needs to store goods during the day before meeting the seller. Since the trade is monitored, the seller is able to verify that a claim to storage offered for goods is valid. To trade with a type 1 seller at night, a buyer needs cash, as in non-monitored trade

sellers do not accept claims to storage. Claims to storage are useless for type 1 sellers, because they do not participate in the morning CM where the storage pays off. Thus, during the afternoon of the day, the buyer acquires nominal money balances m_t and stores k_t units of output and, again assuming take-it-or-leave-it offers at night, solves

$$\max_{m_t, x_t} -\phi_t m_t - k_t + \alpha u(\beta \phi_{t+1} m_t) + (1 - \alpha)[u(\beta R k_t) + \beta \phi_{t+1} m_t].$$

The FOC are

$$-\phi_t + \beta \phi_{t+1}[\alpha u'(x_t) + 1 - \alpha] = 0, \quad (65)$$

$$-1 + (1 - \alpha)\beta R u'(s_t) = 0, \quad (66)$$

where x_t is the quantity traded at night in non-monitored exchange, and s_t the quantity traded in monitored exchange.

Assume that the monetary authority makes lump-sum transfers during the afternoon of the day to buyers. Then the Friedman rule is optimal: the money supply grows at the rate $\beta - 1$ and $\phi_{t+1} / \phi_t = 1/\beta$. This implies from Eq. (65) that $x_t = x^*$ in monetary exchange. However, claims to storage are of no use to buyers, so if a buyer does not meet a type 2 seller, his storage is wasted, even if we run the Friedman rule.

Now consider what happens if banks can accept deposits from buyers, in the form of goods, and use them to acquire money or storage. The bank maximizes the expected utility of its depositors. Since all buyers are identical, consider an equilibrium where all depositors make the same deposit,

$$d_t = \phi_t m_t + k_t, \quad (67)$$

Here, k_t and m_t denote, respectively, storage and money acquired by the bank. If the bank is perfectly diversified, as it will be in equilibrium, it offers agents who wish to withdraw $\hat{m}_t = m_t/\alpha$ dollars, and permits those who do not withdraw to trade claims to $\hat{k}_t = k_t/(1 - \alpha)$ units of storage. Since the bank maximizes the expected utility of the representative depositor, in equilibrium, k_t and m_t solve

$$\max_{m_t, x_t} \left\{ -\phi_t m_t - k_t + \alpha u\left(\frac{\beta \phi_{t+1} m_t}{\alpha}\right) + (1 - \alpha)u\left(\frac{\beta k_t R}{1 - \alpha}\right) \right\}.$$

As above, let x_t denote the quantity of output exchanged during the night in a non-monitored transaction, and s_t the quantity of output exchanged in a monitored transaction. Then, the FOC for an optimum are

$$u'(x_t) = \frac{\phi_t}{\beta \phi_{t+1}}, \quad (68)$$

$$u'(s_t) = \frac{1}{\beta R}, \quad (69)$$

determining x_t and s_t , respectively. Compare Eqs. (65) and (66) with Eqs. (68) and (69). Letting μ denote the money growth rate, we will have $\phi_t/\phi_{t+1} = 1 + \mu$. Therefore, if $\mu > \beta - 1$ then x_t is smaller in the equilibrium without banks than with banks. This is because, without banks, money is held by all buyers but cannot be used in exchange in monitored transactions, as type 2 sellers will not accept it. When $\mu = \beta - 1$, $x_t = x^*$ whether or not there are banks, as there is no opportunity cost to holding money from one day to the next. Note that s_t is always larger with banks than without. This is because, if there are no banks, storage might be wasted if a buyer has a non-monitored meeting at night. Anticipating this, buyers invest less in storage than if the bank implicitly provides insurance.

Thus, banking acts to increase consumption in the night and to eliminate wasted storage, increasing welfare. As in Diamond-Dybvig, there is an insurance role for banks, in that banks allow agents to economize on currency and promote investment in higher yielding assets. However, there is also an efficiency gain, in that storage is not wasted. With banking, the quantity of goods x_t exchanged for money during the night is efficient under the Friedman rule, which by Eq. (68) gives $x_t = x^*$. A policy that we can analyze in this model is Friedman's recommendation for 100% reserve requirements. This effectively shuts down financial intermediation and constrains buyers to holding outside money and investing independently, rather than holding deposits backed by money and storage. We then revert to the outcome without banks, which we know is inferior.

One can also consider the case of aggregate uncertainty, where α_t is a random variable, capturing fluctuations in the demand for liquidity. Assume that α_t is publicly observable, but is not realized until the end of the day, after consumption and production decisions have been made. For convenience, assume α_t is i.i.d. Now, analogous to the optimization problem above, the bank solves

$$\max_{m_t, x_t, \hat{m}_t} \left\{ -\phi_t m_t - k_t + E_t \left[\alpha_t u \left(\frac{\beta \phi_{t+1} (m_t - \hat{m}_t)}{\alpha_t} \right) + \beta \phi_{t+1} \hat{m}_t + (1 - \alpha_t) u \left(\frac{\beta k_t R}{1 - \alpha_t} \right) \right] \right\},$$

where $\hat{m}_t$ is the quantity of money per depositor which is not spent at night.³⁵ The FOC for m_t and k_t are

$$-\phi_t + \beta E_t \left\{ \phi_{t+1} \max \left[1, u' \left(\frac{\beta \phi_{t+1} m_t}{\alpha_t} \right) \right] \right\} = 0, \quad (70)$$

$$-1 + \beta R E_t u' \left(\frac{\beta R k_t}{1 - \alpha_t} \right) = 0. \quad (71)$$

³⁵ It is irrelevant whether this money is withdrawn by the depositor at the end of the day, or left in the bank until the next day and then withdrawn.

Now, letting M_t denote the money stock per buyer, from Eq. (70) the stochastic process for prices $\{\phi_t\}_{t=0}^\infty$ solves

$$-\phi_t + \beta E_t \left\{ \phi_{t+1} \max \left[1, u' \left(\frac{\beta \phi_{t+1} M_t}{\alpha_t} \right) \right] \right\} = 0, \quad (72)$$

given $\{M_t\}_{t=0}^\infty$.

First, suppose that $M_t = M_0(1 + \mu)^t$ with money growth accomplished through lump-sum transfers to buyers in the day. Then, in a stationary equilibrium, we have $\phi_t = \phi_0(1 + \mu)^{-t}$ and from Eq. (72) we get

$$\frac{1 + \mu}{\beta} = E \left\{ \max \left[1, u' \left(\frac{\beta \phi_0 M_0}{(1 + \mu) \alpha} \right) \right] \right\}, \quad (73)$$

which solves for ϕ_0 (dropping t subscripts for convenience). Let $G(\phi_0)$ denote the right-hand side of Eq. (73). We have $G(0) = \infty$, and $G(\phi_0) = 1$ for $\phi_0 \geq x^*(1 + \mu)\bar{\alpha}/\beta M_0$ where $\bar{\alpha}$ is the largest value in the support of the α distribution. Further, $G(\cdot)$ is strictly decreasing and continuous for $0 < \phi_0 < x^*(1 + \mu)\bar{\alpha}/\beta M_0$. Therefore, if $\mu > \beta - 1$ then from Eq. (73) there is a unique solution for ϕ_0 , and there will be realizations of α_t such that the quantity of goods traded in non-monitored meetings at night is less than x^* .

Further, consider a nominal bond that pays off one unit of outside money in the following day, and is exchanged at the end of the current day after α_t becomes known. The nominal interest rate on this bond is given by

$$r_t = \max \left[1, u' \left(\frac{\beta \phi_0 M_0}{(1 + \mu) \alpha_t} \right) \right] - 1.$$

The nominal interest rate fluctuates with α_t . In general, when α_t is large, currency is scarce and the nominal rate is high. States of the world where there are currency shortages and the withdrawal demand at banks is high are associated with high nominal interest rates, as was the case historically during banking panics.

From Eq. (73), note that one optimal monetary policy is $\mu = \beta$ — the standard Friedman rule. Then, for any $\phi_0 \geq \phi^*$, ϕ_0 is an equilibrium price of money at the first date, where

$$\phi^* = \frac{x^*(1 + \mu)\bar{\alpha}}{\beta M_0}.$$

In any of these equilibria, the nominal interest rate is zero for all t and each buyer consumes x^* in non-monitored meetings during the night. Thus, there exists a continuum of equilibria given $\mu = \beta$ and in any of these equilibria there are states of the world where some portion of the money stock is not spent in monetary transactions at night.

There exist other money supply rules that support a zero nominal interest rate. Suppose that we look for a monetary policy rule such that the nominal rate is always zero and all cash is spent at night each period. From Eq. (72), we first require that

$$\frac{\beta\phi_{t+1}M_t}{\alpha_t} = x^*,$$

so that there is efficient trade in all non-monitored meetings in the night and all cash is spent. From this we obtain

$$\frac{M_t}{M_{t+1}} = \frac{\beta E_t(\alpha_t)}{\alpha_{t-1}},$$

which is an optimal policy rule with the characteristics we are looking for. Under this rule, agents in non-monitored meetings at night anticipate there will be a monetary contraction the next day if the demand for liquidity is high. This will tend to increase the value of money in non-monitored transactions, so that efficient trades can be made. This monetary rule is active, acting to accommodate fluctuations in the demand for liquidity, as opposed to the passive (constant money growth) rule that achieves the same result.

As discussed earlier, there exist optimal policy rules here that look nothing like the prescription in Friedman (1968). Typically, achieving a zero nominal interest rate in all states of the world can be implemented through various monetary policies that do not entail constant growth of the money stock.³⁶ More broadly, we think there is a lot to be learned by carefully modeling banking and the interaction with monetary policy as part of larger general equilibrium models of the exchange process. It is inevitable that these models will be somewhat complicated, at least compared to the simplest examples of money being used as a medium of exchange in Section 2. But the payoff to getting the models right is a better understanding of banking and financial intermediation, which seems very hard to dismiss as unimportant or uninteresting in this day and age.

6. FINANCE

The class of models presented here has recently been used to study asset markets. This work is potentially very productive, as it allows one to examine how frictions and policy affect the liquidity of assets, their prices, and the trading volume in these markets. Moreover, although this may come as a surprise to some people who seem to think that financial markets are as close to a frictionless ideal as there is, it is also one of

³⁶ As Lagos (2009) showed, a path for the money stock that implements the Friedman rule needs to only satisfy two weak properties. Roughly, the money stock must go to zero in the limit, and it must grow on average at a rate higher than minus the rate of time preference.

the most natural applications of the search-and-bargaining approach. As Duffie, Gârleanu, and Pederson (2008) stated,

Many assets, such as mortgage-backed securities, corporate bonds, government bonds, US federal funds, emerging-market debt, bank loans, swaps and many other derivatives, private equity, and real estate, are traded in over-the-counter (OTC) markets. Traders in these markets search for counterparties, incurring opportunity or other costs. When counterparties meet, their bilateral relationship is strategic; prices are set through a bargaining process that reflects each investor's alternatives to immediate trade.

Since the models people use to formalize these ideas are closely related to those used in monetary theory, we provide a taste of these applications using the New Monetarist model.³⁷

6.1 Asset trading and pricing

One of the first papers in finance to use the search-and-bargaining approach is Duffie, Gârleanu, and Pederson (2005). They worked with a version of the second-generation monetary models presented in Section 2.2, which means in particular that agents can hold only $a \in \{0, 1\}$ units of an asset. Even under this restriction many interesting results emerge, and it would be worth discussing how they adapt the model for their purposes. However, we present a model capturing similar ideas using our benchmark model where agents can hold any amount $a \in \mathbb{R}_+$ of an asset.³⁸

One should think now of assets as (shares in or claims on) “trees” paying dividends each period in “fruit” as in the standard Lucas (1978) asset-pricing model. In this application, agents will value assets for their yield or dividend, which we denote by γ . Thus, an agent holding a units of the asset has a claim on $a\gamma$ units of “fruit,” where here dividends accrue and are consumed in the DM. Let A be the fixed supply of the asset, and denote its CM price by ϕ , which is constant because we focus on steady states. Then the CM problem is

$$\begin{aligned} W(a) &= \max \{U(X) - H + \beta V(\hat{a})\} \\ X &= H + \phi a - \phi \hat{a}. \end{aligned}$$

As usual, this implies $U'(X) = 1$, $\phi = \beta V'(\hat{a})$, and $W'(a) = \phi$.

In the DM, agents get utility from consuming dividends, subject to preference shocks realized after the CM closes but before the DM opens. Let π_H and $\pi_L = 1 - \pi_H$ be the probability of a high and a low shock, implying utility for agents with a units of

³⁷ Papers that we have in mind in monetary economics include Ferraris and Watanabe (2010); Geromichalos, Licari, and Lledo (2010); Jacquet and Tan (2009); Lagos (2008); Lester et al. (2009); Ravikumar and Shao (2006); and Rocheteau (2009). Contributions more in finance include Duffie et al. (2005, 2008), Lagos and Rocheteau (2009), Lagos et al. (2008), Silveira and Wright (2010), Weill and Vayanos (2008), and Weill (2007, 2008).

³⁸ Lagos and Rocheteau (2009) provided a different extension of Duffie et al. (2005), which also allows $a \in \mathbb{R}_+$, but here we stay closer to our benchmark model.

the asset $u_H(a\gamma)$ and $u_L(a\gamma)$, respectively, with $u'_H(x) > u'_L(x)$ for all x . There is generally gain from trade between an agent who draws the L shock and one who draws H . In the literature, L is often referred to as a liquidity shock, because it stands in for agents needing to sell assets to meet liquidity needs; that is, while the model literally has agents trading claims to “trees” because of changes in their utility from “fruit” it is meant to capture more generally the idea that sometimes one has to sell assets for any number of reasons, including a need for ready cash. Of course, one could say the papers ought to model the need for liquidity more explicitly; we would concur, and people are working on this. In any case, agents in the DM meet bilaterally and at random. Let σ_H be the probability an agent with shock H meets one with shock L , and σ_L the probability that an agent with L meets one with H . In a meeting where one agent has L and the other H , the former transfers q units of the asset to the latter in exchange for a payment p , interpreted as an IOU for p units of X to be delivered in the next CM, assumed again to be perfectly enforced.³⁹

To reduce notation, define the trade surplus for H and for L as

$$\begin{aligned} S_H(a) &= u_H[(a+q)\gamma] - u_H(a\gamma) + \phi q - p \\ S_L(a) &= u_L[(a-q)\gamma] - u_L(a\gamma) - \phi q + p, \end{aligned}$$

using $W'(a) = \phi$. This allows us to write the DM payoff as

$$V(a) = \pi_H \sigma_H S_H(a) + \pi_L \sigma_L S_L(a) + \pi_L u_L(a\gamma) + \pi_H u_H(a\gamma) + W(a).$$

In terms of bargaining, when type H with a meets type L with A , the solution is

$$\max S_H(a)^\theta S_L(A)^{1-\theta}.$$

It is easy to see that this is solved by (q, p) satisfying

$$u'_H[(a+q)\gamma] = u'_L[(A-q)\gamma] \quad (74)$$

$$p = \phi q + (1-\theta)\{u_H[(a+q)\gamma] - u_H(a\gamma)\} + \theta\{u_L(A\gamma) - u_L[(A-q)\gamma]\}. \quad (75)$$

Inserting this p into $S_H(a)$ and $S_L(a)$, and inserting these into $V(a)$, we get

$$\begin{aligned} V(a) &= \pi_H \sigma_H \theta \{u_H[(a+q)\gamma] - u_H(a\gamma) + u_L[(A-q)\gamma] - u_L(A\gamma)\} \\ &\quad + \pi_L \sigma_L (1-\theta) \{u_H[(A+q)\gamma] - u_H(A\gamma) + u_L[(a-q)\gamma] - u_L(a\gamma)\} \\ &\quad + \pi_H u_H(a\gamma) + \pi_L u_L(a\gamma) + W(a), \end{aligned}$$

where we are careful to note that a is the asset position of the individual whose value function we are considering and A is the position of someone he meets.

³⁹ One may recognize this specification as similar to Section 3.5, in the sense that there is no production, but simply exchange between agents with different preference shocks. However, in this application we assume perfect credit.

Differentiating, we get

$$V'(a) = \pi_H \sigma_H \theta u'_H[(a+q)\gamma] + \pi_L \sigma_L (1-\theta) u'_L[(a-q)\gamma] \\ + \pi_H (1 - \sigma_H \theta) u'_H(a\gamma) + \pi_L [1 - \sigma_L (1-\theta)] u'_L(a\gamma) + \phi.$$

For concreteness, consider a matching technology with $\sigma_H = \sigma\pi_L$ and $\sigma_L = \sigma\pi_H$, which is basically the matching technology introduced back in Section 2.1. Then, since $u'_H[(a+q)\gamma] = u'_L[(a-q)\gamma]$, we can write

$$V'(a) = \sigma\pi_H\pi_L u'_H[(a+q)\gamma] + \pi_H(1 - \sigma\pi_L\theta) u'_H(a\gamma) \\ + \pi_L[1 - \sigma\pi_H(1-\theta)] u'_L(a\gamma) + \phi.$$

Substituting this into the FOC from the CM, $\phi = \beta V'(\hat{a})$, then setting $a = A$, $\pi_L = 1 - \pi_H$, and $r = (1 - \beta)/\beta$, we get

$$r\phi = \sigma\pi_H(1 - \pi_H) u'_H[(A+q)\gamma] + \pi_H(1 - \sigma\theta + \sigma\pi_H\theta) u'_H(A\gamma) \\ + (1 - \pi_H)(1 - \sigma\pi_H + \sigma\pi_H\theta) u'_L(A\gamma). \quad (76)$$

We can now describe equilibrium recursively: first find the q that solves Eq. (74); then the DM asset price p solves Eq. (75) and the CM asset price ϕ solves Eq. (76), where it turns out that p and ϕ are independent conditional on q .

Notice from Eq. (76) that the CM asset price per period, $r\phi$, is a weighted average of three terms: the marginal value of the asset when you trade, which is independent of your shock, $u'_H[(A+q)\gamma] = u'_L[(A-q)\gamma]$; the marginal value when you do not trade but have a high shock, $u'_H(A\gamma)$; and the marginal value when you do not trade but have a low shock, $u'_L(A\gamma)$. If agents are risk neutral in the DM, $u(x) = x$, then $r\phi = \gamma$, which means the asset is priced at its fundamental value (the capitalized value of the dividend stream). If agents are risk averse the asset price will adjust for the fact that its value is random. Even if $\sigma = 1$, so there are no fundamental search frictions, in the sense that you always meet someone, if matching is random you could meet the wrong type and so there is risk.

Suppose we set $\pi_L = \pi_H = 1/2$ and change the matching technology, so that every agent with H meets one with L , which means agents always have the opportunity to rebalance their asset holdings in the DM. It is not hard to rework the analysis to get

$$r\phi = \frac{1}{2} u'_H[(A+q)\gamma] + \frac{1-\theta}{2} u'_H(A\gamma) + \frac{\theta}{2} u'_L(A\gamma). \quad (77)$$

In this case there is no risk per se, since everyone gets to rebalance their asset position, and $u'_H[(A+q)\gamma] = u'_L[(A-q)\gamma]$. But due to bargaining power the asset price can be priced differently from its fundamental value. In this special case without search or matching risk we have

$$\frac{\partial r\phi}{\partial \theta} = -\frac{\gamma}{2} [u'_H(A\gamma) - u'_L(A\gamma)] < 0,$$

so increasing the bargaining power of the agent buying the asset in the DM reduces the asset's price in the CM.

Returning to the more general case, with search and matching frictions, Eq. (74) implies a similar result,

$$\frac{\partial r\phi}{\partial \theta} = \sigma\pi_H(1 - \pi_H)[u'_L(A) - u'_H(A)]\gamma < 0.$$

And in terms of the baseline arrival rate, we have

$$\frac{\partial r\phi}{\partial \sigma} = \pi_H(1 - \pi_H)\{\theta[u'_H(A + q) - u'_H(A)] + (1 - \theta)[u'_L(A - q) - u'_L(A)]\}\gamma.$$

Since $u'_H(A + q) < u'_H(A)$ and $u'_L(A - q) > u'_L(A)$, this will be negative for big θ and positive for small θ . The important point is that search and bargaining frictions in the DM affect the asset price in the CM. And in terms of probabilities,

$$\frac{\partial r\phi}{\partial \pi_H} = \sigma(1 - 2\pi_H)[u'_L(A - q) - u'_L(A)]\gamma + (1 - \sigma\theta + 2\sigma\theta\pi_H)[u'_A(A) - u'_L(A)]\gamma,$$

which is ambiguous, in general, but is definitely positive for $\pi_H \leq 1/2$. So the distribution of liquidity shocks in the DM naturally matters for asset CM prices, too.

This setup is similar in spirit to Duffie et al. (2005), even if the details differ. Models like this have been used to study a variety of issues. One application is to introduce middlemen — dealers, or brokers, say — that can buy assets from L types and sell to H types. Weill (2007) studied the behavior of such intermediaries, not only in steady state, but along dynamic transition paths after a crisis. A crisis is modeled as a reduction in π_H , which stands in for the idea that many people want to sell assets while few want to buy. He actually uses a second-generation version of the model, with $a \in \{0, 1\}$. Lagos et al. (2009) use a generalized model with $a \in \mathbb{R}^+$. An interesting question is whether intermediaries provide enough liquidity, in the sense of buying and holding assets while the economy recovers from a crisis. One can use the model to study the effects of various central bank interventions, including the recent Fed policy of buying up certain assets. The analysis and results are too involved to go into detail, but at least we get to illustrate the types of issues people have been studying with these models.

6.2 Capital markets

Here is another way to model asset markets, which is not far from the benchmark monetary model, except for two twists. First, we have assets other than money acting as a medium of exchange; second, the gains from trade come not from producing goods for consumers, but from reallocating capital across producers. Suppose again that there are two types, buyers and sellers, with a mass $1/2$ of each. A buyer's CM utility is $U(X) - H$, while a seller's CM utility is $X - H$. In the CM, in addition to agents being

able to produce X one-for-one with H , agents can also produce it using capital: anyone with k units of capital can produce $f(k)$ units of X , where $f' > 0$, $f'' < 0$, $f'(0) = \infty$, $f'(\infty) = 0$ and $f(0) = 0$. Sellers, but not buyers, also have a technology to convert X into k one-for-one, after the CM closes. Capital produced at t becomes productive at the beginning of the CM at $t + 1$, after which (for simplicity) it depreciates 100%. No one produces or consumes in the DM in this application — it is only a market for asset exchange.

In addition to k , there is a second asset, a , which as in the previous section one can think of as a share in a Lucas “tree.” Now we normalize the supply of “trees” to $A = 1/2$, and assume the dividend y is realized in units of X in the CM. Shares are now used in the DM as a means of payment. Of course, money can be considered an asset with dividend $y = 0$ (and, moreover, the quantity of the asset can be augmented by the government through transfers and taxes). In the DM, each buyer is matched with a seller with probability σ , and similarly for sellers. Again, buyers and sellers do not produce or consume in the DM, and matches represent only opportunities for trading assets. To generate potential gains from trade, we assume that the technology prohibits buyers from holding capital when the CM closes (say because the buyers have left the CM before capital production takes place). Thus, a match in the DM is an opportunity for a buyer to exchange shares for capital.

First, consider the case where $\sigma = 0$, which means we shut down the DM. Then a will be priced according to fundamentals, $\phi_t = \hat{\phi}$ for all t where $\hat{\phi} = \beta y / (1 - \beta)$. In other words, the share price is the present value of future dividends, and the rate of return on shares is equal to the rate of time preference r .⁴⁰ Sellers acquire capital in each CM that they cannot trade, since we are here assuming $\sigma = 0$, so they accumulate only for production. Letting k_{t+1} denote the capital produced by a seller in period t , and k_{t+1}^s , k_{t+1}^b , respectively, the quantities of capital held by each seller and buyer at the beginning of the CM in period $t + 1$, we have $k_{t+1} = k_{t+1}^s = \hat{k}$ and $k_t^b = 0$, where $f'(\hat{k}) = 1/\beta$. As with shares, the return on capital also equals the rate of time preference.

Now consider the case where $\sigma > 0$. If the buyer has a shares and the seller has k units of capital, the buyer can transfer d shares to the seller for k^b units of capital. The generalized Nash bargaining solution is

$$\max_{d, k^b} [f(k^b) - d(\phi_{t+1} + y)]^\theta [f(k - k^b) + d(\phi_{t+1} + y) - f(k)]^{1-\theta},$$

subject to $d \leq a$ and $k^b \leq k$. The second constraint does not bind since $f'(0) = \infty$. Without loss of generality, we will consider equilibria where buyers always exchange all of their shares for capital in the DM if they are matched with a seller and either:

⁴⁰ Define the return on the share r_s by $1 + r_s = (\phi + y)/\phi$. Then $r_s = y/\phi = \beta/(1 - \beta)$ when shares are priced fundamentally.

(i) Sellers hold part of the stock of shares; or (ii) buyers hold the entire stock of shares at the end of the CM. That is, we consider cases where the first constraint binds for buyers, so $d = a$, and k^b solves $a = z(k^b, k)$, where

$$z(k^b, k) = \left(\frac{1}{\phi_{t+1} + \gamma} \right) \left\{ \frac{\phi f'(k^b)[f(k) - f(k - k^b)] + (1 - \theta)f'(k - k^b)f(k^b)}{\theta f'(k^b) + (1 - \theta)f'(k - k^b)} \right\}. \quad (78)$$

Then a buyer's problem in the CM is

$$\max_{k_{t+1}^b} \left\{ -\phi_t z(k_{t+1}^b, k_{t+1}) + \beta [\sigma f(k_{t+1}^b) + (1 - \sigma)z(k_{t+1}^b, k_{t+1})(\phi_{t+1} + \gamma)] \right\}, \quad (79)$$

while a seller's problem is

$$\max_{k_{t+1}} \left\{ -k_{t+1} + \beta \sigma [f(k_{t+1} - k_{t+1}^b) + \beta z(k_{t+1}^b, k_{t+1})(\phi_{t+1} + \gamma)] + \beta (1 - \sigma)f(k_{t+1}) \right\} \quad (80)$$

The FOC's from these problems yield

$$\left[\frac{\phi_{t+1} + \gamma}{\phi_t} \right] \left[\frac{\sigma f'(k_{t+1}^b)}{z_1(k_{t+1}^b, k_{t+1})(\phi_{t+1} + \gamma)} + 1 - \sigma \right] = \frac{1}{\beta}, \quad (81)$$

$$\sigma [f'(k_{t+1} - k_{t+1}^b) + z_2(k_{t+1}^b, k_{t+1})(\phi_{t+1} + \gamma)] + (1 - \alpha)f'(k_{t+1}) = \frac{1}{\beta}. \quad (82)$$

Note that in Eq. (81),

$$\ell(k_{t+1}^b, k_{t+1}) = \frac{f'(k_{t+1}^b)}{z_1(k_{t+1}^b, k_{t+1})(\phi_{t+1} + \gamma)} \quad (83)$$

represents a liquidity premium on shares, analogous to the one in the baseline model. The larger is $\ell(k_{t+1}^b, k_{t+1})$ the greater is the departure of the share price from its fundamental value, and the lower is the return on the asset $(\phi_{t+1} + \gamma)/\phi_t$. When $\ell(k_{t+1}^b, k_{t+1}) = 1$ there is no liquidity premium.

We first look for an equilibrium where some shares are held in equilibrium by sellers at the end of the CM. This implies that $\phi_t = \hat{\phi}$ and shares are priced according to fundamentals. Thus there is no liquidity premium on shares, since they are priced according to how they are valued at the margin by sellers, who do not trade shares in the DM. In this equilibrium where liquidity is not scarce, let $\bar{k}$ denote the quantity of capital produced at the end of the CM by each seller, and $\bar{k}^b$ the quantity of capital carried into the CM by each buyer. Then Eq. (83) implies

$$f'(\bar{k}^b) = \frac{\gamma z_1(\bar{k}^b, \bar{k})}{1 - \beta}. \quad (84)$$

Also, substituting for the price of shares in Eq. (82) gives

$$\sigma \left[f'(\bar{k} - \bar{k}^b) + \frac{\gamma z_2(\bar{k}^b, \bar{k})}{1 - \beta} \right] + (1 - \sigma)f'(\bar{k}) = \frac{1}{\beta}, \quad (85)$$

and we require that the quantity of shares brought by each seller to the DM is $a \leq 1$, or,

$$z(\bar{k}^b, \bar{k}) \leq 1. \quad (86)$$

Thus, an equilibrium where liquidity is not scarce and shares trade at their fundamental value, with no liquidity premium, consists of quantities $\bar{k}$ and $\bar{k}^b$ solving Eqs. (84) and (85), and satisfying the inequality Eq. (86).

Now, consider an equilibrium where liquidity is scarce, in the sense that buyers hold the entire stock of shares at the end of the CM for transactions purposes, and sellers hold zero. Then, in a steady state

$$z(k^b, k) = 1, \quad (87)$$

and we can use Eqs. (81), (82), and (87) to solve for ϕ , k , and k^b .

To consider an extreme case, let $\theta = 1$. Then, from Eq. (78) we have

$$z(k^b, k) = \left(\frac{1}{\phi_{t+1} + \gamma} \right) [f(k) - f(k - k^b)], \quad (88)$$

and we can write Eqs. (81) and (82) as

$$\left[\frac{\phi_{t+1} + \gamma}{\phi_t} \right] \left[\frac{\sigma f'(k_{t+1}^b)}{f'(k_{t+1} - k_{t+1}^b)} + 1 - \sigma \right] = \frac{1}{\beta}, \quad (89)$$

$$f'(k_{t+1}) = \frac{1}{\beta}. \quad (90)$$

Thus from Eq. (90), in any equilibrium, we will have $k_{t+1} = \hat{k}$, the same total quantity of capital as in an equilibrium with $\sigma = 0$ where capital cannot be traded. We get this result as sellers receive no surplus from trading capital when $\theta = 1$.

Now, in an equilibrium where liquidity is not scarce, we will have $\phi = \hat{\phi}$, and so from Eq. (89) we have $k^b = \frac{\hat{k}}{2}$ in a steady state, so capital is efficiently allocated between buyers and sellers in DM trade. Further, from Eq. (86), an equilibrium where liquidity is not scarce exists iff

$$f(\hat{k}) - f\left(\frac{\hat{k}}{2}\right) < \frac{\gamma}{1 - \beta}, \quad (91)$$

Thus, if γ is sufficiently large, the share price is high enough when shares are at their fundamental value that there is efficient trade in the DM and capital is efficiently allocated in DM trade. This efficient allocation occurs because, given $\theta = 1$, there is no holdup problem for buyers. However, a holdup problem for sellers exists, and they tend to underaccumulate capital relative to what is efficient.

Now, consider a steady state equilibrium where liquidity is scarce. Here, from Eqs. (88), (89), and (87), we obtain

$$\frac{\phi}{\phi + \gamma} = \beta \left[\frac{\sigma f'(k^b)}{f'(\hat{k} - k^b)} + 1 - \sigma \right], \quad (92)$$

$$\phi = f(\hat{k}) - f(\hat{k} - k^b) - \gamma, \quad (93)$$

which solve for ϕ and k^f . It is straightforward to show that the solution is unique, and the equilibrium exists iff the solution satisfies $\phi \leq \hat{\phi}$ or

$$f(\hat{k}) - f\left(\frac{\hat{k}}{2}\right) \geq \frac{\gamma}{1 - \beta}.$$

In this equilibrium we have $k^b \leq \hat{k}/2$, so that the allocation of capital between buyers and sellers is inefficient in equilibrium with insufficient liquidity. Further, from Eq. (83), our measure of the liquidity premium is

$$\ell(k^b, k) = \frac{f'(k^b)}{f'(\hat{k} - k^b)},$$

which increases as k^b falls given $\hat{k}$; that is, as capital allocation becomes more inefficient.

From Eqs. (92) and (93), it is straightforward to show that k^b increases with σ and with γ . First, an increase in the frequency of trade, which increases the frequency with which buyers and sellers can more efficiently allocate capital, also increases the efficiency of capital allocation in each trade. Second, an increase in the dividend, which increases the price of shares, results in a more efficient allocation of capital by enhancing the supply of liquidity.

Now, consider the other extreme case where $\theta = 0$ and sellers have all the bargaining power in the DM. Here, from Eq. (78), we have

$$z(k^b, k) = \frac{f(k^b)}{\phi_{t+1} + \gamma}.$$

Then, from Eq. (81), $\phi = \hat{\phi}$. In this case, as the buyer receives no DM surplus, and shares always trade at their fundamental value. From Eq. (82), optimization by sellers implies that k and k^b must satisfy

$$\sigma f'(k - k^b) + (1 - \sigma)f'(k) = \frac{1}{\beta}. \quad (94)$$

However, the quantity of capital that a buyer trading in the DM carries into the next CM, k^b , is indeterminate. We obtain this result since buyers receive no surplus in the DM and therefore are indifferent in equilibrium concerning the quantity of shares they take to the DM. Given that $\phi = \hat{\phi}$, sellers are also indifferent concerning the quantity of shares they carry from one CM into the next CM. We only require that k^b be small enough that the fundamental value of the stock of shares is sufficient to buy k^b in the DM, i.e. $k^b \in [0, \tilde{k}]$, where $\tilde{k}$ solves

$$f(\tilde{k}) = \frac{\gamma}{1 - \beta}$$

With $\theta = 0$, the holdup problem for buyers in the DM is as severe as possible and so the quantity of capital k is in general inefficiently allocated between buyers and sellers in the DM. However, since $\theta = 0$ implies no holdup problem for sellers, then given k^b , Eq. (94) tells us that sellers accumulate capital efficiently. An increase in σ , since it increases the frequency of trade, will in general raise the quantity of capital from Eq. (94), given k^b .

This model captures the idea that assets are potentially valued for more than their simple returns, and in particular the asset price can include a liquidity premium.⁴¹ This seems important in practice, since money is not the only asset whose value depends at least in part on its use in facilitating transactions. For example, T-bills play an important role in overnight lending in financial markets, where they are commonly used as collateral. Potentially, models like this, which allow us to examine the determinants of the liquidity premium, can help to explain the apparently anomalous behavior of relative asset returns and asset prices. See Lagos (2008) for one such application.

A clear message of this application is that asset markets are important for allocation and efficiency. If the yields on liquid assets are low or these assets are hard to trade, this tends to reduce investment in productive capital, and also to result in an inefficient allocation of capital across productive units. Further, bargaining power in asset exchange matters for efficiency as well as prices. Just as in our benchmark monetary model, the greater the bargaining power of buyers the more likely that trades will be efficient in decentralized exchange. Here, greater bargaining power for buyers increases the efficiency with which capital is allocated. However, in contrast to the benchmark model, greater bargaining power for buyers also increases inefficiency in that it tends to reduce investment.

⁴¹ In the special case where m is money, $\gamma = 0$, we can let the stock be augmented by government through lump-sum transfers. The fundamental equilibrium is then the non-monetary equilibrium where $\phi = 0$. There is also a steady-state monetary equilibrium where $\phi > 0$ and $\phi_t/\phi_{t+1} = 1 + \mu$, with μ the money growth rate. In this case there always exists an equilibrium with insufficient liquidity for $\mu > \beta - 1$. In general, the optimal policy is again the Friedman rule $\mu = \beta - 1$.

A model like this is potentially useful for analyzing phenomena related to the recent financial crisis, since it captures a mechanism by which asset exchange and asset prices are important for investment and allocative efficiency. It may seem that to directly address the reasons for credit market problems during a crisis would require models with lending and collateral. However, it is a very short step from a model like the one presented here, where liquid assets are used in exchange, to one where assets serve as collateral in credit contracts. A key feature of our model in this respect is that, if the future payoffs on liquid assets are expected to be low, and one might think now about mortgage-backed securities, then this can reduce investment and cause allocative inefficiency, both of which reduce aggregate output.

7. CONCLUSION

New Monetarists are committed to modeling approaches that are explicit about the frictions that make monetary exchange and related arrangements socially useful, and that capture the relationship among credit, banking, and currency transactions. Ideally, economic models that are designed for analyzing and evaluating monetary policy should be able to answer basic questions concerning the necessity and role of central banking, the superiority of one type of central bank operating procedure over another, and the differences in the effects of central bank lending and open market operations. New Monetarist economists have made progress in understanding the basic frictions that make monetary exchange an equilibrium or an efficient arrangement, and in understanding the mechanisms by which policy can affect allocations and welfare. However, much remains to be learned about many issues, including the sources of short-run non-neutralities and their quantitative significance, as well as the role of central banking.

This chapter takes stock of how the New Monetarist approach builds on advances in the theory of money and theories of financial intermediation and payments, constructing a basis for progress in the science and practice of monetary economics. We conclude by borrowing from Hahn (1973), who went on to become an editor of the previous *Handbook*. He begins his analysis by suggesting “The natural place to start is by taking the claim that money has something to do with the activity of exchange, seriously.” He concludes as follows:

I should like to end on a defensive note. To many who would call themselves monetary economists the problems which I have been discussing must seem excessively abstract and unnecessary. ... Will this preoccupation with foundations, they may argue, help one iota in formulating monetary policy or in predicting the consequences of parameter changes? Are not IS and LM sufficient unto the day? ... It may well be that the approaches here utilized will not in the event improve our advice to the Bank of England; I am rather convinced that it will make a fundamental difference to the way in which we view a decentralized economy.

REFERENCES

- Andolfatto, D., 2010a. Essential interest-bearing money. *J. Econ. Theory*. in press.
- Andolfatto, D., 2010b. On the social cost of transparency in monetary economies. Federal Reserve Bank of St. Louis, Working paper 2010-001A.
- Aiyagari, R., Wallace, N., 1991. Existence of steady states with positive consumption in the Kiyotaki-Wright Model. *Review of Economic Studies* 58, 901–916.
- Aiyagari, R., Wallace, N., 1992. Fiat money in the Kiyotaki-Wright model. *Econ. Theory* 2, 447–464.
- Aiyagari, R., Wallace, N., Wright, R., 1996. Coexistence of money and interest-bearing securities. *J. Monetary Econ.* 37, 397–420.
- Ales, L., Carapella, F., Maziero, P., Weber, W.E., 2008. A model of banknote discounts. *J. Econ. Theory* 142 (1), 5–27.
- Ales, L., Carapella, F., Maziero, P., Weber, E., 2010. A model of banknote discounts. *J. Econ. Theory* 142 (1), 5–27.
- Aliprantis, C., Camera, G., Puzzello, D., 2006. Matching and anonymity. *Econ. Theory* 29, 415–432.
- Aliprantis, C., Camera, G., Puzzello, D., 2007a. Contagion Equilibria in a Monetary Model. *Econometrica* 75, 277–282.
- Aliprantis, C., Camera, G., Puzzello, D., 2007b. Anonymous markets and monetary trading. *J. Monetary Econ.* 54, 1905–1928.
- Araujo, L., 2004. Social norms and money. *J. Monetary Econ.* 51, 241–256.
- Araujo, L., Camargo, B., Minetti, R., Puzzello, D., 2010. The informational role of prices and the essentiality of money in the Lagos-Wright Model. Michigan State University Mimeo.
- Aruoba, S.B., 2009. Money, search and business cycles. University of Maryland, Working paper.
- Aruoba, S.B., Chugh, S., 2008. Optimal fiscal and monetary policy when money is essential. *J. Econ. Theory* in press.
- Aruoba, S.B., Rocheteau, G., Waller, C., 2007. Bargaining and the value of money. *J. Monetary Econ.* 54, 2636–2655.
- Aruoba, S.B., Schorfheide, F., 2010. Sticky prices vs. monetary frictions: An estimation of policy tradeoffs. *American Economic Journal: Macroeconomics*, in press.
- Aruoba, S.B., Waller, C., Wright, R., 2009. Money and capital: A quantitative analysis. University of Maryland, Working paper.
- Azariadis, C., 1993. *Intertemporal macroeconomics*. Blackwell, Oxford, England.
- Ball, L., Mankiw, G., 1994. A sticky price manifesto. Mimeo.
- Basu, S., Fernald, J., 1997. Returns to scale in U.S. production: estimates and implications. *J. Polit. Econ.* 105, 249–283.
- Baumol, W., 1952. The transactions demand for cash: An inventory theoretic approach. *Quarterly Journal of Economics* 66, 545–556.
- Benabou, R., 1988. Search price-setting, and inflation. *Review of Economic Studies* 55, 353–376.
- Bencivenga, V., Camera, G., 2008. Banks, liquidity insurance, and interest on reserves in a matching model of money. Purdue University, Working paper.
- Berentsen, A., 2002. On the distribution of money holdings in a random-matching model. *Int. Econ. Rev.* 43, 945–954.
- Berentsen, A., Menzio, G., Wright, R., 2008. Inflation and unemployment in the long run. *Am. Econ. Rev.* in press.
- Berentsen, A., Molico, M., Wright, R., 2002. Indivisibilities, lotteries, and monetary exchange. *J. Econ. Theory* 107, 70–94.
- Berentsen, A., Monnet, C., 2008. Monetary policy in a channel system. University of Basel and Federal Reserve Bank of Philadelphia, Working paper.
- Berentsen, A., Rocheteau, G., 2002. On the efficiency of monetary exchange: How divisibility of money matters. *J. Monetary Econ.* 49, 1621–1649.
- Berentsen, A., Rocheteau, G., 2004. Money and information. *Review of Economic Studies* 71, 915–944.
- Berentsen, A., Waller, C., 2009. Outside versus inside bonds: A Modigliani-Miller type result for liquidity constrained economies. University of Basel and Federal Reserve Bank of St. Louis Working paper.

- Bernanke, B., Gertler, M., 1989. Agency costs, new worth, and business fluctuations. *Am. Econ. Rev.* 79, 14–31.
- Binmore, K., Osborne, M., Rubinstein, A., 1992. Noncooperative models of bargaining. In: Aumann, R., Hart, S. (Eds.), *Handbook of Game Theory*. North-Holland, Amsterdam.
- Boel, P., Camera, G., 2006. Efficient monetary allocations and the illiquidity of bonds. *J. Monetary Econ.* 53, 1693–1715.
- Brock, W., 1990. Overlapping generations models with money and transactions costs. In: Friedman, B., Hahn, F. (Eds.), *Handbook of Monetary Economics*. North-Holland, Amsterdam.
- Burdett, K., Judd, K., 1983. Equilibrium price dispersion. *Econometrica* 51, 955–970.
- Burdett, K., Mortensen, D., 1998. Wage differentials, employer size, and unemployment. *Int. Econ. Rev.* 39, 257–273.
- Burdett, K., Trejos, A., Wright, R., 2001. Cigarette money. *J. Econ. Theory* 99, 117–142.
- Calvo, G., 1983. Staggered prices in a utility maximizing framework. *J. Monetary Econ.* 12, 383–398.
- Camera, G., Berentsen, A., Waller, C., 2005. The distribution of money balances and the nonneutrality of money. *Int. Econ. Rev.* 46, 465–487.
- Camera, G., Corbae, D., 1999. Money and price dispersion. *Int. Econ. Rev.* 40, 985–1008.
- Camera, G., Reed, R., Waller, C., 2003. Jack of all trades or a master of one: specialization, trade, and money. *Int. Econ. Rev.* 44, 1275–1294.
- Campbell, J., Eden, B., 2007. Rigid prices: evidence from U.S. scanner data. Federal Reserve Bank of Chicago, Working Paper 2005–08.
- Caplin, A., Spulber, D., 1987. Menu costs and the neutrality of money. *Quarterly Journal of Economics* 102, 703–725.
- Cavalcanti, R., Erosa, A., Temzelides, T., 1999. Private money and reserve management in a random-matching model. *J. Polit. Econ.* 107, 929–945.
- Cavalcanti, R., Wallace, N., 1999a. A model of private banknote issue. *Review of Economic Dynamics* 2, 104–136.
- Cavalcanti, R., Wallace, N., 1999b. Inside and outside money as alternative media of exchange. *Journal of Money, Credit, and Banking* 31, 443–457.
- Champ, B., Smith, B., Williamson, S., 1996. Currency elasticity and banking panics: Theory and evidence. *Canadian Journal of Economics* 29, 828–864.
- Chapman, J., Chiu, J., Molico, M., 2009. A tiered settlement system. Bank of Canada, Working paper.
- Chiu, J., Meh, C., 2010. Banking, liquidity, and inflation. *Macroeconomic Dynamics*, in press.
- Chiu, J., Molico, M., 2010. Liquidity, Redistribution, and the Welfare Cost of Inflation. *J. Monetary Econ.* 57, 428–438.
- Chiu, J., Molico, M., 2008. Uncertainty, inflation, and welfare. Bank of Canada, Working paper.
- Clarida, R., Gali, J., Gertler, M., 1999. The science of monetary policy: A New Keynesian perspective. *J. Econ. Lit.* 37, 1661–1707.
- Coles, M., Wright, R., 1998. A dynamic model of search, bargaining, and money. *J. Econ. Theory* 78, 32–54.
- Cooley, T., Hansen, G., 1989. The inflation tax in a real business cycle model. *Am. Econ. Rev.* 79, 733–748.
- Corbae, D., Temzelides, T., Wright, R., 2003. Directed matching and monetary exchange. *Econometrica* 71, 731–756.
- Craig, B., Rocheteau, G., 2008. State-dependent pricing, inflation, and welfare. *European Economic Review* 52, 441–468.
- Curtis, E., Wright, R., 2004. Price setting, price dispersion, and the value of money, or the law of two prices. *J. Monetary Econ.* 51, 1599–1621.
- Deviatov, A., Wallace, N., 2001. Another example in which lump-sum money creation is beneficial. *Advances in Macroeconomics* 1, article 1.
- Diamond, D., 1984. Financial intermediation and delegated monitoring. *Review of Economic Studies* 51, 393–414.
- Diamond, D., Dybvig, P., 1983. Bank runs, deposit insurance, and liquidity. *J. Polit. Econ.* 91, 401–419.
- Diamond, P., 1970. A model of price adjustment. *J. Econ. Theory* 3, 156–168.

- Diamond, P., 1982. Aggregate demand management in search equilibrium. *J. Polit. Econ.* 90, 881–894.
- Diamond, P., 1984. Money in search equilibrium. *Econometrica* 52, 1–20.
- Diamond, P., 1993. Search, sticky prices, and inflation. *Review of Economic Studies* 60, 53–68.
- Dong, M., 2010a. Inflation and Variety. *Int. Econ. Rev.* 51, 401–420.
- Dong, M., 2010b. Inflation and unemployment in competitive search equilibrium. *Macroeconomic Dynamics*, forthcoming.
- Dong, M., Jiang, J., 2009. Money and price posting under private information. Bank of Canada Mimeo.
- Dressler, S., 2009. Money holdings, inflation, and welfare in a competitive market. *Int. Econ. Rev.* in press.
- Dressler, S., 2010. Inflation and welfare dynamics in a competitive market. Mimeo.
- Duffie, D., Gârleanu, N., Pederson, L., 2005. Over-the-counter markets. *Econometrica* 73, 1815–1847.
- Duffie, D., Gârleanu, N., Pederson, L., 2008. Valuation in over-the-counter markets. *Review of Financial Studies* 20, 1865–1900.
- Dutu, R., Julien, B., King, I., 2009. Liquidity Constrained Competing Auctions, working paper.
- Ennis, H., 2001. Search, monetary equilibria, and sunspots. *Macroeconomic Dynamics* 5, 132–142.
- Ennis, H., 2008. Search, money, and inflation under private information. *J. Econ. Theory* 138, 101–131.
- Ennis, H., Keister, T., 2009a. Bank Runs and Institutions: The Perils of Interventions. *Am. Econ. Rev.* 99, 1588–1607.
- Ennis, H., Keister, T., 2009b. Run Equilibria in the Green-Lin Model of Financial Intermediation. *J. Econ. Theory* 144, 1996–2020.
- Ennis, H., Keister, T., 2010. Banking Panics and Policy Responses. *J. Monetary Econ.* 57, 404–419.
- Faig, M., 2006. Divisible money in an economy with villages. Mimeo.
- Faig, M., 2008. Endogenous buyer–seller choice and divisible money in search equilibrium. *J. Econ. Theory* 141, 184–199.
- Faig, M., Huangfu, S., 2007. Competitive search equilibrium in monetary economies. *J. Econ. Theory* 136, 709–718.
- Faig, M., Jerez, B., 2005. A theory of commerce. *J. Econ. Theory* 122, 60–99.
- Faig, M., Jerez, B., 2006. Inflation, prices, and information in competitive search. *Advances in Macroeconomics* 6 (1).
- Faig, M., Li, Z., 2009. The welfare costs of expected and unexpected inflation. *J. Monetary Econ.* 56, 1004–1013.
- Ferraris, L., Watanabe, M., 2010. Liquidity constraints in a monetary economy. *Int. Econ. Rev.* forthcoming.
- Freeman, S., 1985. Transaction Costs and the Optimal Quantity of Money. *J. Polit. Econ.* 93, 146–157.
- Freeman, S., 1988. Banking as the provision of liquidity. *Journal of Business* 61, 45–64.
- Freeman, S., 1996. The payments system, liquidity, and rediscounting. *Am. Econ. Rev.* 86, 1126–1138.
- Friedman, M., 1960. A program for monetary stability. Fordham University Press, New York.
- Friedman, M., 1968. The role of monetary policy. *Am. Econ. Rev.* 58, 1–17.
- Friedman, M., 1969. The optimum quantity of money and other essays. Aldine Publishing Company, New York.
- Friedman, M., 1977. Inflation and Unemployment. *J. Polit. Econ.* 85, 451–472.
- Friedman, M., Schwartz, A., 1963. A Monetary History of the United States, 1867–1960. National Bureau of Economic Research, Cambridge, MA.
- Galenianos, M., Kircher, P., 2008. A model of money with multilateral matching. *J. Monetary Econ.* 55, 1054–1066.
- Geromichalos, A., Licari, J., Lledo, J., 2007. Asset prices and monetary policy. *Review of Economic Dynamics* 10, 761–779.
- Geromichalos, A., Licari, J., Lledo, J., 2010. The role of money in a model with frictional investment markets. Mimeo.
- Golosov, M., Lucas, R., 2005. Menu costs and Phillips curves. *J. Polit. Econ.* 115, 171–199.
- Gomis-Porqueras, P., Peralta-Alva, A., 2009. Optimal monetary and fiscal policies in a search theoretic model of monetary exchange. *European Economic Review*, in press.
- Green, E., Weber, W., 1996. Will the new \$100 bill decrease counterfeiting? *Federal Reserve Bank of Minneapolis Quarterly Review* 20 (Summer), 3–10.

- Green, E., Zhou, R., 1998. A rudimentary random-matching model with divisible money and prices. *J. Econ. Theory* 81, 252–271.
- Guerrieri, V., Lorenzoni, G., 2009. Liquidity and trading dynamics. *Econometrica* 77, 1751–1790.
- Hahn, F.H., 1973. On the foundations of monetary theory. In: Parkin, M., Nobay, A.R. (Eds.), *Essays in Modern Economics*. Barnes & Noble, New York.
- Hansen, G., 1985. Indivisible labor and the business cycle. *J. Monetary Econ.* 16, 309–337.
- He, P., Huang, L., Wright, R., 2005. Money and banking in search equilibrium. *Int. Econ. Rev.* 46, 637–670.
- He, P., Huang, L., Wright, R., 2008. Money, banking, and monetary policy. *J. Monetary Econ.* 55, 1013–1024.
- Head, A., Kumar, A., 2005. Price dispersion, inflation, and welfare. *Int. Econ. Rev.* 46, 533–572.
- Head, A., Kumar, A., Lapham, B., 2008. Market power, price adjustment, and inflation. *Int. Econ. Rev.* in press.
- Head, A., Liu, L., Menzio, G., Wright, R., 2010. Sticky prices. Mimeo.
- Hicks, J.R., 1935. A suggestion for simplifying the theory of money. *Economica* 2, 1–19.
- Hoerova, M., Monnet, C., Temzelides, T., 2007. Public information and monetary policy. Mimeo.
- Hosios, A., 1990. On the efficiency of matching and related models of search and unemployment. *Review of Economic Studies* 57, 279–298.
- Howitt, P., 2005. Beyond search: fiat money in organized exchange. *Int. Econ. Rev.* 46, 405–429.
- Hu, T., Kennan, J., Wallace, W., 2009. Coalition-proof trade and the Friedman rule in the Lagos-Wright Model. *J. Polit. Econ.* 117, 116–137.
- Huggett, M., 1993. The risk-free rate in heterogeneous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control* 17, 953–969.
- Jacquet, N., Tan, S., 2009. The liquidity effects of monetary policy on asset prices. Singapore Management University and National University of Singapore, Working paper.
- Jean, K., Stanislav, R., Wright, R., 2010. On the multiplicity of monetary equilibria Green-Zhou meets Lagos-Wright. *J. Econ. Theory* 145, 392–401.
- Jevons, W., 1875. *Money and the mechanism of exchange*. Appleton, London.
- Johri, A., Leach, J., 2002. Middlemen and the allocation of heterogeneous goods. *Int. Econ. Rev.* 43, 347–361.
- Jones, R., 1976. The origin and development of media of exchange. *J. Polit. Econ.* 84, 757–775.
- Julien, B., Kennes, J., King, I., 2008. Bidding for money. *J. Econ. Theory* 142, 196–217.
- Kahn, C., 2009. Collateral policy in a world of round-the-clock payment. University of Illinois at Urbana-Champaign, Working paper.
- Kareken, J., Wallace, N., 1980. *Models of Monetary Economies*. Federal Reserve Bank of Minneapolis, Minneapolis, MN.
- Katzman, B., Kennan, J., Wallace, N., 2003. Output and price level effects of monetary uncertainty in a matching model. *J. Econ. Theory* 108, 217–255.
- Kehoe, T., Kiyotaki, N., Wright, R., 1993. More on money as a medium of exchange. *Econ. Theory* 3, 297–314.
- Kim, Y., 1996. Money, barter, and costly information acquisition. *J. Monetary Econ.* 37, 119–142.
- Kiyotaki, N., Matsui, A., Matsuyama, K., 1993. Towards a Theory of International Currency. *Rev. Econ. Stud.* 60, 283–307.
- Kiyotaki, N., Moore, J., 2002. Evil is the root of all money. Mimeo.
- Kiyotaki, N., Wright, R., 1989. On money as a medium of exchange. *J. Polit. Econ.* 97, 927–954.
- Kiyotaki, N., Wright, R., 1991. A contribution to the pure theory of money. *J. Econ. Theory* 53, 215–235.
- Kiyotaki, N., Wright, R., 1993. A search-theoretic approach to monetary economics. *Am. Econ. Rev.* 83, 63–77.
- Klenow, P., Kryvtsov, O., 2008. State-dependent or time-dependent pricing: Does it matter for recent U.S. inflation? *Quarterly Journal of Economics* 123, 863–904.
- Klenow, P., Malin, B., 2010. Microeconomic evidence on price-setting. *Handbook of Monetary Economics*, in press.

- Kocherlakota, N., 1998. Money is memory. *J. Econ. Theory* 81, 232–251.
- Kocherlakota, N., Wallace, N., 1998. Incomplete record-keeping and optimal payment arrangements. *J. Econ. Theory* 81, 272–281.
- Krusell, P., Smith, T., 1998. Income and wealth heterogeneity in the macroeconomy. *J. Polit. Econ.* 106, 867–896.
- Lagos, R., 2008. Asset prices and liquidity in an exchange economy. New York University Working paper.
- Lagos, R., 2009. Some results on the optimality and implementation of the Friedman Rule in the search theory of money. *J. Econ. Theory*, in press.
- Lagos, R., Rocheteau, G., 2005. Inflation, output, and welfare. *Int. Econ. Rev.* 46, 495–522.
- Lagos, R., Rocheteau, G., 2007. Search in asset markets: Market structure, liquidity, and welfare. *American Economic Review Papers & Proceedings* 97, 198–202.
- Lagos, R., Rocheteau, G., 2008. Money and capital as competing media of exchange. *J. Econ. Theory* 142, 247–258.
- Lagos, R., Rocheteau, G., 2009. Liquidity in asset markets with search frictions. *Econometrica* 77, 403–426.
- Lagos, R., Rocheteau, G., Weill, P., 2009. Crises and liquidity in over-the-counter markets. Mimeo.
- Lagos, R., Wright, R., 2003. Dynamics, cycles and sunspot equilibria in “genuinely dynamic, fundamentally disaggregative” models of money. *J. Econ. Theory* 109, 156–171.
- Lagos, R., Wright, R., 2005. A unified framework for monetary theory and policy analysis. *J. Polit. Econ.* 113, 463–484.
- Lee, M., Wallace, N., Zhu, T., 2005. Modeling denomination structures. *Econometrica* 73, 949–960.
- Lester, B., 2009. Settlement systems. *The B.E. Journal of Economics* 9, article 17.
- Lester, B., Postlewaite, A., Wright, R., 2009. Information and liquidity. *Journal of Money, Credit and Banking*, in press.
- Lester, B., Postlewaite, A., Wright, R., 2010. Liquidity, information, asset prices and monetary policy. University of Pennsylvania, Working paper.
- Li, V., 1994. Inventory accumulation in a search-based monetary economy. *J. Monetary Econ.* 34, 511–536.
- Li, V., 1995. The optimal taxation of fiat money in search equilibrium. *Int. Econ. Rev.* 36, 927–942.
- Li, Y., 1995. Commodity money under private information. *J. Monetary Econ.* 36, 573–592.
- Li, Y., 1999. Money and middlemen in an economy with private information. *Econ. Inq.* 37, 1–12.
- Li, Y., 2007. Inside money, organized markets, and specialization. *Macroeconomic Dynamics* 11, 388–404.
- Li, Y., Rocheteau, R., 2009. The threat of counterfeiting. *Macroeconomic Dynamics*, in press.
- Li, Y., Rocheteau, R., 2010. Liquidity constraints. Mimeo.
- Liu, L., 2009. Inflation and unemployment: The roles of goods and labor market institutions. Mimeo.
- Liu, L., Wang, L., Wright, R., 2010. The “hot potato” effect of inflation. *Macroeconomic Dynamics*, in press.
- Lucas, R., 1972. Expectations and the neutrality of money. *J. Econ. Theory* 4, 103–124.
- Lucas, R., 1976. Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy* 1, 19–46.
- Lucas, R., 1978. Asset prices in an exchange economy. *Econometrica* 46, 1426–1445.
- Lucas, R., 1980. Equilibrium in a pure currency economy. In: Kareken, J., Wallace, N. (Eds.), *Models of Monetary Economies*. Federal Reserve Bank of Minneapolis, Minneapolis, MN, Minneapolis, MN, pp. 131–145.
- Lucas, R., 2000. Inflation and welfare. *Econometrica* 68, 247–274.
- Lucas, R., Prescott, E., 1974. Equilibrium search and unemployment. *J. Econ. Theory* 7, 188–209.
- Lucas, R., Stokey, N., 1987. Money and interest in a cash-in-advance economy. *Econometrica* 55, 491–515.
- Mankiw, N., 1985. Small menu costs and large business cycles: A macroeconomic model. *Quarterly Journal of Economics* 100, 529–537.
- Menzio, G., Shi, S., Sun, H., 2009. A monetary theory with non-degenerate distributions. Mimeo.

- Martin, F., 2009. On the joint determination of fiscal and monetary policy. Simon Fraser University Working paper.
- Midrigan, V., 2007. Menu costs, multi-product firms and aggregate fluctuations. Mimeo.
- Molico, M., 2006. The distribution of money and prices in search equilibrium. *Int. Econ. Rev.* 47, 701–722.
- Mortensen, D., Pissarides, C., 1994. Job creation and job destruction in the theory of unemployment. *Review of Economic Studies* 61, 397–416.
- Nakamura, Steinsson, J., 2008. Five facts about prices: A reevaluation of menu cost models. *Quarterly Journal of Economics* 123, 1415–1464.
- Nosal, E., 2010. Search, welfare and the hot potato effect of inflation. *Macroeconomic Dynamics*, in press.
- Nosal, E., Rocheteau, G., 2011. Money, payments, and liquidity. MIT Press, Cambridge, MA.
- Nosal, E., Wallace, N., 2007. A model of (the threat of) counterfeiting. *J. Monetary Econ.* 54, 994–1001.
- Ostroy, J., Starr, R., 1990. The transactions role of money. In: Friedman, B., Hahn, F. (Eds.), *Handbook of Monetary Economics*. North-Holland, Amsterdam.
- Pissarides, C., 2000. *Equilibrium unemployment theory*. MIT Press, Cambridge, MA.
- Ravikumar, B., Shao, E., 2006. Search frictions and asset price volatility. University of Iowa, Working paper.
- Redish, A., Weber, W., 2010. Coin sizes and payments in commodity money systems. *Macroeconomic Dynamics*, in press.
- Ritter, J., 1995. The transition from barter to fiat money. *American Economic Review* 85, 134–149.
- Rocheteau, G., 2009. A monetary approach to asset liquidity. University of California-Irvine, Working paper.
- Rocheteau, G., Rupert, P., Shell, K., Wright, R., 2008. General equilibrium with nonconvexities and money. *J. Econ. Theory* 142, 294–317.
- Rocheteau, G., Rupert, P., Wright, R., 2007. Inflation and unemployment in general equilibrium. *Scandinavian Journal of Economics* 109, 837–855.
- Rocheteau, G., Wright, R., 2005. Money in search equilibrium, in competitive equilibrium, and in competitive search equilibrium. *Econometrica* 73, 175–202.
- Rogerson, R., 1988. Indivisible labor, lotteries, and equilibrium. *J. Monetary Econ.* 21, 3–16.
- Rupert, P., Schindler, M., Wright, R., 2001. A generalized search-theoretic model of monetary exchange. *J. Monetary Econ.* 48, 605–622.
- Samuelson, P., 1958. An exact consumption-loan model with or without the social contrivance of money. *J. Polit. Econ.* 66, 467–482.
- Sanches, D., Williamson, S., 2010. Money and credit with limited commitment and theft. *J. Econ. Theory* 145 (4), 1525–1549.
- Sanches, D., Williamson, S., 2010. Adverse selection, segmented markets and the role of monetary policy. *Macroeconomic Dynamics*, in press.
- Schevchenko, A., 2004. Middlemen. *Int. Econ. Rev.* 45, 1–24.
- Schevchenko, A., Wright, R., 2004. A simple search model of money with heterogeneous agents and partial acceptability. *Econ. Theory* 24, 877–885.
- Shell, K., Wright, R., 1993. Indivisibilities, lotteries, and sunspot equilibria. *Econ. Theory* 3, 1–17.
- Shi, S., 1995. Money and prices: A model of search and bargaining. *J. Econ. Theory* 67, 467–496.
- Shi, S., 1996. Credit and money in a search model with divisible commodities. *Review of Economic Studies* 63, 627–652.
- Shi, S., 1997a. Money and specialization. *Econ. Theory* 10, 99–133.
- Shi, S., 1997b. A divisible model of fiat money. *Econometrica* 65, 75–102.
- Shi, S., 2006. Viewpoint: A microfoundation of monetary economics. *Canadian Journal of Economics* 39, 643–688.
- Silveira, R., Wright, R., 2010. Search and the market for ideas. *J. Econ. Theory* 145 (4), 1550–1573.
- Telyukova, I., Visschers, L., 2009. Precautionary demand for money in a monetary business cycle model. University of California San Diego and Simon Fraser University, Working paper.
- Telyukova, I., Wright, R., 2008. A model of money and credit, with application to the credit card debt puzzle. *Review of Economic Studies* 75, 629–647.
- Tobin, J., 1956. The interest-elasticity of transactions demand for cash. *Rev. Econ. Stat.* 38, 241–247.

- Townsend, R., 1987. Economic organization with limited communication. *Am. Econ. Rev.* 77, 954–970.
- Townsend, R., 1989. Currency and credit in a private information economy. *J. Polit. Econ.* 97, 1323–1345.
- Trejos, A., 1999. Search, bargaining, money, and prices under private information. *Int. Econ. Rev.* 3, 679–695.
- Trejos, A., Wright, R., 1995. Search, bargaining, money, and prices. *J. Polit. Econ.* 103, 118–141.
- Velde, F., Weber, W., Wright, R., 1999. A model of commodity money, with applications to Gresham's law and the debasement puzzle. *Review of Economic Dynamics* 2, 291–323.
- Wallace, N., 1980. The overlapping generations model of fiat money. In: Kareken, J., Wallace, N. (Eds.), *Models of monetary economics*. Federal Reserve Bank of Minneapolis, Minneapolis, MN.
- Wallace, N., 1992. Lucas's signal extraction model: A finite-state exposition with aggregate real shocks. *J. Monetary Econ.* 30, 433–447.
- Wallace, N., 1997. Short-run and long-run effects of changes in money in a random matching model. *J. Polit. Econ.* 105, 1293–1307.
- Wallace, N., 1998. A dictum for monetary theory. *Federal Reserve Bank of Minneapolis Quarterly Review* (Winter), 20–26.
- Wallace, N., 2001. Whither monetary economics? *Int. Econ. Rev.* 42, 847–869.
- Wallace, N., Zhou, R., 1997. A model of a currency shortage. *J. Monetary Econ.* 40, 555–572.
- Wallace, N., Zhu, T., 2007a. Pairwise trade and coexistence of money and higher return assets. *J. Econ. Theory* 133, 524–535.
- Wallace, N., Zhu, T., 2007b. Float on a note. *J. Monetary Econ.* 54, 229–246.
- Waller, C., 2009. Dynamic taxation, private information and money. *Federal Reserve Bank of St. Louis, Working Paper* 2009-035.
- Waller, C., 2010. Random matching and money in the neoclassical growth model: Some analytical results. *Working paper, Macroeconomic Dynamics*, in press.
- Weill, P., 2007. Leaning against the wind. *Review of Economic Studies* 74, 1329–1354.
- Weill, P., 2008. Liquidity premia in dynamic bargaining markets. *J. Econ. Theory* 140, 66–96.
- Weill, P., Vayanos, D., 2008. A search-based theory of the on-the-run phenomenon. *Journal of Finance* 63, 1351–1389.
- Wicksell, K., 1967. *Lectures on political economy*. Vol. 2. Money. 1911, second ed. Kelley, New York (E. Classen, Trans.).
- Williamson, S., 1986. Costly monitoring, financial intermediation and equilibrium credit rationing. *J. Monetary Econ.* 18, 159–179.
- Williamson, S., 1987. Financial intermediation, business failures, and real business cycles. *J. Polit. Econ.* 95, 1196–1216.
- Williamson, S., 1999. Private money. *Journal of Money, Credit, and Banking* 31, 469–491.
- Williamson, S., 2007. Search, limited participation, and monetary policy. *Int. Econ. Rev.* 47, 107–128.
- Williamson, S., Wright, R., 1994. Barter and monetary exchange under private information. *Am. Econ. Rev.* 84, 104–123.
- Williamson, S., Wright, R., 2010. New monetarist economics: Methods. *Federal Reserve Bank of St. Louis Review* 92 (4), 265–302.
- Woodford, M., 2003. *Interest and prices*. Princeton University Press, Princeton, NJ.
- Wright, R., 1995. Search, evolution and money. *Journal of Economic Dynamics and Control* 19, 181–206.
- Wright, R., 2010. A uniqueness proof for monetary steady state. *J. Econ. Theory* 145, 382–391.
- Zhou, R., 1997. Currency exchange in a random search model. *Review of Economic Studies* 64, 289–310.
- Zhu, T., 2003. Existence of a monetary steady state in a matching model: Indivisible money. *J. Econ. Theory* 112, 307–324.
- Zhu, T., 2005. Existence of a monetary steady state in a matching model: Divisible money. *J. Econ. Theory* 123, 130–160.

Money and Inflation: Some Critical Issues★

Bennett T. McCallum* and Edward Nelson**

*Carnegie Mellon University

**Federal Reserve Board

Contents

1. Introduction	98
2. The Quantity Theory of Money	99
3. Related Concepts	102
4. Historical Behavior of Monetary Aggregates	104
5. Flawed Evidence on Money Growth-Inflation Relations	108
5.1 Evidence on money demand stability	109
5.2 Evidence with country-average data	111
6. Money Growth and Inflation in Time Series Data	112
6.1 Is long averaging of data required?	113
6.2 Money growth/inflation dynamics in a New Keynesian model	115
6.3 Nominal spending and inflation	121
6.4 Money growth per unit of output and inflation	123
6.5 Time series evidence	125
6.6 Panel data evidence for the G7	129
6.7 Money demand nominal homogeneity	131
7. Implications of A Diminishing Role for Money	134
8. Money Versus Interest Rates In Price Level Analysis	136
8.1 Conditions for excluding money from the analysis	137
8.2 Determinacy and learnability	140
8.3 Fiscal theory of the price level	141
8.4 Money as an information variable	142
9. Conclusions	146
References	148

★ The editors, as well as Jeffrey Fuhrer, Stephanie Schmitt-Grohé, and other participants at the Conference on Key Developments in Monetary Economics (Federal Reserve Board, October 2009), provided useful comments on an earlier draft. We thank Richard Anderson and Fabrizio Orrego for useful discussions and Kathleen Easterbrook for research assistance. The views expressed in this paper are solely the responsibility of the authors, and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System.

Abstract

We consider what, if any, relationship there is between monetary aggregates and inflation, and whether there is any substantial reason for modifying the current mainstream mode of policy analysis, which frequently does not consider monetary aggregates at all. We begin by considering the body of thought known as the “quantity theory of money.” The quantity theory centers on the prediction that there will be a long-run proportionate reaction of the price level to an exogenous increase in the nominal money stock. The nominal homogeneity conditions that deliver the quantity-theory result are the same as those that deliver monetary neutrality, an important principle behind policy formulation. The quantity theory implies a *ceteris paribus* unitary relationship between inflation and money growth. Simulations of a New Keynesian model suggest that we should expect this relationship to be apparent in time series data, with no heavy averaging or filtering required, but with allowance needed for the phase shift in the relationship between monetary growth rates and inflation. While financial innovation can obscure the relationship between monetary growth and inflation, evidence of a money growth/inflation relationship does emerge from the United States time series and G7 panel data. Various considerations suggest that studies of inflation and monetary policy behavior can benefit from including both interest rates and money in the empirical analysis.

JEL classification: E31, E50, E52

Keywords

Monetary Aggregates

Inflation

Interest Rates

Monetary Policy

1. INTRODUCTION

Extensive and well-publicized developments of the past two decades, most of which are amply documented in contributions to the present *Handbook*, have greatly reduced the role of monetary aggregates in basic monetary theory and especially in monetary policy analysis. Thus, as is well known, today’s mainstream approach to monetary policy analysis presumes that policy rules reflect period-by-period adjustments of a short-term interest rate — not any monetary aggregate. In addition, the model of private sector behavior is typically written in a manner that includes no reference to any monetary aggregate; this is an approximation, in economies that possess a medium of exchange, but one that seems to be satisfactory for policy purposes. Consequently, policy models need not refer to monetary aggregates at all, even when the economy in question does utilize a medium of exchange. Since these models are intended to explain behavior of inflation, as well as movements in aggregate demand and the policy interest rate, current analysis typically ignores the relationship between money and inflation.

The task of this chapter is, accordingly, to consider what if any relationship there is between these variables, and whether there is any substantial reason for modifying the current mainstream mode of policy analysis. The chapter outline is as follows.

In Section 2, we begin with some reflections on the body of thought known as the Quantity Theory of Money. Section 3 is then concerned with related theoretical topics, while Sections 4 to 6 consider empirical regularities relating to money growth and inflation. In Section 7 we turn to the implications of a declining demand for a medium of exchange, and in Section 8 we consider analyses of price level determination that posit interest-rate policy rules. Section 9 presents conclusions.

2. THE QUANTITY THEORY OF MONEY

Any exploration of the relationship between money and inflation almost necessarily begins with a discussion of the venerable “quantity theory of money” (QTM). There is, nevertheless, considerable disagreement over the meaning of this body of analysis. Popular treatments, and some textbooks, often begin by associating the QTM with the equation of exchange, $MV = PY$, where M , Y , and P , respectively, denote measures of the nominal quantity of money, real transactions or physical output per period, and the price level, with V then being the corresponding monetary “velocity.” An outline of the equation of exchange is perhaps acceptable as the beginning of an exposition of the QTM. But it would be unfortunate to take the QTM and the equation of exchange as interchangeable. The equation of exchange is an identity — it might appropriately be thought of as a definition of velocity. Being an identity, the equation of exchange is consistent with *any* proposition concerning monetary behavior and, in the absence of restrictions on the behavior of any terms in the equation, cannot be used to characterize a specific monetary theory. To take the QTM as equivalent to the equation of exchange would, consequently, deprive it of any empirical or theoretical content.

That somewhat different meanings are assigned to the QTM by different writers can be seen by consulting the writings of Hume (1752), Wicksell (1915/1935), Fisher (1913), Keynes (1936), Friedman (1956, 1987), Patinkin (1956, 1972), Samuelson (1967), Niehans (1978), and Lucas (1980). In fact, the later writers have had in mind quantities of fiat (paper) money whereas the earlier ones were discussing quantities of metallic money. David Hume’s treatments (1752) considered both the case where an increase in (metallic or paper) money leads to a gradual, proportional rise in prices, and the case of an open economy where the expansion in metallic money results in an export of that money. Nevertheless, for the currently relevant case of fiduciary money, there seems to be one basic proposition characterizing the QTM; that is, one common thread that unites various definitions and applications. This proposition is that *if* a change in the quantity of (nominal) money were exogenously engineered by the monetary authority, then the long-run effect would be a change in the price level (and other nominal variables) of the same proportion as the money stock, with no change resulting in the value of any real variable.¹ This proposition pertains to “long-run” effects; that is, effects that would occur

¹ This statement concerns effects of the single postulated exogenous change.

hypothetically after all adjustments are completed. In real time, there will always be changes occurring in tastes or technology before full adjustment can be effected, so no experiment of this kind can literally be carried out in actual economies. Furthermore, in most actual economies the monetary authority does not conduct monetary policy to generate exogenous changes in the stock of money, so nothing even approximating the hypothetical experiment is ever attempted in reality.

Does the foregoing imply that no statement with empirical content can be made about the QTM? We suggest not; the essential point is that the basic QTM proposition given earlier holds in a model economy if, and only if, the model exhibits the property known as long-run “neutrality of money.” Indeed, the latter concept is defined to satisfy the stated proposition. Accordingly, we argue that the QTM amounts to the claim that actual economies possess the properties that imply long-run monetary neutrality.

This position is closer to that of Patinkin (1972) than that of Friedman (1972a), in their celebrated exchange, since Friedman (1956, 1972a) preferred to regard the QTM as a proposition exclusively about the demand function for money. Other expositions of Friedman (1987) did, however, treat the QTM as centering on the distinction between the nominal quantity of money (whose path is implied by the choices of the monetary authority) and the real quantity of money (whose path is determined by the choices of the private sector). The model property that separates the determination of the real and nominal quantities of money corresponds to the long-run monetary neutrality property. Friedman’s emphasis on the demand function for money is therefore reconcilable with an identification of the QTM with monetary neutrality, in the sense that price homogeneity of the money demand function is crucial for long-run monetary neutrality.²

Indeed, long-run monetary neutrality is dependent on homogeneity properties holding across the private sector’s main behavioral relations. Basically, private agents’ objective functions and technology constraints should be formulated entirely in terms of real variables — there is no concern by rational private agents for the levels of nominal magnitudes.³ Then implied supply and demand equations will also include only real variables; they will be homogenous of degree zero in nominal variables.⁴ Since supply and demand relations can be estimated econometrically, the QTM has empirical content for structural modeling. It requires that all supply and demand equations have the stated homogeneity property. These equations, if properly formulated, are structural relations that do not

² In addition, Friedman (1956) argued that an infinite interest elasticity of the demand function for money is inconsistent with the quantity theory. This constitutes a further overlap of Friedman’s conception of the QTM and that used here, as an infinite interest elasticity must be ruled out to produce the monetary neutrality result.

³ The government’s tax regime might imply that budget constraints cannot be written entirely in real terms. For simplicity, we abstract from this case.

⁴ Note that in this (standard) case, the monetary authority must follow a rule that depends upon some nominal variable. Otherwise, nominal indeterminacy will prevail — the model will fail to determine the value of any nominal variable. This is substantially different from the type of “indeterminacy” featured in the recent literature, which is the existence of more than one dynamically stable rational-expectations solution.

depend upon the policy rule in effect.⁵ Their validity or invalidity therefore has nothing to do with the operating procedures of the monetary authority. The QTM does not, consequently, have anything to do with “the exogeneity of money” in actual practice. In particular, it does not matter whether the central bank is using an interest rate or a monetary aggregate (or, say, the price of foreign exchange) as its instrument variable.

One of the relations in any complete model for a monetary economy is a demand function for real money balances. As noted previously, one condition for long-run neutrality to prevail is that this function must relate the demand for real balances only to real variables (usually including a real rate of return differential that is the opportunity cost of holding money⁶ and a real transactions quantity). The money demand relation then implies that the steady-state inflation rate will equal the steady-state rate of growth of the money stock minus a term pertaining to the rate of growth of output or real transactions. An exogenous change (if it somehow occurred) in the rate of growth of the money stock would, therefore, induce a change of the same magnitude in the inflation rate unless it induced a change in the *rate of growth* of real transactions or the real interest differential. Neither of these possibilities seems likely, so the QTM essentially implies that steady-state inflation rates move one-for-one with steady-state money growth rates.

The earlier exposition of the QTM, in terms of private reactions to an exogenous policy action, would appear at first glance to leave out what is widely regarded as an important policy implication of the QTM. Many observers have noted that the QTM rules out autonomous factors such as increases in the prices of specific types of good (such as food or energy) from being sources of sustained movements in prices. The position is that, by holding the money stock constant in the face of an increase in the price of a specific good, the monetary authority can prevent total nominal spending, and thus the aggregate price level, from undergoing a sustained increase. A stress with the critical importance of monetary “accommodation” in price level determination underlies Samuelson’s (1967) characterization of the QTM and is embedded in many textbook treatments (e.g., Mishkin, 2007). In fact, this element is encompassed by the QTM definition previously given. Although our statement focused on a policy-induced monetary increase, the process described in the wake of that increase involves a price level reaction that is complete once prices have restored their proportional relation to money. A model in which prices are unrestrained by the extent of monetary accommodation would imply that an initial price level increase can trigger an indefinite price level spiral. Thus, our QTM definition, although expressed in terms of exogenous policy actions, involves restrictions on model behavior that imply that the monetary policy response to nonpolicy shocks is crucial in determining the repercussions for price level behavior of those shocks.

⁵ Here we have in mind behavioral relations, for example, Euler equations.

⁶ This differential is the difference between the real, and nominal, rates of return on money and interest-bearing assets. For simplicity, we assume that money is, like actual currency, not interest-bearing, in which case the differential equals the nominal interest rate.

3. RELATED CONCEPTS

Other concepts, related to but distinct from the QTM's long-run monetary neutrality, deserve brief mention. The first of these is the *superneutrality* of money. The QTM proposition, with its implication that steady-state inflation rates move one-for-one with steady-state money growth rates, does not imply that different maintained money-growth (and inflation) rates have no lasting effect on real variables. In particular, it does not rule out permanent effects on *levels* of output, consumption, real interest rates, and so forth. A higher inflation rate, for example, typically implies an increased nominal interest rate and therefore an elevated spread between the rates of return on money and securities. Such a change raises the interest income foregone when holding real money balances, so rational agents will reduce the fraction of their assets held in the form of money. In many cases, the implied type of portfolio readjustment will lead to changes in the steady-state capital/labor and capital/output ratios, which are key real variables.

In the case where *no* change in real variables occurs with altered steady-state inflation rates, the economy is said to possess the property of "superneutrality." From what has been said, however, it should be clear that superneutrality should not be expected to hold in economies in which money provides transactions-facilitating services, as it does normally in most actual economies. It is plausible that the departures from superneutrality in practice will be small, for reasons discussed in McCallum (1990). Thus, for example, a shift in the steady-state inflation rate from 0% (per annum) to 5% might imply a fall in the steady-state real rate of interest of perhaps only about 0.04%.⁷ Superneutrality will therefore be a property that holds approximately.

One of the variables that is insensitive to alternative ongoing inflation rates when superneutrality holds is the real rate of interest (e.g., the one-period real rate). The absence of superneutrality, on the other hand, implies that a change in the steady-state inflation rate may change the steady-state real rate of interest. It should be noted that such a change is entirely consistent with the so-called "Fisher equation," which in its linearized form may be written as $r_t = R_t - E_t \pi_{t+1}$ (with π being the net rate of inflation). The latter should be thought of as an identity; that is, as a definition of r_t .⁸ The literature arguably contains some confusion on this matter, with some writers treating the Fisher equation as a behavioral equation that separates nominal from real variables, going on to claim that the Fisher equation is contradicted if an altered inflation rate produces a (steady-state) shift in the real interest rate. In the Sidrauski-Brock model, the steady-state real rate of interest is indeed independent of the steady-state rate of inflation, but the same feature is *not* true in a typical overlapping-generations model, even though the Fisher equation holds in both models (see McCallum, 1990).

There is another widely used concept involving long-run relationships, a distinct property in its own right but sometimes incorrectly regarded as part and parcel of

⁷ For this calculation, involving specific assumptions about functional forms and quantitative magnitudes, see McCallum (2000a, pp. 876–879).

⁸ Actually, the exact discrete-time expression is $(1 + R_t) = (1 + r_t)(1 + E_t \pi_{t+1})$.

superneutrality. This is the “natural rate hypothesis” (NRH), introduced by Friedman (1966, 1968) and refined by Lucas (1972). Friedman’s version of this hypothesis states that differing steady-state inflation rates will not keep output (or employment) permanently high or low relative to the “natural-rate” levels that would prevail in the absence of nominal price stickiness. Lucas’s version is stronger; it states that there is no monetary policy that can permanently keep output (or employment) away from its natural-rate value, not even an ever-increasing (or ever-decreasing) inflation rate. Note the distinction between these concepts and superneutrality: an economy could be one in which superneutrality does not obtain, in the sense that different permanent inflation rates lead to different steady-state levels of capital and thus natural levels of output, but the economy would nevertheless satisfy the natural-rate hypothesis.

The validity of the NRH, or Friedman’s weaker version called the “accelerationist” hypothesis, was a subject of considerable debate starting in the late 1960s. Lucas (1972) and Sargent (1971) pointed out that the initial tests (such as those of Solow, 1969) were inconsistent with rational expectations, and later evidence favored the NRH, which by the early 1980s had become integrated even into Keynesian treatments (see, e.g., Gordon, 1978, or Baumol and Blinder, 1982). In the last decade and half, however, what is in effect an overturning of this consensus has occurred, thanks to the widespread adoption of the Calvo (1983) specification of nominal price adjustment. The basic discrete-time form of the Calvo specification implies that in any period only a fraction of sellers may make price adjustments, with all others compelled to hold their nominal prices at their prior values. This assumption leads to the following economy-wide relationship, in which π_t is inflation, γ_t is the log of output, and $\bar{\gamma}_t$ the natural (i.e., flexible-price) level of output:

$$\pi_t = \beta E_t \pi_{t+1} + \kappa(\gamma_t - \bar{\gamma}_t). \quad (1)$$

Here $\kappa > 0$ and β is a discount factor satisfying $0 < \beta < 1$. If we take this relation as referring to level of inflation, it implies a steady-state relationship between inflation and the (constant) output gap; that is., each value of $E[\pi_t]$ is associated with its own constant value of $\gamma_t - \bar{\gamma}_t$. The Calvo adjustment scheme consequently fails to satisfy even the accelerationist hypothesis, still less the stronger NRH. A minimal step toward remedying this situation would be to replace Eq. (1) with something like the following:

$$\pi_t - \pi = \beta(E_t \pi_{t+1} - \pi) + \kappa(\gamma_t - \bar{\gamma}_t), \quad (2)$$

as in Yun, 1996, or Svensson, 2003. Here π represents the steady-state inflation rate under an existing policy rule, assumed to be one that admits a steady-state inflation rate. A relationship such as Eq. (2) would prevail if those sellers who are not given an opportunity (in a given period) to reset their prices optimally, have their prices rise at the trend rate (rather than holding them constant). Equation (2) would imply that on average $\gamma_t - \bar{\gamma}_t$ is zero, thereby satisfying the accelerationist hypothesis, Friedman’s weaker

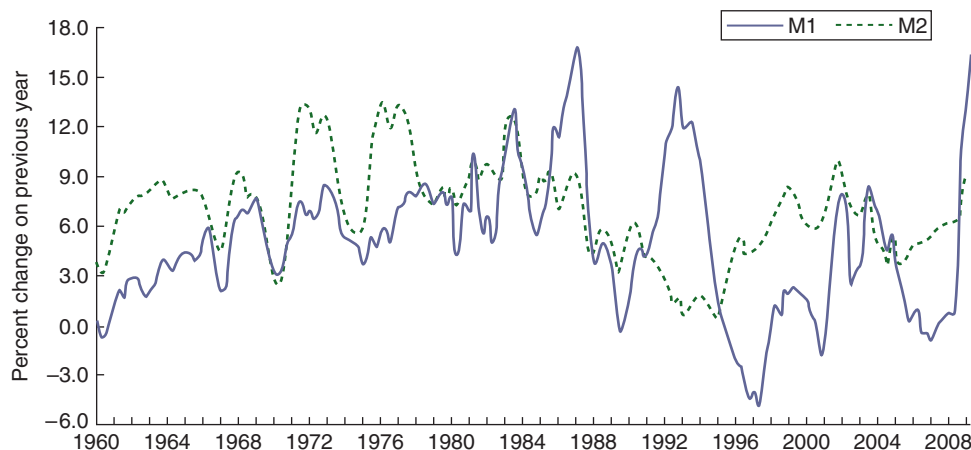


Figure 3.1 Growth in M1 and M2.

version of the NRH. (Even so, specification Eq. (2) does not imply the stronger Lucas version, which pertains to inflation paths more general than steady states.)⁹

4. HISTORICAL BEHAVIOR OF MONETARY AGGREGATES

Some perspective on the behavior of monetary aggregates in the United States is provided by Figure 3.1, which plots quarterly observations on four-quarter growth rates of M1 and M2 since 1959. The modern M1 and M2 series were introduced by the Federal Reserve Board in 1980 (with some minor redefinitions thereafter). These series replaced narrower official definitions of each series.¹⁰ Despite their broader coverage, the pre-1980 growth rates of the modern definitions of M1 and M2 closely match those of the prior definitions. A partial demonstration of this fact is given in Figure 3.2, which plots growth in annual averages of the former M1 aggregate against the corresponding growth in the modern M1 series.¹¹

On the choice between M1 and M2 definitions, Friedman and Schwartz (1970, pp. 2, 92) stated: “important substantive conclusions seldom hinge on which definition is used . . . We have tried to check many of our results to see whether they depend

⁹ A more ambitious step is to utilize a formulation in which price setters choose, in each period, an optimal price and also an optimal rate of increase to pertain in future periods in which no other adjustment is permitted. A recent analysis of a case of this type has been developed by Juillard, Kamenik, Kumhof, and Laxton (2008).

¹⁰ See Hafer (1980) on the differences between old and new monetary aggregate definitions, and Anderson and Kavajecz (1994) on the history of money stock estimates in the United States. Anderson and Kavajecz credited Abbot (1962) with the invention of the “M1” label. The label “M2” for a broad definition that includes time deposits dates at least to Friedman and Meiselman (1963).

¹¹ The source for the data on old M1 used in Figure 2 is Lothian, Cassese, and Nowak (1983); the vintage of the M1 series tabulated there is close to that used by Lucas (1980).

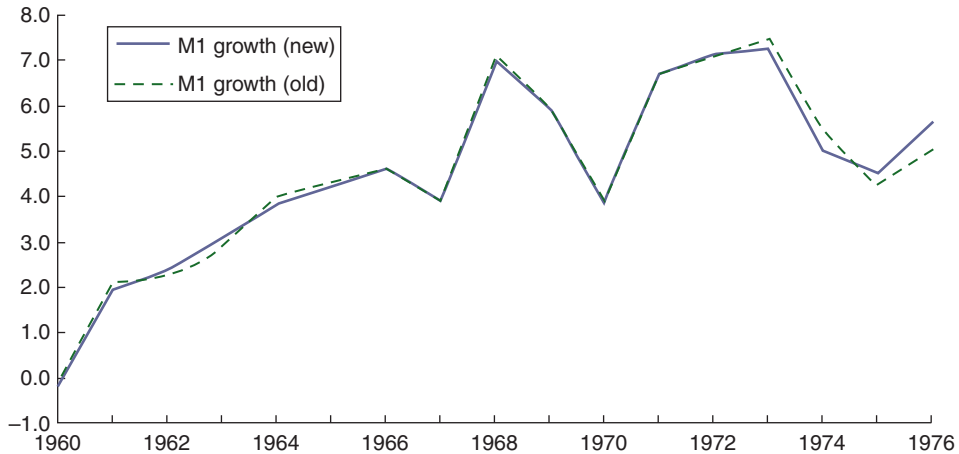


Figure 3.2 Pre-1980 and new definition of M1 (annual averages, percent change).

critically on the specific definition used. Almost always, the answer is that they do not ...”¹² This conclusion has not proved to be durable. For much of the period since 1970, the M1 and M2 series have moved differently. Regulation Q was cited as a factor promoting discrepancies between M1 and M2 growth in the 1960s and 1970s. But the abolition of Regulation Q did not bring an end to the discrepancies between M1 and M2 growth. On the contrary, the deregulated environment prevailing since the early 1980s seems to have perpetuated the differences in the behavior of the rates paid on M1 and non-M1 M2 deposit balances. The result has been an intensification of the discrepancies between the growth rates of the M1 and M2 aggregates.

A change in interest paid on the deposits included in a monetary aggregate (and so a rise in the own-rate on money), holding constant the interest rates on securities, tends to change the *real* demand for that aggregate. Whether this affects the growth rate of the nominal quantity of money depends on the operating procedure of the monetary authority. When the Federal Reserve uses an interest-rate instrument, it must acquiesce to the implications for money growth of its interest-rate choices. Consequently, the discrepancies between M1 growth and M2 growth in practice frequently reflect the different opportunity costs associated with the two aggregates.

Discussions of the effect of financial deregulation and innovation on the behavior of monetary aggregates often include the claim that the advent of payment of interest on M1 deposits has greatly changed the character of M1.¹³ While this argument appears

¹² Similarly, Meltzer (1969, p. 97) stated, “I don’t know of any period in which there would be a substantial difference ... using one rather than the other definition of money as an indicator of monetary policy.”

¹³ For example, the discussion in Lucas (2000, p. 270) suggested that U.S. demand deposits formerly could not bear interest, but now can do so. Many similar statements by other authors could be cited.

to be important for the analysis of the international experience with deregulation,¹⁴ it has limited validity for the United States. The prohibition of interest on demand deposits has in fact never been lifted in the United States. The M1 series, as redefined in 1980, does include, in addition to currency, travelers' checks, and demand deposits, the category of other checkable deposits (OCDs); that is, certain nondemand, checkable deposits that can legally bear interest. The OCD component of M1 rose relative to the demand deposit portion of M1 during most of the 1980s, suggesting that the interest return on OCDs had some attraction to bank customers. But, on the whole, it seems that explicit interest on M1 deposits has not proved to be a major factor affecting portfolio decisions. Convention, surviving regulations, and continuing differences in the transactions services provided by M1 funds compared to non-M1 M2, have all meant that the rate of return on M1 deposits has rarely been attractive relative to other deposit rates even in the era of deregulation.

The fall in M1 velocity in the 1980s has occasionally been attributed to the payment of interest on M1. But M1 velocity movements up to the late 1980s appear to be well captured by the declining opportunity cost of holding money as recorded in market interest rates, without recourse to an explanation that involves a changing own-rate on M1 (Hoffman & Rasche, 1991; Lucas, 1988; Stock & Watson, 1993).

Generally speaking, therefore, the *whole* of M1 is interest sensitive, and a rise in securities market interest rates promotes flows out of M1 balances. By contrast, from the late 1970s onward, the proportion of non-M1 M2 deposits bearing market-related interest rates rose considerably, standing at over 60% by early 1982 (Gramley, 1982). The overall interest sensitivity of M2 arises primarily from the fact that the rates on several classes of deposit, such as retail certificates of deposit, within M2 adjust to securities market interest rates only with a delay.

A different means through which financial innovation affects M1 behavior has proved to be much more significant in practice. The innovation that banks have favored has not made M1 deposits more attractive vehicles, but rather made it easier to shift between M1 deposits and interest-bearing deposits that are outside M1 but included in M2. "Sweeps" programs allow routine transfers, at the banks' initiative, between M1 deposits and non-M1 deposits. An embryonic version of this arrangement developed during the 1970s in the form of automatic transfer services (ATS; see Hafer, 1980), but extensive adoption of retail sweep deposit programs on the part of banks did not take effect until January 1994 (Anderson, 2003). The arrangement is attractive to depositors because of the better returns on non-M1 M2 deposits, and appeals to banks as a means of avoiding the more onerous reserve requirement on M1 deposits. The resulting portfolio behavior is believed to have created variations in M1 that have little macroeconomic meaning, with Anderson (2003, p. 1) arguing, "*Retail-deposit sweep programs are only accounting changes: they do not affect the amounts of*

¹⁴ For example, the table of rates on M1 deposits in the UK provided in Hendry and Ericsson (1991, p. 876) indicates that UK transactions deposits went from non-interest-bearing at the start of 1984 to earning 7.5% annual interest rates on average at the end of the year.

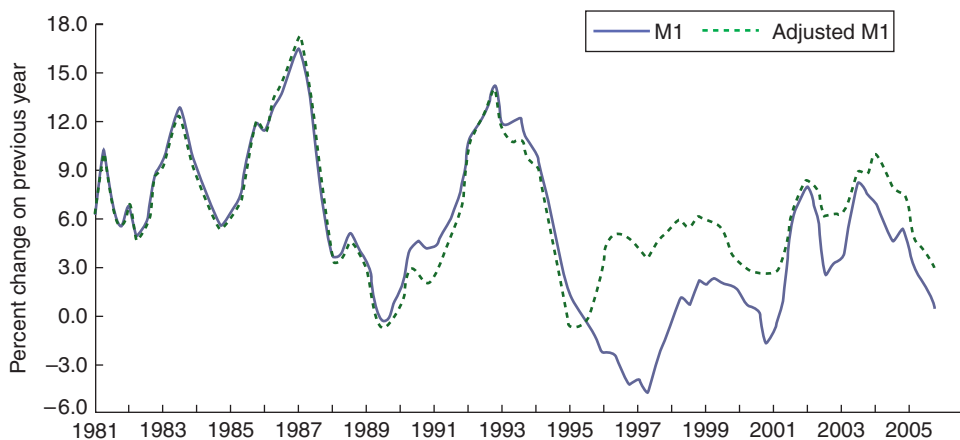


Figure 3.3 Growth in M1 and adjusted M1.

transaction deposits that banks' customers perceive themselves to own." (Italics in original.) A series of studies (including Cynamon, Dutkowsky, & Jones, 2006; Dutkowsky, Cynamon & Jones, 2006; Jones, Dutkowsky, & Elger, 2005) has attempted to correct the U.S. monetary aggregates for the effect of the sweep program. Figure 3.3 plots growth in M1 against growth in an adjusted M1 series. The deposits component of this adjusted series, following Ireland (2009), is based on replacing M1 deposits after 1993 with the Cynamon-Dutkowsky-Jones M1 deposit series that corrects for sweeps. In addition, the adjusted series used in Figure 3.3 subtracts Federal Reserve Board estimates (available from 1964 onward) of U.S. currency held abroad, as reported in the flow of funds. We see from Figure 3.3 that these adjustments, on balance, lead to a more moderate decline in M1 growth during the late 1990s.

Figure 3.4 plots the velocities of M1 and M2. As is well known, the combination prevailing before the early 1980s was of an upward-trending M1 velocity and a stationary M2 velocity. As is also well known, M1 velocity underwent a major break in trend after 1981. (The apparent resumption of an upward M1 velocity trend in the late 1990s is largely illusory, reflecting the sweeps programs.) The presentation of both series on the same scale in Figure 3.4 means that M2 velocity appears very stable over the whole sample. But on closer inspection there emerge several notable shifts in the series — including a fall in M2 velocity with the introduction of money market deposit accounts in 1983 Q1, followed by a major velocity rise in the mid-1990s,¹⁵ and a decline, not fully reversed, that occurred during the monetary policy easing and international turmoil of 2001–2002.

¹⁵ The behavior of M2 demand during the 1990s has been the subject of numerous studies, including Duca (1995), Lown, Peristiani, and Robinson (1999), and Carlson, Hoffman, Keen, and Rasche (2000).

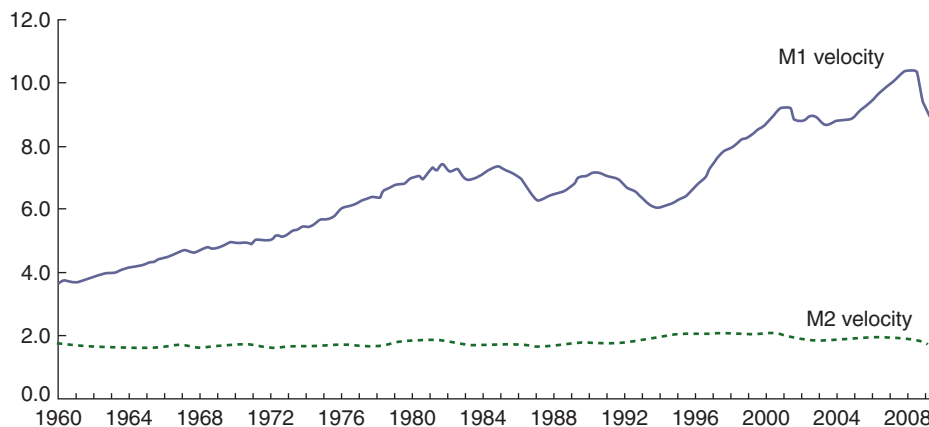


Figure 3.4 Quarterly values of M1 and M2 velocity.

One argument that has been advanced to explain the stability of M2 velocity is that the sweeps program tends to produce variations in M1 that cancel within M2. Beyond this more or less mechanical basis for favoring M2, it is also possible that might be a preferable definition even from the perspective of standard theories of money demand. While the M1 definition was intended to capture the concept of transactions balances, some of the non-M1 components of M2, such as money market deposit accounts, might be used routinely for performing transactions. In that case, the medium-of-exchange concept of money might better be represented by M2. Dorich (2009) argued that M2 should be used as the empirical measure of transactions money, and Reynard (2004) did so excluding one class of M2 deposit (namely, small time deposits, in recent years about one-seventh of M2). Arguing somewhat against the use of M2-type series as measures of transactions money, at least for studies using long sample periods, are the empirical results coming from the Divisia procedure, which Lucas (2000) argued is the best way to construct monetary aggregates. The Divisia procedure produces a series that downweights much of the non-M1 component of M2, and leads to quite different behavior of M2 and Divisia M2 during key episodes in the 1970s and 1980s (see Barnett & Chauvet, 2008).

5. FLAWED EVIDENCE ON MONEY GROWTH-INFLATION RELATIONS

A number of test procedures have been widely advanced as yielding evidence — pro or con — regarding quantity-theory relations between money growth and inflation. Two of the most prominent test procedures, however, are conceptually flawed. These are procedures based on: (i) determination of long-run money demand stability; (ii) regressions of inflation on money growth (or scatterplots of the series) using cross-country averages. We discuss each in turn.

5.1 Evidence on money demand stability

Quantity-theory relations between money growth and inflation do not depend on constancy of all parameters in an estimated money demand function, nor on cointegration among the components of the money demand function. To see this, let us write down a standard money demand equation:

$$\log (M/P)_t = c_0 + c_1 \log (Y_t) - c_2 R_t + c_3 t + e_t \quad (3)$$

where c_1 and c_2 are positive. This is the typical specification (possibly with aggregate consumption C_t substituting for aggregate output Y_t) that would emerge from utility analysis (e.g., Lucas, 1988, 2000; McCallum & Goodfriend, 1987), other than our inclusion of the $c_3 t$ term. This linear trend term is designed to capture smooth progress in payments technology, which we will take as exogenous.¹⁶ If the financial system develops in a way that allows agents to economize on their money holdings over time, then $c_3 < 0$. With a unitary income elasticity and a stationary nominal interest rate, the trend term implies a rising trend in velocity; that is, real balances grow at a slower rate than real income.

Money demand and cointegration studies are often motivated by the claim that money demand stability is a condition for the existence of quantity-theory relations between money growth and inflation. Lucas (1980), however, rejected the alleged dependence of a money growth/inflation link on money demand stability. There are several reasons to support Lucas' position. For example, a unit root in e_t , the money demand shock in Eq. (3), would be considered a violation of dynamic stability in the money demand function, implying no cointegration and, by some definitions, money demand instability; but it would imply a first-difference relation,

$$\Delta \log (M/P)_t = c_1 \Delta \log Y_t - c_2 \Delta R_t + c_3 + \Delta e_t \quad (4)$$

and hence a unitary money growth/inflation relationship, conditional on other variables. In particular, with stationary R_t behavior,

$$E[\Delta \log M_t] = E[\pi_t] + c_3 + c_1 E[\Delta \log Y_t] \quad (5)$$

so that there is on average a one-for-one relation between money growth, adjusted for output growth, and inflation. Hence, as argued by McCallum (1993), lack of cointegration between the levels of money (or money per unit of output) and prices is not a problematic result for the quantity theory.

Likewise, a change in the intercept term in the money demand function would permanently shift the relationship between the levels of money and prices, but would, once the shift to the new intercept was complete, wash out entirely from the first-differenced

¹⁶ It has been argued, we think correctly, that payments technology tends to develop more rapidly during periods of relatively high inflation. But if these shifts in the pace of innovation are due to policy, then the changes are more accurately treated as "endogenous" and so are separate from those captured by the trend term.

money demand function, which is the underpinning of the money growth/inflation relationship. Furthermore, a one-time shift in the long-run interest semielasticity of money demand, such as has been argued by Ireland (2009) to have occurred in recent years in the case of M1 demand, does not affect the longer term relation between money growth and inflation, provided ΔR_t averages zero. Summing up, while the *price level homogeneity* of the money demand function is crucial for delivering quantity-theory relations, instability in several other aspects of the long-run money demand relation does not preclude a close relation between money growth and inflation.

It should furthermore be clear that, as Lucas (1980) also argued, money demand stability is consistent with a weak relationship between inflation and monetary growth. The case of M1 in the United States is perhaps the best example. As noted earlier, long-run M1 demand behavior up to the late 1980s appeared explicable via a standard demand function for money. But the M1 growth/inflation relationship seemed to break down in the early 1980s. The discrepancy between M1 growth rates and inflation is attributable to the sustained change in the opportunity cost of holding money. The ΔR_t term in Eq. (4), instead of averaging zero, was negative on average, and this declining opportunity cost of holding money promoted a recovery of real money balances. To be sure, a tendency toward nonzero ΔR_t was not exceptional by post-war standards. The ΔR_t term had been on average positive in the 1950s, 1960s, and 1970s. This led Barro (1982) to dispute the way that contributions of velocity growth to inflation were typically characterized in presentations of the quantity theory. These expositions tended to treat velocity growth arising from interest-rate increases as a “one-time” factor, affecting the price level but not the trend of prices. Barro (1982) pointed out that, with R_t in practice trending upward, the contribution that velocity growth made to U.S. inflation, when measuring money with the M1 definition, turned out to be substantial. The contribution of ΔR_t to velocity growth over these decades was, however, steady enough that it did not prevent a close correlation between inflation and prior monetary growth. After 1981, the trend of R_t turned downward. But the actual decline in R_t , and associated fall in velocity, came in spurts. For example, the decline in the federal funds rate that took place in the second half of 1982 was almost entirely reversed in the course of the Federal Reserve’s tightening over most of 1983 and 1984; but in 1985 and 1986, interest rates fell to levels not seen since the early 1970s. Thus, instead of the interest-rate decline contributing to a more or less constant difference between M1 growth and inflation, it affected M1 velocity growth markedly in specific periods, notably mid-1982 to mid-1983 and 1985–1986, essentially wiping out the correlation between inflation and money growth once these periods were incorporated into calculations.

The downward trend in nominal interest rates continued in the 1990s and 2000s, with both the real interest rate and the expected-inflation component declining. While financial developments such as sweeps have undoubtedly contributed to distortions to both M1 growth and M1 demand, one should not expect a close money growth/inflation relation even in the absence of such distortions, because of the uneven but substantial shifts in the opportunity cost of holding money.

The fact that there is no close mapping between stability of money demand and closeness of the money growth/inflation relationship is the reason we do not review studies of money demand in this chapter. We will, however, next discuss the available evidence on the income elasticity of money demand, which does have bearing on the money growth/inflation relationship, and on the nominal homogeneity of money demand.

5.2 Evidence with country-average data

One popular way of scrutinizing putative quantity-theory relations is to construct per-country average observations on money growth and inflation, for use in scatterplots or in regressions (possibly with panel data) of inflation on money growth. When high double-digit inflation countries are included, scatterplots of annual averages of money growth and inflation tend to bring out an impressive relation (see, e.g., Friedman, 1973, p. 18; Lucas, 1980, Figure 3.1; and McCandless & Weber, 1995, Chart 1). Results for countries that have experienced average inflation in single digits tend to be more mixed. For example, Issing, Gaspar, Angeloni, and Tristani (2001, p. 11) displayed, for a set of “low-inflation” countries, a scatter of mean money growth and inflation rates; they treat the QTM as implying a unitary slope for the plot, and fail to reject this slope restriction. De Grauwe and Polan (2005), on the other hand, found a poor relation between averages of money growth and inflation for low-inflation countries, although much stronger results have been reported in an exercise by Frain (2004) using the same sources for data as De Grauwe and Polan.

Favorable or unfavorable, these results using cross-country data are flawed as evidence on the quantity theory (see Nelson, 2003). A limiting case brings out the point. Consider two countries, A and B , in both of which there is no change in real income or nominal interest rates over time, and assume no money demand shocks. Then the first-differenced money demand equation implies that the money growth/inflation correlation is perfect in each country; that is, $\Delta \log M_{it} = \Delta \log P_{it} + c_{3i}$, for $i = A, B$. But the noninflationary rate of money growth will *not* be identical across countries, except in the special case of identical trends in payment technology, $c_{3A} = c_{3B}$. The flaw in tests of the quantity theory based on cross-country averages is that they impose a constant c_3 value across each country — in essence, a common trend to velocity across countries.

Studies of money growth and inflation across countries have rarely recognized this point; an exception is Parkin (1980, p. 172), who correctly noted for six major countries that “there is virtually no association between averages of inflation and money growth,” owing not to the absence of a within-country money growth/inflation link, but to “different trend changes in the demand for M1 balances arising from financial innovations.”^{16a} The point is of crucial quantitative significance when it comes to studying low-inflation countries. For example, Germany had lower inflation than the United States from 1962–1979: 3.7% CPI inflation in Germany, 4.9% in the United States. But M1 growth over 1962–1979 averaged 8.3% in Germany (with 4.6%

^{16a} Another early discussion recognizing this point appeared in Citibank (1979)

growth in M1 per unit of output) and 5.3% in the United States (1.4% growth in per-unit terms). An approach that focused on these cross-country averages would suggest that inflation was not closely related to money growth. But, in each country, inflation was highly correlated with prior M1 growth over the 1962–1979 period, with time series evidence supporting an approximately unitary relation. The cross-country approach neglects the different velocity trends across countries and fails to bring out the money growth/inflation relation that is obtainable from time series evidence.¹⁷

Admittedly, under very high inflation conditions, the trend in velocity due to exogenous improvements in payments technology is typically swamped by other factors: the inflation rates associated with rapid rates of money growth are large relative to the exogenous velocity trend.¹⁸ This accounts for the fact that money growth/inflation correlations computed from cross-country averages often look impressive despite the flaws inherent in this type of evidence.

6. MONEY GROWTH AND INFLATION IN TIME SERIES DATA

In this section we consider the time series relationship between money growth and inflation. Our contention is that while the static, contemporaneous relationship between monetary growth and inflation is weak, it is not the case that the only horizon at which the relationship becomes significant is at the very long run. Rather, inflation is strongly, though not at all perfectly, correlated with monetary growth of the immediately preceding years. This is the case whether one is considering quantitative experiments with standard models or drawing on evidence from historical time series data.

In taking this position, we are challenging a view that has been widely expressed in the literature, both by critics and advocates of the use of money in monetary policy analysis. For example, while affirming the use of money in policy analysis, Assenmacher-Wesche and Gerlach (2007) do so subject to the qualifier (p. 535) that “money growth and inflation are closely tied only in the long run.” That position could be taken as supportive of Svensson’s (1999, p. 215) criticism that “this long-run correlation is irrelevant at the horizon relevant for monetary policy.”

Svensson’s claim that a very long-run relationship lacks any policy relevance seems doubtful, since policymakers *are* concerned with very long-term inflation expectations. But the more general notion that quantity-theory considerations only “bite” at very

¹⁷ For studies that use reserves or the monetary base as the empirical measure of money (such as Haldane, 1997), a further factor that can distort comparisons across countries is a failure to adjust for changes in reserve requirements. McCallum and Hargraves (1995) provided illustrations of the historical importance of this factor.

¹⁸ In cases of hyperinflation, trends in velocity may continue to reflect developments in financial processes, but it would no longer be appropriate to treat this development as taking place smoothly and exogenously. Steep trends in velocity can emerge as holders of money balances make more intensive efforts to reduce the fraction of their assets in the form of money. These trends tend to reinforce the money growth/inflation correlation, but also to push the slope describing their relationship away from unity; the induced reaction of velocity growth leads to a more than one-for-one reaction of inflation to monetary growth.

long horizons does seem to reduce the QTM's relevance for monetary policy decisions. In questioning this notion, it is useful to consider first the practice of taking long moving averages of data in studying the quantity theory, and accordingly we do so in Section 6.1. Then we turn to the time series relationship between money growth and inflation, both in quantitative models (Section 6.2) and in historical data (Sections 6.3 to 6.6). We finally consider evidence for the United States pertaining to the QTM's nominal homogeneity proposition (Section 6.7).

6.1 Is long averaging of data required?

We noted above that an implication of the QTM is that steady-state money growth rates and steady-state inflation rates are linked one-for-one, once allowance is made for output growth. Lucas (1980, 1986) argued that, in studying time series of a particular country, this steady-state relation can be brought out by taking long moving averages of monetary growth and inflation. Lucas (1986, p. S405) went on to say, "Without such averaging, the quantity theory ... does not provide a serviceable account of comovements in money and inflation." The argument that taking long moving averages of time series is the way to recover close money growth/inflation relations is also advanced in empirical studies such as Dewald (2003).

One objection to this procedure, which is not the criticism on which we focus here, is examined in detail by Sargent and Surico (2008). The interpretation of coefficient estimates in a regression of inflation (or its moving average) on a moving average of monetary growth will depend on whether past quarters' money growth rates (which enter the calculation of the moving average) are actually standing in for expectations of future money growth. If that is so, then the coefficient estimate associated with the average-money-growth term will not tend to 1.0 even in an environment where the quantity theory is valid; it will be a function of the policy rule parameters, for the same reason as that discussed in the literature on the natural rate hypothesis.

Sargent and Surico (2008) explored the behavior of the coefficient on the money growth term in moving-average regressions from simulations of a variety of models. Some of the models and parameter values contemplated do deliver large departures from a unitary money growth/inflation relation, and hence serve as one argument against the moving-average approach.¹⁹ But the practical relevance of their results for monetary policy models used in practice is open to question. Even under the conditions contemplated by Sargent and Surico (2008), the coefficient on average money growth does tend to unity if long-term inflation is a unit root process, as it is assumed to be in Smets and Wouters (2007) and Woodford (2008), for example. Moreover, as detailed next, when we simulate a standard New Keynesian model with a standard interest-rate rule, the money growth/inflation relation is approximately unitary even when money growth and inflation are stationary.²⁰

¹⁹ The *unconditional* means of inflation and monetary growth, however, retain a unitary relationship with one another.

²⁰ Additional grounds for questioning the applicability of the Sargent-Surico argument to actual money growth/inflation combinations are offered in Benati (2009).

Table 3.1 M1 Growth/ CPI Inflation Relationship Using Different Degrees of Time Aggregation, United States, 1955–1975

Dependent variable	Explanatory variable	Sample period	Coefficient on money growth term	R ²
Annual inflation	Annual money growth	1955–1975	0.515 (0.236)	0.200
Five-year moving average of inflation	Five-year moving average of money growth	1960–1975	0.832 (0.134)	0.732
Annual inflation	Annual money growth lagged two years	1955–1975	0.809 (0.178)	0.518
Annual inflation	Annual money growth lagged two years	1960–1975	0.829 (0.214)	0.517

Note: The annual data underlying the regressions are for four-quarter growth rates of M1 and the CPI for the second quarter of the year.

Our criticism of the moving-average procedure is somewhat different. Time averaging is advertised as a means of allowing for lags — especially by McCandless and Weber (1995) — but in practice it may do so poorly. In particular, long averaging does not appear in practice to deliver any greater improvement in fit of the QTM than would be obtained by retaining the non-averaged time series data.

To see this, consider the data Lucas (1980) used in studying the United States. He used second-quarter observations for M1 growth and CPI inflation for 1955–1975. Using the modern vintage of CPI data and the Lothian-Cassese-Nowak (1983) data on old M1 (which are close to the data used by Lucas), and taking the four-quarter log differences for each second-quarter observation, we present three regressions in Table 3.1. The first regresses inflation on money growth for 1955–1975. This was the relationship that Lucas characterized as loose and that motivated his use of moving averages. The second regression replaces the annual data with (overlapping) five-year averages of the data (the average for 1956–1960 was the first observation, 1957–1961 the second, and so forth, for a total of 16 observations). The third and fourth regressions return to the annual data (with sample periods 1955–1975 and 1960–1975, respectively), but instead of specifying inflation as a function of the current year's money growth, they regress inflation on money growth two years earlier.

Taking moving averages does have the effect of moving the coefficient on money growth from significantly below unity to above 0.80 — insignificantly different from unity. But so too does the procedure of retaining the annual data while replacing current money growth with lagged money growth. It is clear that the improvement in the performance of the QTM as one moves to heavily averaged data is no better than that delivered by a time series calculation that allows for an interval between movements in money growth and in inflation.

We suggest that this result is not special to Lucas' example. On the contrary, the timing relationships between money growth, nominal income growth, and inflation mean that

similar results are likely to show up using other sample periods and other countries. Replacing a regression of inflation on money growth with moving averages of the same series changes the right-hand-side variable from current money growth to an average of current, prior, and future money growth terms. But movements in money growth tend on average to *lead* movements in inflation — a regularity noted even in classic contributions on the quantity theory by Hume (1752) and Wicksell (1915/1935), and stressed in the monetarist literature, especially by Milton Friedman from 1970 onward (e.g., Friedman, 1972b, 1987). It is a regularity that continues to be found in studies using more recent data (see Batini & Nelson, 2001; Christiano & Fitzgerald, 2003; Leeper & Roush, 2003; Dotsey and King, 2005). In the terminology of spectral analysis, there is a phase shift in the relationship between monetary growth and inflation.

Superficially, time-averaging might seem to go in the right direction in allowing for this phase shift, as the averaging introduces prior money growth into the right-hand-side monetary term. But it is an inadequate approach if inflation regularly follows money growth. A regression of time-averaged inflation on time-averaged money growth still implies a relationship between inflation and money growth that is on average contemporaneous; future money growth rates enter the right-hand-side expression with the same weight as lagged rates.

Thus, taking long moving averages of time series data seems an undesirable means of extracting the relationship between monetary growth and inflation. It is preferable to continue to use nonaveraged time series data, and to allow for lags explicitly instead of implicitly.

What about the argument that long averages help remove measurement error? We have much sympathy with the view that there are substantial problems with the measurement of money, and have noted that these are likely to distort the relationship between monetary growth and inflation. But this is not, as far as we can see, a low-frequency versus high-frequency data issue *per se*; it seems unrealistic to expect that measurement problems matter only for the cyclical relationship and wash out of the long-run relationship.

6.2 Money growth/inflation dynamics in a New Keynesian model

In our discussion of U.S. time series data on money growth and inflation, it may be instructive to consider the relationship between money growth and inflation that emerges from quantitative experiments with a structural model of a kind often used in monetary policy analysis. We deploy a New Keynesian model, appended by a money demand function. The New Keynesian model is standard, other than featuring date- $t-1$ calculations for the expectations terms that appear in the IS and Phillips curves. The use of lagged expectations in the spending and pricing relations follows Svensson and Woodford (2005), and yields a simplified version of the more elaborate representation of inertia specified in Rotemberg and Woodford (1997). Accordingly, in place of Eq. (2), the Phillips curve takes the form:

$$\pi_t = \beta E_{t-1} \pi_{t+1} + \kappa (E_{t-1} [\gamma_t - \bar{\gamma}_t]). \quad (6)$$

This Phillips curve arises from an environment where those firms changing prices in the current quarter (i.e., period t) make decisions on the basis of the prior quarter's (i.e., period $t - 1$) information set.

The IS equation is

$$\gamma_t = E_{t-1}\gamma_{t+1} - \sigma(E_{t-1}[R - E_{t-1}\pi_{t+1}]) + e_{\gamma t}. \quad (7)$$

Here $\sigma > 0$, and $e_{\gamma t}$ is an IS shock. We retain the money demand function (4), so portfolio decisions are based on realized output and interest rates.

To complete the model, we assume that monetary policy follows, up to a white noise shock, a Taylor (1993) rule with smoothing:

$$R_t = \rho_R R_{t-1} + (1 - \rho_R)(\phi_\gamma \gamma_t + \phi_\pi \pi_t) + e_{Rt}. \quad (8)$$

We set the parameters as follows: $\beta = 0.99$, $\kappa = 0.024$, $\sigma = 0.5$, $\rho_R = 0.8$, $\phi_\gamma = 0.125$, $\phi_\pi = 1.5$, $c_1 = 1$, $c_3 = 0$.²¹ The money demand interest semielasticity c_2 is kept to 4, corresponding to the value suggested for the business cycle frequency by King and Watson (1996). We assume that the nonpolicy shocks (IS, money demand, and natural output shocks) are AR(1) processes, each with autoregressive parameter 0.95 and innovation standard deviation of 0.5%. The monetary policy shock is treated as white noise, as previously noted, with standard deviation 0.2%.

We solve the model and compute impulse responses. Figure 3.5 plots the responses to a unit monetary policy shock of money growth, inflation, nominal interest rates, and nominal income growth (Δx , defined as $\pi + \Delta y$). The monetary policy shock lowers the nominal interest rate and leads to an immediate rise in money growth. Because of the delays implied by the lagged-expectation terms, real spending (not shown) and inflation react with a delay to interest-rate movements. Thus money growth leads inflation in the responses, even though the term that drives inflation (i.e., the sum of current and expected future output gaps) is wholly forward-looking.

Figure 3.6 plots the model response to a unit IS shock. Again, money growth reacts ahead of inflation.

Figure 3.7 plots responses to a (positive) potential output shock. This shock reduces inflation after a one-period delay, while the policy loosening triggered in response serves to brake the decline in inflation. The contemporaneous money growth/inflation relation is negative in this case, and the decline in inflation precedes an eventual decline in money growth. These patterns contrast with the lead of money growth over inflation observed in

²¹ The value of κ utilized here is the baseline value employed by Woodford (2003), and is in turn derived from the estimates of Rotemberg and Woodford (1997). The policy-rule parameters imply responses to inflation and detrended output equal to those in Taylor (1993), albeit spread out by interest-rate smoothing. The smoothing parameter value of 0.8 is standard. The choice of an IS slope of $\sigma = 0.5$ is modest relative to values often used in the literature, and is used here as the model lacks other features (such as habit formation) that could moderate the short-term response of aggregate demand to monetary policy actions.

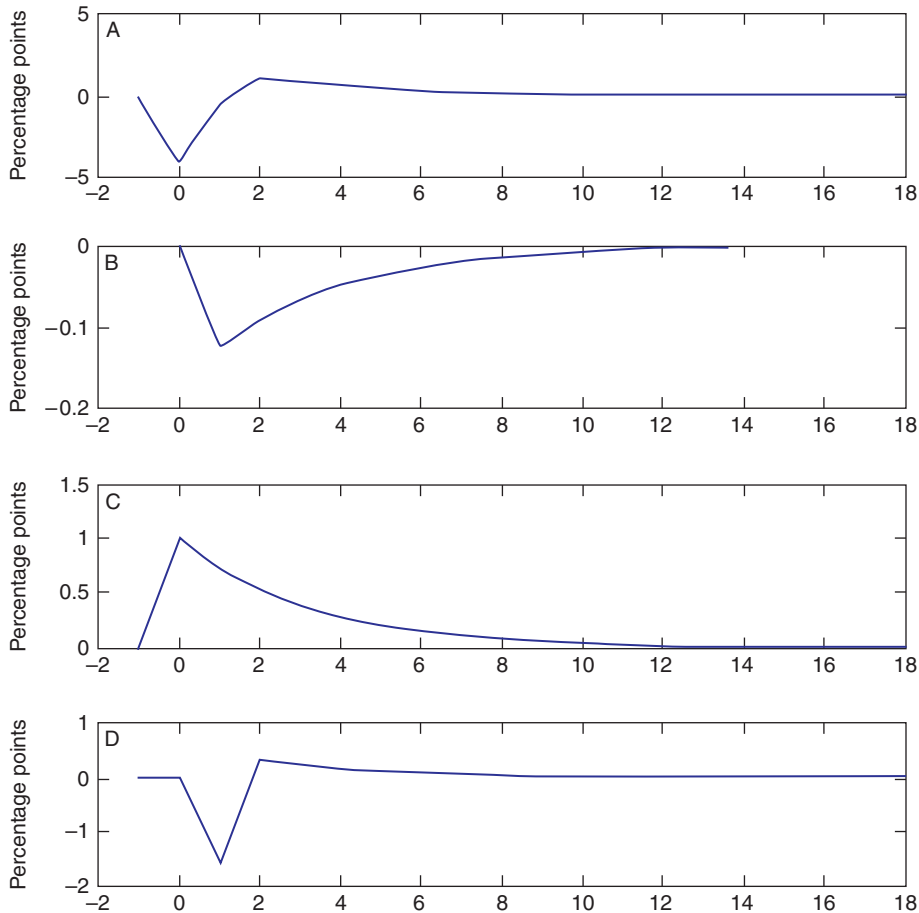


Figure 3.5 Responses to a monetary policy shock, New Keynesian model: (A) Δm response to policy shock, (B) π response to policy shock, (C) R response to policy shock, and (D) Δx response to policy shock.

the previous responses. On the other hand, nominal income growth/inflation relation also differs from those previously depicted, as nominal income growth does not begin to decline until after the decline in inflation. This may suggest that the set of reactions associated with this shock is relatively unimportant empirically, since, as we discuss later, the average tendency in the data is for nominal income growth to lead inflation.

Four aspects of the overall results are worth cataloging. First, money growth and inflation seem to be closely related — indeed, they seem to enjoy an approximately unitary relationship. This is despite the fact that the responses describe dynamics rather than steady-state relations. This standard New Keynesian model suggests that a great deal of the relationship between money growth and inflation is manifested at the business cycle frequency.

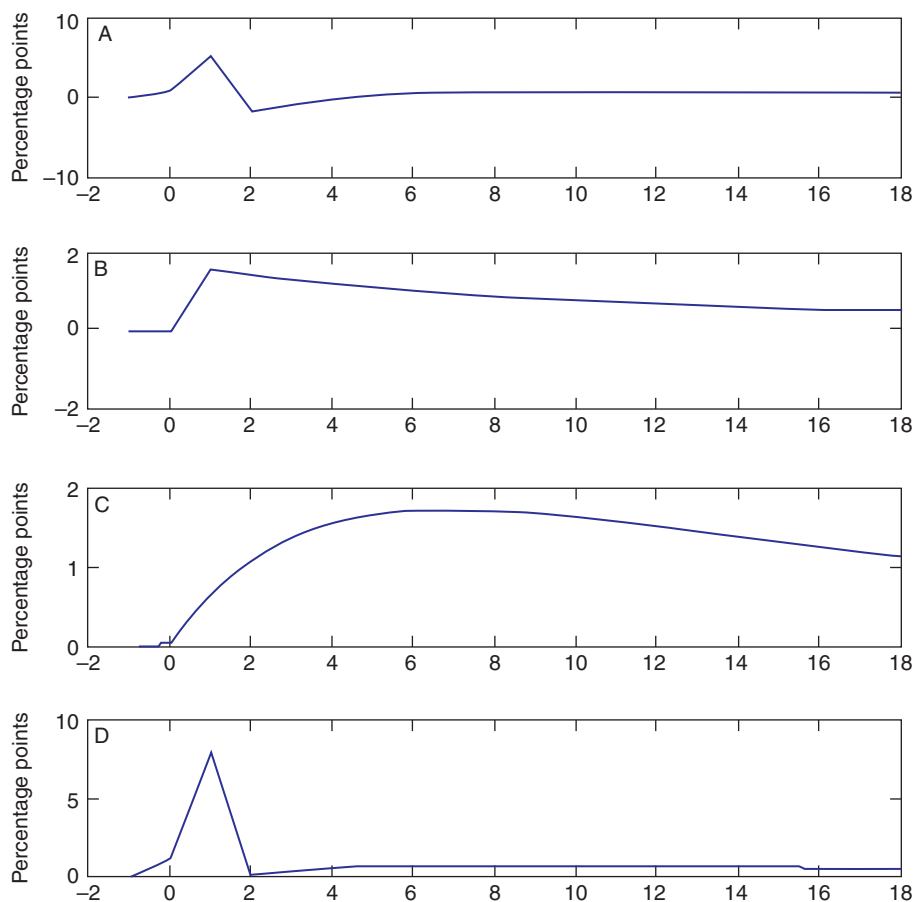


Figure 3.6 Responses to an IS shock, New Keynesian model: (A) Δm response to IS shock, (B) π response to IS shock, (C) R response to IS shock, and (D) Δx response to IS shock.

Second, money growth tends to have a contemporaneous or leading relation with inflation in this model. The Lucas (1980) approach to extracting quantity-theory relations can be thought of as implying a dependence of inflation on a two-sided distribution (i.e., both lags and leads) of money growth rates. The earlier responses suggest that in practice the future-money terms are less important for the study of the relation between inflation and money growth. This is despite the fact that, in the model, inflation is forward-looking when expressed in terms of the output gap. The decision delays built into the model confer on money a leading relationship. Also note that while, in principle, following a shock that raises the level of money, the proportionality between money and prices can be restored by a return of the money stock to its original level, that is not how the proportionality is principally restored for the shocks we consider. Rather, for IS and policy shocks, prices tend to move in the wake of the shift in money in a manner that restores the original level of real balances.

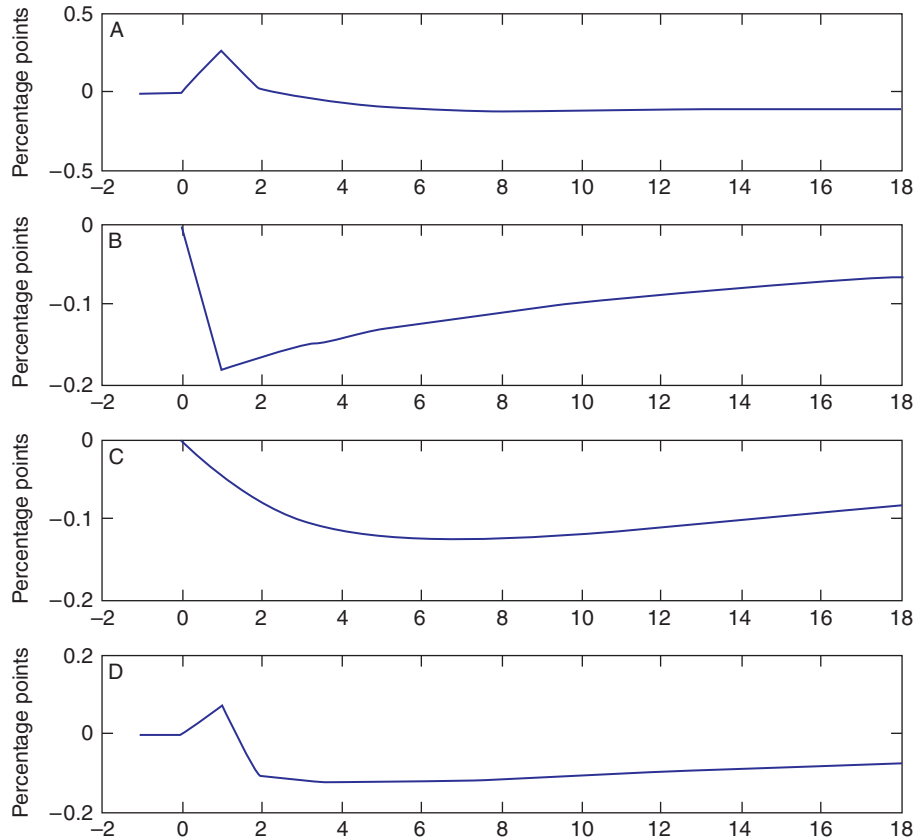


Figure 3.7 Responses to a natural output shock, New Keynesian model: (A) Δm response to natural output shock, (B) π response to natural output shock, (C) R response to natural output shock, and (D) Δx response to natural output shock.

Third, the results with this model are not consistent with the notion that a policy rule that takes stabilizing actions against inflation is likely to have the effect of wiping out the money growth/inflation relation. The reasons this argument, which has appeared widely since the 1960s, does not appear relevant are, first, that the delays built into the model prevent complete stabilization of inflation, and, second, the inflation response coefficient of 1.5 implied by the Taylor rule still leaves some muted variation in inflation, which in turn has its counterpart in muted variation in monetary growth.

Fourth, while none of the responses depict the experiment we referred to in our definition of the QTM, that is, an exogenous change in the money stock, they have several features common with the QTM experiment; the shocks contemplated in Figures 3.5 to 3.7 produce permanent changes in the levels of nominal money and prices, but only temporary movements in output and interest rates, and feature the level of money and prices being restored to their original proportional relationship with one another.

Table 3.2 Second-Moment Results, New Keynesian model
Correlation of inflation and lag k of money growth

	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$				
	0.419	0.435	0.395	0.361	0.329				
Regressions of inflation on money growth									
	Coefficient on lag of money growth								
	0	1	2	3	4	5	6	Sum	R ²
Static regression	0.235	—	—	—	—	—		0.235	0.167
Distributed-lag regression	0.166	0.179	0.169	0.149	0.121	0.089	0.062	0.935	0.579

These results reinforce the suggestion that quantity-theory relations should be recoverable from business-cycle data; that recovering the relation between inflation and money growth mainly involves looking at the relation between inflation and prior, not future, money growth; and that environments in which policymakers follow a firm interest-rate rule tend to deliver traditional quantity-theory patterns in the reduced-form behavior of money and prices.

We consider the relationship further by computing a selection of second-moment statistics. Table 3.2 displays the correlations between inflation and (current and prior) monetary growth that emerge from simulations of the preceding model; specifically, the correlations tabulated are the averages of the correlations that arose from 100 simulated data series of 200 observations in length. The results indicate that money growth and inflation are positively correlated in the model, with money growth leading inflation by a quarter.

We further report average coefficient estimates and R^2 statistics that arise from (averages of) regressions of inflation on money growth in the simulated data. A static regression delivers a coefficient on money growth of only 0.24. But when the regression specification includes lags of money growth, the coefficient sum rises to above 0.90. Thus in this model the unitary relation between the two series, in principle visible completely only in the very long run, appears to be almost entirely recoverable from a reduced-form distributed-lag regression.

We have also considered an alternative New Keynesian model that replaces the Phillips curve with a curve based on indexation to lagged inflation. Equation (6) is replaced by:

$$\pi_t - \gamma\pi_{t-1} = \beta(E_{t-1}[\pi_{t+1} - \gamma\pi_t] + \kappa(E_{t-1}[y_t - \bar{y}_t]) + e_{\pi t}. \quad (9)$$

Other than the dating of expectations to $t-1$, this specification follows Giannoni and Woodford (2002, Eq. 2.1), whose specification allowed for the dynamic indexation scheme advocated by Christiano, Eichenbaum, and Evans (2005). We assume partial

Table 3.3 Second-Moment Results, New Keynesian Model with Indexation
Correlation of inflation and lag k of money growth

	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$				
	0.398	0.420	0.379	0.343	0.307				
Regressions of inflation on money growth									
	Coefficient on lag of money growth								
	0	1	2	3	4	5	6	Sum	R ²
Static regression	0.318	—	—	—	—	—		0.318	0.254
Distributed-lag regression	0.241	0.205	0.187	0.154	0.117	0.075	0.042	1.022	0.593

Note: All numbers reported in the tables are the averages across 100 stochastic simulations of output computed from time series of 250 generated data points.

indexation (i.e., $\gamma = 0.2$). The indexation feature, when combined with a stabilizing policy rule, tends to compress the variation of inflation. To compensate for this, we raise the output gap elasticity (κ) to 0.15.

The second-moment results are given in Table 3.3. Here the correlation again is highest when money growth leads inflation, and the coefficient on money growth rises sharply when lags of money growth are included in regressions for inflation. The coefficient sum here is very near to 1.0, so it is again the case that once allowance is made for lags, reduced-form regressions tend to convey the unitary relationship between money growth and inflation implied by the QTM.

Thus fortified by these model results, let us now examine some examples of the empirical relation between money growth and inflation.

6.3 Nominal spending and inflation

Our contention that a relationship between money growth and inflation exists at the business cycle frequency does not rest on any claim that money appears in the structure of the IS or Phillips curves that describe spending and pricing decisions. Neither New Keynesian nor monetarist analyses imply the presence of money in the structural IS and Phillips curve equations, even though quantity-theory relations do prevail in models featuring these equations. The relationship in time series data between money growth and inflation rather is one that arises indirectly from the interaction of several equations. Indeed, since as Lucas (1986, p. S405) observed, “a change in money does not automatically cause prices to move equiproportionally in any direct sense,” one important function of models of monetary policy analysis is to spell out the *indirect* process that tends to produce an equiproportionate relation between prices and money. This was seen in the preceding experiments with the New Keynesian model, where no money terms appeared in the system other than in the money demand relations, yet the model dynamics generated a close-to-unitary time series relationship between inflation and monetary growth.

In particular, the relationship between money growth and inflation is dependent on a relationship between nominal spending growth and inflation. Looseness in the relationship between monetary growth and nominal GDP growth will tend to imply a loose money growth/inflation relationship too. There is also a dynamic complication, for nominal spending growth tends empirically to exhibit timing relationships with its two components (real GDP growth and inflation) that should be taken into account when attempting to determine the money growth/inflation relationship. We state these two regularities before considering their implications for the study of monetary growth. The first regularity is that *nominal and real spending move together in the short run*: In their study of U.S. monetary history, Friedman and Schwartz (1963, p. 678) observed that “real income tends to vary over the cycle in the same direction as money income does . . .” This observation holds true for U.S. data beyond the period covered by Friedman and Schwartz. McCallum (1988, p. 176) reported a correlation above 0.8 for 1954–1985 quarterly changes in U.S. nominal and real GNP.²² Likewise, Brown and Darby (1985, p. 192) concluded from a study of annual data for several major countries that, contemporaneously, “the course of money income is much more closely related to that of real income than of price,” while Woodford (2003, p. 188) noted “the persistence of the real effects of disturbances to nominal spending.” This regularity is that *inflation tends to follow nominal income growth*: The regularity, consistent with but not implied by the first, means that inflation rates tend to be more closely related to prior nominal income growth than to same-period nominal income growth. This phenomenon was noted for the United States by Nelson (1979, p. 1308) who stated, “An important conclusion is that the price level is very slow to respond to changes in nominal income.”²³ It is illustrated for several major countries in Table 3.4, which presents correlations of inflation with current and prior nominal GDP growth, for two inflation series (i.e., computed from the GDP deflator and the CPI), using annual data for selected sample periods.

Table 3.4 documents a pronounced tendency for nominal income growth to have a better correlation with the following year’s inflation than with current inflation. The lagged character of this relation is especially notable in the case of deflator inflation, a series that is biased toward having a close contemporaneous correlation with nominal GDP growth because of their connection via an identity.

The full-sample correlations for the UK in Table 3.4 would appear to contradict the claim that nominal income growth leads inflation, but in fact do not do so. For most of the first quarter of 1974, the UK government imposed restrictions on days worked as an energy-conservation measure. As a result, recorded rates of both nominal and real UK

²² Likewise, the correlation between quarterly real GDP growth and quarterly nominal GDP growth for the United States for the period 1954 Q3–2009 Q2 is 0.82. This calculation, like those in Tables 1 to 7, uses log-differences to measure percentage changes.

²³ Gordon (1988, p. 24) also takes note of this phenomenon.

Table 3.4 Correlations of inflation and nominal income growth (Inflation in year t , nominal income growth in year $t-k$)

		GDP deflator inflation		CPI inflation	
		$k = 0$	$k = 1$	$k = 0$	$k = 1$
Germany	1957–1998	0.587	0.753	0.209	0.443
Germany	1980–1998	0.544	0.767	0.182	0.461
Japan	1959–2008	0.837	0.829	0.720	0.795
Japan	1980–2008	0.843	0.851	0.716	0.770
United States	1959–2008	0.624	0.708	0.541	0.709
United States	1980–2008	0.574	0.661	0.505	0.662
United Kingdom	1957–2008	0.923	0.834	0.893	0.845
United Kingdom	1980–2008	0.862	0.902	0.767	0.893
United Kingdom	1957–1972	0.785	0.860	0.761	0.808
United Kingdom	1977–2008	0.911	0.929	0.841	0.917

GDP growth were artificially low in 1974, and nominal GDP growth did not peak until the inflation peak of 1975. Correlations for the UK omitting the mid-1970s observations reestablish a lead of nominal income growth over inflation, as Table 3.4 also shows.

6.4 Money growth per unit of output and inflation

What do these two regularities imply for the relationship between money growth and inflation? The principal implication is that, while money growth's correlation with inflation can be thought of as a by-product of the connection between monetary growth and nominal spending growth, money growth is likely to have different timing relations with the other two nominal aggregates.

With a unitary income elasticity, the demand for money function provides a connection of money to nominal income. As we have seen, the empirical relation between growth rates in nominal income and in prices seems to be close, but with nominal GDP growth tending to lead inflation. Taking these points together leads us to the implication that, when money growth is closely related to inflation, it is usually also closely related to nominal income growth. But different lags are relevant in each case; in annual data, money growth tends to be most closely related to the current year's nominal income growth, but its maximum correlation with inflation is typically with inflation one or more years later. Consequently, there are problems with the procedure of adjusting money growth for output growth to obtain a measure of inflationary pressure. Over

long periods, such an adjustment is appropriate, but over short periods, money growth adjusted for output growth may be an inferior indicator to money growth proper.

If correlations of money growth per unit of output growth and inflation are actually roundabout measures of the association between nominal income growth and monetary growth, they fail to capture the lead of money growth over inflation. That this is not simply a hypothetical issue brought out by considering data for M1 growth and inflation in the United States in the 1960s and 1970s (Figures 3.8 and 3.9). The raw M1 growth data clearly led movements in inflation; but adjusting for output growth delivers merely a contemporaneous money growth/inflation relationship.

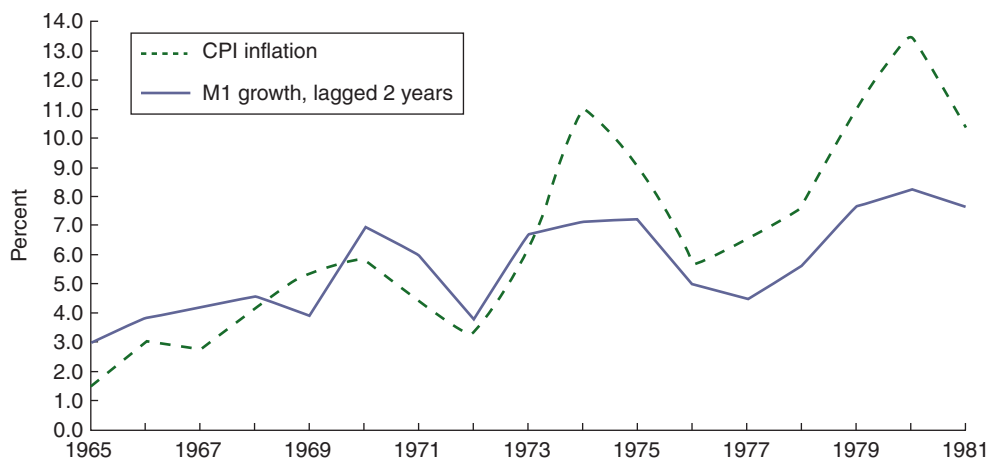


Figure 3.8 CPI inflation in the 1970s and M1 growth two years earlier, United States.

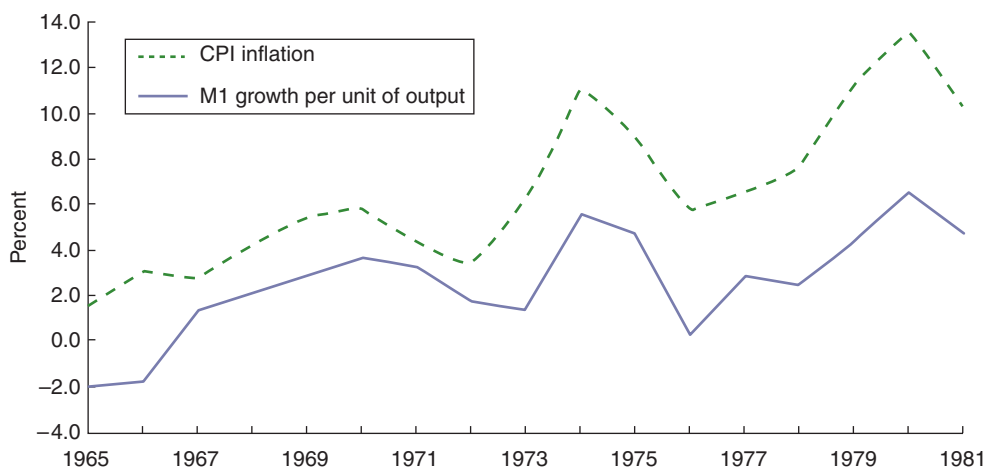


Figure 3.9 CPI inflation and M1 growth per unit of output in the 1970s, United States.

Another problem inherent in comparisons of inflation with output-adjusted money growth is that the short-run non-neutrality of money may disguise the inflationary pressure implied by a given amount of money growth. In the late 1970s, for example, loose U.S. monetary policy led to rapid growth in both money and output. The strength of output disguised the longer term weakness in output implied by the productivity slowdown, and indeed led some observers to contend that productivity from 1975 onward had returned to its pre-1973 rate of increase (see, e.g., Blinder, 1979, p. 67; McNees, 1978, p. 56). Subtracting output growth from monetary growth in these years gave false comfort; the picture thus conveyed suggested that policy settings were not as inflationary as they actually were.

6.5 Time series evidence

With this background, let us turn to reduced-form evidence on the relationship between money growth and inflation. We focus on annual data as they provide a convenient means of allowing for the possibility of lags between money growth and inflation of more than a year. We consider first the case of Japan, whose monetary experience illustrates several of the points noted above. Our data for Japan's M1 growth and CPI inflation are constructed from annual averages of data from *International Financial Statistics*. Regressions of inflation on money growth using this data set are reported in Table 3.5.

We consider first the sample period 1959–1989. A static regression of inflation on money growth delivers an insignificant and low coefficient estimate. But this reflects not the absence of a relation in the time series data, but the failure to allow for lags; adding lags one to three of M1 growth to the specification has the effect of raising the R^2 from 0.08 to 0.58. The sum of estimated coefficients on monetary growth is, however, only 0.44. The post-1973 slowdown in Japan's real growth rate, which necessarily lowered the noninflationary rate of monetary growth, appears to be having a major impact on the results. Adding an intercept dummy, D74, equal to 1.0 after 1973 greatly improves the fit and interpretability of the regression, with the coefficient sum on money growth now 0.825 and insignificantly different from unity, and the coefficient on the dummy suggesting a rise in the inflation rate for given money growth (and a corresponding slowdown in potential growth, assuming a unit income elasticity of money demand) of 5.8%.

We also present results with money growth per unit of output as the explanatory variable. For the coefficient sum on money growth, the results that allow for lags closely agree with the results using M1 growth. The intercept dummy does not appear in the regressions because the per-unit term already adjusts for the slowdown in potential.

Results deteriorate when the sample period is 1959–2008. The post-1973 intercept dummy no longer seems to capture the growth slowdown well, and, while inclusion of lags of money growth raises the coefficient sum on money growth, the sum is still only 0.4. The results with per-unit money growth are poorer. The decade of the 1990s is not a decade for

Table 3.5 Regressions for CPI Inflation in Japan
Sample Period: 1959–1989

Monetary variable	Coefficients						R ²	SEE	DW
	Lag								
	0	1	2	3	Sum	D74			
M1 growth	0.178 (0.110)	—	—	—	0.178 (0.110)	—	0.082	0.040	0.84
M1 growth	−0.279 (0.119)	0.214 (0.131)	0.311 (0.131)	0.196 (0.117)	0.441 (0.093)	—	0.582	0.028	1.35
M1 growth	0.067 (0.133)	0.278 (0.108)	0.313 (0.107)	0.166 (0.095)	0.825 (0.126)	0.058 (0.015)	0.735	0.023	1.46
M1 growth per unit of output	0.413 (0.125)	—	—	—	0.413 (0.125)	—	0.274	0.035	1.06
M1 growth per unit of output	0.171 (0.120)	0.106 (0.125)	0.217 (0.123)	0.317 (0.118)	0.811 (0.122)	—	0.639	0.026	1.48

Sample Period: 1959–2008

Monetary variable	Coefficients						R ²	SEE	DW
	Lag								
	0	1	2	3	Sum	D74			
M1 growth	0.177 (0.081)	—	—	—	0.177 (0.081)	—	0.091	0.038	0.58
M1 growth	−0.058 (0.100)	0.165 (0.122)	0.120 (0.122)	0.188 (0.101)	0.415 (0.089)	—	0.365	0.033	0.66
M1 growth	−0.061 (0.116)	0.164 (0.124)	0.119 (0.124)	0.187 (0.103)	0.409 (0.128)	−0.001 (0.015)	0.365	0.034	0.67
M1 growth per unit of output	0.131 (0.097)	—	—	—	0.131 (0.097)	—	0.037	0.040	0.46
M1 growth per unit of output	0.074 (0.120)	0.043 (0.135)	0.064 (0.134)	0.128 (0.119)	0.309 (0.135)	—	0.108	0.039	0.45

Note: A constant term was also included in all equations.

which the non-neutral effects of monetary policy average out; adjusting money growth for output growth worsens money growth as an indicator of inflation under these circumstances.

Do the full-sample results refute the quantity theory, or indicate a lack of practical usefulness in understanding inflation behavior? We would argue not. The collapse of nominal interest rates during the 1990s in Japan led to a series of permanent increases in real money demand that distorted the money growth/inflation correlation, as they did in the United States in the 1980s. From the viewpoint of the quantity theory, the trend in the opportunity cost of holding money in Japan during the 1990s is a solid basis for expecting surges in money growth that never have a counterpart in inflation — particularly for a very interest-elastic aggregate like M1. That trend has left an indelible impression on the Japanese data, one that is unlikely to go away even with the taking of long averages. Nevertheless, an interest-rate trend is not something that can be confidently extrapolated. Once the economy has completely adjusted to a permanent decline in interest rates, the quantity theory suggests that the underlying unitary relation between money growth and inflation should become more evident.

Let us now consider the reduced-form relation between money growth and inflation in the United States. Table 3.6 presents regressions of inflation on money growth. Consider first the results with M1 as the measure of money. For the 1963–1979 sample, the coefficient sum on lags 0–3 of M1 growth is significant and very large. Indeed, it is well above unity. Allowing for the post-1973 growth slowdown via an intercept dummy brings the money-growth coefficient sum close to unity. But extending the sample period to 1989 destroys this result, making the sum negative. The 1963–1989 regression result supports earlier evidence suggesting the breakdown of bivariate M1/inflation relations in the United States once observations from the 1980s are included in estimation (see, e.g., Friedman & Kuttner, 1992). As noted previously, this deterioration reflected the protracted recovery of real M1 balances in response to permanent declines in U.S. nominal interest rates in the 1980s.

Adding the years 1990–2008 to the sample seems to restore some significance to M1 growth, but the coefficient sum is far below unity, and the explanatory power of the regression is low.

Moving to M1 growth per unit of output produces a near-unit sum on money growth for 1963–1979. But it makes the money growth/inflation relation contemporaneous for the reasons discussed previously. There is a deterioration in the relation in the 1980s (not as great as the deterioration using M1 growth, because rapid output growth in 1983 and 1984 makes inflation in those years easier to reconcile with M1 behavior)^{23a} and a further fall in the coefficient sum as 1990–2008 data are included.

The use of money per unit of output in the preceding regressions implicitly entailed an assumption of a unitary income elasticity of money demand; otherwise, it would not

^{23a} Siegel (1986, p. 12) presents a related finding.

Table 3.6 Regressions for U.S. CPI Inflation using M1

Monetary variable	Sample period	Coefficients								R ²	SEE	DW
		Lag					Sum	D74				
		0	1	2	3							
M1 growth	1963–1979	−0.085 (0.309)	0.624 (0.336)	0.995 (0.348)	0.214 (0.270)	1.748 (0.262)	—	0.831	0.014	1.07		
M1 growth	1963–1979	−0.103 (0.180)	0.507 (0.198)	0.877 (0.204)	−0.037 (0.166)	1.244 (0.184)	0.026 (0.005)	0.947	0.008	2.20		
M1 growth	1963–1989	−0.271 (0.287)	0.044 (0.354)	0.003 (0.397)	−0.136 (0.355)	−0.359 (0.435)	0.038 (0.017)	0.278	0.029	0.42		
M1 growth	1963–2008	0.094 (0.169)	0.021 (0.246)	0.127 (0.249)	0.143 (0.169)	0.384 (0.150)	—	0.156	0.025	0.46		
M1 growth	1963–2008	0.097 (0.170)	0.015 (0.247)	0.136 (0.251)	0.120 (0.173)	0.368 (0.151)	0.007 (0.009)	0.169	0.026	0.44		
<i>M1 growth relative to output:</i>												
Δ log (M1/Y)	1963–1979	0.925 (0.264)	0.030 (0.268)	0.065 (0.255)	0.125 (0.230)	1.145 (0.210)	—	0.735	0.018	1.00		
Δ log (M1/Y)	1963–1989	0.398 (0.245)	−0.064 (0.282)	0.225 (0.306)	−0.095 (0.259)	0.464 (0.236)	—	0.184	0.030	0.59		
Δ log (M1/Y)	1963–2008	0.273 (0.139)	−0.087 (0.194)	0.109 (0.195)	0.080 (0.137)	0.376 (0.127)	—	0.188	0.025	0.52		
Δ log (M1/Y ^{0.5})	1963–1979	0.925 (0.541)	0.294 (0.548)	0.474 (0.497)	−0.161 (0.445)	1.532 (0.281)	—	0.715	0.018	1.04		
Δ log (M1/Y ^{0.5})	1963–1989	0.242 (0.353)	−0.025 (0.438)	0.312 (0.516)	−0.062 (0.417)	0.468 (0.290)	—	0.108	0.031	0.50		
Δ log (M1/Y ^{0.5})	1963–2008	0.234 (0.175)	−0.095 (0.268)	0.155 (0.271)	0.105 (0.174)	0.399 (0.141)	—	0.167	0.025	0.51		

Note: A constant term was included in all equations.

be appropriate to impose a unit weight on output growth in constructing a “money growth relative to output” series. For Japan, a unitary long-run elasticity of real M1 demand appears to have empirical support (Rasche, 1990), and many econometric studies for U.S. real M1 demand also support a unitary income elasticity (see Hoffman & Rasche, 1991; Lucas, 1988). There is some evidence, however, that the long-run income elasticity of M1 demand is better characterized empirically as 0.5 rather than 1.0 (see, e.g., Ball, 2001). That being so, the “money growth relative to output” concept relevant for discussions of inflation should be measured as $\Delta \log(M1) - 0.5 \Delta \log Y$ rather than money growth per unit of output, $\Delta \log(M1) - \Delta \log Y$.

Results imposing the alternative income elasticity of 0.5 appear as the final three regressions of Table 3.6. The results agree closely with those that used a unit weight on output growth, with similar equation standard errors and comparable performances across different sample periods. In addition, as before, the expression of money growth in relative-to-output terms makes the coefficient on current money the dominant term in the sum of coefficients.

In Table 3.7 we present regressions of CPI inflation on M2 growth. The results help explain why many researchers (such as Benati, 2009) prefer to use that aggregate in empirical studies rather than M1.²⁴ In the regressions with a post-1973 intercept dummy, the coefficient sum on M2 growth changes little as the sample is extended from 1979 to 2008, and has 1.0 within its confidence interval throughout. Not all is well with the M2/inflation relation; for example, the regression standard error rises as the sample is extended, and residual serial correlation is substantial. But the greater resilience of the M2 results in response to the addition of more recent years’ data supports two points stressed earlier: that a filter is not required to establish a relation between money growth and inflation,²⁵ and that, while measurement problems with money are undoubtedly significant in practice, many of the discrepancies that arose between M1 growth and inflation, especially those prior to 1994, are attributable to the substantial interest sensitivity of M1 balances, rather than to measurement problems with M1.

6.6 Panel data evidence for the G7

We now consider panel data evidence, using annual observations on CPI inflation and monetary growth for the G7. To avoid some of the problems associated with the effect of disinflation on M1 behavior, the monetary series we use is an M2-type aggregate. The sample period is 1958–2008 for four of the seven economies; for those members

²⁴ The M2 series corresponds to the annual average of the M2 series plotted in Figure 3.1, but with an adjustment corresponding to the introduction of money market deposit accounts in 1983, using, as in Batini and Nelson (2001), an adjustment that follows Friedman (1988), which in turn agrees with the estimate of the effect in Small and Porter (1989).

²⁵ Assenmacher-Wesche and Gerlach’s (2007) treatment of the U.S. data with a low-frequency filter does not deliver a point estimate on M2 growth closer to unity than we obtain in Table 3.5 using unfiltered annual data.

Table 3.7 Regressions for U.S. CPI Inflation using M2

Monetary variable	Sample period	Coefficients								
		Lag					D74	R ²	SEE	DW
		0	1	2	3	Sum				
M2 growth	1963–1979	0.119 (0.235)	0.193 (0.259)	0.550 (0.269)	0.680 (0.229)	1.543 (0.363)	—	0.714	0.018	0.77
M2 growth	1963–1979	−0.050 (0.263)	0.027 (0.281)	0.395 (0.287)	0.333 (0.346)	0.705 (0.731)	0.026 (0.020)	0.753	0.018	0.60
M2 growth	1963–2008	0.006 (0.185)	−0.053 (0.241)	0.206 (0.239)	0.522 (0.181)	0.682 (0.152)	—	0.417	0.021	0.48
M2 growth	1963–2008	0.060 (0.185)	−0.050 (0.237)	0.240 (0.237)	0.474 (0.181)	0.723 (0.152)	0.012 (0.008)	0.449	0.021	0.51
<i>M2 growth relative to output:</i>										
$\Delta \log (M2/Y)$	1963–1979	0.340 (0.229)	−0.208 (0.240)	0.324 (0.249)	0.651 (0.235)	1.107 (0.263)	—	0.657	0.020	0.98
$\Delta \log (M2/Y)$	1963–2008	0.364 (0.133)	−0.074 (0.153)	0.120 (0.151)	0.309 (0.133)	0.718 (0.148)	—	0.387	0.022	0.74

Note: A constant term was included in all equations.

of the G7 that are now part of the euro area (France, Italy, and Germany), we consider data only for the pre-euro period 1958–1998.

We present several estimated specifications in Table 3.8. In all cases these are panel regressions of inflation on money growth that impose common slopes across countries. We also consider, however, cases where the intercept is allowed to vary across countries. The first regression in Table 3.8 is a static regression with a single intercept imposed. This delivers a coefficient of money growth of 0.387, which is highly significant but well below unity and the regression itself has only mild explanatory power. Introducing lags of monetary growth raises the coefficient sum to about 0.50. We argued above that it is *not* an implication of the quantity theory that the intercept term in panel regressions is constant across countries. In the remaining regressions, we relax this restriction by moving to a fixed-effects specification. This change in specification does significantly reduce the equations' residual standard error, but seems initially to leave the slopes at their previous fairly low estimated values. But once we allow, as we did in our previous time series regressions, for breaks in the intercept term after 1973 to take the secular decline in real GDP growth into account, the fixed-effects panel regressions exhibit much higher slope estimates than their single-intercept counterparts. For example, the regression including lags 0, 1, and 2 of money growth and allowing for cross-country variation in intercepts with a break in intercepts after 1973, produces a coefficient sum of 0.692 and an R^2 of 0.614, compared to values of 0.494 and 0.354, respectively, in the single-intercept case.

Our stochastic simulations with the New Keynesian models indicated that, when the QTM holds, distributed-lag regressions of inflation on money growth tend to generate a coefficient on money growth close to unity, but perhaps somewhat lower than unity: in our simulation of the New Keynesian model, a coefficient sum of about 0.90. The empirical panel regression, on the other hand, delivers a coefficient sum of about 0.70. This is also close to the coefficient sum we obtained in similar specifications for estimated on U.S. time series data using M2. This perhaps suggests that empirical difficulties with finding a satisfactory measure of money, while not eliminating the relationship between money growth and inflation in the data, are responsible for reducing the coefficient on money growth in this type of time series regression by about 0.20 or 0.25.

6.7 Money demand nominal homogeneity

Our definition of the quantity theory does not associate the quantity theory closely with propositions about the money demand function. We have, however, insisted that zero degree homogeneity with respect to nominal variables is a property of money demand implied by the quantity theory — the demand is for *real* balances, in terms of real determinants.²⁶ We now consider U.S. M1 and M2 demand further in this light.

²⁶ The nominal interest rate in this context measures the real opportunity cost of holding real money balances, as it reflects the difference between the real rates of return on money and interest-bearing assets.

Table 3.8 G7 Panel Regressions for CPI Inflation

	Single intercept		Multiple intercepts with breaks in 1974		Multiple intercepts	
	Lags of money growth used		Lags of money growth used		Lags of money growth used	
	0	0–2 ^a	0	0–2 ^a	0	0–2 ^a
Coefficient on money growth	0.614 (0.040)	0.494 (0.039)	0.377 (0.039)	0.509 (0.038)	0.514 (0.043)	0.692 (0.040)
R ²	0.227	0.354	0.314	0.454	0.463	0.614
SEE	0.035	0.032	0.033	0.029	0.029	0.025

Note: Observations in all regressions is 327. Data consist of annual observations for 1958–2008 (1958–1998 for France, Italy, and Germany).

^aEstimate reported is coefficient sum.

The nominal homogeneity restriction implies $g_1 = 0$ in the relation:

$$\Delta \log V = g_0 + g_1 \pi_t + g_2 \Delta OPP_t + u_t,$$

where V is velocity, defined as nominal GDP divided by nominal money, and OPP_t is the opportunity cost of the relevant aggregate. We measure OPP_t for M1 by the federal funds rate (annual average) and OPP_t for M2 by the spread between the federal funds rate and the M2 own-rate.²⁷ A money demand relation can be cast as a velocity relation (with no separate real income term) if the money demand function has a unitary income elasticity, a property often found for M2 demand and, as noted earlier, also a common finding for M1. Note that this recasting of the relationship as a velocity relation means a change in sign when interpreting the coefficient on the interest rate: a negative money demand interest semielasticity implies a positive velocity interest semielasticity.

Given the definition of velocity, the natural price series to use in testing the nominal homogeneity restriction is the GDP deflator. For completeness, however, we also present results using CPI inflation. We express the relation in first differences rather than levels to allow for the likely presence of permanent money demand shocks, which produce nonstationarity in velocity and imply that levels of real money and real income are not cointegrated (see McCallum, 1993).

Because GDP deflator inflation and velocity growth have a definitional relation with one another, measurement errors in inflation may produce a correlation between inflation and velocity growth. These errors would tend to bias tests in the direction of rejecting nominal homogeneity. To protect against this bias, we estimate by instrumental variables, with two lags of each series (velocity growth, inflation, and first difference of opportunity cost) serving as instruments.

Estimates, using annual data, are presented in Table 3.9 for M1 and in Table 3.10 for M2. We consider the full sample (starting in 1962 for the M1 velocity estimation, a year later for M2), and results for samples starting in 1980. Because of the increased importance of sweeps for M1 behavior after 1993, we also present results for the 1962–1993 period in the case of M1 velocity.

Observing the point estimates and standard errors for g_1 , we see that nominal homogeneity of money demand is not rejected irrespective of the inflation series used, the definition of money chosen, or sample period considered. As the final rows in Tables 9 and 10 show, this continues to be the case if we relax the assumption of a unitary income elasticity of money demand.²⁸ Thus, nominal homogeneity of money demand, a fundamental aspect of the quantity theory, appears to be consistent with the U.S. data.

²⁷ The M2 own-rate is a standard variable in M2 demand studies published since the 1980s (e.g., Small and Porter, 1989). We use annual averages of the series available from the Federal Reserve Bank of St. Louis' FRED site.

²⁸ A positive coefficient on real income growth in these estimates implies an income elasticity of money demand below unity.

Table 3.9 Tests of Nominal Homogeneity of M1 Demand
Dependent variable: Log-difference in M1 velocity

Sample period	Coefficients on:					
	GDP deflator	CPI	ΔOPP	$\Delta \log Y_t$	SEE	DW
1962–2008	−0.010 (0.240)	—	0.505 (0.535)	—	0.034	0.76
1962–1993	0.301 (0.260)	—	0.520 (0.543)	—	0.032	0.73
1980–2008	−0.344 (0.401)	—	1.427 (0.745)	—	0.037	0.77
1962–2008	—	0.016 (0.209)	0.595 (0.519)	—	0.034	0.76
1962–1993	—	0.199 (0.221)	0.574 (0.536)	—	0.032	0.73
1980–2008	—	−0.224 (0.352)	1.339 (0.672)	—	0.037	0.75
1962–2008	−0.152 (0.329)	—	0.498 (0.556)	−0.328 (0.475)	0.037	0.82
1962–1993	0.497 (0.376)	—	0.306 (0.672)	0.318 (0.462)	0.031	0.70
1962–2008	—	−0.162 (0.328)	0.612 (0.576)	−0.386 (0.534)	0.036	0.82
1962–1993	—	0.399 (0.381)	0.246 (0.567)	0.361 (0.554)	0.031	0.68

Note: Instrumental variables estimates are reported in the tables. Instruments are a constant and two lags of each variable, including dependent variable. “GDP deflator” and “CPI” refer to log differences of these variables.

7. IMPLICATIONS OF A DIMINISHING ROLE FOR MONEY

Benjamin Friedman (1999, 2000) suggested that technological improvement in the financial sector raises the prospect of the virtual obsolescence of central bank money. In terms of the subject matter of this chapter, the scenario that Friedman envisages is consistent with continuing quantity-theory relations between inflation and money growth, provided that the latter refers to growth in deposit-inclusive measures of money. But both deposit creation and market interest rates would become disconnected, in this scenario, from central bank actions, with associated loss of central bank control over nominal spending.²⁹ Friedman’s argument did not involve the complete disappearance of money, but instead a state of affairs in which the role of base money diminishes to the point where central banks’ ability to influence aggregate demand in a dependable fashion would be in jeopardy. Reactions to these conjectures include those of Goodhart (2000) and Woodford (2000, 2001). In the following paragraphs we attempt to outline the main contours, and evaluate the merits, of the debate.

Base money includes, of course, both currency and bank reserves. Goodhart (2000) argued convincingly that private sector demand for currency will persevere for the

²⁹ King (1999) advanced similar arguments.

Table 3.10 Tests of Nominal Homogeneity of M2 Demand
Dependent variable: Log-difference in M2 velocity

Sample period	Coefficients on:					
	GDP deflator	CPI	ΔOPP	$\Delta \log Y_t$	SEE	DW
1963–2008	0.130 (0.147)	—	0.634 (0.372)	—	0.021	1.23
1980–2008	−0.006 (0.237)	—	1.076 (0.860)	—	0.025	0.96
1963–2008	—	0.172 (0.125)	0.732 (0.360)	—	0.020	1.24
1980–2008	—	0.036 (0.179)	1.394 (0.653)	—	0.020	0.87
1963–2008	0.048 (0.193)	—	0.740 (0.421)	−0.290 (0.290)	0.023	1.39
1963–2008	—	0.049 (0.185)	0.802 (0.437)	−0.299 (0.320)	0.022	1.40

Note: Instrumental variables estimates are reported in the tables. Instruments are a constant and two lags of each variable, including dependent variable. “GDP deflator” and “CPI” refer to log differences of these variables.

foreseeable future, in part because of the anonymity conferred on currency transactions. In principle, the interest elasticity of currency demand gives central banks scope to manipulate interest rates without departing from their traditional policy of providing the amount of currency that the public demands at prevailing income and interest rates. But this would constitute a departure from the standard central bank practice of focusing on interbank transactions as the means through which to affect interest rates.

One part of Friedman’s (2000) argument is that technological progress makes it possible for buyers to make payments through accounts, bank or nonbank, that are not subject to reserve requirements. The existence of such arrangements is widely accepted by all participants in the debate. Woodford (2000) argued convincingly that the magnitude of required reserves is irrelevant. After all, several central banks do not rely on reserve requirements in their arrangements for setting interest rates. Overnight interest rates in these economies are typically controlled by means of “channel” arrangements, involving standing facilities that set both a floor and a ceiling on overnight rates. These rates apply to the operational reserve balances, useful for settlement purposes, which financial intermediaries hold with the central bank. An arrangement consistent with the channel system involves central bank payment of interest on reserves, including excess reserves; this possibility is discussed by Woodford (2000, 2001) and Goodfriend (2002). The arrangement offers the promise of securing a positive demand for central bank money in a technologically advanced financial system. If settlement reserves with the central bank are held by banks, along with overnight securities, then the interest rate on the latter will equal the sum of the interest rate paid on reserve balances plus the marginal service yield provided by these balances. By adjusting the interest paid

on reserves, the central bank can exert a dominant influence on the overnight interest rate. The Federal Reserve introduced interest payments on reserves in October 2008.

One element of Woodford's optimistic discussion of the prospects for monetary policy in an economy with a negligible medium of exchange should be read with special care. That part includes his statement that "the unit of account in a purely fiat system is *defined* in terms of the liabilities of the central bank" (Woodford, 2000, p. 257). His subsequent discussion pertains to the unit of account (UOA) as so defined. But in many analyses the UOA is defined instead as the unit in which prices are quoted in most transactions; see, for example, Niehans (1978, pp. 118–119) and Jevons (1875, p. 5).³⁰ Now certainly the liabilities of the central bank would be a favored candidate for the role of UOA under this second meaning for an economy with no medium of exchange (MOE), but there is no necessity that it be the one that prevails. Goods prices will, in a market economy, be quoted in terms of the medium that market participants find most convenient. Just as central bank currency can be supplanted by some other candidate medium of exchange if its supply is managed too badly (e.g., under hyperinflation conditions), the central bank's contender for the MOA can conceivably lose the competition to a rival medium. And it is the UOA in terms of the MOA actually prevailing in market transactions that is of macroeconomic importance; it is stickiness of those prices used in actual transactions that is relevant for the definition of real rates of interest that influence aggregate demand. Woodford (2000, pp. 257–258) understood this point—indeed, made it explicitly himself. But the tone of his discussion is, we suggest, made considerably more optimistic (in the sense at hand) in its impression by his choice of definition.

8. MONEY VERSUS INTEREST RATES IN PRICE LEVEL ANALYSIS

The diminishing role for money provides a natural point of departure for a discussion of recent approaches to the analysis of price level determination and monetary policy operating procedures. The trend of professional work in recent years can be put in context by juxtaposing two observations from earlier decades: Patinkin's (1972, p. 898) statement that "one of the primary tasks of monetary theory is indeed to explain the determination of the wage and price levels," and Gowland's (1991, p. 122) observation that "the term 'monetary policy' seems inappropriate in a model without money." The recent literature can be thought of as embracing the first observation while rejecting the second. Monetary policy analysis remains concerned with explaining price level determination, but it has become prevalent in the course of such explanations to omit reference to monetary aggregates. In particular, the "cashless" and "neo-Wicksellian"

³⁰ Terminologically, the UOA is some specified quantity (e.g., 0.484 ounces) of the *medium of account* (MOA), (e.g., gold). Wicksell (1915/1935, p. 7) used the term "measure of value" to refer to the medium of account and mentions the convenience of having the MOA coincide with the MOE.

treatment in Woodford (2003) represents a crystallization of a framework in which the central bank manipulates interest rates and in which there may be no medium of exchange, with price level variations still capable of being influenced by deviations of the real interest rate from the natural rate of interest.

8.1 Conditions for excluding money from the analysis

The result that no reference to money arises when working out inflation behavior is not special to the analysis of cashless economies. It holds whenever the money stock appears in the money demand equation but not in the IS or spending equations, monetary policy rule, or Phillips curve. In New Keynesian models that feature a transactions technology or money in the utility function, there are two principal requirements for obtaining solution expressions for inflation and the output gap does not require considering money stock behavior. These conditions are (i) the assumed monetary policy rule does not feature a response to money (real or nominal) or monetary growth and (ii) the utility or transaction cost function is separable across money and consumption.

The exclusion of money from the IS and Phillips curves, in turn, is not special to New Keynesian analysis; on the contrary, it was typical in prior monetary analysis. In that earlier analysis, it was usually also the case that monetary policy effects on spending were specified as working through interest rates, making it possible, when studying interest-rate rules, to treat the system excluding money as self-contained, the money demand equation then standing alone, with money becoming a “residual” variable.³¹

What is different in the modern literature is that the cases where money can be neglected have been formalized as the two conditions previously given, and these conditions appear to have become accepted as realistic assumptions for policy analysis and empirical work. To a far greater extent than previously, the literature has focused on interest-rate or targeting rules in which money does not appear. Moreover, a number of studies have argued that utility can be treated as approximately separable across money and other variables.³²

The limiting case of no medium of exchange would, in our terms, indeed be a non-monetary economy; there would be no monetary policy, literally defined. Nevertheless, as discussed above, there would be scope for different types of policy measures

³¹ Consider, for example, these descriptions that appeared in the older literature on IS-LM and macroeconomic models, respectively. Brown (1965, p. 308): “The reader may have noted that there has been no mention thus far of the market for money. This has been done deliberately to indicate that with a theory of asset prices, we can regard the market for money as a residual.” Next, Ando (1981, pp. 349–350): “Influence from the quantity of money supply to both [output and prices] goes through the short-term interest rate almost exclusively . . . The MPS model may be thought of as being block triangular . . .”

³² On the latter, see Woodford (2003), McCallum (2001a), and Ireland (2004).

regarding price level behavior, with the price level being regarded as some general index of prices in terms of the unit of account.

If we do not adopt a literally cashless model — so that a positive demand for money exists — but we absorb the message that the separable-utility case as realistic, is there a useful role left for money in monetary policy formation? Or is it satisfactory to have interest rates as the sole monetary policy variable in the analysis?

In answering these questions, one should note that the shift toward analyses that ignore or downplay money largely reflects a change in empirical judgments. In the era in which monetary aggregates were used as guides to policy, policymakers expressed the view that — although monetary policy actions did work on spending via interest rates, and the authorities did typically employ a short-term nominal interest rate as their policy instrument — it was a more straightforward matter to establish money/inflation relations than it was to establish connections between policy-rate actions and subsequent inflation movements. For example, the Reserve Bank of New Zealand (1985, p. 627) stated that the “empirical linkages between interest rates and inflation are less well established than the linkages between monetary growth and inflation.” Similarly, Federal Reserve Governor Henry Wallich (1985, p. 40) argued that the “impact on inflation of a given level of interest rates, nominal or real . . . is far less predictable” than the relationship between inflation and prior monetary growth.

These statements can be interpreted as reflecting doubts about the reliability of empirical estimates of the natural rate of interest. At any point, there is an actual level of the real short-term interest rate and there exists a natural value of that rate which by definition is consistent with price stability.³³ Likewise, at any time there will be an observed rate of monetary growth and there will be a noninflationary growth rate of money corresponding to the rate at which money would grow if the real short-term interest rate were at its natural level. Predominant reliance on monetary-aggregate data in policymaking in these circumstances could reflect a judgment that estimates of the noninflationary rate of monetary growth are more reliable than estimates of the natural rate of interest. Conversely, the shift in recent decades toward policy frameworks that relied less on monetary aggregate data likely reflects a judgment that estimates of the natural rate of interest are more reliable than estimates of the noninflationary rate of monetary growth.

Interest in a Wicksellian approach to price level analysis showed some signs of reviving at a policy level in the early 1990s (e.g., Kohn, 1990), but has exploded in recent years in light of Woodford’s (2003) emphasis on the role of the natural rate of interest in dynamic stochastic general equilibrium models. We have not contrasted

³³ This does not imply that policies that tend to keep the real rate close to the natural real rate of interest, and thereby avoids output gaps, are necessarily associated with price stability. But from Eq. (2) a policy that prevents output gaps does tend to prevent π_t from deviating from the steady-state or “target” inflation rate π .

Wicksellian and quantity-theory approaches up to this point because, provided that a medium of exchange is present, the two are compatible, being in essence alternative ways of viewing the same process, as is acknowledged by Woodford (2003, p. 53).³⁴ Wicksell (1915/1935), one might note, emphasized the money stock adjustments that were implied by the banking system's variations in interest rates, although he also considered a "pure credit" economy. And in dynamic general equilibrium models, the money demand function that implies a connection between steady-state money growth and inflation comes from the same private sector optimization that delivers the IS and Phillips curves that Woodford uses.

To facilitate the discussion, consider the following variant of the model of Section 6, written now without the E_{t-1} operators in the IS and Phillips curve relations, and including a Phillips curve shock term, so as to conform even more closely to the mainstream model of recent years:

$$\gamma_t = E_t \gamma_{t+1} + b_0 - b_1(R_t - E_t \pi_{t+1}) + v_t \quad b_1 > 0 \quad (10)$$

$$\pi_t = \beta E_t \pi_{t+1} + \kappa(\gamma_t - \bar{\gamma}_t) + u_t \quad 0 < \beta < 1, \kappa > 0 \quad (11)$$

$$R_t = \mu_0 + \mu_1 \pi_t + \mu_2(\gamma_t - \bar{\gamma}_t) + e_t \quad \mu_1 > 1, \mu_2 \geq 0 \quad (12)$$

Here γ_t is log output, π_t is inflation, $\bar{\gamma}_t$ is the flexible-price (natural- rate) value of γ_t , and R_t is the one-period interest rate controlled by the central bank.³⁵

The basic point relating to the interest-rate policy rule is, as is very well known, that with $\bar{\gamma}_t$ taken for simplicity as exogenous, this system is complete; that is, it suffices to determine values of the system's endogenous variables γ_t , π_t , and R_t . Consequently, if the economy includes medium-of-exchange money with a demand function of the form

$$m_t - p_t = c_0 + c_1 \gamma_t - c_2 R_t + e_t \quad (13)$$

with m_t being the log of nominal money balances, the latter serves only to describe how much (high-powered) money needs to be supplied by the central bank to implement its policy rule (Eq. 12). Thus a shift in the parameters of Eq. (13) would, if there were no change in the other structural Eqs. (10)–(12), have no effect on the behavior of the key variables γ_t , π_t , and R_t . It is true that the crucial absence from Eq. (10) of any term involving m_t depends upon the assumption of separability of the relevant underlying function describing the way in which money facilitates transactions. But, as noted previously, analyses by Woodford (2003, pp. 111–123), McCallum (2000a, 2001a), and Ireland (2004) indicated that taking account of a plausible degree of nonseparability would have a negligible effect on the behavior of the key variables. Consequently,

³⁴ Of course, it may still be the case that policy rules with interest-rate and money-stock instruments may tend to have different properties.

³⁵ Also, v_t , u_t , and e_t are exogenous shocks.

the omission of money from policy analysis involving the standard model is not a *prima facie* reason to doubt the validity of studies that incorporate such an omission.³⁶

8.2 Determinacy and learnability

Recently, however, a major challenge to the validity of the current mainstream approach has been put forth by Cochrane (2007), who strongly questioned its basic economic logic, arguing that one standard presumption — namely that “determinacy” of a rational expectations (RE) equilibrium suffices to imply that stable inflation behavior will be generated when the Taylor principle is satisfied — is incorrect. His point is that New Keynesian models such as that expressed in Eqs. (10)–(12) are typically consistent with the existence of RE paths with explosive inflation rates (in addition to one or more stable paths) that normally do *not* imply explosions in real variables relevant for transversality conditions. Consequently, the usual logic does not imply the absence of explosive inflation. This point is (we believe) correct, but it does not (we contend) justify Cochrane’s negative conclusions about New Keynesian analysis. As argued in McCallum (2009), there is a different criterion that is logically satisfactory for the purpose at hand. This is the requirement that, to be plausible, a RE solution should satisfy the property of least-squares *learnability* of the type featured in the work of Evans and Honkapohja (2001). Adoption of this criterion amounts to a requirement of feasibility, with respect to available information, of a candidate equilibrium and accordingly should be attractive to analysts concerned with actual monetary policy. In the class of New Keynesian models discussed by Cochrane, it transpires that the learnability criterion singles out the standard New Keynesian solution as the only plausible equilibrium. In this respect, it serves to justify in principle a large portion of current mainstream monetary analysis.³⁷

We now argue, nevertheless, that there is one respect in which a money stock growth rule is distinctly preferable to an interest rate rule, at least when analyzed in the context of a typical linear model that includes a standard money demand function.³⁸ In particular, it is the case that for *nonactivist* rules a money-growth rule (i.e., a constant money growth rate) with standard parameterization leads to a unique and stable RE equilibrium that is *learnable* — in the least-squares sense researched extensively by Evans and Honkapohja (2001) — whereas a constant interest rate rule does not give rise to any learnable RE equilibrium. The latter fact is fairly well known from writings by Woodford (2003, pp. 264–268), Bullard and Mitra (2002), and others.

³⁶ More generally, Woodford (2003) was careful to demonstrate, in several places, that recognition of MOE money would not overturn conclusions developed in the context of cashless models. For a recent defense of his position concerning the practical applicability of Wicksellian analysis, which considers various possible objections, see Woodford (2008).

³⁷ We say “in principle” because the theoretical coherence of a model does not guarantee its empirical validity.

³⁸ Such functions are obtainable by either transactions-cost or money-in-utility-function reasoning, and appear frequently in the work of Woodford (2003), despite his emphasis on “cashless” economies.

To demonstrate that a well-behaved RE equilibrium is, by contrast, learnable with a nonactivist money growth rule, we proceed as follows.

Consider the standard linearized New-Keynesian model of Eqs. (10)–(12) but in which there is, for simplicity, full price flexibility so that in each period output y_t equals its flexible-price, natural-rate value $\bar{y}_t$. We measure all real variables as deviations from their natural-rate values so $c_0 = b_0 = \bar{y}_t = 0$ for all t and, after substituting in the identity $\pi_t = p_t - p_{t-1}$, the model can be written as

$$0 = 0 - b_1(R_t - E_t p_{t+1} + p_t) + v_t \quad (10')$$

$$m_t = m_{t-1} + \Delta m \quad (14)$$

together with money demand Eq. (13). Equation (14) is the money supply rule. Then substitution of Eq. (13) into Eq. (10') yields

$$0 = b_1[(1/c_2)(m_t - p_t - c_1 0) + E_t p_{t+1} - p_t] + v_t. \quad (15)$$

Inserting the money supply rule (Eq. 14) and rearranging we then obtain

$$p_t = a[k + E_t p_{t+1} + (1/c_2)m_{t-1} + (1/b_1)v_t], \quad (16)$$

where k is a constant, and $a = c_2/(1 + c_2)$ satisfies the inequalities $0 < a < 1$. Here both m_{t-1} and v_t are exogenous so the system has a single nonexplosive solution that is learnable.³⁹ Thus a nonactivist money growth rule leads to a well-behaved RE equilibrium in which the inflation rate equals the money growth rate minus a term reflecting technical progress (which equals zero in the example above).⁴⁰

To this argument it might be objected that for practical purposes it is the monetary base that the central bank can actually control, whereas the medium-of-exchange aggregate is what appears in the money demand equation in our model. That is true, so our analysis should also include a random component reflecting the semitechnological relationship between the two. But ignoring this distinction in our analysis is analogous to our treatment of interest rates in which we ignore the difference between the overnight interest rate typically controlled by a central bank and the longer-maturity market rates that are likely relevant for aggregate demand.

8.3 Fiscal theory of the price level

Probably the most drastic conceptual challenge to today's mainstream analysis, and also to traditional views concerning the relationship between money growth and inflation, has come *not* from empirical findings or the foregoing arguments, but from an intricate, elusive,

³⁹ This conclusion follows readily from results presented in Evans and Honkapohja (2001, pp. 201–204) and Bullard and Mitra (2002), among others.

⁴⁰ Simulations with a few numerical parameter values suggest that these results continue to prevail when the flexible-price assumption is replaced with a standard Calvo price-adjustment relationship.

and controversial doctrine known as the “fiscal theory of the price level” (FTPL), which was developed primarily by Leeper (1991), Sims (1994), Woodford (1994, 1995), and Cochrane (1998). We have not attempted to survey this topic in general because it is covered extensively in Volume 3B of the *Handbook* by Canzoneri, Cumby, and Diba (CCD; 2010). We find admirable, in most respects, their excellent and thorough treatment. We believe, nevertheless, that ultimate disagreements concerning the FTPL result basically from differing strategies for responding to multiplicities of rational expectations solutions, as suggested in McCallum (2001b).⁴¹ The most satisfactory means of dealing with such multiplicities seems to be, once again, provided by analysis of the learnability of the various solutions, following procedures of the type developed and expounded most prominently in the treatise of Evans and Honkapohja (2001). To date the most extensive application of these techniques to the FTPL is that of Evans and Honkapohja (2007). We have discussed these results at some length in McCallum and Nelson (2005), which concluded that (i) several of the phenomena implied by the FTPL are actually consistent with traditional monetarist doctrine⁴² and that (ii) our study’s main messages for policy are that “central banks can control inflation irrespective of fiscal policy and that detailed coordination between monetary and fiscal authorities is not needed for effective macroeconomic policy” (2005, p. 581). The second of these conclusions, which we continue to support, constitutes a partial disagreement with the CCD emphasis on the necessity of monetary–fiscal coordination.

8.4 Money as an information variable

We have argued above that from a purely theoretical perspective it has become very widely understood that analysis that posits an interest rate instrument and ignores monetary aggregates is coherent under the assumption that any absence of separability in the transactions–cost function that expresses the MOE (i.e., transactions facilitating) role of money⁴³ is mild enough to be of negligible effect. That does not settle the issue of whether an interest rate or monetary aggregate rule would perform more satisfactorily in practice, or whether it is desirable for a central bank with an interest rate instrument to ignore entirely monetary aggregates. In this regard it deserves mention that Woodford (2008) has recently developed the “OK to ignore” position in considerable detail, arguing that several claims by others for the usefulness of monetary aggregates are actually based on the behavior of credit (rather than monetary) aggregates.

A different approach to this issue is developed in McCallum (2000b), which employs counterfactual historical comparisons of the type utilized by Stuart (1996) and Taylor (1999). This type of analysis proceeds by contrasting actual settings of potential

⁴¹ This statement does not constitute a claim that the approach taken in that paper is satisfactory; see our discussion in McCallum and Nelson (2005) and references therein.

⁴² In which case the FTPL does not provide a fundamentally different, and hence challenging, approach to price level determination.

⁴³ Or nonseparability of the money-in-utility-function, if that modeling approach is taken.

instrument variables during important historical time spans with the values that would have been specified by particular rules in response to prevailing conditions. Discrepancies between rule-specified and actual values are then evaluated, in light of *ex post* judgments concerning macroeconomic performance during the span studied, to yield tentative conclusions concerning the merits and demerits of the various rules. Of particular interest is whether major policy mistakes, judged *ex post*, might have been prevented by adherence to some of the candidate rules and not others.⁴⁴ The study in question considered both interest-rate and monetary base instrument rules, with alternative target variables also being examined in each case. The countries considered were Japan, the UK, and the United States, over the years 1962–1998 (1972–1998 for Japan). Periods of major policy mistakes are taken to be 1965–1979 for the United States, 1970–1979 and the mid-to-late 1980s for the UK, and 1989 onward for Japan. By and large the rules with a monetary base instrument seemed to perform somewhat better than those with an interest instrument. The most clear-cut conclusion of the analysis, however, is that the rules' messages are more dependent upon which instrument, rather than which target, is used.⁴⁵ This is to us a surprising result. McCallum (2000b, p. 77) suggested that it can perhaps be understood “as resulting from the necessity of specifying a reference value, relative to which instrument settings are implicitly compared, in representing policy tightness or ease. For rules to be sufficiently simple, these reference-value specifications must themselves be simple, but different implicit assumptions about macroeconomic behavior are thereby built into the rule.” We believe that it is still the case that more work of this type needs to be conducted.

It is worth dwelling further on reasons monetary aggregates might give different, and in some circumstances more accurate, signals from those coming from interest-rate rules. One possibility is that money growth could contain valuable information on a key unobserved variable, the natural rate of interest. If one considers the variables that appear in the standard money demand function (Eq. 12), money does not appear to be promising as a variable whose fluctuations will shed light on variation in the natural rate. One of the arguments in the money demand function is the short-term interest rate. If this corresponds to the policymaker's policy instrument, it is directly observable and policymakers have no need to consider money in keeping track of that variable. Another variable in the money demand function is the money demand shock, which is usually interpreted as uninteresting noise which in and of itself is not a source of actual or prospective fluctuations in output and inflation. A third variable is the scale variable, current real income. Friedman (1975, p. 444) argued that, with money data

⁴⁴ An important extension of this type of research, involving “real-time data,” has been developed valuably by Orphanides (2003a,b).

⁴⁵ Target variables considered include (i) a “hybrid” linear combination of inflation and the output gap, as in the Taylor rule, (ii) variants of (i) with different detrending procedures, (iii) nominal GDP growth rate, (iv) a smoothed version of the latter, and (v) variants of strict inflation targeting rules.

arriving more promptly than GDP data, and perhaps less subject to revision, fluctuations in real money balances could convey information about current real GDP fluctuations (see also Friedman, 1990). Notwithstanding a recent flurry of interest in the potential informational role of money arising from this route (see, e.g., Coenen, Levin & Wieland, 2005; Dotsey & Hornstein, 2003), it seems unpromising. Indeed, the studies of money as an indicator of current GDP have been overtaken by events. Unofficial but widely watched indices of “monthly real GDP” now exist in the United States and other countries, and the advent of these series has made much headway into the problem of delays in official GDP releases. Their prevalence and success cast doubt on the need to look at money for the purpose of tracking current GDP.^{45a}

A more promising possibility is that money reveals fluctuations in variables that matter for future aggregate demand developments, and may do so in a way that goes beyond the information recorded in current output and nominal interest rate variations. There are episodes in the historical U.S. experience in which money growth seemed to exhibit this property. For example, during the credit controls episode of 1980, both monetary growth and short-term interest rates fell abruptly. Looking solely at interest rates, Bordo, Erceg, Levin, and Michaels (2007) interpreted this period as one of extreme monetary policy ease; likewise, the estimated monetary policy shock coming from Smets and Wouters’ (2007) dynamic general equilibrium model (estimated without money stock data) finds 1980 Q2 to have featured the most expansionary monetary policy shock in post-war U.S. history. By contrast, estimating a monetary policy shock series from a VAR that does include money (M1), Blanchard and Watson (1986) found that 1980 Q2 featured one of the most contractionary monetary policy shocks in U.S. post-war history. The mid-1980 economic downturn suggests that the interpretations of monetary policy tightness that make use of monetary aggregates are the correct ones, and that evaluations based on standard interest rates are unreliable.

It is tempting to conclude that the reason that money growth accurately reflected the severity of aggregate demand conditions during the 1980 episode is because of the accounting relations between deposit and (bank) credit creation. If this were the case, then the value of monetary aggregates as an indicator over this period might simply be a by-product of their connection via identities to more fundamental credit aggregates. But the details of monetary behavior over this period provide evidence against this explanation. The credit control episode was associated with greater weakness in M1 growth than in M2 or M3 growth, yet it is the broader aggregates that have closer accounting connections with bank credit series.

A different explanation for the information contained in money growth during the 1980 episode does not rely on accounting connections between money and credit. Instead, it relies on the nature of the monetary policy transmission process. This process involves the

^{45a} Goodhart (1983, p. 50) was an early skeptic of money’s promise as an indicator of current output.

adjustment of a wide range of asset prices to monetary injections. In a standard monetary policy model such as the New Keynesian models used earlier, the effect of monetary injections on these asset prices can be summarized by the reaction of the policy rate. If, however, alternative non-money financial assets differ in their short-run substitutability for money balances, then money demand could depend on a vector of opportunity cost variables rather than a single short-term interest rate. This could create circumstances under which, when important interest rates other than the riskless short-term rate fluctuate, these fluctuations will be recorded in real money balances. Different interest rates tend to move together over longer periods, so the money demand Eq. (13) will remain a valid description of longer term portfolio behavior. But the short-run divergences between different interest rates could give rise to occasions where the real quantity of money demanded fluctuates for given values of current income and the riskless short rate. These fluctuations may in turn signal future movements in real and nominal aggregate demand. This insight, emphasized in much of the monetarist literature on the transmission mechanism, may be relevant to understanding the value of M1 as an indicator in the credit controls episode. Studies of M1 demand in 1980 generally find that the credit controls episode is associated with large estimated residuals for conventional money demand equations (see Gordon, 1984; Hafer & Thornton, 1986; Hein, 1982). This is *prima facie* evidence that an important source of variation in real money balances over this period was not found in contemporaneous real income or short-term interest rates.

Perfect substitution between non-money assets, implying single-interest-rate specifications such as a money demand equation like (13), remains a convenient assumption for monetary policy analysis. But there are likely to be occasions where keeping track of distinct interest rates, and assessing the associated monetary policy options, is essential. McCallum (2000a), for example, argues that a realistic policy option for a central bank in an open economy when the policy rate has reached zero is to manipulate the nominal exchange rate via large-scale, unsterilized foreign exchange intervention. Such an option arises from a theoretical framework in which Treasury bills and base money are perfect substitutes at the zero lower bound, but money and foreign exchange are not.

This informational role of money arising from an environment of imperfect asset substitution can be expressed in terms of the natural rate. Let aggregate demand depend on a vector of real yields beside the real policy rate. Fluctuations in these real yields will affect the level of the real policy rate consistent with maintaining aggregate demand conditions conducive to price stability. They can thus be considered factors that affect the natural rate of interest.⁴⁶ Let the opportunity cost of money demand consist of a vector of nominal yields, the nominal counterpart of the real interest rates that matter for aggregate demand.⁴⁷ With nominal yields and real yields moving together in the

⁴⁶ To be specific, the nonpolicy rates would have a negative relationship with the natural rate.

⁴⁷ For simplicity we ignore the own-rate on money.

short run, variations in real yields beside the real policy rate will be recorded in fluctuations in the real quantity of money demanded. And because these real-rate fluctuations are a source of movement in the natural rate of interest, real money variation provides information on variation in the natural interest rate. In the instance of the United States in 1980 previously mentioned, imposition of the credit controls can be thought of as increasing the degree of monetary restriction for a given setting of the real policy rate (i.e., raising nonpolicy rates relative to the policy rate) and reducing the natural rate of interest. The fall-off in monetary growth during 1980 accurately reflected this fall in the natural rate.

Other periods also provide further apparent instances where monetary and real developments not recorded in the policy rate conferred information power onto money. For example, in the early 1990s in the United States, the real and nominal federal funds rate fell substantially, but inflationary pressure and aggregate demand conditions were weak. Consistent with this development, empirical estimates of the natural rate of interest show a protracted decline in the first half of the 1990s to low levels (see Laubach & Williams, 2003, Figures 3.1 and 3.2). Some commentators (e.g., King, 1993) have noted that the weak money growth rates observed over this period in the United States and other countries gave a more accurate picture of economic prospects than did the low levels of policy rates, and have conjectured that the low money growth rate reflected variation in unobserved nonpolicy rates. This would be consistent with the informational role for money sketched above. Once again, it is tempting to suggest that the value of money over this period was due to money growth's correlation with credit growth. But again there exists evidence against this interpretation. In the UK and the United States, monetary base growth tends to have a fairly weak year-to-year connection with the measures of private credit creation; but indicators of policy stance derived from monetary base growth give a signal of a sharp tightening in the early 1990s (McCallum, 2000b). Thus the signal about policy stance in the early 1990s was different from the signal coming from short-term interest rates, and at the same time did not appear to be a by-product of arithmetical connections between money and credit.

9. CONCLUSIONS

This chapter has considered what, if any, relationship there is between monetary aggregates and inflation, and whether there is any substantial reason for modifying the current mainstream mode of policy analysis, which frequently does not consider monetary aggregates at all. The quantity theory, as we have defined it, centers on the prediction that there will be a long-run reaction of prices to an exogenous increase in the nominal money stock. The fact that policymakers in practice do not set money growth rates

exogenously does not rob the quantity theory of empirical content.⁴⁸ Likewise, the observation that policymakers frequently are concerned with price behavior at horizons shorter than the very long run does not deprive the quantity theory of policy significance. On the contrary, the nominal homogeneity conditions that deliver the quantity-theory result are the same as those that deliver monetary neutrality, an important principle behind policy formulation. Furthermore, the quantity theory implies a *ceteris paribus* unitary relationship between inflation and money growth. After allowing for lags, this unitary relationship tends to emerge from examination of time series; it does not appear to be the case that replacing the time series with long averages of the data is a necessary or particularly valuable step in recovering that relationship.

Our discussion has not disputed the position that financial innovation can obscure the relationship between monetary growth and inflation. What is needed, however, is a sense of proportion. We believe that too much of the reaction to problems in measuring money has taken the form of abandoning the analysis of monetary aggregates, and too little has taken the form of more careful efforts at improved measurement. The problems of measurement associated with monetary aggregates have parallels in the measurement and estimation problems that occur with policy analysis that excludes money. Frameworks that include interest rates as the sole monetary variable in the analysis must, for example, grapple with the fact that the natural rate of interest is unobserved. Any shift in the natural real rate of interest will modify the consequences for inflation of a specified interest-rate policy. Such a shift in the natural interest rate would call not for leaving interest rates out of the analysis, but for more intense efforts at estimating the natural rate. Moreover, since the connections of interest rates and monetary growth to inflation are clouded by the presence of an imperfectly observed series (especially the natural rate, in the case of interest rates; financial innovations, in the case of money), studies of inflation and monetary policy behavior can benefit from including both interest rates and money in the empirical analysis.

APPENDIX: DATA SOURCES

The panel regressions in Section 6.6 use annual-average data for G7 economies on growth rates in the CPI and in an M2-type series. The source for CPI data was the Federal Reserve Bank of St. Louis FRED site for the United States, Bank of England and Office for National Statistics for the UK (for which we used the RPIX series where available and RPI otherwise), and *International Financial Statistics (IFS)* for the remaining countries. *IFS* data were also the source for the nominal GDP data used in Table 3.4. Sources for money data were as follows:

⁴⁸ Note that if policymakers set their monetary instrument — be it an aggregate or an interest rate — actively in response to the state of the economy, then they would not be setting it exogenously.

Canada: Annual averages of M2 series, constructed from Lothian, Cassese, and Nowak (1983) for 1955–1968; *IFS* 1969–2008.

France: Annual average of M2 constructed from the second M2 series in Lothian, Cassese, and Nowak (1983) for 1955–1968; *IFS* for 1969–1998.

Germany: Annual average of M2 data from Lothian, Cassese, and Nowak (1983) for 1955–1967; International Money Fund (1983) for 1968–1980; *IFS* for 1981–1998.

Italy: Annual average of Lothian, Cassese, and Nowak (1983) M2 data for 1955–1967; IMF (1983) data on M2 for 1968–1975; *IFS* data on “M2, national definition” for 1976–1998.

Japan: Annual average of M2 data constructed from *IFS* 1955–2007; G10 database for 2008.

UK: Annual average of M1 data 1955–1982, spliced into annual average of Bank of England series “Retail M4” (also known as M2) for 1983 onward (Source: Bank of England Web site). Source for M1 data is Capie and Webber (1985) for 1955–1963, Hendry and Ericsson (1991) for 1964–1982.

United States: Annual average of M2 series (from FRED), adjusted for 1983 Q1 break. Pre-1959 M2 data are annual averages of the Federal Reserve series tabulated in Lothian, Cassese, and Nowak (1983).

European countries: We plotted our money growth data, constructed as described above, against those constructed by Benati (2009; and supplied by Luca Benati) and found few differences. We also plotted the German money growth data against Bundesbank data on M2 supplied by Christina Gerberding and we verified that our series was similar.

REFERENCES

- Abbot, W.J., 1962. Revision of money supply series. Federal Reserve Bulletin 48, 941–951.
- Anderson, R.G., 2003. Retail deposit sweep programs: Issues for measurement, modeling and analysis. Federal Reserve Bank of St. Louis, Working Paper 2003–026A.
- Anderson, R.G., Kavajecz, K.A., 1994. A historical perspective on the Federal Reserve’s monetary aggregates: Definition, construction and targeting. Federal Reserve Bank of St. Louis Review 76, 1–31.
- Ando, A., 1981. On a theoretical and empirical basis of macroeconomic models. In: Kmenta, J., Ramsey, J.B. (Eds.), Large-scale macroeconomic models. North-Holland, Amsterdam, pp. 329–369.
- Assenmacher-Wesche, K., Gerlach, S., 2007. Money at low frequencies. Journal of the European Economic Association 5, 534–542.
- Ball, L.M., 2001. Another look at long-run money demand. J. Monetary Econ. 47, 3–44.
- Barnett, W.A., Chauvet, M., 2008. International financial aggregation and index number theory: A chronological half-century empirical overview. University of Kansas, Manuscript.
- Barro, R.J., 1982. United States inflation and the choice of monetary standard. In: Hall, R.E. (Ed.), Inflation: causes and effects. University of Chicago Press, Chicago, pp. 99–110.
- Batini, N., Nelson, E., 2001. The lag from monetary policy actions to inflation: Friedman revisited. International Finance 4, 381–400.
- Baumol, W.J., Blinder, A.S., 1982. Economics: principles and policy, second ed. Harcourt Brace Jovanovich, New York.

- Benati, L., 2009. Long-run evidence on money growth and inflation. European Central Bank, Working Paper No. 1027.
- Blanchard, O.J., Watson, M.W., 1986. Are business cycles all alike? In: Gordon, R.J. (Ed.), *The American business cycle: continuity and change*. University of Chicago Press, Chicago, pp. 123–182.
- Blinder, A.S., 1979. Economic policy and the great stagflation. Academic Press, New York.
- Bordo, M., Erceg, C., Levin, A., Michaels, R., 2007. Three great American disinflations. Federal Reserve Board, International Finance Discussion Paper No. 2007–898.
- Brown, A.J., Darby, J., 1985. World inflation since 1950: An international comparative study. Cambridge University Press, Cambridge, UK.
- Brown, C.V., 1965. A theory of interest rates or asset prices? *Scottish Journal of Political Economy* 12, 297–308.
- Bullard, J.B., Mitra, K., 2002. Learning about monetary policy rules. *J. Monetary Econ.* 49, 1105–1129.
- Calvo, G.A., 1983. Staggered prices in a utility-maximizing framework. *J. Monetary Econ.* 12, 383–398.
- Canzoneri, M.B., Cumby, R.E., Diba, B., 2010. The interactions between monetary and fiscal policy. In: Friedman, B.M., Woodford, M. (Eds.), *Handbook of monetary economics IIIB*, Elsevier/North-Holland, Amsterdam, Chapter 17 of this volume.
- Capie, F., Webber, A., 1985. A monetary history of the United Kingdom, 1870–1982, volume 1: Data, sources, methods. Allen and Unwin, London.
- Carlson, J.B., Hoffman, D.L., Keen, B.D., Rasche, R.H., 2000. Results of a study of the stability of cointegrating relations comprised of broad monetary aggregates. *J. Monetary Econ.* 46, 345–383.
- Christiano, L.J., Eichenbaum, M., Evans, C., 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. *J. Polit. Econ.* 113, 1–45.
- Christiano, L.J., Fitzgerald, T.J., 2003. Inflation and monetary policy in the twentieth century. *Federal Reserve Bank of Chicago Economic Perspectives* 27, 22–45.
- Citibank, 1979. The dollar: Why the market can't be conned, *Citibank Monthly Economic Letter* (February) 75, 12–15.
- Cochrane, J.H., 1998. A frictionless view of U.S. inflation. *NBER Macroeconomics Annual* 13, 323–384.
- Cochrane, J.H., 2007. Inflation determination with Taylor rules: A critical review. NBER, Working Paper No. 13409.
- Coenen, G., Levin, A.T., Wieland, V., 2005. Data uncertainty and the role of money as an information variable for monetary policy. *European Economic Review* 49, 975–1006.
- Cynamon, B.Z., Dutkowsky, D.H., Jones, B.E., 2006. Redefining the monetary aggregates: A clean sweep. *Eastern Economic Journal* 32, 661–673.
- De Grauwe, P., Polan, M., 2005. Is inflation always and everywhere a monetary phenomenon? *Scandinavian Journal of Economics* 107, 239–259.
- Dewald, W.G., 2003. Bond market inflation expectations and longer-term trends in broad monetary growth and inflation in industrial countries, 1880–2001. European Central Bank, Working Paper No. 253.
- Dorich, J., 2009. Resurrecting the role of real money balance effects. Bank of Canada, Manuscript.
- Dotsey, M., Hornstein, A., 2003. Should a monetary policymaker look at money? *J. Monetary Econ.* 50, 547–579.
- Dotsey, M., King, R.G., 2005. Implications of state-dependent pricing for dynamic macroeconomic models. *J. Monetary Econ.* 52, 213–242.
- Duca, J.V., 1995. Should bond funds be added to M2? *Journal of Banking and Finance* 19, 131–152.
- Dutkowsky, D.H., Cynamon, B.Z., Jones, B.E., 2006. U.S. narrow money for the twenty-first century. *Econ. Inq.* 44, 142–152.
- Evans, G.W., Honkapohja, S., 2001. Learning and expectations in macroeconomics. Princeton University Press, Princeton, N.J.
- Evans, G.W., Honkapohja, S., 2007. Policy interaction, learning, and the fiscal theory of prices. *Macroeconomic Dynamics* 11, 665–690.
- Fisher, I., 1913. *The purchasing power of money*, second ed. Macmillan, New York.
- Frain, J.C., 2004. Inflation and money growth: Evidence from a multi-country dataset. *The Economic and Social Review* 35, 251–266.

- Friedman, B.M., 1975. Targets, instruments, and indicators of monetary policy. *J. Monetary Econ.* 1, 443–473.
- Friedman, B.M., 1990. Targets and instruments of monetary policy. In: Friedman, B.M., Hahn, F.H. (Eds.), *Handbook of monetary economics II*, Elsevier/North Holland, Amsterdam, pp. 1186–1230.
- Friedman, B.M., 1999. The future of monetary policy: The central bank as an army with only a signal corps? *International Finance* 2, 321–338.
- Friedman, B.M., 2000. Decoupling at the margin: The threat to monetary policy from the electronic revolution in banking. *International Finance* 3, 261–272.
- Friedman, B.M., Kuttner, K.I., 1992. Money, income, prices and interest rates. *Am. Econ. Rev.* 82, 472–492.
- Friedman, M., 1956. The quantity theory of money: A restatement. In: Friedman, M. (Ed.), *Studies in the quantity theory of money*. University of Chicago Press, Chicago, pp. 3–21.
- Friedman, M., 1966. Comments. In: Shultz, G.P., Aliber, R.Z. (Eds.), *Guidelines: Informal controls and the market place*. University of Chicago Press, Chicago, pp. 55–61.
- Friedman, M., 1968. The role of monetary policy. *Am. Econ. Rev.* 58, 1–17.
- Friedman, M., 1972a. Comments on the critics. *J. Polit. Econ.* 80, 906–950.
- Friedman, M., 1972b. Have monetary policies failed? *Am. Econ. Rev.* 62, 11–18, (Papers and Proceedings).
- Friedman, M., 1973. *Money and economic development*. Praeger, New York.
- Friedman, M., 1987. Quantity theory of money. In: Eatwell, J., Milgate, M., Newman, P. (Eds.), *The new Palgrave: a dictionary of economics* 4, Macmillan, London, pp. 3–20.
- Friedman, M., 1988. Money and the stock market. *J. Polit. Econ.* 96, 221–245.
- Friedman, M., Meiselman, D., 1963. The relative stability of monetary velocity and the investment multiplier in the United States, 1897–1958. Commission on Money and Credit. In: *Stabilization policies*. Englewood Cliffs, NJ, Prentice Hall, pp. 165–268.
- Friedman, M., Schwartz, A.J., 1963. *A monetary history of the United States, 1867–1960*. Princeton University Press, Princeton.
- Friedman, M., Schwartz, A.J., 1970. *Monetary statistics of the United States*. Columbia University Press, New York.
- Giannoni, M.P., Woodford, M., 2002. Optimal interest-rate rules II: applications. NBER, Working Paper No. 9420.
- Goodhart, C.A.E., 1983. Comment. In: Meek, P. (Ed.), *Central bank views on monetary targeting* (New York: Federal Reserve Bank of New York), pp. 46–50.
- Goodfriend, M., 2002. Interest on reserves and monetary policy. *Federal Reserve Bank of New York Economic Policy Review* 8, 77–84.
- Goodhart, C.A.E., 2000. Can central banking survive the IT revolution? *International Finance* 3, 189–209.
- Gordon, R.J., 1978. *Macroeconomics*. Little, Brown and Company, Boston.
- Gordon, R.J., 1984. The short-run demand for money: A reconsideration. *Journal of Money, Credit and Banking* 16, 403–434.
- Gordon, R.J., 1988. Postwar developments in business-cycle theory: An unabashedly New-Keynesian perspective. In: Oppenländer, K.H., Poser, G. (Eds.), *Contributions of business cycle surveys to empirical economics*. Gower Publishing Co., Aldershot, U.K, pp. 21–50.
- Gowland, D., 1991. *Money, inflation and unemployment*, second ed. Prentice Hall, New York.
- Gramley, L.E., 1982. Financial innovation and monetary policy. *Federal Reserve Bulletin* 68, 393–400.
- Hafer, R.W., 1980. The new monetary aggregates. *Federal Reserve Bank of St. Louis Review* 62, 25–32.
- Hafer, R.W., Thornton, D.L., 1986. Price expectations and the demand for money: A comment. *Rev. Econ. Stat.* 68, 539–542.
- Haldane, A.G., 1997. Designing inflation targets. In: Lowe, P. (Ed.), *Monetary policy and inflation targeting*. Reserve Bank of Australia, Sydney, pp. 74–112.
- Hein, S.E., 1982. Short-run money growth volatility: Evidence of misbehaving money demand? *Federal Reserve Bank of St. Louis Review* 64, 27–36.

- Hendry, D.F., Ericsson, N.R., 1991. Modeling the demand for narrow money in the United Kingdom and the United States. *European Economic Review* 35, 833–886.
- Hoffman, D.L., Rasche, R.H., 1991. Long-run income and interest elasticities of money demand in the United States. *Rev. Econ. Stat.* 73, 665–674.
- Hume, D., 1752. Of money. In: Hume, D. (Ed.), *Political discourses*. Fleming, Edinburgh.
- International Monetary Fund, 1983. IFS supplement on money. International Monetary Fund, Washington, DC.
- Ireland, P.N., 2004. Money's role in the monetary business cycle. *Journal of Money, Credit and Banking* 36, 969–983.
- Ireland, P.N., 2009. On the welfare cost of inflation and the recent behavior of money demand. *Am. Econ. Rev.* 99, 1040–1052.
- Issing, O., Gaspar, V., Angeloni, I., Tristani, O., 2001. *Monetary policy in the euro area*. Cambridge University Press, Cambridge, UK.
- Jevons, W.S., 1875. *Money and the mechanism of exchange*. Henry S. King & Co, London.
- Jones, B.E., Dutkowsky, D.H., Elger, Thomas, 2005. Sweep programs and optimal monetary aggregation, <http://ideas.repec.org/s/eee/jbfin.html> *Journal of Banking and Finance* 29, 483–508.
- Juillard, M., Kamenik, O., Kumhof, M., Laxton, D., 2008. Optimal price setting and inflation inertia in a rational expectations model. *Journal of Economic Dynamics and Control* 32, 2584–2621.
- Keynes, J.M., 1936. *The general theory of employment, interest and money*. Macmillan, London.
- King, M.A., 1993. The Bundesbank: A view from the Bank of England. *Bank of England Quarterly Bulletin* 34, 269–273.
- King, M.A., 1999. Challenges for monetary policy: New and old. In: *New challenges for monetary policy*. Federal Reserve Bank of Kansas City, Kansas City, MO, pp. 11–57.
- King, R.G., Watson, M.W., 1996. Money, prices, interest rates and the business cycle. *Rev. Econ. Stat.* 78, 35–53.
- Kohn, D., 1990. Making monetary policy: Adjusting policy to achieve final objectives. In: Norton, W.E., Stebbing, P. (Eds.), *Monetary policy and market operations*. Reserve Bank of Australia, Sydney, pp. 11–26.
- Laubach, T., Williams, J.C., 2003. Measuring the natural rate of interest. *Rev. Econ. Stat.* 85, 1063–1070.
- Leeper, E.M., 1991. Equilibria under “active” and “passive” monetary and fiscal policies. *J. Monetary Econ.* 27, 129–147.
- Leeper, E.M., Roush, J.E., 2003. Putting “M” back in monetary policy. *Journal of Money, Credit and Banking* 35, 1217–1256.
- Lothian, J.R., Cassese, A., Nowak, L., 1983. Data appendix. In: Darby, M.R., Lothian, J.R. (Eds.), *The international transmission of inflation*. University of Chicago Press, Chicago, pp. 525–718.
- Lown, C.S., Peristiani, S., Robinson, K.J., 1999. What was behind the M2 breakdown?. Federal Reserve Bank of New York, Staff Report No. 83.
- Lucas Jr., R.E., 1972. Econometric testing of the natural rate hypothesis. In: Eckstein, O. (Ed.), *The econometrics of price determination*. Board of Governors of the Federal Reserve System, Washington, DC, pp. 50–59.
- Lucas Jr., R.E., 1980. Two illustrations of the quantity theory of money. *Am. Econ. Rev.* 70, 1005–1014.
- Lucas Jr., R.E., 1986. Adaptive behavior and economic theory. *Journal of Business* 59, S401–S426.
- Lucas Jr., R.E., 1988. Money demand in the United States: A quantitative review. *Carnegie-Rochester Conference Series on Public Policy* 29, 137–167.
- Lucas Jr., R.E., 2000. Inflation and welfare. *Econometrica* 68, 247–274.
- McCallum, B.T., 1988. Robustness policies of a rule for monetary policy. *Carnegie-Rochester Conference Series on Public Policy* 29, 173–204.
- McCallum, B.T., 1990. Inflation: theory and evidence. In: Hahn, F.H., Friedman, B.M. (Eds.), *Handbook of monetary economics* 2, Elsevier/North-Holland, Amsterdam, pp. 963–1012.
- McCallum, B.T., 1993. Unit roots and economic time series: some critical issues. *Federal Reserve Bank of Richmond Economic Quarterly* 79, 13–33.
- McCallum, B.T., 2000a. Theoretical analysis regarding a zero lower bound on nominal interest rates. *Journal of Money, Credit, and Banking* 32, 870–904.

- McCallum, B.T., 2000b. Alternative monetary policy rules: a comparison with historical settings for the United States, the United Kingdom, and Japan. *Federal Reserve Bank of Richmond Economic Quarterly* 86, 49–79.
- McCallum, B.T., 2001a. Monetary policy analysis in models without money. *Federal Reserve Bank of St. Louis Review* 83, 145–160.
- McCallum, B.T., 2001b. Indeterminacy, bubbles, and the fiscal theory of price level determination. *J. Monetary Econ.* 47, 19–30.
- McCallum, B.T., 2009. Inflation determination with Taylor rules: Is new-Keynesian analysis critically flawed?, *J. Monetary Econ.* 56, 1101–1108.
- McCallum, B.T., Goodfriend, M., 1987. Demand for money: Theoretical studies. In: Eatwell, J., Milgate, M., Newman, P. (Eds.), *The new Palgrave: A dictionary of economics* 1, Macmillan, London, pp. 775–781.
- McCallum, B.T., Hargraves, M., 1995. A monetary impulse measure for medium-term policy analysis. *Staff Studies for the World Economic Outlook* 52–70, September.
- McCallum, B.T., Nelson, E., 2005. Monetary and fiscal theories of the price level: the irreconcilable differences. *Oxford Review of Economic Policy* 21, 565–583.
- McCandless, G.T., Weber, W.E., 1995. Some monetary facts. *Federal Reserve Bank of Minneapolis Quarterly Review* 19, 2–11.
- McNees, S.K., 1978. The current business cycle in historical perspective. *New England Economic Review* 60, 44–59.
- Meltzer, A.H., 1969. Tactics and targets: Discussion. In: *Controlling monetary aggregates*. Federal Reserve Bank of Boston, Boston, pp. 96–103.
- Mishkin, F.S., 2007. *The economics of money, banking, and financial markets*, eighth ed. Addison-Wesley, Boston.
- Nelson, C.R., 1979. Recursive structure in U.S. income, prices and output.. *J. Polit. Econ.* 87, 1307–1327.
- Nelson, E., 2003. The future of monetary aggregates in monetary policy analysis. *J. Monetary Econ.* 50, 1029–1059.
- Niehans, J., 1978. *The theory of money*. Johns Hopkins University Press, Baltimore, MD.
- Orphanides, A., 2003a. Historical monetary policy analysis and the Taylor rule. *J. Monetary Econ.* 50, 983–1022.
- Orphanides, A., 2003b. The quest for prosperity without inflation. *J. Monetary Econ.* 50, 633–663.
- Parkin, M., 1980. Oil push inflation? *Banca Nazionale del Lavoro Quarterly Review* 33, 163–185.
- Patinkin, D., 1956. *Money, interest and prices: An integration of money and value theory*. Row, Peterson, Evanston, IL.
- Patinkin, D., 1972. Friedman on the quantity theory and Keynesian economics. *J. Polit. Econ.* 80, 883–905.
- Rasche, R.H., 1990. Equilibrium income and interest elasticities of the demand for M1 in Japan. *Bank of Japan Monetary and Economic Studies* 8, 31–58.
- Reserve Bank of New Zealand, 1985. Monetary policy: Some questions and answers. *Reserve Bank of New Zealand Bulletin* 48 (November), 626–629.
- Reynard, S., 2004. Financial market participation and the apparent instability of money demand. *J. Monetary Econ.* 51, 1297–1317.
- Rotemberg, J.J., Woodford, M., 1997. An optimization-based econometric framework for the evaluation of monetary policy. *NBER Macroeconomics Annual* 12, 297–346.
- Samuelson, P.A., 1967. Money, interest rates and economic activity: Their interrelationship in a market economy. American Bankers Association. In: *Proceedings of a symposium on money, interest rates and economic activity*. American Bankers Association, New York, pp. 40–60.
- Sargent, T.J., 1971. A note on the “accelerationist” controversy. *Journal of Money, Credit and Banking* 3, 50–60.
- Sargent, T.J., Surico, P., 2008. Monetary policies and low-frequency manifestations of the quantity theory. Bank of England, External MPC Unit Discussion Paper No. 26.

- Siegel, D.F., 1986. The relationship of money and income: the breakdowns in the 70s and 80s, Federal Reserve Bank of Chicago Economic Perspectives 10, 3–15.
- Sims, C.A., 1994. A simple model for study of the determination of the price level and the interaction of monetary and fiscal policy. *Econ. Theory* 4, 381–399.
- Small, D.H., Porter, R.D., 1989. Understanding the behavior of M2 and V2. *Federal Reserve Bulletin* 75, 244–254.
- Smets, F., Wouters, R., 2007. Shocks and frictions in US business cycles: A Bayesian DSGE approach. *Am. Econ. Rev.* 97, 586–606.
- Solow, R.M., 1969. Price expectations and the behavior of the price level. Manchester University Press, Manchester, U.K.
- Stock, J.H., Watson, M.W., 1993. A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica* 61, 783–820.
- Stuart, A., 1996. Simple monetary policy rules. *Bank of England Quarterly Bulletin* 36, 281–287.
- Svensson, L.E.O., 1999. Monetary policy issues for the eurosystem. *Carnegie-Rochester Conference Series on Public Policy* 51, 79–136.
- Svensson, L.E.O., 2003. What is wrong with Taylor rules? Using judgment in monetary policy through targeting rules. *Journal of Economic Literature* 41, 426–477.
- Svensson, L.E.O., Woodford, M., 2005. Implementing optimal policy through inflation-forecast targeting. In: Bernanke, B.S., Woodford, M. (Eds.), *The inflation-targeting debate*. University of Chicago Press, Chicago, pp. 19–92.
- Taylor, J.B., 1993. Discretion versus policy rules in practice. *Carnegie-Rochester Conference. Series on Public Policy* 39, 195–214.
- Taylor, J.B., 1999. An historical analysis of monetary policy rules. In: Taylor, J.B. (Ed.), *Monetary policy rules*. University of Chicago Press, Chicago, pp. 319–341.
- Wallich, H.C., 1985. U.S. monetary policy in an independent world. In: Ethier, W., Marston, R.C. (Eds.), *International financial markets and capital movements: A symposium in honor of Arthur I. Bloomfield*. International Finance Section, Princeton, N.J, pp. 33–44.
- Wicksell, K., 1915. *Lectures on political economy* 2, Macmillan, London (R.F. Kahn, Trans.) 1935.
- Woodford, M., 1994. Monetary policy and price level determinacy in a cash-in-advance economy. *Econ. Theory* 4, 345–380.
- Woodford, M., 1995. Price level determinacy without control of a monetary aggregate. *Carnegie-Rochester Conference Series on Public Policy* 43, 1–46.
- Woodford, M., 2000. Monetary policy in a world without money. *International Finance* 3, 229–260.
- Woodford, M., 2001. Monetary policy in the information economy. *Economic policy for the information economy*. Federal Reserve Bank of Kansas City, Kansas City.
- Woodford, M., 2003. *Interest and prices: foundations of a theory of monetary policy*. Princeton University Press, Princeton.
- Woodford, M., 2008. How important is money in the conduct of monetary policy? *Journal of Money, Credit and Banking* 40, 1561–1598.
- Yun, T., 1996. Nominal price rigidity, money supply endogeneity, and business cycles. *J. Monetary Econ.* 37, 345–370.

This page intentionally left blank

PART *Two*

Foundations: Information and Adjustment

This page intentionally left blank

Rational Inattention and Monetary Economics

Christopher A. Sims
Princeton University

Contents

1. Motivation	156
2. Information Theory	157
2.1 Shannon's definition of mutual information	157
2.2 Channels, capacity	158
2.3 Coding	159
3. Information Theory and Economic Behavior	160
3.1 The Gaussian case	161
3.2 Some qualitative conclusions, based on Gaussian-linear-quadratic examples	163
3.2.1 <i>Rational inattention smooths responses and injects signal-processing noise</i>	163
3.2.2 <i>Rational inattention solutions are a special case of rational expectations with noisy observations</i>	164
3.2.3 <i>Rational inattention creates correlation across initially independent sources of uncertainty</i>	166
3.2.4 <i>Rationally inattentive agents react more slowly to slowly moving components of an aggregate</i>	167
3.2.5 <i>Losses from imperfect information processing are small, implying that even small information costs are likely imply substantial imprecision in reactions to signals</i>	168
3.3 Contrast with Mankiw-Reis formulation	168
3.4 Beyond LQ	169
3.5 General equilibrium	170
4. Implications for Macroeconomic Modeling	171
4.1 Be more relaxed about microfoundations for dynamics	171
4.2 Local expansions?	172
5. Implications for Monetary Policy	174
5.1 A critique of rational expectations policy evaluation	174
5.2 Monetary policy transparency	175
6. Directions for Progress	176
7. Conclusion	178
References	180

Abstract

Rational inattention theory is economic theory that recognizes that people have finite information-processing capacity, in the sense of Shannon and engineering information theory. This approach is still in the early stages of development, but it promises to provide a unified explanation for some

of the frictions and delays that are important in dynamic macroeconomics and finance. In this chapter we introduce the basic ideas of information theory, show how it can be introduced formally into dynamic optimization problems, discuss existing applications of the approach, and indicate some of its implications for macroeconomic modeling and monetary policy.

JEL classification: E10, E31, E50, C50, C61

Keywords

Information Theory
Rational Inattention

1. MOTIVATION

Everyone ignores or reacts sporadically and imperfectly to some information that they “see.” I page through the business section of the *New York Times* most mornings, “seeing” charts and tables of a great deal of information about asset markets. I also most days look at *ft.com*’s charts of within-day movements of oil prices, stock indexes, and exchange rates once or twice. But most days I take no action at all based on the information I have viewed. In fact, if you asked me a half hour after I looked at the paper or the Web site what the numbers were I had viewed, I would usually be able to give at best a rough qualitative answer — unless there was some strikingly unusual data. If I were continually dynamically optimizing, I would be making fine adjustments in portfolio, spending plans, bill payment delays, and so forth, based on this information. It is intuitively obvious why I do not — the benefits of such continuous adjustment would be slight, and I have more important things to think about.

One might think that if we were to recognize that people do not use some freely available information, we would have to abandon optimizing-agent models of behavior. Some would be happy with this conclusion, but optimizing-agent models have served economic science well, so it is worthwhile asking whether it is possible to construct optimizing-agent models that are consistent with people not using freely available information. “Rational inattention” models introduce the idea that people’s abilities to translate external data into action are constrained by a finite Shannon “capacity” to process information. Such models explain why some freely available information is not used, or imperfectly used.

Another appeal of such models is that they imply sluggish and erratic response of all types of behavior to external information. In macroeconomic data we see few examples of variables that respond promptly to changes in other variables. Keynesian models recognize inertia in prices, but in their simpler forms translate this inertia in prices into prompt and strong responses of quantities to policy and to other disturbances. This implication of Keynesian models can be softened or eliminated by the introduction of adjustment costs, but such costs are usually modeled one variable at a time and have little support in either intuition or formal theory. A rational inattention approach implies pervasive inertial and erratic behavior, and implies connections across variables in the degree and nature of the inertia.

Studies of transactions prices of individual products, which have proliferated in recent years as electronic cash registers have become common, show that prices tend to stay

constant for extended periods of time, and to jump back and forth among a few specific price points when they do change. This pattern of discretely distributed prices is hard to reconcile with most existing theories of price sluggishness. Yet, although this pattern was not part of the initial inspiration for rational inattention modeling, it has turned out that it is an implication of the rational inattention approach under fairly broad conditions.

In hopes that the reader is now interested in the topic, we turn to the basic mathematics of information theory.

2. INFORMATION THEORY

2.1 Shannon's definition of mutual information

Suppose we are sending the message “yes” and want to quantify how much information is contained in that message. Shannon’s measure of information flow starts from the insight that the amount of information in that message depends on what other messages might have been sent instead. If the recipient of the message was already sure that the message was going to be “yes,” no information is transmitted, and indeed no message should have been sent. If the recipient knew the message would be either “yes” or “no” and was unsure which, a small amount of information would be involved, and it would be easy to send it reliably. But if the recipient knew in advance only that the message would be some English language word, the message would contain much more information and would be much more difficult to send reliably. Shannon’s idea was that the information transmitted ought to be measured by how much the uncertainty of the recipient is reduced by receipt of the message.¹

When two random objects, say X and Y , have a joint distribution with a probability density function $p(x, y)$ Shannon’s definition makes the mutual information between them

$$I(X, Y) = E[\log p(X, Y)] - E\left[\log \left(\int p(X, y) dy\right)\right] - E\left[\log \left(\int p(x, y) dx\right)\right].$$

That is, the information between X and Y is the difference between the expected value of the log of the joint pdf of X and Y and the sum of the two expected values of the logs of the marginal pdf’s of X and Y . This measure has some easily verified appealing properties. It is zero when X and Y are independent, and it is always non-negative. If we have a sequence of observations, say on Y and on Z , we would like the information about X in seeing Z , then Y , to be the same as that in seeing Y , then Z . Thus we would like $I(X, Y)$, calculated from the joint distribution for X and Y , plus $I(X, Z | Y)$, calculated using the joint pdf of X and Z conditional on Y , to be the same as $I(X, Z)$ plus $I(X, Y | Z)$. It turns out that these simple properties are restrictive enough to leave

¹ Here we can only sketch the basic ideas of information theory. More complete treatments are in, for example, Cover and Thomas (1991) or MacKay (2003).

us with essentially only the Shannon measure of mutual information. The “essentially” is needed because we have not specified the base of the log function in the definition. The usual base is 2, in which case the unit of information is a “bit,” while sometimes it is convenient to use base e , in which case the unit is called a “nat.”²

Besides these intuitively appealing properties, the Shannon measure stands out for its proven usefulness in communications engineering. These days, most people are familiar with the idea that they can have fast or slow Internet connections, that there is a measure for the speed (megabits or megabytes — 1 byte = 8 bits — per second), and that the measure does not depend on either the content of the messages being sent (music, text, pictures) or on the physical details of the connection (fiber optic, cable, DSL, etc.).

We should note that the symmetric definition given above is equivalent to

$$I(X, Y) = E[E[\log(q(X|Y))]] - E[\log(h(X))],$$

where $h(X) = \int p(X, y)dy$ is the marginal pdf of X and

$$q(X|Y) = p(X, Y) / \int (p(x, Y)) dx$$

is the conditional pdf of $X | Y$. The quantity $-E[\log(h(X))]$ is called the **entropy** of the random variable X , so that this form of the definition of $I(X, Y)$ makes it the expected reduction in entropy of X from seeing Y . The symmetry of the first definition makes it clear that the expected reduction in entropy of Y from seeing X is the same as the expected reduction in the entropy of X from seeing Y .

2.2 Channels, capacity

Shannon defined a channel as a description of possible inputs and of conditional distributions of inputs given outputs. For example, an ideal telegraph line could send a “dot” or a “dash” (the inputs) and produce a dot at the other end when the input was a dot, and a dash when the input was a dash. A more interesting channel would be a noisy telegraph line, in which the dot or dash input reproduces itself in the output only with probability 0.6, otherwise producing the opposite. In this latter channel, in other words, the probability of error is 0.4 with each transmission. Or a channel might be able to send arbitrary real numbers x drawn from a distribution with variance no greater than 1, producing in the output $y \sim N(x, \sigma^2)$.

The channel only defines conditional distributions of outputs Y given inputs X . The mutual information between inputs and outputs depends also on the distribution of the inputs. If we choose the distribution of the inputs to maximize the mutual information between inputs and outputs, the channel transmits information at its **capacity**. The ideal telegraph key makes the distribution of inputs given outputs degenerate, with all probability on the true value of the input. A discrete distribution with probability

² See Bierbrauer (2005, Chapter 8) for further discussion of the uniqueness.

one on a single point has entropy $0(0 \log(0) + 1 \log(1))$, with the convention that $0 \log(0) = 0$, the limiting value of $a \log(a)$ as $a \downarrow 0$). The information flow is maximized if the input makes dots and dashes equally probable, in which case it is one bit per time period. The noisy telegraph key also has maximal mutual information between input and output when the dashes and dots are equiprobable in the input. Then the information flow rate is 0.029 bits per time period. The channel with Gaussian noise has maximal information flow rate when the input is distributed as $N(0, 1)$, in which case the information flow rate is $-\frac{1}{2} \log_2((\sigma^2/(1 + \sigma^2)))$ bits per time period. When the noise is as variable as the input, so $\sigma^2 = 1$, for example, the rate is 0.5 bits per time period.

2.3 Coding

It is a relatively familiar idea these days that one can take information in various forms and transmit it via an Internet connection. Many of these connections naturally take “ones” and “zeros” (commonly called bits, although this is not exactly the same as the information theory use of that term) as input, and computer disk files represent any kind of information as a pattern of bits. The well-known ASCII code maps each number or upper or lower case letter into a pattern of seven bits. Pictures can be mapped into bit patterns that describe pixels — color intensity amounts at specific points in the picture. This kind of translation of diverse types of information into bits is coding.

But there are many possible ways to map letters and numbers or picture descriptions into bits. Text translated into ASCII codes generally does not emerge with serially uncorrelated bit patterns or with equal numbers of 0s and 1s, and as a result is not ideal input for our ideal telegraph key. There are algorithms that translate such inefficiently coded files into more efficiently coded ones; for example, the zip (for general files) and jpeg (for image files) compression schemes that most computer users have encountered. These compression algorithms produce patterns of zeros and ones that are more nearly i.i.d. and mean 0.5, and become smaller files. The shrinking of these files is equivalent to making them transmit more quickly through an ideal telegraph key.

The coding theorem of information theory states that regardless of the nature of the input we wish to transmit, it can be “coded” so that it is sent with arbitrarily low error rate at arbitrarily close to the channel capacity transmission rate. To get an idea of what coding is and of the meaning of the theorem, suppose we are sending a simple bit-mapped graph of a few black and white lines. The graph has been scanned into a 100×100 grid of pixels, and the file we wish to send is the 100 rows of pixels, one row at a time. With a 0 representing white and a 1 representing black, most of the file will be zeros. Our channel is a perfect telegraph key. Say 2% of the file is 1s. If we simply send the raw file through the channel, it will take 10,000 time periods, one for each pixel. But we could instead transform the file so that a 0 now represents the sequence 000, while 1001 represents 001, 1010 represents 010, and so forth. (Note we end up not using 1000 at all.) Then $0.98^3 = 0.94$ of our three-pixel blocks will be represented by a single 0 in the output, while 0.06 of them will be represented by four-element

sequences. On average, our three-pixel blocks will take $0.94 \times 1 + 0.06 \times 4 = 1.18$ time periods to transmit, so the whole file will take $10000 \times 1.18/3 = 3934$ time periods to transmit. If we think of the file as drawn from a collection of files that have i.i.d. sequences of zeros and ones with probability 0.02 of a one, the entropy of the file is 10000 $(0.02\log_2(0.02) + 0.98\log_2(0.98)) = 1414$ bits.³ If we use the proposed coding, then, we would be sending $1414/3934 = 0.36$ bits per time period, whereas as we have already noted the channel capacity is 1 bit per time period. To get closer to the channel capacity would require more elaborate codes, for example using blocks longer than three.⁴

This example may also help in understanding an important and possibly confusing fact: Even though our ideal telegraph line transmits without error and at a finite rate, a channel that takes continuously distributed input cannot transmit without error unless it has infinite capacity. Suppose input X can be any real number, and output Y simply equals X . Consider our earlier 10000-pixel graphic file. If we take its sequence of zeros and ones and put a decimal point in front of it, it becomes the binary representation of a real number between zero and one. We could then send it through our channel in a single time period without error, a rate of 1414 bits per time period. And of course the same idea would work no matter how large the file, so there is no upper bound on the transmission rate.

The coding theorem is not constructive. Given a channel and a type of message to be sent, finding a way to code it so it can be sent at close to capacity is generally difficult and has generated a substantial literature in engineering.

Our example of coding illustrates another complication that we will be mostly ignoring: coding introduces delay. We showed how to send a file that is mostly zeros by sending the message in blocks. But to do this we need to wait until we have a full block to transmit, which generates some delay. How much delay depends on the nature of the channel and of the message; that is, on properties of the channel and message beyond the channel capacity and the entropy of the message. We ignore coding delay for two reasons. We are at the stage in applications to economic behavior of trying to avoid discussing the physical characteristics of people as information channels, and coding delay is likely to be small — the proportional gap between channel capacity and actual transmission rate decreases at least at the rate $1/n$, where n is the block length of the coding (Cover & Thomas, 1991, Section 5.4).

3. INFORMATION THEORY AND ECONOMIC BEHAVIOR

The idea of rational inattention is to introduce into the theory of optimizing agents an assumption that their translation of observed external random signals into actions must represent a finite rate of information flow; that is, economic agents are finite-capacity channels.

³ If we were really considering only graphics files with black and white line art, the zeros and ones would not actually be i.i.d. (because the ones occur in mostly continuous lines), so the entropy would be smaller and faster transmission possible.

⁴ A longer-block coding example is in the Appendix to my 1998 paper.

Before we proceed to discussing rational inattention models, we should note that these models do not subsume or claim to replace all previous economic models of costly information. In statistical decision theory it is possible to quantify the utility value of observing a random variable, and if the problem includes a budget constraint, to convert this value into a dollar equivalent. This kind of “value of information” applies when there is some physical cost to acquiring the observation such as commissioning a marketing survey, drilling a test well, and so forth. This kind of information cost has nothing to do with the number of bits of information acquired by observing the random variable. Finding whether a test well indicates oil is present may cost thousands of dollars, yet provide only the answer to a yes-or-no question; that is, no more than one bit of information. Rational inattention theory provides no guidance on whether drilling a test well is a good idea. Where it might provide guidance is in explaining why an executive in the oil company, having had a report on the test well on her desk along with other reports about routine matters, might after “looking at” all the reports seem to know the test well report in detail, while having only a vague idea of what was in the other reports. The test well report was important to her job, the others less so, so the others are absorbed less precisely.

Notice also that in the examples that follow the information flow rate is lower than any reasonable guess as to the actual Shannon capacities of humans. It is probably most natural to think of an abstract economic agent as having a shadow value of capacity rather than a fixed capacity bound, because economic optimizations represent only a tiny part of the information-processing that people do. To get realistic delay and noisiness in reactions to information in models where economic decision making is the only reason to process information, we need to postulate very low Shannon capacity, yet at small costs of capacity we find optimizing agents use little of it. This reflects the well-known fact brought out by Akerlof and Yellen (1985) that in the neighborhood of an optimum, modest deviations from fully optimal choices are likely to have very small consequences. People may use economic information at a low rate not because they could not possibly use it more precisely, but because the benefits of doing so would be small and there are other important uses of information-processing capacity.

3.1 The Gaussian case

Rational inattention models are easiest to handle when random variables are all jointly normal. The entropy of a k -dimensional $N(\mu, \Sigma)$ random vector is $\frac{1}{2}(\log(2\pi) + \log |\Sigma| + k)$. This means that the mutual information between two jointly normally distributed random vectors X and Y is half the difference between the log of the unconditional covariance matrix of Y and the log of the residual covariance matrix for a regression of Y on X . It depends only on the correlation matrix of X and Y ,

not on the levels of the variances themselves. If X and Y are each one-dimensional, their mutual information is just $-\frac{1}{2} \log(1 - \rho^2)$, where ρ is the correlation of X with Y :

$$\begin{aligned} X, Y \sim N(\mu, \Sigma) \Rightarrow I(X, Y) &= \frac{1}{2} (-\log |\Sigma| + \log(\text{Var}(X)) + \log(\text{Var}(Y))) \\ &= -\frac{1}{2} \log(1 - \rho_{XY}^2). \end{aligned}$$

Joint normality of a signal Y and an action X is a strong assumption, because rational inattention theory naturally takes the distribution of Y as given and then, based on the loss function and the information constraint, implies a joint distribution for X and Y . Generally, even with Y normally distributed, the information-constrained optimal joint distribution for Y and X is not normal. A comforting result is that there is a form for the loss function that implies joint normality as the optimal form of the joint distribution.

A general static information-constrained decision problem can be formulated as follows:

$$\begin{aligned} \max_{f(\cdot)} \left\{ E[U(X, Y)] = \int U(x, y) f(x, y) dx dy \right\} \text{ subject to} \\ \int f(x, y) dy = g(y) \text{ all } y \\ f(x, y) \geq 0 \text{ all } x, y \\ I(X, Y) = \int \log(f(x, y)) f(x, y) dx dy \\ - \int \log\left(\int f(x', y) dx'\right) f(x, y) dy dx - \int \log(g(y)) g(y) dy \leq \kappa, \end{aligned}$$

where X is the choice variable, g is the given marginal pdf of Y and κ is the maximum information flow rate between Y and X . The objective function is linear in the object of choice (f) and the constraint set is convex, so the problem has a unique maximal value for the objective function. A closely related formulation (actually applied in the examples we will take up) assumes that capacity is variable, at a cost. The left-hand side of the information constraint then appears in the objective function, multiplied by the cost, rather than in a separate constraint.

It may be puzzling that the agent is modeled as choosing a joint distribution rather than as simply choosing X . The problem could be formulated equivalently by saying that the agent chooses an observation $Z = h(Y, \zeta)$, where ζ is a random variable independent of Y and h is an arbitrary (measurable) function. The information constraint is $I(Z, Y) \leq \kappa$ and the agent chooses also a function $d(\cdot)$ and sets $X = d(Z)$. Here the choice of information and the setting of X are separated, which may perhaps be easier to understand. But this formulation is equivalent to the one in terms of choosing f , and has the disadvantage that the same solution $f(\cdot)$ can generally be characterized with many different $d(\cdot)$, $h(\cdot)$ pairs.

At points in X, Y -space where $f(x, y) > 0$, the first-order conditions for an optimum require

$$U(x, y) = \lambda(\log(f(x, y)) - \log\left(\int f(x, y)dy\right)) - \mu(y), \quad (1)$$

where λ is the Lagrange multiplier on the information constraint and $\mu(y)$ is the Lagrange multiplier on the constraint that defines the marginal distribution of Y . This condition can be rearranged to read

$$p(y|x) = M(x)e^{\frac{1}{\lambda}U(x, y)}. \quad (2)$$

If $U(\cdot)$ is quadratic, then the conditional distribution of $Y | X$ is normal at all points x, y where $f(x, y) > 0$. Suppose f is nonzero everywhere, the range of X and Y is unbounded, and the given marginal distribution of Y is $N(\psi, \Delta)$. The exponential part of Eq. (2) will then be proportional to some conditional normal distribution, say $N(\Phi(x), \Sigma(x))$, where Φ is linear in x . We know, though, that for a given $U, \Sigma(x)$, and therefore the normalizing constant $M(x)$, in fact does not depend on x . Therefore the mutual information flow between Y and X does not depend on the conditional mean $\Phi(x)$. Now suppose we choose a normal marginal distribution for X , say $N(\theta, \Omega)$. To match the given marginal distribution for Y , we will have to pick $\Omega = \Delta - \Sigma$. Σ is determined up to a scale factor proportional to λ by $U(\cdot)$. The information flow constraint will require $\log |\Delta| - \log |\Sigma| = 2\kappa$, which pins down λ . Certainty equivalence requires that there is a linear function $d(\cdot)$ that defines the optimal choice of $x = d(y)$ when there is no uncertainty about y , and that $x = d(\Phi(x))$.

3.2 Some qualitative conclusions, based on Gaussian-linear-quadratic examples

The Appendix at the end of the chapter describes how to solve general linear-quadratic optimal control problems. Here we apply the method laid out there to some simple examples that provide insight into the economic implications of rational inattention.

3.2.1 Rational inattention smooths responses and injects signal-processing noise

Suppose P_t is an asset price and X_t is some action an agent takes in response to the asset price. Suppose that in the absence of an information constraint the optimal way to set X_t is to set $X_t = P_t$. If P is a Gaussian stochastic process then, unless it is constant, $P_{t+s} | \{P_s, s < t\}$, the distribution of P_{t+s} given the history of P up to time t , is a Gaussian random variable.

If the optimal choice of X without an information constraint would be $X_t = \alpha P_t$, it is impossible to implement this choice under rational inattention, because it makes knowledge of X_{t+s} completely resolve the continuously distributed uncertainty about P_{t+s} , which as we have already observed implies an infinite information flow rate.

And it is not enough simply to add noise. Suppose $X_t = \theta P_t + \varepsilon_t$. Continuously traded asset prices tend to behave like Wiener processes over small time intervals. In particular, the variance of $P_{t+\delta} - P_t$ decreases linearly with δ and price changes over nonoverlapping time intervals are independent. If ε_t also has this character, then the correlation of $X_{t+\delta} - X_t$ with $P_{t+\delta} - P_t$ tends to some nonzero level as δ shrinks. But that means that the mutual information between $P_{t+\delta} - P_t$ and $X_{t+\delta} - X_t$ tends to a constant as δ shrinks. Thus given a fixed time interval we can, by slicing it up into arbitrarily small subintervals, convey arbitrary amounts of information in the fixed time interval.

It is common to represent continuous time Gaussian processes as stochastic differential equations, of the form

$$dy_t = g(y)dt + h(y)dW_t, \quad (3)$$

where dW_t is a vector of independent white noise processes. The kind of argument we have given earlier implies that if y consists of two components, $y = (x, z)$, and if $h(y)$ is full rank, then for the rate of information flow between z and x to be finite, $h(y)$ must be block diagonal, with blocks corresponding to x and z . This implies that over short time intervals x and z each have variation dominated by their own disturbance process. The component of, say, x that is related to the z shock process must be “more differentiable” than the component related to the shock process of x , so that the variance of changes in x can be dominated by the own-shocks at small time intervals.

3.2.2 Rational inattention solutions are a special case of rational expectations with noisy observations

Consider this very simple dynamic tracking problem. We have a target process y_t that is a first-order univariate autoregression, and we wish to keep our action x_t close to it, with quadratic losses. We can tighten our variance for y_t before we choose x_t by paying an information cost of λ per nat. Formally,

$$\max_{x_t, \sigma_t} \frac{1}{2} E \left[\sum_{t=0}^{\infty} \beta^t \left((y_t - x_t)^2 + \lambda \log \left(\frac{\rho^2 \sigma_{t-1}^2 + \nu^2}{\sigma^2} \right) \right) \right] \text{ subject to} \quad (4)$$

$$y_t = \rho y_{t-1} + \varepsilon_t, \quad (5)$$

where $\nu^2 = \text{Var}(\varepsilon_t)$, σ_t^2 is the variance, after information collection, at t for y_t , and therefore $\rho^2 \sigma_{t-1}^2 + \nu^2$ is the variance for y_t based on time $t - 1$ information, before collecting information at time t .

It is clear that it will be optimal to make x_t always the expectation of y_t given information at t , so we can reduce the problem to one in which the only choice variable is σ_t^2 :

$$\max_{\sigma_t} \sum_{t=0}^{\infty} \beta^t \left(\sigma_t^2 + \lambda \log \left(\frac{\rho^2 \sigma_{t-1}^2 + \nu^2}{\sigma_t^2} \right) \right). \quad (6)$$

This problem can be solved by standard methods, and it has a solution in which σ_t^2 is constant at some finite value. As one might expect, $\sigma_t^2 \rightarrow 0$ as $\lambda \rightarrow 0$. Also, $\sigma_t^2 \rightarrow \infty$ as $\lambda \rightarrow \infty$. This latter result brings out the fact that we have ignored to this point the requirement that $\sigma_t^2 \leq \rho^2 \sigma_{t-1}^2 + \nu^2$. That is, one cannot improve the objective function by “forgetting” previously known information about y . So the full solution is that if the solution to the unconstrained problem implies violation of this forgetting constraint, no information is collected and uncertainty about y is allowed to grow. If the variance of uncertainty about y grows to the point where it exceeds the variance of y in the unconstrained solution, the “no-forgetting” constraint ceases to bind and the solution path begins to follow the unconstrained solution.

Considered as a univariate process, x_t inherits the properties of y_t . This is a general characteristic of rational inattention (and other noisy-observation rational expectations) dynamic optimizations: relative to the decision-relevant information set, the decision variables have the same dynamic structure as the decision variables in the problem with no information-processing constraint. (Here the no-constraint solution would just be $x_t = y_t$.) It is easy to see that, denoting information available to the decisionmaker at time t by $\mathcal{I}_t$, $E[x_t | \mathcal{I}_{t-1}] = E[E[y_t | \mathcal{I}_t] | \mathcal{I}_{t-1}] = \rho x_{t-1}$, so that x_t is an AR process with the same parameter as y . But even though the best predictor of x_t from its own past is ρx_{t-1} , this is generally not the best predictor of x_t from the joint past of y and x .

What then is the joint times series behavior of x_t and y_t in the unconstrained solution? The prediction error for y_t based on information available to the decisionmaker at time $t-1$ is $y_t - \rho x_{t-1}$. The choice of x_t will be based on an improved estimate of this error, and since everything is jointly Gaussian, we can write

$$x_t = \rho x_{t-1} + \theta(y_t - \rho x_{t-1}) + \xi_t, \quad (7)$$

where ξ_t is pure time- t information-processing error and therefore uncorrelated with $\{y_{t-s}, s \geq 0\}$ or with $\{x_{t-s}, s \geq 1\}$. This lets us derive a joint autoregressive representation of (y, x) as

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} \rho & 0 \\ \theta\rho & (1-\theta)\rho \end{bmatrix} \begin{bmatrix} y_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ \theta\varepsilon_t + \xi_t \end{bmatrix}. \quad (8)$$

This implies the moving average representation

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \sum_{s=0}^{\infty} \begin{bmatrix} \rho^s & 0 \\ \theta\rho^s \sum_{u=0}^s (1-\theta)^u & \rho^s (1-\theta)^s \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \theta\varepsilon_t + \xi_t \end{bmatrix}. \quad (9)$$

Notice that if the time unit were very small, we would expect ρ to be near one and, to be consistent with small information flow over small time intervals, θ to be near zero.

Then the second diagonal component of the sequence of weighting matrices in Eq. (9) is the weights on the noise component, and the lower left off-diagonal component is the weights on the part of x that is related to y . We see that as our reasoning above implied, the systematic part of x has small weight ($\theta\rho$) on the initial shock, but that the weight rises smoothly, nearly linearly at first, as we go to more distant lags of the shock. The noise component responds immediately, and the weights decline rapidly — it is less serially correlated than y , while the systematic part of x is much more serially correlated than y .

Note also that this solution is exactly what we would have obtained if we simply postulated that the optimization has to be based on observing at each t a noisy indicator variable $z_t = y_t + \zeta_t$. The variance of ζ would determine the corresponding value of θ in the previous expressions, and $\theta\zeta_t = \xi_t$. What is added by the derivation from rational inattention is (i) that the rational inattention theory predicts that θ and the variance of ξ_t will vary systematically if ν^2 (the variance of ε_t) or λ changes and (ii) we can show that the normal distribution for the “measurement error” is actually what an agent will optimally choose with this objective function. If we made y multivariate we would have still further implied restrictions on the relation of information processing error to underlying disturbance processes and to the objective function.

We were able to solve this problem in two steps. First we recognized that, regardless of the error variance, it was going to be optimal to set $x_t = E[y_t \mid \mathcal{I}_t]$. That allowed us to convert the problem into one that involved only choice of error variance matrices. This two-step process is possible generically in LQ rational inattention problems: First solve for the optimal function relating control variables to states, using certainty equivalence. Then use that solution to find the objective function value as a function of the sequence of error variance matrices alone. The first stage is a standard LQ control problem. The second stage is nonlinear, but deterministic.

Finally, observe that we had to take account of the $\sigma_t^2 \leq \rho^2 \sigma_{t-1}^2 + \nu^2$ constraint, and this slightly complicated our solution. In a multivariate problem the corresponding constraint is that the time- $(t - 1)$ covariance matrix for the state at t minus the post-observation covariance matrix must be a positive semidefinite matrix. Imposing this constraint, when it is necessary to do so, can be much more complicated than imposing it in a univariate problem.

3.2.3 Rational inattention creates correlation across initially independent sources of uncertainty

In our LQ dynamic tracking problem that reduces to Eq. (6), suppose there is no serial correlation, that is, $\rho = 0$. Then the solution is obviously just $\lambda = \sigma^2$. But now add the complication that in fact $y_t = \sum_i z_{it}$, where $z_{it} \sim N(0, w^2)$, independent across t and i . Brief reflection makes it clear that this complication is no complication at all. For optimally choosing x in the face of information process costs, all that matters is that

$y_t \sim N(0, nw^2)$, where n is the number of elements in z . Note, though, that this implies that even if the vector z is freely observable, it will be optimal to collect information only about $\sum_i z_{it}$. The variance of any linear combination $\ell'z_t$ of the z_{it} 's that is uncorrelated with $\mathbf{1}'z_t$ will not be reduced, no matter how low the information cost parameter λ . This implies that the conditional distribution of z_t after an observation has been taken will be of the form

$$\omega^2(I - \alpha(1/n) \mathbf{1}_{n \times n}),$$

where $\alpha = 1$ when $\lambda = 0$ and $\alpha \rightarrow 0$ as $\lambda \uparrow nw^2$. Even though the uncertainty about z_t was uncorrelated across elements of the z_t vector to start with, it optimally becomes negatively correlated across i after information processing.

While this point may seem obvious, taking account of it can complicate analysis. It can be attractive for analytic convenience to assume that uncertainty is constrained to be reduced to keep the correlation structure⁵ of the z 's the same before and after observations are taken. This amounts to discarding one of the important insights from rational inattention theory, however, and should be seen as a last resort at best.

3.2.4 Rationally inattentive agents react more slowly to slowly moving components of an aggregate

A very stylized model of pricing behavior might have a monopolist trying to match prices to a linear function of costs, with quadratic losses. Suppose cost is the sum of two components, one fast-moving, a univariate autoregression with lag coefficient (for example) 0.4, and another slow-moving component with lag coefficient 0.95. Suppose we make the innovation variances to these two components independent of each other and pick them so that the unconditional variances of the two components are equal. We also assume future costs are discounted at the rate β .

Formally, the problem is

$$\min_{p, \Sigma} E \left[\sum_{t=0}^{\infty} \beta^t (\mathbf{1}' \Sigma_t \mathbf{1} + \lambda (\log(|\Omega_{t-1}|) - \log(|\Sigma_t|))) \right] \text{ subject to} \quad (10)$$

$$\Omega_t = \rho \Sigma_t \rho' + \nu \quad (11)$$

$$\Omega_t - \Sigma_t \text{ positive semi-definite,} \quad (12)$$

where our example numbers make

$$\rho = \begin{bmatrix} .95 & 0 \\ 0 & .4 \end{bmatrix}, \nu = \begin{bmatrix} .0975 & 0 \\ 0 & .86 \end{bmatrix} \quad (13)$$

⁵ More precisely, the eigenvectors of the covariance matrix of z .

and λ is the cost of information. As might be intuitively clear, since the maximizer cares only about the sum of the two components, when information costs are low he will choose to make the variances of the components conditional on his information roughly equal and negatively correlated. Since the innovation variance for the slow-moving component is smaller, it is optimal not to track the innovation variance of that component closely, but rather to allow uncertainty about that component to cumulate until it approaches that in the fast-moving component. With $\beta = 0.9$ and $\lambda = 1$, our example makes the optimal choice

$$\Sigma_t = \begin{bmatrix} 0.373 & -0.174 \\ -0.174 & 0.774 \end{bmatrix}, \quad (14)$$

from which we see that the post-observation variance of the fast-moving component is 8% smaller than its innovation variance, while that of the slow-moving component is nearly four times larger than its innovation variance. When we relax the information constraint by setting $\lambda = 0.1$, we find instead

$$\Sigma_t = \begin{bmatrix} 0.318 & -0.300 \\ -0.300 & 0.380 \end{bmatrix}. \quad (15)$$

“News” about the fast-moving component is perceived fairly promptly, while there is little immediate reaction to news about the slow-moving component. The uncertainty about the two components is ex post negatively correlated, reflecting the fact that the monopolist cares only about the sum of the two components and chooses to have imprecise knowledge about how the sum is allocated across components. And as the information constraint is relaxed, it is applied more to the fast-moving than to the slow-moving component.

3.2.5 Losses from imperfect information processing are small, implying that even small information costs are likely imply substantial imprecision in reactions to signals

In these examples, information-processing noise increases linearly with variance. The standard deviation of information processing noise therefore increases very rapidly with information processing costs in the neighborhood of zero processing costs. Though our examples have not been realistically calibrated, when models are realistically calibrated (e.g., Luo, 2008) small information costs lead to low optimal information flow rates and substantial effects on dynamic behavior.

3.3 Contrast with Mankiw-Reis formulation

In an influential paper Mankiw and Reis (2002) proposed a way to model inertial behavior that they call “sticky information.” They discuss their approach in their contribution to this *Handbook* in Chapter 5 (Mankiw & Reis, 2010). Their work is

motivated by some of the same insights that motivate the rational inattention approach. They postulate that agents update their information only at regular intervals that are either fixed or (in later work) variable at a cost. At an information update, agents formulate plans for the period until the next update and stick with those plans. This implies delay and imprecision in response to variation in market signals, just as does rational inattention.

Their formulation is somewhat easier to incorporate into standard macro models, but it is quite different in many of its implications from rational inattention, and it takes us less far along the road away from ad hocery. At updates, agents see all the random variables that define the state of the economy, which are generally taken to be continuously distributed, without error, which as we have seen implies an infinite information flow rate. In a rational inattention setting, no continuously distributed external source of random variation is ever perceived without error, even with a lag. Under rational inattention, delays in reacting to information depend on the amount of serial correlation and the size of disturbances to the external variable; when the external variable moves slowly and varies little, delays in reacting to it can be very long. Under sticky information, there is no such connection of the nature of the external variation to the amount of delay in reacting to it.

Rational inattention, as we have seen, has rich implications about how information from multiple sources is perceived and about how the relative precision of information about different variables depends on loss functions and on the stochastic structure of the external variation being tracked. Sticky information implies no theory about relative precision or delay in observation of different variables. It can allow for differences across variables by allowing for the rate of information collection to be different for different variables, but such formulations are less tractable.

Sticky information implies a different approach to possible microeconomic empirical verification of the theory. It suggests that we would want to examine how often firms or individuals change “plans” for behavior and use these frequencies as an index of the effects of information constraints. Rational inattention, on the other hand, implies that behavior may continually but imprecisely be reacting to external signals, even when information effects are strong. As we will see below, outside the linear-quadratic Gaussian framework rational inattention can imply behavior that changes only at discrete intervals, yet at the same time imply that imprecise knowledge of the state prevails as much at change dates as at other dates.

3.4 Beyond LQ

Sims (2006), Matějka (2008, 2009), and Matějka and Sims (2009) take up models in which objective functions are not necessarily quadratic and supports of distributions are not necessarily unbounded. This necessarily takes us out of the realm of certainty

equivalence and Gaussian distributions. Probably the most interesting result emerging from this work is that solutions often imply a discrete distribution for agent actions, even when the external uncertainty is continuously distributed. The result is the outcome of numerical calculations in most of these papers, but Matějka and Sims (2009) provided an analytic result covering a fairly broad category of models. They show that if (i) the objective is to maximize $U(|x - \gamma|)$, with U having its maximum at zero, (ii) U is analytic on the entire real line, and (iii) the given marginal distribution of γ has bounded support, then with any cost on mutual information between x and γ , the marginal distribution of x is optimally concentrated on a finite set of points.

This kind of result is interesting, because microeconomic data on product prices show not only that prices stay constant over moderately long time intervals, but also that when they change they often jump among a finite set of values (Eichenbaum, Jaimovich, & Rebelo, 2008). There are a number of models in the literature that can explain why prices might change only occasionally, but none that explain why, when they do change, they should move among a discrete set of values. Rational inattention provides an explanation.

If rational inattention is playing even a partial role in determining price-setting behavior, it casts into doubt interpretations often placed on price micro data. Rationally inattentive price setters do not fully adapt to all available information each time they change their prices. Intervals between price changes are therefore nearly irrelevant in determining the degree to which pricing responses to external information (e.g., monetary policy) are delayed or incomplete.

3.5 General equilibrium

Up to this point we have been discussing models of the behavior of individuals reacting to “external” information sources. In modeling an entire economy, or even a market, we must consider interacting agents. This raises special difficulties, as standard market equilibrium models assume prices adjust to clear markets. In a model of a competitive market, prices are usually taken as “external” to both suppliers and purchasers, and it is assumed that both sides of the market see and react to the price. That is how markets are assumed to clear. But in reality prices vary stochastically. If both sides of the market react to market prices with rational inattention, then neither side is reacting precisely and immediately. Prices therefore cannot play their usual market-clearing role.

There are a few models in the literature that consider markets with rationally inattentive agents. They do so by allocating variables to agents, with each variable a decision variable for one type of agent and an external signal to others. For example, Matějka (2009) considered a market with a monopolistic seller choosing prices subject to an information constraint on tracking costs. In a companion paper (Matějka, 2008) he considered a monopolistic price setter facing consumers who face an information

constraint on tracking prices. Maćkowiak and Wiederholt (2009a) set out a complete dynamic stochastic general equilibrium model with pervasive rational inattention, but they too allocated each variable to a unique agent type as a choice variable. Because this allocation is apparently somewhat arbitrary, they examined variants of their model with different allocations.

Such models are reasonable starts on the project of introducing rational inattention into equilibrium models, but probably we need to go further. In many markets, for example any with continuous trading among many buyers and sellers, the allocation of a price variable to one type of agent as a choice variable does not make sense. We instead see special institutions or types of market participants—e.g. retailers, wholesalers, market-makers, and inventories—that allow markets to function without infinite attention from most participants. Recently we have had in asset markets specialist high-frequency traders that process market information at a high rate, using powerful computers. Conventional economic theory, with all agents continuously optimizing using all available information, finds it difficult to explain the role of these specialized economic roles and institutions. At this point, rational inattention has not provided any theory for these institutions and roles either, but it seems to be a promising starting point for such a theory.

Another issue that arises in bringing rational inattention to equilibrium models is that the rational inattention models of individual behavior have nothing to say about properties of information processing error other than its conditional distribution given decision choices. Consider commuters who regularly drive past several gas stations on the way to work. They might not usually pay much attention to gas prices, stopping at stations randomly, or at some customary station, but if one station cut prices sharply, they might, after a day or two, notice and take advantage of the low price. Which day they noticed might be random and uncorrelated across the commuters. On the other hand, many of them might talk to each other, or the local newspaper might run a story on the unusual behavior of gas prices, in which case the timing of their reaction to the price, while no less “noisy,” might be highly correlated across commuters. Information-processing noise that is independent across agents will average out in macroeconomic behavior, whereas highly correlated information processing noise will become an additional source of macroeconomic randomness.

4. IMPLICATIONS FOR MACROECONOMIC MODELING

4.1 Be more relaxed about microfoundations for dynamics

Rational inattention models are difficult to work with and there remain serious substantive issues about how to formulate such models as equilibrium systems. Nonetheless, from the kinds of qualitative results we have described in previous sections, there are some important implications for modeling practice. Rational inattention is

a potential explanation for much of the inertia we see in economic behavior, yet its implications are in many respects quite different from those of other hypotheses about the sources of inertia. This suggests that for the time being it may not be a good idea to insist on specific microeconomic stories about the sources of inertia. Invocation of rational expectations micro-theory models to justify constraints on model dynamics may be a mistake. Use of such microeconomic stories to justify welfare evaluations of alternative policies may also be a mistake. On the other hand, resorting to models that pay no attention to the pervasive inertia and noisiness that we actually observe in dynamic economic behavior would be an even bigger mistake.

We should recognize that many aspects of economic behavior will show slow and erratic adjustment in the direction predicted by optimizing theory, without insisting that agents react as quickly and precisely as rational expectations dynamics would suggest. A promising route forward in this respect is represented by the work of DelNegro, Schorfheide, Smets, and Wouters (2007). They lay out a method for using a rational expectations equilibrium model to generate a prior distribution for the form of a structural vector autoregression (SVAR). The SVAR is left free to match the dynamics in the actual data, to the extent that these data have a strong message about the dynamics, while aspects of the model about which these data do not speak strongly conform to the equilibrium model. Since data generally have much weaker information about long run than about short-run dynamics, this has the effect of putting emphasis on the equilibrium model for the long run, and on the data for the short run. Their method could arguably be improved,⁶ but it seems a step in the right direction and has already been widely applied.

4.2 Local expansions?

Most of the work in economics that applies rational inattention has focused on the linear-quadratic Gaussian case. This fits well with the fact that most of the use of economic equilibrium models fitted to data have entailed working with their local expansions, often just linear expansions, about deterministic steady states. There is a reason for caution, here, however. Working with low-order local expansions of a nonlinear equilibrium model is justified under reasonable regularity conditions when the initial conditions are close to the steady state and the scale of disturbance variation is small.⁷ But in models with a fixed cost of information, like Eqs. (4) and (5), as we let the scale of random variation in the disturbances shrink, information collection goes to zero before disturbances have gone to zero. That is, there is generally a level of random variation so small it is optimal for no information at all to be collected.

⁶ See my comments on the paper in the same issue of the journal.

⁷ See Kim, Kim, Schaumburg, and Sims (2008) for one such set of conditions.

This paradox does not arise if the problem is formulated with fixed Shannon capacity rather than a fixed cost of information processing. As we have already argued, though, it is more appealing to think of people as applying a small part of their full information processing capacity to monitoring economic signals, with a stable shadow price on that processing capacity, than to suppose that they have a fixed capacity constraint.

To end up with a model that is well approximated as linear-quadratic and Gaussian we must think of the scale of economic disturbances to the model as “small,” and at the same time think of the shadow price of Shannon capacity as small. As documented in every application of rational inattention, to get interesting and realistic effects on dynamics requires that information about individual economic variables be processed at a rate of a few bits per month or quarter. Variations in processing rate in that range probably are realistically modeled as having a stable opportunity cost to individuals.

It might seem that the fact that, as we discussed in Section 3.4, optimal behavior of capacity-limited agents often implies discrete behavior would undermine the validity of local LQ Gaussian expansions. This is not necessarily true, however. While it is true that, with initial uncertainty truncated-Gaussian and a quadratic loss function, behavior will emerge as discretely distributed, the number of points in the discrete distribution grows larger as the truncation points become larger in absolute value relative to the standard deviation of the initial uncertainty. The discretely distributed behavior becomes distributed over a finely spaced grid of many points, and its distribution becomes close in the metric of convergence in distribution to a Gaussian distribution, despite remaining discrete.

Although we have presented no formal argument proving this, it does seem then that using local linear expansions of models with rational inattention and maintaining Gaussian assumptions on randomness can be justified. But the conditions that justify this should be kept in mind. In periods of economic disruption — hyperinflations or financial crises, for example — stochastic disturbances are large and people may devote a large fraction of their information-processing capacity to tracking economic signals. In some markets, particularly financial markets, there are some people whose full time job consists of tracking price signals and making trades. The behavior of those people, and hence those markets, are probably not well approximated by linear-quadratic Gaussian rational inattention models, although implications of rational inattention may be even more important in studying the short-term dynamics of such markets than in most macroeconomic applications. At the other extreme, we should bear in mind that it is possible for optimal behavior to imply ignoring variation in some economic signals, because the information costs of attending to it at all do not justify the returns from doing so.

5. IMPLICATIONS FOR MONETARY POLICY

5.1 A critique of rational expectations policy evaluation

One of the main insights about policy from rational expectations theory has been the “rational expectations critique of econometric policy evaluation.” This is the point that, because the stochastic process followed by the economy changes when macroeconomic policy changes, private sector agents’ forecasting rules also change with economic policy. This implies that to project the long-run effects of a policy change, one must calculate the change induced in the stochastic process, accounting for the fact that private sector forecasting rules change.

But in a standard rational expectations model agents forecast optimally, no matter how small or smooth the stochastic fluctuations in the economy. Agents respond to optimal forecasts with the same coefficients, regardless of whether the forecasts are oscillating strongly or are hardly changing.⁸ Agents with rational inattention, though, will respond with more delay and information-processing error — or may not respond at all — to fluctuations that are small and therefore relatively unimportant to them. This implies that rational expectations models estimated from periods of stability will imply large adjustment costs, and that these models are more than likely to be unreliable in tracking behavior when policy or exogenous shocks become much more volatile.

There is, in other words, a “rational inattention critique of rational expectations policy evaluation.” The rational expectations critique of econometric policy evaluation has sometimes been interpreted to mean that use of econometric model conditional forecasts in policy formation is pointless or misleading, as this sort of exercise seldom accounts explicitly for endogenous shifts in expectation-formation in reaction to changed policy rules. As I have argued elsewhere (Sims, 1987), this is a mistake. Most real-time policymaking is in the nontrivial task of implementing a policy rule that changes little, if at all. A correctly identified model can make useful conditional projections of the effects of policy choices without separately identifying the part of its effects that arise from shifts in expectation-formation rules. On the other hand, when we contemplate major changes in policy, we should keep in mind possible rational expectations effects on forecasting rules.

These same points apply to rational inattention. Usually, the effects of rational inattention on the economy’s dynamics take a stable form, so that we can project the effects of policy actions without an explicit model of how rational inattention affects those dynamics. But when there are major shifts in policy or in the nature of

⁸ Strictly speaking, this is true only in a linear or linearized rational expectations model, but the point that coefficients do not shrink when shocks become small in a rational expectations model, while they do shrink as shocks become small in a rational inattention model, remains valid.

exogenous disturbances, we should keep in mind that apparent inertia in historical data from less turbulent times could change character as people shift their attention.

5.2 Monetary policy transparency

Central bankers sometimes have the impression that financial markets and the press misinterpret or overinterpret their policy announcements. The U.S. Federal Reserve makes brief written policy statements after each of the periodic open market committee meetings. The wording of these statements sometimes changes only slightly from one meeting to the next, and the changes in wording are the subject of close analysis by financial market participants and the press. This is sometimes seen as a reason for being parsimonious about handing out information. If small amounts of information produce overreactions in financial markets, after all, wouldn't large amounts of information produce even worse overreactions? And if sophisticated financial experts misinterpret information, wouldn't increased transparency produce even worse misinterpretation by the general public?

A rational inattention perspective suggests that this reasoning has it backwards. Financial market participants are likely to attend to every nuance of whatever information about its policy that the central bank supplies. If the central bank supplies little information, financial experts will make their own estimates of what lies behind the policy statements and will inevitably make some mistakes. Ordinary people will most likely pay little attention to even simple policy announcements, and they will react sluggishly — in effect simplifying the policy statement through their own information-processing filters — whether the information supplied is dense and complex or simple. This might suggest that there is no harm in simply providing detailed information about policy, and as a first approximation this is indeed what rational inattention theory would suggest. Once we recognize that it is inevitable that complex information will be perceived by the public with delay and error, there is an argument for guiding the simplification of the policy message. In effect, by providing its own simplified summary of a more detailed description of policy, the central bank can do some of the work of “coding” the policy statement into a form that the public can track more directly.

Most inflation-targeting banks provide policy statements called inflation reports at regular intervals, and these often have a two-tiered format. A simple and brief characterization of policy and the state of the economy starts the report, and more detail is provided in later pages. This seems like the right approach: a short, low-information-content summary to guide people who will give the announcement only slight attention, together with detail for those who have reason to read it closely.

Some central banks (e.g., those of New Zealand, Norway, and Sweden) have begun providing information about expected future time paths of policy rates. One argument

against this practice has been that it could undermine central bank credibility. The public might focus on, say, a projected interest rate one year ahead, and become disillusioned when, inevitably, the forecast turned out to be inaccurate. But central banks that have taken this course have done so in the context of detailed, regularly updated, inflation reports, of which interest rate forecasts are only one element, and often not the most newsworthy one. Interest rate forecasts are usually displayed as “fan charts” that inhibit their interpretation as simple numerical targets. Since people are unlikely to have loss functions that make minor deviations of forecast from actual interest rates important to them, they are unlikely to focus narrow attention on interest rate point forecasts when these are just one part of a richer presentation of information.

6. DIRECTIONS FOR PROGRESS

We have by now examples of research applying Shannon information-theoretic ideas in a number of directions in economics and finance. One of the earliest was Woodford (2002), which cited rational inattention theory as motivation for considering a model in which agents perceive the state of the economy imprecisely. In later work, Woodford (2009) used information theory more formally, while combining it with other sources of inertia. In finance, Mondria (2005), van Nieuwerburgh and Veldkamp (2004), and Peng and Xiong (2005), for example, have applied information-theoretic ideas. We have already noted the work of Luo (2008), Matějka (2009, 2008), and Matějka and Sims (2009). Luo and Eric Young have a series of papers that apply a rational inattention permanent income framework to, among other things, asset pricing and the current account, a recent example being Luo, Nie, and Young (2010). Maćkowiak and Wiederholt (2009a,b) have worked out a partial equilibrium model of producers pricing in response to multiple sources of cost variation and, later, a complete dynamic stochastic general equilibrium model in which interacting agents of different types face information processing constraints.

All of these papers are worthwhile efforts, but all make compromises to keep the modeling problem tractable. Only the Mondria paper and my early paper (Sims, 2003) consider models with a multivariate state variable and recognize the point made in Section 3.2.3 that rational inattention induces ex post correlation of uncertainty across initially independent state variables. Some deal with problems in which the state is one-dimensional, while others, like those of Peng and Xiong (2005), van Nieuwerburgh and Veldkamp (2004), and Maćkowiak and Wiederholt (2009b), impose ex post independence on initially independent states as a matter of convenience. In their paper, Maćkowiak and Wiederholt (2009a) recognized this limitation on their approach, and tried to allow for it by experimenting with what amounts to rotations of the state space. In a multivariate problem, ex post correlation is induced by the fact that agents will want to collect information only about certain dimensions

of variation in the state. By reducing uncertainty in those dimensions, they induce correlation of remaining uncertainty in other dimensions. But if the state vector can be redefined via a linear transformation so that the components about which agents do not collect information are distinct “state variables,” there will be no induced ex post correlation. Maćkowiak and Wiederholt’s (2009a) approach is therefore a step in the right direction, although there is no way within their framework to verify that they have checked all relevant rotations of the state vector.

As we have already noted, competitive markets, in which prices are equilibrium phenomena not controlled by any one optimizing agent, raise difficult issues for rational inattention modeling. In macro models, in which it has become conventional to postulate prices set by monopolistically competitive firms, this is not directly an issue. But in finance models, where asset prices are not realistically treated as set by monopolists, it is a serious difficulty. The most interesting models would involve market participants who see the market price only via a capacity-limited channel, but if all agents are so limited, the usual competitive market-clearing mechanisms are not available. Finance models that have attempted to model market equilibrium, like Mondria (2005), have therefore tended to make schizophrenic compromises, assuming that some external signals (e.g., market prices) are perceived without error, while others are subject to a capacity constraint.

Recent instabilities in asset markets and their macroeconomic consequences have generated renewed interest by economists in trying to understand liquidity. Gorton and Metrick (2009) provided suggestive evidence that economizing on information-processing requirements created demand for some types of securities before the crash, and the loss in liquidity of these securities as their information-processing requirements increased was a major source of disruption during the crash. It seems likely that insights from information theory can help us understand these phenomena, and there are economists working in this direction, although not with any citable research output to this point.

In modeling asset markets particularly, moving beyond the linear-quadratic Gaussian framework seems important. Even if risky assets have yields with Gaussian distributions, the optimal portfolio problem in the presence of risk aversion is not linear-quadratic, and apparently has not yet been solved, even numerically, under a rational inattention assumption. The result will not be ex post Gaussian uncertainty about yields, and the nature of the induced non-Gaussianity would be interesting to explore. My own work on the two-period savings problem (2005, 2006) and Matějka’s (2008) previously cited work focus primarily on two-period problems. Matějka considered a very simple dynamic problem. Tutino (2009) took up a fully dynamic savings problem without assuming normality, but was constrained by computational considerations to work within a fairly small, discrete probability space. Much, therefore, remains to be done in this area.

7. CONCLUSION

Rational inattention has cast a critical light on much existing financial and macroeconomic modeling, suggesting that the now-standard technical apparatus of rational expectations could easily give misleading conclusions. At the same time, formally incorporating rational inattention into macroeconomic and financial models is an immense technical challenge. While the modest progress to date on these technical challenges may be discouraging, we might take comfort in the fact that rational expectations were seen as imposing immense technical challenges at the outset, so that it took decades for them to become a regular part of policy modeling.

APPENDIX

General Linear-Quadratic Control with an Information Cost

Consider the problem

$$\max_{X_t, \hat{Y}_t, \Sigma_t} E \left[\sum_{t=0}^{\infty} \beta^t (Y_t' A Y_t + Y_t' B X_t + X_t' C X_t - \lambda H_t) \right] \quad (16)$$

subject to

$$Y_{t+1} = G_1 Y_t + G_2 X_t + \varepsilon_{t+1} \quad (17)$$

$$H_t = \frac{1}{2} (\log |M_t| - \log |\Sigma_t|) \quad (18)$$

$$M_{t+1} = \Omega + G_1 \Sigma_t G_1' \quad (19)$$

$$\varepsilon_t | \{Y_s, X_{s,s} < t\} \sim N(0, \Omega) \quad (20)$$

$$M_t - \Sigma_t \text{ positive semi-definite} \quad (21)$$

$$Y_t | \mathcal{I}_t \sim N(\hat{Y}_t, \Sigma_t) \quad (22)$$

$$\{X_t, X_{t-1}, \dots\} \subset \mathcal{I}_t. \quad (23)$$

Then by the law of iterated expectations we can rewrite the objective function as

$$E \left[\sum_{t=0}^{\infty} \beta^t (\text{trace}(\Sigma_t A) + \hat{Y}_t' A \hat{Y}_t + \hat{Y}_t' B X_t + X_t' C X_t - \lambda H_t) \right], \quad (24)$$

where $\hat{Y}_t$ is $E[Y_t | \{X_t, X_{t-1}, \dots\}]$. Since H_t depends on Σ_t and Σ_{t-1} , but not on any values of X or $\hat{Y}$, the objective function is the sum of two pieces, one a function of only the X and $\hat{Y}$ values, the other depending only on Σ_t and M_0 .

We can also rewrite the dynamic constraint (Eq. 17) as a constraint in terms of $\hat{Y}$:

$$\hat{Y}_{t+1} = G_1 \hat{Y}_t + G_2 X_t + \xi_{t+1} \quad (25)$$

$$\text{with } \xi_t = \hat{Y}_t - Y_t + G_1(Y_{t-1} - \hat{Y}_{t-1}) + \varepsilon_t. \quad (26)$$

The error term ξ_t in this equation has two components in addition to the original disturbance ε_t , both of which are uncorrelated with any element of $\mathcal{I}_{t-1}$. The first, $\hat{Y}_t - Y_t$ is minus the error of prediction of Y_t based on the larger information set $\mathcal{I}_t$, and is therefore uncorrelated with anything in $\mathcal{I}_{t-1}$. The second is a linear function of the error in the best predictor of Y_{t-1} based on $\mathcal{I}_{t-1}$, and is therefore also uncorrelated with anything in $\mathcal{I}_{t-1}$. Thus the problem has as one component a conventional linear-quadratic stochastic control problem:

$$\max_{X_t, \hat{Y}_t} E \left[\sum_{t=0}^{\infty} \beta^t (\hat{Y}'_t A \hat{Y}_t + \hat{Y}'_t B X_t + X'_t C X_t) \right]$$

subject to Eq. (25). This can be solved for the optimal linear relation between X_t and Y_t using certainty equivalence, since the variances of disturbances do not affect the solution.

While the solution of the embedded linear quadratic control problem does not depend on the disturbance variances, the value function for the problem does, in general. We will not try to present a general solution method here. However, in the examples considered in this paper, because they are “tracking problems,” the value function for the linear quadratic problem is trivial. The optimal certainty-equivalent solution makes X and Y match perfectly and delivers zero losses. Thus the terms in the objective function involving $\hat{Y}$ and X drop out, leaving the deterministic problem

$$\max_{\Sigma_t} \sum_{t=0}^{\infty} \beta^t (\text{trace}(\Sigma_t A) - \lambda H_t) \quad (27)$$

subject to

$$H_t = \frac{1}{2} (\log |M_t| - \log |\Sigma_t|) \quad (28)$$

$$M_t = \Omega + G_1 \Sigma_{t-1} G_1' \quad (29)$$

$$M_t - \Sigma_t \text{ positive semi-definite.} \quad (30)$$

For this $\sum_t$ part of the problem, the first order condition, if we ignore the positive-definiteness constraint (Eq. 30), is

$$A = \beta \lambda G_1 M_{t+1}^{-1} G_1' - \lambda \Sigma_t^{-1}. \quad (31)$$

If the positive-definiteness constraint does not bind, this is (after using Eq. 29) to eliminate M_{t+1}) a nonlinear equation in $\sum_t$ that can be solved by standard methods. A starting point for a solution, therefore, will generally be to solve this equation and check whether in fact Eq. (30) is satisfied by the solution value of $\sum_t$ and the initial M_0 . If so, the problem is solved. If not, in the univariate case, the solution is still straightforward, because the model is implying that even when no information is collected, so X_t is just a constant, the contribution of additional information is less than its cost. It is possible that with no information collected M_t will grow to the point where it exceeds the optimal $\sum_t$, after which $\sum_t$ remains constant at its optimal value. In the general case, though, we have to treat the solution for $\sum_t$ as a constrained nonlinear deterministic dynamic programming problem.

Even in the simple two-dimensional tracking problem of Section 3.2.4, the positive-definiteness constraint binds. The problem can be solved by making the Cholesky decomposition of $M_t - \sum_t$ the solution parameter, using a Cholesky decomposition constrained to be of a fixed, less than full, rank, and applying the chain rule to convert the first-order conditions with respect to $\sum$ in Eq. (31) to FOCs with respect to the new parameters.

Note some implications of this general treatment. In tracking problems in which information enters the objective function with a fixed cost per bit, the optimal solution will eventually imply a constant $\sum_t$. That is, the uncertainty about the state will not vary with the level of the state variable. Also, when information costs are low enough and initial uncertainty large enough, the solution will move immediately to its steady-state value. And finally, in a multivariate problem it can happen that $M_t - \sum_t$ is only positive semidefinite, not positive definite, implying that information is optimally collected only about certain dimensions of uncertainty about the state vector.

REFERENCES

- Akerlof, G.A., Yellen, J.L., 1985. Can small deviations from rationality make significant differences to economic equilibria? *Am. Econ. Rev.* 75 (4), 708–720.
- Bierbrauer, J., 2005. Introduction to coding theory, discrete mathematics and its applications. Chapman and Hall/CRC, Boca Raton, FL.
- Cover, T.M., Thomas, J.A., 1991. Elements of information theory. Wiley-Interscience, Hoboken, NJ.
- DelNegro, M., Schorfheide, F., Smets, F., Wouters, R., 2007. On the fit and forecasting performance of New Keynesian models. *Journal of Business and Economic Statistics* 25 (2), 123–162.
- Eichenbaum, M., Jaimovich, N., Rebelo, S., 2008. Reference prices and nominal rigidities. Northwestern University and Stanford University, Discussion paper NBER Working paper 13829.
- Gorton, G.B., Metrick, A., 2009. Securitized banking and the run on repo. National Bureau of Economic Research. Working Paper 15223. <http://www.nber.org/papers/w15223>.
- Kim, J., Kim, S., Schaumburg, E., Sims, C., 2008. Calculating and using second order accurate solutions of discrete time dynamic equilibrium models. *Journal of Economic Dynamics and Control* 32 (11), 3397–3414.
- Luo, Y., 2008. Consumption dynamics under information processing constraints. *Review of Economic Dynamics* 11 (2), 366–385.

- Luo, Y., Nie, J., Young, E.R., 2010. Robustness, information-processing constraints, and the current account in small open economies. University of Hong Kong. Discussion paper. <http://yluo.weebly.com/uploads/3/2/1/4/3214259/carbri2010h.pdf>.
- MacKay, D.J.C., 2003. Information theory, inference, and learning algorithms. Cambridge University Press, Cambridge, UK.
- Maćkowiak, B., Wiederholt, M., 2009a. Business cycle dynamics under rational inattention. European Central Bank and Northwestern University. Discussion paper. <http://faculty.wcas.northwestern.edu/mwi774/RationalInattentionDSGE.pdf>.
- Maćkowiak, B., Wiederholt, M., 2009b. Optimal sticky prices under rational inattention. *Am. Econ. Rev.* 99 (3), 769–803.
- Mankiw, N.G., Reis, R., 2002. Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips Curve*. *Quarterly Journal of Economics* 117 (4), 1295–1328.
- Mankiw, N.G., Reis, R., 2010. Imperfect information and aggregate supply. In: Friedman, B.M., Woodford, M. (Eds.), *Handbook of monetary economics*. Elsevier/North-Holland, Amsterdam in press.
- Matějka, F., 2008. Rationally inattentive seller: sales and discrete pricing. PACM, Princeton University. Discussion paper http://www.pacm.princeton.edu/publications/Matejka_F_2008-wp.pdf.
- Matějka, F., 2009. Rigid pricing and rationally inattentive consumer. Princeton University, Discussion paper.
- Matějka, F., Sims, C., 2009. Discrete actions in information-constrained tracking problems. Princeton University Discussion paper.
- Mondria, J., 2005. Financial contagion and attention allocation. Princeton University, Discussion paper.
- Peng, L., Xiong, W., 2005. Investor attention, overconfidence and category learning. Princeton University, Discussion paper.
- Sims, C.A., 1987. A rational expectations framework for short-run policy analysis. In: Barnett, W.A., Singleton, K.J. (Eds.), *New approaches to monetary economics*. Cambridge University Press, Cambridge, UK, pp. 293–308.
- Sims, C.A., 2003. Implications of rational inattention. *Journal of Monetary Economics* 50 (3), 665–690.
- Sims, C.A., 2006. Rational inattention: Beyond the linear-quadratic case. *Am. Econ. Rev.* 96 (2), 158–163.
- Tutino, A., 2009. The rigidity of choice: Lifetime savings under information-processing constraints. Princeton University. Ph.D. thesis. <http://docs.google.com/fileview?id=0B7CdO9AORsjcNwYwZmM1MWEtNDZiNi00NzQzLTgzOTItZmNiM2IzOWQ3MDhh&hl=en>.
- van Nieuwerburgh, S., Veldkamp, L., 2004. Information acquisition and portfolio under-diversification. New York University: Stern School of Business, Discussion paper.
- Woodford, M., 2002. Imperfect common knowledge and the effects of monetary policy. In: Aghion, P., Frydman, R., Stiglitz, J., Woodford, M. (Eds.), *Knowledge, information, and expectations in modern macroeconomics: In honor of Edmund S. Phelps*. Princeton University Press, Princeton, NJ. <http://www.columbia.edu/mw2230/phelps-web.pdf>.
- Woodford, M., 2009. Information-constrained state-dependent pricing. *Journal of Monetary Economics* 56 (S), 100–124.

This page intentionally left blank

Imperfect Information and Aggregate Supply[☆]

N. Gregory Mankiw and Ricardo Reis

Harvard University
Columbia University

Contents

1. Introduction	184
2. The Baseline Model of Aggregate Supply	186
2.1 The starting elements	186
2.2 The solution to the consumer's problem	188
2.3 The full-information equilibrium	188
2.4 The imperfect information equilibrium	190
3. Foundations Of Imperfect-Information and Aggregate-Supply Models	191
3.1 What to choose and plan?	191
3.2 Menu costs	193
3.3 Real rigidities	195
3.4 Strategic complementarities	195
4. Partial and Delayed Information Models: Common Predictions	196
4.1 Nonvertical aggregate supply	197
4.2 Persistence	200
4.3 A digression on sticky prices	203
4.4 Two sources of shocks	205
5. Partial and Delayed Information Models: Novel Predictions	207
5.1 Delayed information and time-varying disagreement	207
5.2 Partial information and optimal transparency	211
6. Microfoundations of Incomplete Information	213
6.1 Inattentiveness	213
6.2 Rational inattention	215
7. The Research Frontier	217
7.1 Merging incomplete information and sticky prices	217
7.2 Heterogeneity in the frequency of information adjustment	218
7.3 Optimal policy with imperfect information	219
7.4 Other choices with imperfect information	220
7.5 DSGE models with imperfect information	221
8. Conclusion	222
References	223

[☆] We are grateful to students at Columbia University and Faculdade de Economia do Porto for sitting through classes that served as the genesis for this survey, and to Stacy Carlson, Benjamin Friedman, John Leahy, and Neil Mehrotra for useful comments.

Abstract

This paper surveys the research in the past decade on imperfect information models of aggregate supply and the Phillips curve. This new work has emphasized that information is dispersed and disseminates slowly across a population of agents who strategically interact in their use of information. We discuss the foundations on which models of aggregate supply rest, as well as the microfoundations for two classes of imperfect information models: models with partial information, where agents observe economic conditions with noise, and models with delayed information, where they observe economic conditions with a lag. We derive the implications of these two classes of models for the existence of a nonvertical aggregate supply, the persistence of the real effects of monetary policy, the difference between idiosyncratic and aggregate shocks, the dynamics of disagreement, and the role of transparency in policy. Finally, we present some of the topics on the research frontier in this area.

JEL classification: D8, E1, E3

Keywords

Inattention
Monetary Policy
Phillips Curve

1. INTRODUCTION

In his Nobel Prize lecture, George Akerlof (2002) said, “Probably the single most important macroeconomic relationship is the Phillips curve.” He is surely right that this relationship has played a central role in many business cycle theories over the past half century. At the same time, however, the Phillips curve has also been controversial and enigmatic.

As originally proposed by Phillips (1958), the eponymous curve entered macroeconomics as an empirical regularity — a mere correlation between a measure of inflation and a measure of economic activity. But soon thereafter, starting with Samuelson and Solow (1960), it was used to fill a need within macroeconomic theory. It explained how the Keynesian short run with sticky prices evolved in the classical long run with flexible prices. Today, in mainstream textbooks, the Phillips curve — or, equivalently, the aggregate supply relation — is the key connection between real and nominal variables. It explains why monetary policy, and aggregate demand more broadly, has real effects.

Once economists recognized the Phillips curve as a key relationship, they quickly started wondering what microeconomic foundation gave rise to this macroeconomic correlation. Friedman (1968) and Phelps (1968) suggested that imperfect information was the key. In the short run, some agents in the economy are unaware of some economic conditions, and this lack of knowledge gives rise to a short-run Phillips curve that, crucially, disappears in the long run.

This emphasis on imperfect information gave rise to more formal treatments of the Phillips curve and, more broadly, to the rational expectations revolution of the 1970s.

Lucas (1972) formalized these ideas in a model in which some agents observe the prices of the goods they produce but not, contemporaneously, the prices of the goods they purchase. Because of this imperfect information, when households observe prices, they face a signal extraction problem to sort out movements in relative prices from movements in the overall price level. The result of this temporary confusion is a short-run Phillips curve.

Following Lucas, a large literature on imperfect information models developed. Some of it was empirical. Barro (1977), for instance, presented results suggesting that the distinction between anticipated and unanticipated movements in money was in fact crucial for explaining the real effects of money. Some of it was theoretical. Townsend (1983), for instance, emphasized how, under imperfect information, people can have different information and thus different expectations, and so forecasting the forecasts of others could be a central element of economic dynamics. In the 1990s, however, this literature went into hibernation. Other theories, including real business cycle models and new Keynesian sticky-price models, took center stage in discussions of economic fluctuations.

This chapter reviews the literature from the 2000s that revives imperfect information as a key to understanding aggregate supply and the Phillips curve. This work differs from the older work in three important, related ways. First, in the new models, information disseminates slowly rather than being perfectly revealed after some brief delay. The older literature assumed that the only obstruction to full information was the unavailability of data, whereas the new work starts from the realization that even when data on aggregates are available, it takes time and resources for people to process this information so they will only gradually incorporate it into their actions. Second, the new work places a greater emphasis on the heterogeneity of expectations that comes with dispersion of information. It is the interaction between agents that are differentially informed that generates new theoretical questions. Third, whereas the older literature had limited strategic interactions, in the new work they take center stage.¹

We start Section 2 by presenting a general equilibrium model of aggregate supply that allows for imperfect information. The model is deliberately simple and, but for one linearization, can be solved exactly in closed form. At the same time, it is quite general; many more complicated models have a similar reduced-form. Section 3 presents the foundations for most models of aggregate supply, including those that rely on imperfect information, introducing fundamental concepts such as menu costs and real rigidities.

Section 4 presents the two approaches to imperfect information models that we will study: partial and delayed information. Under partial information, individuals observe economic conditions subject to noise, whereas under delayed information, they observe conditions subject to a lag. We derive the common implications of these two approaches for three questions: the existence of a nonvertical aggregate supply curve, the persistence of the real effects of monetary policy, and the difference between idiosyncratic and aggregate shocks. We also compare imperfect information to the other leading model of aggregate supply, sticky prices.

¹ Hellwig (2006) gave an alternative short survey of some of the topics covered in this chapter, and Veldkamp (2009) provided a book-length treatment of many other recent applications of imperfect-information models.

Section 5 presents two implications of these two models that have led to new questions and data analysis. Delayed information models make sharp predictions for the dynamics of disagreement and have led to the use of survey data, while partial information models have shed new light on the debate over whether policy should be transparent.

Section 6 looks at the microfoundations of the two approaches. Recent work on “rational inattention” (surveyed in Chapter 4 by Sims in this *Handbook*) has been used to justify the assumption of partial information. In turn, models of “inattentiveness” have provided a microfoundation for delayed information models.

Section 7 discusses more recent work that has taken these new approaches to imperfect information in different directions. These include the merging of imperfect information with sticky prices, the study of optimal policy, and the integration of these models with more conventional dynamic stochastic general equilibrium models. Section 8 presents conclusions.

2. THE BASELINE MODEL OF AGGREGATE SUPPLY

We start with a model of monopolistic competition in general equilibrium, which is now standard in the study of monetary policy.²

2.1 The starting elements

To focus on the behavior of aggregate variables, we assume that there are complete insurance markets where all individual risks can be diversified. It takes only a small step to further assume that there is a representative agent that maximizes a utility function with a convenient functional form:

$$\max_{\{[C_{it}, L_{it}]_0^1, B_t\}_{t=0}^\infty} E \left\{ \sum_{t=0}^{\infty} \beta^t \left[\ln C_t - \int_0^1 \left(\frac{L_{it}^{1+1/\psi}}{1 + 1/\psi} \right) di \right] \right\}. \quad (1)$$

The representative consumer has full information, $E(\cdot)$ denotes the statistical expectations operator, and $\beta \in (0, 1)$ is the discount factor. There are many varieties of labor in this economy, referring to different skills and occupations, and the labor supplied by each is denoted by L_{it} , where i lies in the unit interval. ψ is the common Frisch elasticity of labor supply. Aggregate consumption, C_t , is a Dixit-Stiglitz aggregator of the consumption of many varieties of goods, also indexed by i , where $\gamma > 1$ governs the elasticity of substitution across varieties:

$$C_t = \left(\int_0^1 C_{it}^{(\gamma-1)/\gamma} di \right)^{\gamma/(\gamma-1)}. \quad (2)$$

² Blanchard and Kiyotaki (1987) presented an early example. Gali (2008) gave a recent textbook presentation on these models in the context of aggregate supply.

The budget constraint at each date t is

$$\int_0^1 P_{it} C_{it} di + B_t \leq \int_0^1 W_{it} L_{it} di + B_{t-1}(1 + R_t) - T_t + P_t \int_0^1 X_{it} di. \quad (3)$$

On the left-hand side are the uses of funds: spending on goods' varieties that each sells for P_{it} dollars, and saving an amount B_t in one-period bonds. On the right-hand side are the sources of funds. The first term is labor income, where W_{it} is the dollar wage that the i^{th} variety of labor earns. The second term is the return on savings, where R_t is the nominal interest rate. The two other terms are T_t , government lump-sum taxes, and X_{it} , the real profits from firm i .

There is a continuum of firms, where firm i hires labor variety i in a competitive market, taking W_{it} as given, but is the monopolistic supplier of good variety i . The maximand of each firm is its perceived real profits, as given by:

$$X_{it}(\cdot) = \hat{E}_{it}[(1 + \tau)P_{it}Y_{it}/P_t - W_{it}H_{it}/P_t], \quad (4)$$

where τ is a sales subsidy and Y_{it} is output produced using H_{it} units of labor. Because it is a monopolist, the firm takes into account that sales equal market demand, $Y_{it} = C_{it}$, together with the production function:

$$Y_{it} = A_{it}H_{it}. \quad (5)$$

Productivity A_{it} is stochastic and we denote its aggregate component by $A_t = \int A_{it} di$.

Note that the expectations of the firm are represented by the operator $\hat{E}_{it}(\cdot)$, which does not have to coincide with the full-information statistical operator $E(\cdot)$. If the firm had full information, then there would not be an expectation in expression (4) because all variables are known at date t when the firm makes its choices. The focus of this chapter is on the consequences of firms not having full information and having to form expectations of current prices, wages, and productivity.

The market-clearing conditions are $L_{it} = H_{it}$ in the labor market and $B_t = 0$ in the bond market. Fiscal policy simply taxes the consumer to pay for the sales subsidy: $T_t = \tau \int P_{it} Y_{it} di$. Monetary policy ensures that nominal income,

$$N_t = P_t Y_t, \quad (6)$$

follows an exogenous stochastic process. We refer to these shocks to N_t as “demand” shocks, while changes in productivity are “supply” shocks. We do not model the way in which monetary policy achieves the path for N_t , which may be directly via the money supply together with a cash-in-advance constraint in the consumer's problem, or via a nominal interest-rate rule with a very large response to deviations of N_t from $P_t Y_t$. Chapter 24 by Friedman and Kuttner (2010) in Volume 3B of this *Handbook* discusses these modeling and implementation issues.

2.2 The solution to the consumer's problem

Because the consumer's utility function is time separable and the aggregator across varieties is homothetic, the consumer problem breaks into two stages. In the first stage, for a given total consumption C_t , the consumer minimizes total spending subject to the constraint in Eq. (2). The solution to this problem delivers the demand function for each variety:

$$C_{it} = C_t (P_{it}/P_t)^{-\gamma}, \quad (7)$$

and the definition of the static cost-of-living price index:

$$P_t = \left(\int_0^1 P_{it}^{1-\gamma} di \right)^{1/(1-\gamma)}, \quad (8)$$

with the property that $\int P_{it} C_{it} di = P_t C_t$.

In the second stage, the consumer solves the intertemporal problem of choosing aggregate consumption and labor supply to maximize Eq. (1) subject to the sequence of budget constraints in Eq. (3). The solution is characterized by a Euler equation and a continuum of labor supply equations at each date in time:

$$1 = \beta E_t[(1 + R_{t+1})P_t C_t / P_{t+1} C_{t+1}], \quad (9)$$

$$C_t L_{it}^{1/\psi} = W_{it}/P_t. \quad (10)$$

These conditions describe the consumer's decisions under both full information and imperfect information on the part of firms.

2.3 The full-information equilibrium

We first solve the model under the assumption of full information. In this special case, the firms' expectations $\hat{E}_{it}(\cdot)$ are identical to the full-information statistical operator $E(\cdot)$. This case is a standard benchmark against which we will compare the model with imperfect information.

Turning to the firm's problem, under full information, maximizing Eq. (4) subject to Eqs. (5) and (7) has a simple solution:

$$P_{it} = \left[\frac{\gamma}{(\gamma - 1)(1 + \tau)} \right] \left(\frac{W_{it}}{A_{it}} \right). \quad (11)$$

Firm i sets a price equal to a fixed markup over marginal cost, which equals the wage rate divided by labor productivity.

Combining all of the equations from Eqs. (7)–(11), a few steps of algebra show that in equilibrium:

$$p_{it} = p_t + \mu + \alpha(\gamma_t - a_{it}). \quad (12)$$

We have followed the convention that variables in small letters equal the natural logarithm of the same variable in capital letters. This equation states that the price of each firm increases one-to-one with the aggregate price level. The constant in this equation, $\mu = \ln[\gamma/(\gamma - 1)(1 + \tau)]/(1 + \gamma/\psi)$, reflects the markup. It is zero if price exactly equals marginal cost; more generally, it depends on the substitutability of the goods' varieties and the magnitude of the sales subsidy. Finally, the third term in the equation reflects the facts that higher output and consumption raise the marginal disutility of working and lower the marginal utility of consumption, thereby raising wages, marginal costs, and prices, while higher productivity lowers marginal costs and, therefore, prices.

The elasticity of the firm's price with respect to output is α , which equals $(\psi + 1)/(\psi + \gamma)$. This elasticity will play an important role, so let us pause and gauge its likely size. Because γ is greater than one, α must be smaller than one; α increases with the Frisch elasticity of labor supply and falls with the goods' elasticity of demand. Estimates of the labor supply elasticity ψ using micro data tend to be around 0.2, while macro estimates are closer to 1. Micro estimates of the goods' demand elasticity γ are around 4, while macro estimates are around 10.³ Therefore, α lies somewhere between 0.12 and 0.4. Our baseline preferred values are $\psi = 0.5$ and $\gamma = 7$, leading to $\alpha = 0.2$.

The monetary policy rule in Eq. (6) is exactly log-linear:

$$n_t = p_t + \gamma_t, \quad (13)$$

but the price index in Eq. (8) is not. It has a simple log-linear approximation around the point where all prices are the same:

$$p_t = \int_0^1 p_{it} di. \quad (14)$$

This is the only approximation that we make in the full-information case.

Combining equations Eqs. (12)–(14) gives the full-information equilibrium for output and prices:⁴

$$y_t^F = a_t - \mu/\alpha, \quad (15)$$

$$p_t^F = n_t - a_t + \mu/\alpha. \quad (16)$$

We are now in a position to define the object of our study: the *aggregate supply curve*. This is a map in (γ, p) space that comes from varying the demand shock n_t . With full information, aggregate supply is vertical, as output is independent of monetary policy.⁵ It shifts to the right when productivity increases, and to the left if markups rise.

³ See Rogerson and Wallenius (2009) and Chetty (2009) for a discussion of micro and macro elasticities of labor supply, and Kimball and Shapiro (2008) for recent macro estimates. For macro estimates of the elasticity of goods' substitution see Hall (1988) and Basu and Fernald (1997), while for micro estimates see Broda and Weinstein (2006).

⁴ The model also has solutions for nominal interest rates, hours worked, and consumption of different varieties, which can be derived using the equilibrium conditions. We do not focus on these.

⁵ Mathematically, the slope of the aggregate supply curve is defined as $(\partial y_t / \partial n_t) / (\partial p_t / \partial n_t)$.

The Pareto optimum in this economy has output equal to productivity, which is ensured by $\mu = 0$ or a constant subsidy $\tau = 1/(\gamma - 1)$, and we will assume this case from now onwards (but most conclusions do not depend on this simplification).

2.4 The imperfect information equilibrium

Now consider the case in which firms have imperfect information about economic conditions. The consumer optimality conditions are still given by Eqs. (9)–(10). For the firm though, optimal prices now satisfy:

$$\hat{E}_{it} \left[\left(\frac{P_{it}}{P_t} \right)^{-\gamma} \left(\frac{Y_t}{P_t} \right) \right] = \left[\frac{\gamma}{(\gamma - 1)(1 + \tau)} \right] \hat{E}_{it} \left[\left(\frac{P_{it}}{P_t} \right)^{-\gamma-1} \left(\frac{W_{it}}{A_{it}P_t} \right) \left(\frac{Y_t}{P_t} \right) \right] \quad (17)$$

If the firm has full information, this reduces to Eq. (11). Log-linearizing Eq. (17) around the nonstochastic case and using the assumption that $\mu = 0$ delivers the solution:

$$p_{it} = \hat{E}_{it}[p_t + \alpha(\gamma_t - a_{it})] \quad (18)$$

The term inside the expectations is the nominal marginal cost of the firm. The firm must form expectations of the aggregate price level, output, and idiosyncratic productivity, because these are the three determinants of marginal costs. In this simple model, the firm would only have to see the wage it is paying its workers and their productivity to exactly measure marginal cost, but in the far more complicated reality that the model is trying to capture, firms find it quite difficult to precisely measure their own marginal cost, as evidenced by the large sums spent every year in accounting systems and consultants.⁶

Equation (18) reflects the certainty-equivalence result that prices with imperfect information equal the expected price under full information in Eq. (12). Here it follows because a linearization of the optimality conditions is equivalent to a quadratic approximation of the objective function.⁷ This property has been used at least since Simon (1956) to make problems of incomplete information easier, and we will often (but not always) rely on it. The imperfect information equilibrium is defined as the values of γ_t and p_t such that Eqs. (13), (14) and (18) hold. To complete the model, the only ingredient that needs to be added is a specification of how firms form expectations.⁸

⁶ A more realistic model would also take into account that production and delivery lags imply that the firm must make many decisions based on future marginal costs, so that forming expectations is unavoidable.

⁷ This result will also hold exactly if all variables are log-normal, but now with a different expression for μ .

⁸ While the previous model is quite simple, it is also quite general. As Woodford (2003) showed, assuming that the preferences of the representative consumer are $u(C_t) - \int v(L_{it})di$ each period or that the production function is $Y_{it} = A_{it}f(H_{it})$ leads to the same reduced-form after a log-linearization around a nonstochastic steady state. The only change is that the parameter α now depends on the curvature of these functions at the steady state, but reasonable calibrations lead to values not far from the 0.2 that we will work with.

3. FOUNDATIONS OF IMPERFECT-INFORMATION AND AGGREGATE-SUPPLY MODELS

If the firm has neither limits to its rationality nor any constraints on its ability to process information, then more information is better. The firm can always freely dispose of the information, and in general the ability to make more accurate forecasts will allow it to make decisions that yield higher expected profits.⁹ To justify why people do not have full information therefore requires the presence of some information or rationality cost, k .¹⁰ The cost can be real resources or utility losses, may be variable or fixed, and may even be implicit in the form of shadow multipliers on an information constraint. Section 6, on the microfoundations of imperfect information, is devoted to models of these costs. In this section, we discuss the choices that these information costs generate.

3.1 What to choose and plan?

With full information, we can think of the firm as either choosing the quantity of output to produce or the price to set. Choosing one of them instantly determines the other via the demand function. For instance, if the firm chooses its price, then using its information on aggregate output and the price level, it knows exactly the amount of output it will produce. With imperfect information, these two options are no longer equivalent. If the firm that chooses a price does not know aggregate output and the price level, it will not know how much output it will end up producing and selling at that price. An important component of an imperfect information model is the decision variable of the agent.

Reis (2006a) endogenized this choice by letting the firm choose *ex ante* its decision variable. If the firm chooses a plan for the price it charges, replacing the constraints into Eq. (4), its expected profits are

$$\bar{X}_{it}^P = \max_{p_{it}} \hat{E}_{it}[(1 + \tau)P_{it}^{1-\gamma}P_t^{\gamma-1}Y_t - W_{it}A_{it}^{-1}P_{it}^{\gamma-1}Y_t]. \quad (19)$$

A firm that instead chooses a plan for the output it produces expects to earn:

$$\bar{X}_{it}^Y = \max_{Y_{it}} \hat{E}_{it}[(1 + \tau)Y_{it}^{1-1/\gamma}Y_t - W_{it}A_{it}^{-1}P_t^{-1}Y_{it}]. \quad (20)$$

Assuming there is no cost differential between planning prices and planning quantities, the firm will choose a price plan if $\bar{X}_{it}^P \geq \bar{X}_{it}^Y$ and a quantity plan otherwise.

⁹ It is possible that, even though each firm individually is better off with more information, in equilibrium all are worse off. Hirshleifer (1971) is a classic example where the private return to inventors of racing to obtain information before others exceeds its social value.

¹⁰ There have been a few attempts at measuring this information cost directly. The most notable is Zbaracki et al. (2004). By following a large industrial firm for a year, they measured the information costs of changing prices to be as large as 1% of revenue.

To see what this decision entails, assume that all firms have full information, so the aggregate equilibrium is the full-information one described in Section 2.2 with $Y_t = A_t$ and $P_t = N_t/A_t$, and consider the marginal firm i that is choosing between price and quantity plans. Three cases highlight the different considerations at play. First, assume that there are no supply shocks ($A_{it} = 1$) and only demand N_t is stochastic, so that on aggregate $Y_t = 1$. In this case, manipulating Eq. (17) shows that the quantity plan involves choosing $Y_{it} = 1$, which is the full-information optimum. Quantity plans are preferred in this case, as the configuration of shocks makes the optimal quantity independent of news. Second, consider the case where monetary policy targets prices by the rule $N_t = A_t$, which ensures that on aggregate $P_t = 1$ and $Y_t = A_t$. Now, the optimal price for the marginal firm is $P_{it} = 1$, which can be achieved by a price plan since it requires no knowledge of news. Therefore, the price plan is preferred. Finally, consider the case where the A_{it} are idiosyncratic, with no aggregate shocks ($N_t = A_t = 1$). Some algebra shows that in this case, the firm is indifferent between price and quantity plans. Intuitively, with idiosyncratic productivity shocks, only the firm's idiosyncratic marginal costs are random. The demand for its good is fixed, so picking a price sets a quantity, and vice versa, so the two give the same expected profits.

More generally, consider the case where the demand is an arbitrary function, $Y_{it} = Q(P_{it}, s_{it})$, with shocks s_{it} , while marginal costs are constant.¹¹ Then, a second-order approximation of the real profits under the two plans around the nonstochastic means of the shocks reveals that price plans are preferred if:

$$Q_s Q_{ps} + \left(-\frac{Q_s}{2Q_p} \right) Q_{pp} \leq 0 \quad (21)$$

To understand this result, consider the case depicted in Figure 1 of a monopolist producing with zero marginal costs and facing a linear demand with slope one and additive shocks. Linear demand means that Q_{pp} is zero, and additive shocks that Q_{ps} is also zero, so Eq. (21) states that the firm should be indifferent between price and quantity plans. To see this graphically, the optimal price and quantity are Q^* and P^* if the shock equals its expected value, and because of the assumptions, the line segments from Q^* to O and from P^* to O are of the same length. If there is a positive shock to demand, then with a price plan, the new equilibrium will be at A , whereas with a quantity plan it will be at B . Because OA and OB have the same length, the firm is indifferent, confirming the mathematics.

Consider now the case where the shock hits the slope of the demand curve, so that when it shifts out, it becomes flatter. In this case, $Q_s Q_{ps} < 0$ so the result says that price plans are preferred. To see this graphically, note that OC is longer than OB so profits under a price plan are higher. Finally, say that when the demand curve shifts out, its

¹¹ For the case with a general cost function, see Reis (2006a).

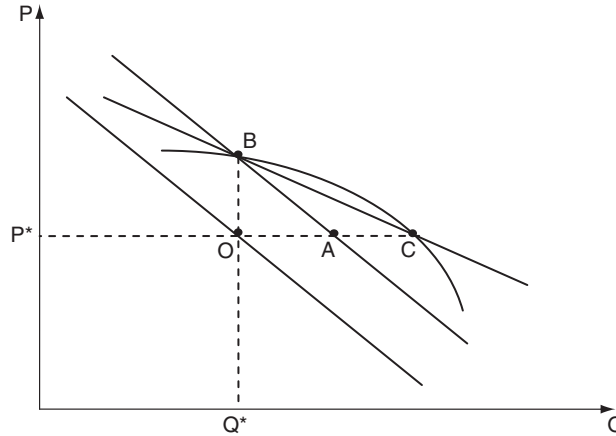


Figure 1 Choosing between price and quantity plans.

slope on the horizontal dislocation is unchanged ($Q_{ps} = 0$) but the demand curve is now concave ($Q_{pp} < 0$). Again, because OC is longer than OB , price plans are preferred.

In the end, either price planning or quantity planning may be optimal for a firm facing imperfect information. But the determinants of this choice, like the shape of the firm's demand curve and the influence of the shocks on demand, are measurable so the theory provides sharp answers to guide the construction of models and can be tested using data.

3.2 Menu costs

Consider the following question: If everyone has full information, will the marginal firm facing information costs k wish to pay this cost to obtain information? If the answer is no, then with these information costs the full information outcome is not a Nash equilibrium.

This question is another way to pose the issue examined by Mankiw (1985) and Akerlof and Yellen (1985). Figure 2 plots the profit function for a marginal, imperfectly informed firm in a full-information economy, using Eq. (4), the functional forms in Section 2, and the extra assumption that there are only aggregate demand shocks, which are zero mean i.i.d log-normal with standard deviation σ . On the vertical axis are the profits with imperfect information relative to profits with full information, and on the horizontal axis is the standard deviation of the aggregate demand shock. Noticeably, the profit function is flat at the certainty case, so even a small cost k implies that the firm does not want to obtain information even for relatively high σ . Numerically, a cost of 1% of profits in the nonstochastic case leads to optimal individual

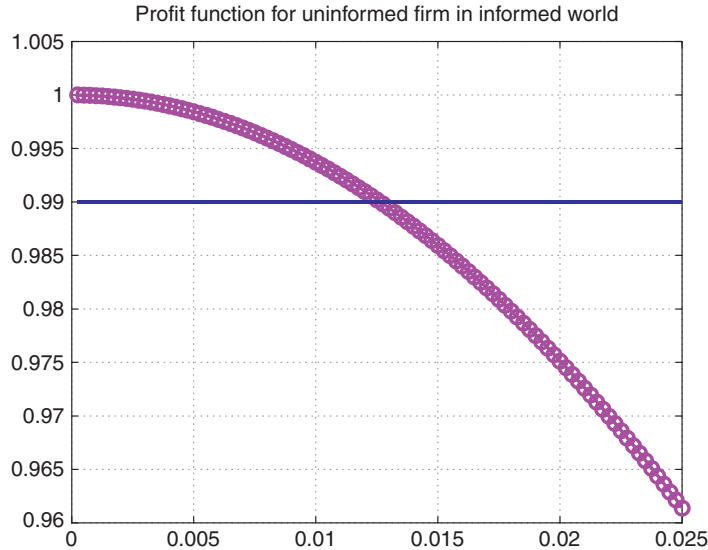


Figure 2 Profits if inattentive while all other firms are fully informed.

inattentiveness for $\sigma \leq 0.0125$. In post-war United States, the standard error of nominal quarterly GDP growth is 0.01, which from the other perspective implies that as long as k exceeds 0.63% of profits, the firm will wish to become inattentive, and full information is not a Nash equilibrium.

This point can be made more generally using second-order log approximations. The firm's profits in Eq. (4), $X_{it}(p_{it} - p_t, \cdot)$, depend on the price it charges together with the other exogenous variables. With full information, the optimal choice is $p_{it} = p_t$, whereas, without information, the optimal p_{it} is some value p_{it}^* . The firm will choose to stay inattentive if:

$$X(0, \cdot) - X(p_{it}^* - p_t, \cdot) \leq k. \quad (22)$$

A second-order approximation around $p_{it}^* = p_t$ yields:

$$-X_p(0, \cdot)(p_{it}^* - p_t) - 0.5X_{pp}(0, \cdot)(p_{it}^* - p_t)^2 \leq k \quad (23)$$

The crucial insight, similar to that in Mankiw (1985), is that $X_p(0, \cdot) = 0$, since this is the necessary condition for the full-information price choice. Moreover, for small shocks to nominal income, p_{it}^* is close to p , and the second squared term is tiny. Even if k is a small cost of getting the information for updating a price menu, condition (23) will likely hold. This result is rooted in the envelope theorem: close to the maximum the profit function is flat, so small shocks have a second-order impact on profits. Hence, small informational costs may be sufficient to explain the failure of price setters to be fully informed.

3.3 Real rigidities

While the previous result shows that it is unlikely for full information to be a Nash equilibrium, the opposite question remains: Is an equilibrium where all are uninformed a Nash equilibrium? The answer to this question is closely related to the concept, emphasized by Ball and Romer (1990), of real rigidities.

Focusing on the case with only demand shocks so the profit function is $X(p_{it} - p_t, n_t)$, then $X(0, 0)$ are the profits without any shock to nominal income, $X(0, n_t)$ are the profits if the firm remains inattentive like all the other firms in the economy, and $X(p_{it}^*(n_t), n_t)$ are the profits if it obtains information, where $p_{it}^*(n_t)$ is the optimal price in this case as a function of the state of demand. Imperfect information will be a Nash equilibrium if:

$$X(p_{it}^*(n_t), n_t) - X(0, n_t) \leq k. \quad (24)$$

A second-order approximation of the expression on the left-hand side of Eq. (24) for n_t close to 0 yields:

$$0.5[X_{pp}(0, 0)(\partial p_{it}^*/\partial n_t) + 2X_{pn}(0, 0)](\partial p_{it}^*/\partial n_t)n_t^2 \leq k. \quad (25)$$

Because $p_{it}^*(n_t)$ is implicitly defined by the optimality condition $X_p(p_{it}^*(n_t), n_t) = 0$, the implicit function theorem gives the derivative: $\partial p_{it}^*/\partial n_t = -X_{pn}(0, 0)/X_{pp}(0, 0)$. But going back to the solution for $p_{it}^*(n_t)$ in Eq. (12), note that this is just the definition of the parameter α . Using it in the expression above gives the final condition:

$$-0.5\alpha^2 n_t^2 X_{pp}(0, 0) \leq k. \quad (26)$$

Note that if α is small, this condition is more likely to be satisfied.¹²

Ball and Romer (1990) labeled the parameter α an index of real rigidities. In particular, a smaller α means more real rigidity. Note that α is a “real” parameter in that it depends on the properties of the real profit function. Ball and Romer’s (1990) insight was that this real parameter influences the economy’s nominal rigidity. Their result carries over to this setting: The more real rigidity there is, the more likely it will be that imperfect information on the part of price setters is a Nash equilibrium.¹³

3.4 Strategic complementarities

A concept closely related to real rigidity is the concept of strategic complementarity. Combining the expression for desired prices with the exogenous process for nominal income yields:

¹² The second-order condition for the optimum requires that $X_{pp}(0, 0)$ is negative.

¹³ There are different mechanisms to generate real rigidities (Romer, 2008), as well as some challenges like the common finding that real rigidities induce firms to want to adjust more frequently in response to idiosyncratic shocks (Dotsey & King, 2005).

$$p_{it} = \hat{E}_{it}[(1 - \alpha)p_t + \alpha n_t - \alpha a_{it}]. \quad (27)$$

Cooper and John (1988) interpreted this expression as the best response by firm i to the other firms' actions, captured by the sufficient statistic p . Taking this game-theoretic perspective to the equilibrium of the model, $\alpha < 1$ implies that pricing decisions are strategic complements. That is, if other firms raise their prices, then firm i wishes to raise its price as well.

Strategic complementarities are important because with heterogeneity of information, there will be some firms that are better informed than others. If pricing decisions are strategic complements, then the better-informed firms will still not want to change their prices by much to keep them in line with the less-informed firms. Strategic complementarities therefore ensure that the aggregate supply curve is not too steep, so there is significant monetary non-neutrality. One illustration of this role is that two influential articles that found very steep aggregate supply curves (Chari, Kehoe & McGrattan, 2000; Golosov & Lucas, 2007) both chose parameters that make α larger than one.

It is not entirely surprising that the same parameter α and condition $\alpha < 1$ are important for both real rigidities and strategic complementarities, even though these concepts start from different places. If the informed firm i does not want to change its price p_{it} much after a shock because it knows the other uninformed firms will not, then it will typically also be the case that the profit gain from obtaining information and changing p_{it} is small. Because of these similarities, the concepts of real rigidity and strategic complementarity are often used interchangeably in this literature, and we will do so in this chapter as well.¹⁴

4. PARTIAL AND DELAYED INFORMATION MODELS: COMMON PREDICTIONS

Having set out the basic framework in Section 2 and examined some foundational issues in Section 3, we now consider two models of imperfect information that have commanded attention in recent years. We call these the partial information model and the delayed information model.

Both of these models assume that people form expectations optimally but with incomplete information. The difference is the nature of the incompleteness. The delayed information model assumes that only a share λ of firms have up-to-date information, while the remaining have old information from previous periods. The partial information model assumes that firms observe a noisy signal with a relative

¹⁴ With strategic complementarities comes the scope for multiple equilibria. Ball and Romer (1989) characterized the equilibrium multiplicity in their model with full information, while Morris and Shin (1998, 2001) and Heineman (2000) did it with partial information and Hellwig and Veldkamp (2009) with delayed information.

precision τ . Both models introduce just one new parameter, λ or τ , which can be interpreted as an index of informational rigidities. By maintaining the assumption of optimal behavior subject to these new informational constraints, the tools used to solve these models are familiar to economists accustomed to rational expectations models.

To present the essence of these two approaches, consider our baseline model with only aggregate demand shocks that follow a random walk, so $n_t = n_{t-1} + v_t$, with v_t normally distributed with mean zero and variance σ^2 .¹⁵ Combining Eqs. (13), (14), and (18), the equilibrium price level solves the equation:

$$p_t = \int_0^1 \hat{E}_{it}[\alpha n_t + (1 - \alpha)p_t] di. \quad (28)$$

The overall price level in the economy is an average across firms of their expectation of their optimal prices, which in turn are a weighted average of the level of demand n_t and the price level p_t .

From this equation we can examine several features of imperfect information models that apply in both variants. First, we will show how incomplete information generates a nonvertical aggregate supply curve in a simple model where all information gets revealed after one period. Next, we will introduce gradual revelation of information to understand the persistence of the real effects of aggregate demand shocks. After a brief detour to compare imperfect information with sticky prices, we finally will consider the effects of idiosyncratic productivity shocks.

4.1 Nonvertical aggregate supply

Consider first the delayed information model. In this model, λ of agents have full information, so their subjective expectation of the contemporaneous values of aggregate demand and the price level coincides with the actual values of these variables. Suppose, for now, that the remaining $1 - \lambda$ do not observe current shocks but do have full information on all variables one period before and form expectations optimally given this information. The equation describing the equilibrium for the price level becomes:

$$p_t = \lambda[\alpha n_t + (1 - \alpha)p_t] + (1 - \lambda)E_{t-1}[\alpha n_t + (1 - \alpha)p_t]. \quad (29)$$

The key tool to solve this class of models is the “innovations representation” of the equation, sometimes also called the Wold representation. In particular, by rearranging terms, we can write the equation as:

$$p_t - E_{t-1}(p_t) = \alpha\lambda[n_t - E_{t-1}(n_t)] + (1 - \alpha)\lambda[p_t - E_{t-1}(p_t)] + \alpha E_{t-1}(n_t - p_t). \quad (30)$$

¹⁵ There is nothing special about the random walk beyond making the algebra slightly easier. The tools laid out below would apply to most other linear stochastic processes.

Now, with the exception of the last term on the right-hand side, all other terms are uncorrelated innovations, and therefore have an expectation of zero as of the previous period. Taking expectations at $t-1$ of both sides of the equation shows that $E_{t-1}(p_t) = E_{t-1}(n_t) = n_{t-1}$, so the last term is zero. Solving for the innovation in prices as a function of the innovation in aggregate demand yields:

$$p_t = \left[\frac{\alpha\lambda}{1 - (1-\alpha)\lambda} \right] (n_t - n_{t-1}) + n_{t-1}, \quad (31)$$

$$y_t = \left[1 - \frac{\alpha\lambda}{1 - (1-\alpha)\lambda} \right] (n_t - n_{t-1}). \quad (32)$$

Variation in the expected level of aggregate demand n_{t-1} leads to proportional changes in prices and no effect on output. However, shocks to aggregate demand $n_t - n_{t-1}$ increase both output and prices: the aggregate supply is no longer vertical.¹⁶ The slope of the aggregate supply curve falls with both α and λ ; that is, the stronger are informational or real rigidities, the flatter is the aggregate supply curve. Intuitively, uninformed firms do not adjust their price in response to a positive aggregate demand shock, which causes their sales to rise. Therefore, more uninformed firms lead to stronger monetary non-neutrality. In turn, for lower values of α , firms that do become informed want to set their prices closer to those of the uninformed firms, which leads their sales to rise and aggregate output to increase by more.

Now consider the partial information model. This model assumes that all firms have noisy signals of the state of aggregate demand. They observe $z_{it} = n_t + \varepsilon_{it}$, where the noise ε_{it} is independent across firms and time, normally distributed, has zero mean, and variance equal to σ^2/τ . The parameter τ plays the same role as λ did in the delayed-information model: a higher τ means fewer informational rigidities because it implies that z_{it} is a more accurate signal of n_t .

A key feature of the partial information model is the absence of common knowledge. In particular, because each firm's signal is its private information, it cannot credibly transmit it to anyone else, so no one knows what others in the economy know.¹⁷ There is a role for higher order beliefs, as each firm must form a belief of what other firms believe, as well as of what other firms believe that the firm believes, and so on. A consequence of this is that the law of iterated expectations does not hold in aggregate: in particular $\int \hat{E}_{it}[\int \hat{E}_{it}(\cdot) di] di \neq \int \hat{E}_{it}(\cdot) di$, or the second-order average belief is not equal to the first-order one. Successively taking expectations from the perspective of agent i , and averaging over all the agents, Eq. (28) becomes:

¹⁶ We can also write the equilibrium in terms of an expectations-augmented Phillips curve as in Friedman (1968): $\Delta p_t = E_{t-1}(\Delta p_t) + [\alpha\lambda/(1-\lambda)]y_t$.

¹⁷ In contrast, in the delayed-information model, a share λ of firms have full information and know exactly what the other firms know.

$$p_t = \alpha \sum_{j=1}^{\infty} (1 - \alpha)^{j-1} \bar{E}_t^{(j)}(n_t). \quad (33)$$

We used the notation $\bar{E}_t(\cdot) \equiv \int \hat{E}_{it}(\cdot) di$, $\bar{E}_t^{(2)}(\cdot) \equiv \int \hat{E}_{it}[\int \hat{E}_{it}(\cdot) di] di$, and so on, as well as the limiting condition that average infinite-order beliefs do not explode to infinity (which can be verified later).

The crucial tool to solve the partial information problem is the signal extraction formula. In particular, it is a standard result in statistics that:

$$\hat{E}_{it}(n_t) = E_t(n_t | z_{it} = n_t + \varepsilon_{it}) = E_{t-1}(n_t) + \left(\frac{\tau}{1 + \tau} \right) [z_{it} - E_{t-1}(n_t)]. \quad (34)$$

This equation gives up the first-order belief — the expectation of aggregate demand. The second-order belief is the expectation of others' expectations of aggregate demand. This is found by averaging Eq. (34) over all firms and then taking the expectation of the resulting expression, which yields:

$$\hat{E}_{it}^{(2)}(n_t) = E_{t-1}(n_t) + \left(\frac{\tau}{1 + \tau} \right)^2 [z_{it} - E_{t-1}(n_t)]. \quad (35)$$

In this equation, each firm is using the signal it obtains to forecast other firms' signals and thus their expectations of demand. Note the signal obtains a smaller weight in this second-order belief than it did in the first-order belief. More generally, iterating over these two steps delivers the j -th order belief:

$$\bar{E}_t^{(j)}(n_t) = E_{t-1}(n_t) + [\tau / (1 + \tau)]^j [n_t - E_{t-1}(n_t)]. \quad (36)$$

Combining this expression with Eq. (33) gives the solution:

$$p_t = \left(\frac{\alpha \tau}{1 + \alpha \tau} \right) (n_t - n_{t-1}) + n_{t-1}, \quad (37)$$

$$\gamma_t = \left(1 - \frac{\alpha \tau}{1 + \alpha \tau} \right) (n_t - n_{t-1}). \quad (38)$$

Comparing with the solution for the delayed information model in Eqs. (31) and (32), one finds that the models make very similar predictions. In particular, stronger real and nominal rigidities again imply a flatter aggregate supply curve. The intuition is clear in the limits. If $\tau = 0$, the signal is useless, the firms have no information on the current shocks, and prices are unchanged, so the aggregate supply curve is horizontal. If instead $\tau \rightarrow \infty$, firms have full information and the aggregate supply curve is vertical. In between, better information implies that prices adjust by more so the curve is steeper. The role of real rigidities arises because the smaller is α , the more firms want to

charge what other firms are charging, and the more weight each gives to what others are thinking. In other words, more real rigidity gives a larger role to higher order beliefs. The higher the order of the beliefs, the closer they are to $E_{t-1}(n_t)$ and the less they respond to the signal. Thus, in the partial information model, as in the delayed information model, more real rigidity means greater monetary non-neutrality.

4.2 Persistence

So far, because all information becomes known after one period, aggregate demand shocks moved output for only one period. Now, we relax this assumption by assuming that firms have imperfect information that may last for an extended span of time.

In Mankiw and Reis (2002), we proposed a model of persistent delayed information, which we called the sticky-information model. We assumed that every period, a fraction λ of firms gets independently drawn from the population and receives full information.¹⁸ At any date, there will be a share $\lambda(1-\lambda)^j$ of firms that last updated their information j periods ago. With this exponential distribution, the equilibrium price level now solves the equation:

$$p_t = \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-j}[\alpha n_t + (1-\alpha)p_t]. \quad (39)$$

The innovations representation for aggregate demand is $n_t = \sum_{k=0}^{\infty} v_{t-k}$, where the v_{t-k} are the uncorrelated innovations. Since the n_{t-k} are the only shocks in the model, it is a good guess that the innovations representation for the price level will depend on them as well: $p_t = \sum_{k=0}^{\infty} \varphi_k v_{t-k}$. Solving the model is to solve for the φ_k unknown coefficients. While the approach of the previous section will not work, a slight extension of it does: the method of undetermined coefficients.¹⁹ It relies on two observations: first, that $E_{t-j}(p_t) = \sum_{k=j}^{\infty} \varphi_k v_{t-k}$ and likewise for n_t , and second that the φ_k must be the same for all possible realizations of the shocks. Equation (39) then imposes the conditions:

$$\varphi_k = \lambda \left[\alpha \sum_{j=0}^k (1-\lambda)^j + (1-\alpha) \varphi_k \sum_{j=0}^k (1-\lambda)^j \right], \quad (40)$$

for every $k = 0, 1, \dots$. These equations yield the model's solution:

¹⁸ An allegory for this model is to think of each firm having a stochastic alarm clock that every period rings with probability λ making it “wake up” and see what is going on.

¹⁹ There is a long tradition of using this method to solve macroeconomic models with rational expectations. See Taylor (1985) for an early review. More recently, Mankiw and Reis (2007), Reis (2009b), and Meyer-Gohde (2010) developed general algorithms to solve sticky-information models with many equations and variables.

$$p_t = \sum_{k=0}^{\infty} \left[\frac{\alpha[1 - (1 - \lambda)^{k+1}]}{1 - (1 - \alpha)[1 - (1 - \lambda)^{k+1}]} \right] v_{t-k}, \quad (41)$$

$$y_t = \sum_{k=0}^{\infty} \left[\frac{(1 - \lambda)^{k+1}}{1 - (1 - \alpha)[1 - (1 - \lambda)^{k+1}]} \right] v_{t-k}. \quad (42)$$

On impact, a positive aggregate demand shock still leads to an increase in both prices and output, and stronger real and informational rigidities still enhance the response of output and attenuate the response of prices. Figure 3 plots the impulse responses of both output and inflation over time, with $\lambda=0.25$, so that firms update their information on average once per year.²⁰ Output only approaches zero asymptotically as the share of firms that have learned about the shock goes to 1, and the half-life of the shocks is one-and-a-half years. The response of inflation is also delayed with two

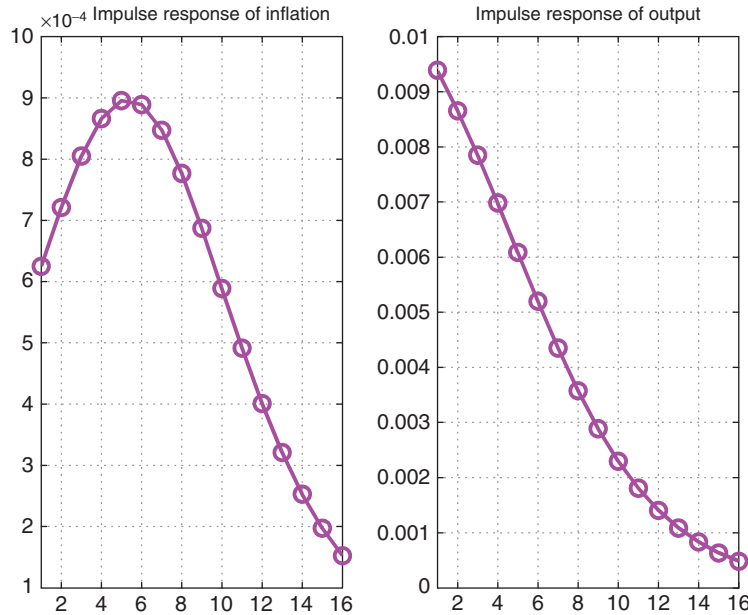


Figure 3 Impulse response of inflation and output to nominal demand shocks with delayed information.

²⁰ Khan and Zhu (2006) and Döpke, Dovern, Fritsche, and Slacaleck (2008a) econometrically estimated Phillips curves with sticky information and found $\lambda = 0.25$ for the United States, France, Germany, and the UK, while $\lambda = 0.5$ for Italy.

properties that have been emphasized in the empirical literature: (i) it is hump-shaped, and (ii) it peaks after output.²¹

Let us turn now to the partial information model. Its dynamic version is due to Woodford (2002), who called it the imperfect common knowledge model. This model assumes that each firm receives a private signal z_{it} of aggregate demand, just as before, but now never gets to learn what past aggregate demand was. As it receives new signals, the firm not only forms an expectation of the present circumstances, but also revises its views on the past. Therefore, as in the sticky-information model, all firms will eventually become informed about the value of a shock today.

The approach of the last section only works if we assume that after some large number of periods, shocks become common knowledge. Hellwig (2008) and Lorenzoni (2009, 2010) take this route and let the number of periods become larger and larger to obtain an approximation to the solution. Woodford (2002) instead proposed an alternative guess-and-verify method using dynamic signal-extraction tools.²² The guess is that:

$$p_t = (1 - \theta)p_{t-1} + \theta n_t. \quad (43)$$

Writing this guess, together with the random-walk for nominal demand and the signal z_{it} in vectors, gives:

$$\begin{pmatrix} n_t \\ p_t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \theta & 1 - \theta \end{pmatrix} \begin{pmatrix} n_{t-1} \\ p_{t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ \theta \end{pmatrix} v_t \Rightarrow s_t = Ms_{t-1} + cv_t, \quad (44)$$

$$z_{it} = \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} n_t \\ p_t \end{pmatrix} + \varepsilon_{it} \Rightarrow x_{it} = es_t + \varepsilon_{it}. \quad (45)$$

Here we have defined the new matrices and vectors s_t , M , c , and e to write the problem as a state-space system. The dynamic version of the signal extraction formula in Section 4.1 is the Kalman filter:

$$E_{it}(s_t) = ME_{it-1}(s_{t-1}) + \kappa[z_{it} - eME_{it-1}(s_{t-1})], \quad (46)$$

where $\kappa = (\kappa_1, \kappa_2)'$ is a 2×1 vector of Kalman gains (e.g., Hamilton, 1995, Chapter 13). Integrating this expression over all agents, and using Eq. (44) then leads to:

$$\bar{E}_t(s_t) = \kappa e Ms_{t-1} + (M - \kappa e M) \bar{E}_{t-1}(s_{t-1}) + \kappa e cv_t. \quad (47)$$

²¹ Coibion (2006) thoroughly described the features of the sticky-information model that generate hump shapes in inflation.

²² Other approaches to solving partial information models are Amato and Shin (2006), who truncated the problem going backwards at some date, Rondina (2008), who used the Wiener-Kolmogorov formulae for signal extraction, and Kasa (2000) who attacked the problem in the frequency domain.

Next, note that Eq. (33) implies that $p_t = \alpha \bar{E}_t(n_t) + (1 - \alpha) \bar{E}_t(p_t)$. Using Eq. (47) to replace for the average expectations of n_t and p_t , and performing the matrix algebra operations, this equation for the price level becomes:

$$p_t = (1 - \theta)p_{t-1} + [\alpha\kappa_1 + (1 - \alpha)\kappa_2]n_t + [\theta - \alpha\kappa_1 - (1 - \alpha)\kappa_2]\bar{E}_{t-1}(n_{t-1}). \quad (48)$$

This verifies the original guess in Eq. (43) and shows that $\theta = \alpha\kappa_1 + (1 - \alpha)\kappa_2$. The expressions for the Kalman gains are messy, but one can show that θ is the positive solution of the quadratic equation:

$$\theta^2 + \alpha\tau\theta - \alpha\tau = 0. \quad (49)$$

The partial information model again has similar predictions to its delayed information counterpart. There is still an upward-sloping aggregate supply curve, and the larger the indices of real and informational rigidities then the larger and more persistent the effects of nominal demand on output. Figure 4 has the impulse responses, and while the one for output is similar to Figure 3, the one for inflation has a significant difference: there is no hump-shape.²³ While the absence of hump shapes is not a generic property of the partial information model (they appear with other stochastic processes for aggregate demand), this case shows that the two models are not observationally equivalent. With good enough data, we would be able to distinguish between them.

4.3 A digression on sticky prices

The main alternative to models of imperfect information and aggregate supply are models based on sticky prices. Indeed, in much of the recent business-cycle literature, the norm for explaining price adjustment is some version of the Calvo (1983) model. A full comparison of these approaches is beyond the scope of this chapter. But, because we have just been discussing persistence, it is worth noting one specific comparison regarding the dynamics of inflation. This particular difference between the approaches, at least in their simplest form, has motivated some recent work on imperfect information.

The Calvo model can be viewed as a special case of the sticky-information model in which the plan that firms set for prices must consist of a single number for all dates. Therefore, when a firm chooses its plan, it sets a price that is optimal on average over the duration of the plan. The optimal price to set at the adjustment date is then a weighted average of the expected optimal price at all dates in the future. This leads to front-loading: changes in expected future conditions affect prices today. This

²³ The value of τ was set to 0.005 so that the impact response of output is the same as in the delayed information model. The standard deviation of the noise is therefore fourteen times the standard deviation of the shock to demand. Whether this is realistic or not is hard to say; finding a direct empirical counterpart to the signal-to-noise ratio in partial information models is a standing challenge.

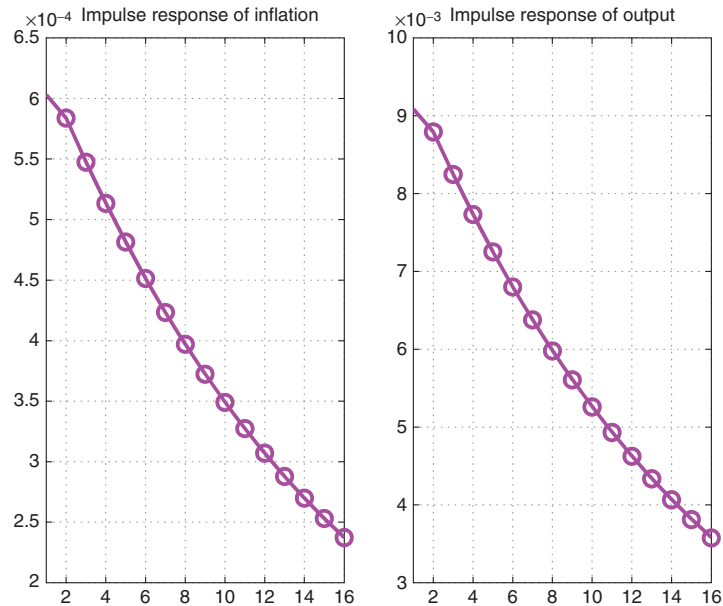


Figure 4 Impulse response of inflation and output to nominal demand shocks with partial information.

front-loading is the source of many empirical problems of the sticky-price model, described by Mankiw (2001), Mankiw and Reis (2002), Rudd and Whelan (2007), and others.

The first problem comes from trend inflation. The weighted average that gives the optimal adjustment price will be too high relative to the optimum today, and too low relative to the optimum in the future, so that even if there is full information, the long-run aggregate supply curve will not be vertical. The second problem is that prices and inflation will jump in response to news today about future circumstances. In the data, however, estimated impulse responses of inflation to shocks are very sluggish and often hump-shaped. Ball (1994) put this problem in an elegant way: if the monetary authorities announce today a disinflation for the future, Calvo price-setters will cut their prices immediately, leading to a boom in economic activity. The experience of almost all disinflations in the OECD refutes this prediction.

Various solutions to the problems of the Calvo model have been suggested. Perhaps firms choose not prices but price deviations from a trend or target price index. Or perhaps firms automatically index their prices to past inflation. Or perhaps a fraction of firms follow simple rules of thumb when setting prices.

While these modifications of the Calvo model solve some of its empirical shortcomings, they come with two problems of their own. First, by assuming

backward-looking behavior in ways that are not observed in the micro data, they effectively renounce the enterprise of microfoundations. Firms do not seem to index their prices, nor does such indexation follow from even boundedly rational behavior; if the goal is to just add whatever it takes to fit the macro data, then one might as well do this from the start, in the tradition of good reduced-form work. Imperfect information, in contrast, is a theory of optimal forward-looking behavior that does not imply front-loading and therefore does not require these fixes to avoid its counterfactual implications.

The second problem with these fixes was highlighted by Reis (2006a). The sticky-information model can account not only for the persistent inflation of post-war United States, but also for the serially uncorrelated inflation of the pre-war era. The reason is that incompletely informed but optimizing agents adjust their behavior to the different monetary policies of those two periods. The many hybrid versions of the Calvo model, by rigging in automatic persistence to fight the front-loading behavior of the model, cannot fit the data from different policy regimes if their key parameters (such as the degree of automatic indexation or the share of rule-of-thumb agents) are truly structural and therefore invariant to the policy regime.

In addition to the Calvo model, there is another strand of models with sticky prices, in which firms choose every instant whether to change their prices subject to a fixed cost. These are sometimes called state-dependent models. An important difference between these models and models of imperfect information is the role of what is called the selection effect. In state-dependent pricing models, only those firms whose current price is very far from their optimal price will choose to adjust. Thus, when firms adjust, they do so by a large amount. This selection effect means that substantial movements in the overall price level can be consistent with many firms not adjusting at all. As a result, the aggregate supply curve can be very steep and the effects of monetary policy very small and transient. By contrast, with imperfect information, firms do not know for sure what their optimal price would be. Therefore, this selection effect is mitigated, and all else equal, aggregate demand shocks have larger and more persistent effects.

Despite the problems of models with full information and sticky prices in fitting the aggregate data, the fact remains that most prices in the economy change infrequently. A more promising route than comparing sticky prices with imperfect information is instead to develop models that merge the two approaches. There is already some exciting work in this area, which we review in Section 7.

4.4 Two sources of shocks

The models of imperfect information can also take into account many sources of information. In this section we show how by reintroducing the shocks to idiosyncratic productivity A_{it} . For simplicity, we revert to the assumption of Section 4.1 that information becomes known to all after one period.

One approach to deal with multiple shocks is to assume, following Mankiw and Reis (2006), that there is still only one source of information. In particular, in the delayed information model, there is a single parameter λ , and when firms obtain information, they observe both the aggregate and the idiosyncratic shocks. In the partial information model, the corresponding assumption is that there is only one noisy signal. Because the firm wants to set a price proportional to its nominal marginal cost, it would want its piece of information to be a single signal of this variable. If we restrict signals to exogenous variables, the component of nominal marginal cost is $n_t - a_{it}$ (see Eq. 18), so the firm would choose to observe a noisy signal z_{it} on this.²⁴

Following the same steps as in Section 4.1, the solution for output and prices is exactly the same as in Eqs. (31) and (32) and (37) and (38). Imperfect information on idiosyncratic shocks leads to more mistakes in the prices set by uninformed firms, but these are uncorrelated with the mistakes due to aggregate demand shocks. Even though the losses in profits from lack of information increase, the predictions for the slope of the aggregate supply curve are unchanged.²⁵

An alternative approach, following Carroll and Slacalek (2007) and Mackowiak and Wiederholt (2009), is to assume that there are two sources of information. In terms of the delayed information model, this would imply that the share of firms receiving news about aggregate demand (call it λ'') is different from the share of firms receiving information about idiosyncratic productivity (say λ'). In the partial information model, the precision of information on the two shocks might be different, leading to two separate indices of informational rigidity, τ'' and τ' .

Working through this version of the model, it is straightforward to show that again the same aggregate equilibrium holds, and that it is the rigidity of aggregate information, λ'' or τ'' , that affects the aggregate supply curve. One virtue of this extension is that it is possible to have firms that are well informed about their local conditions, while being misinformed about the aggregate. Moreover, because the firm cares about marginal cost, which depends on $n_t - a_{it}$, if idiosyncratic shocks are much more volatile than aggregate shocks, firms will try to obtain more accurate information on a_{it} rather than on n_t . Since the benefits of obtaining more information on the more volatile idiosyncratic shocks are always larger than the benefits of more information on the aggregate shocks, as long as the cost of the two types of signals is the same, firms will get more information on the idiosyncratic conditions.²⁶

²⁴ Models where agents receive signals from endogenous variables, like prices, are much harder to solve and so have unfortunately been little explored so far. Angeletos and Werning (2006) are an exception, but their focus was on the uniqueness of equilibrium.

²⁵ One difference from the model in of Section 4.1 is that now, like in Lucas (1973), the ratio of the variances of the aggregate and idiosyncratic shocks will affect the slope of aggregate supply.

²⁶ Mackowiak and Wiederholt (2009) showed this result for the partial information model using the rational inattention approach to model the costs of information. The same result holds for the delayed information model using the inattentiveness microfoundation, as long as improving the accuracy of information on each of the two shocks has the same cost.

The virtue of allowing for two sources of information is that it is then possible to have individual prices being quite volatile in response to the closely monitored idiosyncratic productivity shocks, while at the same time aggregate prices are sluggish in response to poorly observed nominal demand shocks.²⁷ Mackowiak and Wiederholt (2009) and Nimark (2008) emphasized this point to match the large and frequent price changes that we observe in the micro data. Klenow and Willis (2007) found support in the micro data for the proposition that price changes only slowly incorporate past aggregate information on nominal demand.

5. PARTIAL AND DELAYED INFORMATION MODELS: NOVEL PREDICTIONS

Beyond addressing the long-standing questions about the slope of the aggregate supply curve and the persistence of economic fluctuations, the two classes of imperfect information models also have generated a variety of new applications.

5.1 Delayed information and time-varying disagreement

Mankiw, Reis, and Wolfers (2003) emphasized the predictions of the sticky-information model for disagreement. In this model, without news, everyone would have the same information and would make the same forecasts of the future. In response to news, some people learn about it and revise their forecasts, while others remain uninformed, so there is disagreement. As more people become informed, and more news happens, different groups emerge with different forecasts. In the delayed information model, disagreement is therefore an endogenous variable that comoves with the other endogenous variables in response to the shocks. This prediction can be tested using survey data on people's expectations.²⁸

The most reliable large data sets on people's expectations concern inflation. The Michigan Survey of Consumer Attitudes and Behavior asks a cross-section of 500 to 700 members of the general public every month, the Livingston Survey collects the forecasts of 48 professional economists twice a year, and the Survey of Professional Forecasters surveys 34 professional forecasters every quarter. These surveys have long-term series (starting in 1946, 1946, and 1968 for the three surveys, respectively), and they expend considerable effort making sure that the respondents provide answers on a particular common measure of inflation. While some care is always warranted in

²⁷ One feature of reality that these models ignore is that information on some variables may be easier to obtain and understand than information on other variables.

²⁸ In contrast, while there is also disagreement in the partial information model, it is always equal to the exogenous variance of the signals.

interpreting the results from surveys, these are perhaps the best available measures of disagreement.²⁹

In the delayed information model, define disagreement as the cross-sectional standard deviation of inflation expectations:

$$D_t = \sqrt{\lambda \sum_{i=0}^{\infty} (1-\lambda)^i [E_{t-i}(\Delta p_{t+1}) - \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-j}(\Delta p_{t+1})]^2}. \quad (50)$$

Taking the solution for the price level in Eq. (41) in Section 4.2, and using again the method of undetermined coefficients, a few steps of algebra show that this expression equals:

$$D_t = \sqrt{\sum_{i=0}^{\infty} \left\{ \lambda (1-\lambda)^i [1 - \lambda (1-\lambda)^i] \sum_{k=i}^{\infty} (\varphi_k - \varphi_{k-1}) \right\} v_{t-i}^2}. \quad (51)$$

Figure 5 plots the impulse response of disagreement to a one-standard deviation shock to aggregate demand. On impact, disagreement increases by almost 0.12%, and at its peak by 0.18%. Mankiw, Reis, and Wolfers (2003) found that in the data disagreement

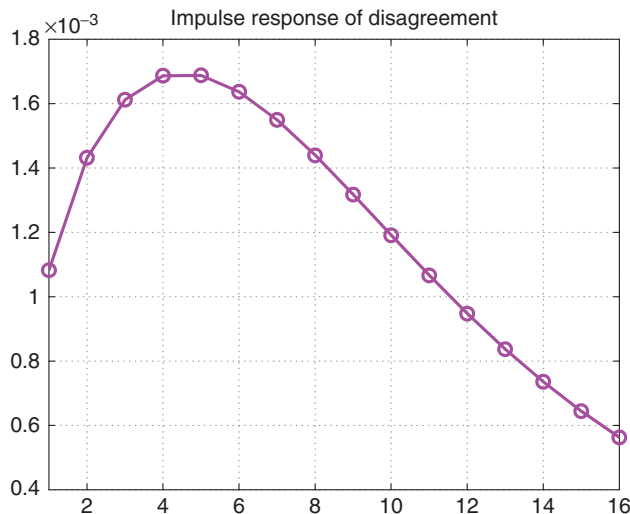


Figure 5 Impulse response of disagreement with delayed information.

²⁹ There is fairly strong evidence that survey expectations are reliable and have useful information. Ang, Bekaert, and Wei (2007) found that the median inflation expectation is the best available forecaster of inflation beating all econometric alternatives. Inoue, Kilian, and Kiraz (2009) confirmed that the surveys are backed by actions, by finding that household consumption growth responds to their perceived real interest rate, using their reported expectation of inflation, and this is stronger the higher the education of the household.

is indeed positively related with recent changes in inflation and output, and Coibion and Gorodnichenko (2008) found a positive relation between disagreement and oil price shocks, but were unable to statistically pin down the sign of the relationship between disagreement and other measures of shocks. Finally, Branch (2007) found that the sticky-information model can match many features of the distribution of inflation expectations in the Michigan survey.

Carroll (2003) took a different approach that emphasized the distinction between professional forecasters (in the Survey of Professional Forecasters) and households (in the Michigan survey). He assumed that professionals have close to perfect information, while households have very sticky information. He found that, just as the sticky-information model predicts, the expectations of households gradually converge to the expectations of professionals.³⁰

Mankiw, Reis, and Wolfers (2003) noted that the model's predictions are broadly consistent with the U.S. experience in the first half of the 1980s, following the Volcker disinflation. As monetary policy contracted, inflation and output fell, while disagreement increased substantially. Moreover, as shown in Figure 6, disagreement in the data moved in striking agreement with the model. It is noticeable that the distribution of inflation expectations went from its usual bell shape to a bimodal distribution for a little over a year, as some people seemed to have updated their expectations while others had not.³¹

Taking the unconditional expectation of D_t in Eq. (49), we obtain a prediction for the average amount of disagreement. For our baseline parameters, $\alpha = 0.2$, $\lambda = 0.25$, $\sigma = 0.01$, predicted disagreement is 0.5%. This predicted value is well below the disagreement we observe in the data, but this discrepancy may be expected for at least two reasons. First, D_t in Eq. (51) ignores other sources of shocks, and in particular, aggregate productivity shocks. Second, there is more heterogeneity in the real world than just the differences in information sets that the model emphasizes.

Other empirical work using survey data has typically been supportive of imperfect information models more generally. In particular, Curtin (2009) added new questions to the Michigan survey asking people about their knowledge of current inflation. He found that knowledge of the present is as imperfect as forecasts of the future — a feature of the world that is perhaps the very essence of imperfect information models.

Moving forward, imperfect information models face the difficulty that sometimes slight changes in the information structure can change their predictions significantly. Berkelmans (2009) and Hellwig and Venkateswaran (2009) introduced multiple shocks

³⁰ Döpke, Dovern, Fritsche, and Slacalek (2008b) confirmed Carroll's findings for France, Germany, Italy, and the UK.

³¹ Dovern, Fritsche, and Slacalek (2009) found that in the G7, countries with more independent central banks have less disagreement about inflation and nominal variables.

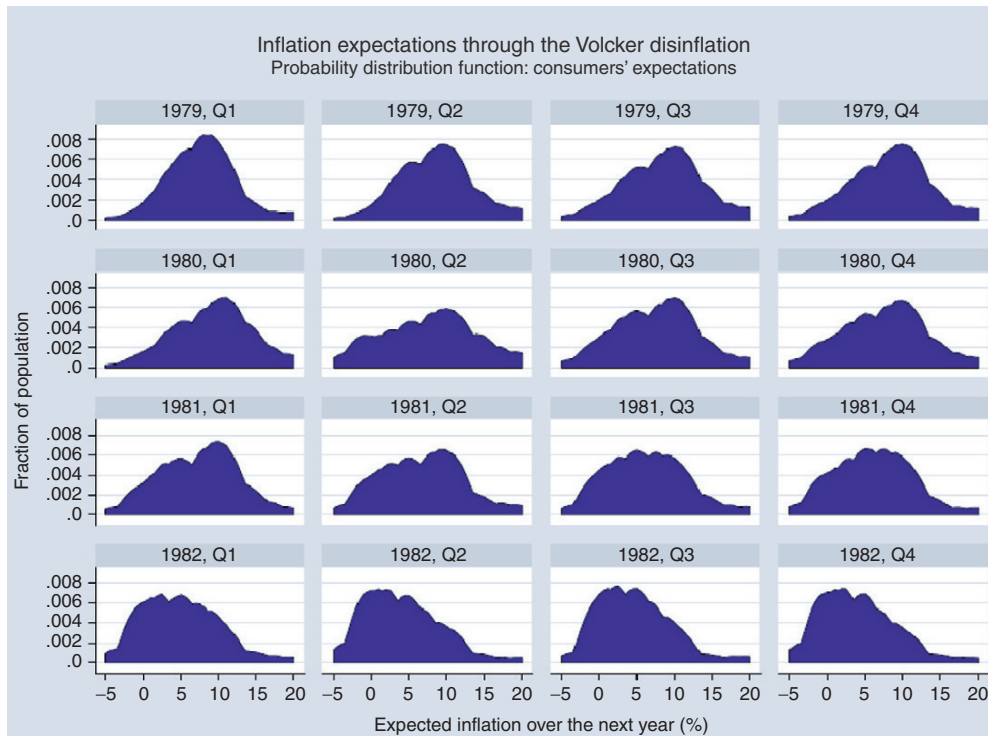


Figure 6 Disagreement during the Volcker years. (From Mankiw, Reis, & Wolfers, 2003, with permission.)

in partial information models, similar to the one in Section 4.2, and found that the impulse responses of inflation to demand shocks can be quite different depending on which combinations of other shocks were allowed. Another example is the different predictions for the choice of portfolios by investors with rational inattention reached by van Nieuwerburgh and Veldkamp (forthcoming) and Mondria (forthcoming) from small differences in the specification of the available signals. One way out of this problem is to use data that directly disciplines the modeling of information. There is a wealth of data asking people about their expectations, and using these data in novel ways offers, in our view, the biggest promise in empirical work on imperfect information models in the near future.

Moreover, most of the work previously described tries to explain the data on expectations using the data on aggregate variables. There is much less work attempting to explain macroeconomic variables using expectations data. We expect that this will be a fruitful topic of research in the years to come.

5.2 Partial information and optimal transparency

A classic issue is the role of transparency in monetary policy. Typically, economists have argued that more clarity on the part of central banks is desirable. As Morris and Shin (2002) have emphasized, partial information models provide some novel insights to study the optimal degree of transparency.

Within the 1-period partial information model of Section 4.1, assume that beyond the private signal z_{it} , there is also a public signal $m_t = n_t + v_t$, where v_t is normal, has mean zero, and variance σ^2/ω . One interpretation of this public signal is that it is a policy announcement by the central bank. The parameter ω measures the precision of the public signal. If the authority is maximally transparent, then $\omega \rightarrow \infty$, whereas a completely opaque central bank makes no announcements, which corresponds to $\omega = 0$.

Given its two signals, the firm's optimal forecast now is

$$\hat{E}_{it}(n_t) = E_{t-1}(n_t) + \left(\frac{\tau}{1 + \tau + \omega} \right) [z_{it} - E_{t-1}(n_t)] + \left(\frac{\omega}{1 + \tau + \omega} \right) [m_t - E_{t-1}(n_t)]. \quad (52)$$

Averaging over this and iterating on the expectations as before gives the solution:

$$p_t = \left(\frac{\alpha\tau + \omega}{1 + \alpha\tau + \omega} \right) [n_t - E_{t-1}(n_t)] + \left(\frac{\omega}{1 + \alpha\tau + \omega} \right) v_t + E_{t-1}(n_t), \quad (53)$$

$$\gamma_t = \left(\frac{1}{1 + \alpha\tau + \omega} \right) [n_t - E_{t-1}(n_t)] - \left(\frac{\omega}{1 + \alpha\tau + \omega} \right) v_t. \quad (54)$$

Once again, when the aggregate supply curve is nonvertical and flatter the stronger the real and informational rigidities. The public signal has two effects. First, as with the private signal, the more precise the public signal, the steeper the aggregate supply curve. Second, shocks to this common information now generate fluctuations in prices and output. In particular, if the central bank's announcement misleads firms into believing aggregate demand is higher than it actually is ($n > 0$), they will raise prices and output will fall.

Imperfection of information creates welfare losses relative to the first best through two channels. First, because output would be constant with full information, any output variability is costly to the risk-averse consumer. Second, because all firms are identical, any price dispersion reflects a misallocation of resources. Using the equilibrium solution, these two measures are

$$Var_{t-1}(\gamma) \equiv E_{t-1}(\gamma^2) = \frac{(1 + \omega)\sigma^2}{(1 + \omega + \alpha\tau)^2}, \quad (55)$$

$$Var_i(p_{it}) \equiv \int_0^1 (p_{it} - p_t)^2 di = \frac{\alpha^2 \tau \sigma^2}{(1 + \omega + \alpha\tau)^2}. \quad (56)$$

The first best is achieved here with maximal transparency, which occurs as $\omega \rightarrow \infty$. In this case, there is complete information. This case, however, is arguably of limited relevance, as the central bank can never be completely clear or completely certain that all agents in the economy will perfectly process the information it provides.

The more relevant question is whether, at the margin, increased transparency is good or bad. An improvement in transparency (higher ω) unambiguously lowers the cross-sectional dispersion of prices. As firms have more precise common information, they coordinate more. However, more transparency has an ambiguous effect on output volatility. The reason is that with a more precise public signal, firms on the one hand decide to rely less on their private signals, undermining the information they reveal, and on the other hand are now exposed to fluctuations because of the public signal mistakes. Because of strategic complementarities, each firm would like the other firms to respond more to their private signals than they do, as this aggregates and reveals information. Increased transparency may exacerbate the inefficient use of information by firms and could potentially reduce welfare.

Depending on the relative weight of output and price stabilization in the policymaker's objective function, there may be a range of ω between 0 and some positive value, where raising ω actually lowers welfare. While complete transparency is the global optimum, if there is an upper bound on the precision of the public signals, it may be best to be less transparent than this upper bound. By picking different parameters, Morris and Shin (2002) argued this case is likely, while Svensson (2006) argued it was not.

Within the context of the specific aggregate supply model that we consider, Roca (2006) provided the unambiguous answer for all parameter values. He posited that a natural utilitarian measure of social welfare is Woodford's (2003) second-order approximation of the utility of the representative agent. In this case, the relative weight on the cross-sectional dispersion of prices vis-à-vis the variance of output is equal to γ/α . Because the elasticity of substitution across varieties is positive, $\gamma > 1$, simple algebra shows that this condition is sufficient for welfare to increase with higher transparency.

Outside of this particular model of aggregate supply, Angeletos and Pavan (2007) provided a general characterization of the inefficiency in using information, and a set of conditions for transparency to increase or decrease welfare. Amador and Weill (2008) recovered the Morris and Shin (2002) result that transparency may be harmful by assuming that agents must distinguish between productivity and monetary shocks and use the distribution of prices in the economy to learn. Reis (2010) studied the optimal timing for releasing information by policymakers, asking how far in advance (if at all) should changes in policy be announced.

One conclusion from this literature seems robust: Increased transparency may reduce the incentive for people to rely on and thus reveal private information. The effects of this behavior on welfare, however, are more ambiguous and may depend

on the particulars of the model. This literature has already succeeded in showing that the case for transparency is not as clear cut as it may have seemed just a decade ago. The hope is that future work using these tools and insights may lead to a better understanding of how authorities may wish to communicate with the public, a long-standing question in economics.

6. MICROFOUNDATIONS OF INCOMPLETE INFORMATION

So far we have discussed two models of aggregate supply built on the assumption of imperfect information, but we have not addressed a more foundational question: Why is information imperfect in the first place? The theory of “inattentiveness” proposed by Reis (2006a,b) has been used to justify delayed information, while the theory of “rational inattention” proposed by Sims (2003) has been used to justify why firms would have partial information.

6.1 Inattentiveness

For a firm to set a price reflecting the current state of the world requires incurring at least three costs. There is a cost of acquiring information, in the sense of obtaining all of the relevant bits of data that are informative. There is a cost of absorbing information, in the sense of interpreting all of this information and translating it into the sufficient statistics for the price decision. And finally, there is a cost of processing information in the sense of computing the map from the sufficient statistics to the optimal action on prices. The cost of acquiring information may be small, and may have fallen in this information age, but the costs of absorbing and processing information may be large and arguably higher today than in the past. In Reis (2006a), these various costs were modeled as a fixed cost that the firm has to pay whenever it wants to acquire information and become attentive. If it does not pay the cost, the firm remains inattentive, following a predetermined plan that may not be best for the current circumstances.³²

Letting the costs of planning be denoted by the fixed amount k , and the value of a firm at date t that has just obtained information on the random-walk demand shock by $V(n_t)$ then the Bellman equation for this problem is

$$V_i(n_t) = \max_d E_t \left\{ \sum_{s=0}^{d-1} \beta^s \max_{p_{i,t+s}} [X_i(p_{i,t+s}, \cdot)] - \beta^d k + \beta^d V_i(n_{t+d}) \right\}, \quad (57)$$

where d is the number of time periods between information acquisition. The solution will be a function $d(n_t)$, so that while price adjustment is time-dependent, in that it does

³² The assumption that attention is an all-or-nothing affair is extreme, but it could be relaxed. For instance, the model could be extended to allow firms to observe some information when inattentive.

not depend on the state of the world at the date of the adjustment, it is recursively state-dependent, since it depends on the state of the world at the last adjustment date.³³ In principle, this result should make it possible to distinguish between this model of inattentiveness and partial information models. Testing whether the fraction of firms adjusting their plans today does not depend on news today would test the inattentiveness model. However, because the fraction of adjusters depends on the past state of the world, and since most relevant variables are very persistent, in practice these tests will have little power.

This problem can be solved numerically, but to obtain an analytic solution, we make three simplifications. First, we work with a quadratic log-approximation of the profit function. This is the certainty equivalent approximation, and it implies that $X(p_{i,t+s}) = -\Xi(p_{i,t+s} - p_{i,t+s}^*)^2$ where Ξ is a scalar, and that the inner maximization has the solution $p_{i,t+s} = E_i(p_{i,t+s}^*)$. Second, we ignore the fact that d must be an integer and proceed to take derivatives and solve equations as if d could be a number in the real line. This approximation is not too damning; using quantum calculus, we could dispense with this assumption and obtain similar results. Third, we ignore strategic considerations by focusing on the $\alpha = 1$ case.³⁴

The first step to solving the problem is to realize that, using these assumptions:

$$E_t \left[\sum_{s=0}^{d-1} \beta^s \max_{p_{i,t+s}} [X_i(p_{i,t+s}, \cdot)] \right] = \sigma^2 \Xi \left[\frac{1 - d\beta^{d-1} + (d-1)\beta^d}{(1-\beta)^2} \right]. \quad (58)$$

Expected profits (at the time when the firm is making its pricing plan) do not depend on the state of aggregate demand. Thus, under these special conditions, the value function is a constant, and the optimal inattentiveness does not depend on n_t . It then follows from the problem in Eq. (57) that the necessary optimality conditions imply that d^* maximizes:

$$\frac{-\sigma^2 \Xi [1 - d\beta^{d-1} + (d-1)\beta^d] - (1-\beta)^2 \beta^d k}{1 - \beta^d}. \quad (59)$$

Using the implicit function theorem, it is straightforward to show that there is a unique positive d^* that is of order $\sqrt{k}$, and that it increases with k and falls with σ^2 . Therefore, inattentiveness is first-order long with second-order costs of planning, increases the more costly it is to plan, and falls as the world becomes more volatile.

³³ The issue of time versus state dependence is important, because the latter comes with a selection effect that greatly reduces the real impact of nominal shocks (Caballero & Engel, 2007; Golosov & Lucas, 2007).

³⁴ Reis (2006a) provided two alternatives: an alternative case with an exact analytical solution by setting the problem in continuous time and assuming an isoelastic profit function, and a general approximate solution to the problem using perturbation theory.

The assumption of the sticky-information model that information arrives as a Poisson process, implying an exponential distribution of uninformed price-setters, is harder to justify. Reis (2006a) provided some conditions under which it holds, but they are quite strict.³⁵ Carroll (2006) proposed an alternative, arguing that information spreads like a virus in a population with the rate of arrival of information λ being the analogous of an infection contact rate. However, this idea has not yet been formalized.

6.2 Rational inattention

Chapter 4 by Sims in this *Handbook* reviews rational inattention theory in detail, so here we limit ourselves to its link to partial information models of aggregate supply. We start with a brief introduction to the two key concepts of rational inattention.

The first concept is entropy. For a variable n_t in the real line with probability density function $f(n_t)$, its entropy is

$$H(n_t) = - \int f(n_t) \ln(f(n_t)) dn_t. \quad (60)$$

Entropy is analogous to variance in that it measures uncertainty, is non-negative, and equals zero if n_t is certain.³⁶ The second concept is mutual information, defined as:

$$I(n_t; z_{it}) = H(n_t) - H(n_t | z_{it}). \quad (61)$$

The information that the signal z_{it} has on the variable n_t is therefore the reduction in entropy that results from having the conditional distribution of n_t on z_{it} instead of the unconditional distribution of n_t .

The rational inattention problem for a price-setting firm consists of picking the signals to maximize profits subject to the constraint on the amount of information it processes:

$$\max_{f(n_t | z_{it})} [\max_{p_{it}} X(p_{it}, n_t)] \text{ subject to } : I(n_t; z_{it}) \leq k. \quad (62)$$

While this seems like a standard constrained maximization, several features make the problem unique. First, note that the choice variable is a conditional probability

³⁵ Dupor and Tsuruga (2005) examined the predictions of a sticky-information model in which all firms are inattentive for the same amount of time N and are perfectly staggered in their adjustment dates, so the distribution of inattentiveness is uniform. It turns out that the comparison between this model and the more standard model depends on how the two models are calibrated. If the mean duration of inattentiveness at the time of adjustment is the same for the two models ($N = 1/\lambda$), then demand shocks are less persistent with a uniform distribution than with an exponential. But if, instead, the average age of plans within the economy at any moment is set to be the same ($0.5(N + 1) = 1/\lambda$), then the two models yield similar dynamics. Dixon and Kara (2006) argued that the latter is the better calibration.

³⁶ Entropy has some appealing properties including its link to the notion of information, data compression, and descriptive complexity (Cover & Thomas, 1991), although it has been strongly criticized as a measure of risk (Aumann & Serrano, 2008).

density function, not a scalar. Another way of stating this is that the signal is $z_{it} = n_t + \varepsilon_{it}$, and we are choosing the distribution of ε_{it} . Second, nothing in the structure of the problem guarantees that the solution is a known distribution or even that it has a smooth density.³⁷ Third, the constraint is that there is a fixed finite capacity k , so the firm is unable to expend resources in obtaining more capacity (e.g., more managers or consultants) even if the benefits from doing so were very large. Fourth, note that this is not an intertemporal problem (unlike in the inattentiveness theory) because it is assumed that the firm cannot trade capacity over time. In the theory of rational inattention, it is assumed that agents can only observe some signals of the world every period, but cannot choose to pay more attention at certain times.³⁸

Because this is a hard problem, three approaches have been followed in the literature to solve it. One approach is to solve the problem numerically (Sims, 2006). This work is still in its infancy as it seems that the needed numerical tools are not in the standard economist's toolkit.³⁹

A second approach is to constrain the set of admissible signal distributions to known distributions (van Nieuwerburgh and Veldkamp, forthcoming, Mondria, forthcoming). In particular, it is often assumed that the signals must be normally distributed, since then the functional problem reduces to choosing a single parameter, the variance of the noise.⁴⁰ In particular, using the definition of mutual information in Eqs. (60) and (61) and the density of the normal distribution, a few steps of algebra show that the information constraint becomes a constraint on the precision of the signal:

$$0.5 \ln(1 + \tau) \leq k. \quad (63)$$

Because more precise signals raise expected profits, it is clear that this constraint will always bind at the optimum. Therefore, expression (63) holds as an equality, and it gives the optimal precision of signals τ as a monotonic function of information capacity k . Firms with higher capacity have more precise signals.

A third approach is to solve for the optimal distribution for some special cases of the profit function. One natural and simple case is when the profit function is quadratic, $X(n_{t+s}) = -\Xi(p_{it+s} - p_{it+s}^*)^2$, and nominal income is normally distributed. In this case, one can show that the optimal distribution function for the errors is the normal distribution. This is the only case where the exact analytic solution is known.

³⁷ In fact, Matejka (2008) found that rational inattention problems typically have discontinuous solutions with point-mass distributions.

³⁸ An exception is Moscarini (2004), who formulated a rational inattention problem in terms of choosing the discrete dates at which to observe continuously arriving information.

³⁹ Recent developments in the numerical solution of rational inattention models are in Matejka (2008), Lewis (2009), and Tutino (2009).

⁴⁰ It is an elementary result that the signal will be unbiased since changes in the mean have no effect on entropy and the firm would not benefit from any such bias.

7. THE RESEARCH FRONTIER

There has been much other recent work on imperfect information with implications for aggregate supply and the effect of aggregate demand. We review some of this work in this section.

7.1 Merging incomplete information and sticky prices

When one looks at the price path for many goods, three features stand out.⁴¹ First, prices change all the time, on average every three to four months. Second, many of these changes follow what seem like predetermined patterns that simple algorithms can spot; the actual resetting of price plans reflecting new information seems to occur less often than once a year.⁴² Third, in a plot of prices over time, there are many horizontal segments, reflecting short-lived intervals when nominal prices are unchanged.

The first two features match the predictions of imperfect information models, and sticky-information models in particular. The prevalence of what some researchers call predictable “sales” are precisely the price plans in these models and, as found by Klenow and Willis (2007), these plans seem to only slowly incorporate available information. The third feature is puzzling to these models, because there is no reason why the predetermined plan would involve the exact same price over an interval of time. There are some attempts at explaining the prevalence of these prices using imperfect information, but a more common answer has been the presence of physical costs of changing prices in addition to the information costs, leading to sticky prices.⁴³

Bonomo and Carvalho (2004, forthcoming) assumed that the cost of changing price plans included both an information cost and a physical price-adjustment cost. Thus, when firms update their information, they are constrained to pick a plan where a single price is chosen, unlike the time-varying plans in the sticky-information model. In a stationary environment, the result is the Calvo model of price adjustment, derived here as a special case of sticky information. The advantage of this information-interpretation of the Calvo model is that it leads to an endogenous choice of the frequency of price adjustment, along similar lines to the inattentiveness theory in Section 6.1.

⁴¹ See Chapter 6 in this *Handbook* (Klenow & Malin, 2010), and the recent work of Eichenbaum, Jaimovich, and Rebelo (2008).

⁴² The once-a-year adjustment matches the survey responses in Blinder, Canetti, Lebow, and Rudd (1998), suggesting that perhaps firm managers were responding to how often they adjusted their price plans, rather than the actual prices. This is plausible since many of the predetermined changes look like sales.

⁴³ Matejka (2008) showed that the optimal distribution of signals from a particular rational inattention problem has point masses so that a discrete set of signals and prices are chosen. Bergen, Chen, Levy, and Ray (2008) documented that price increases tend to be small while declines tend to be large, and after ruling out other explanations concluded in favor of information-based theories. Knotek (2008) found that “convenient” prices are more likely in locations where transactions must be made quickly.

Another approach is taken by Dupor, Kitamura, and Tsuruga (2010), who merge sticky information with the Calvo model of price adjustment. They assume that every period, each firm has a random chance of updating its information (as in the sticky-information model), while an independent random event determines whether the firm can reset its price (as in the Calvo model). They find that this model empirically dominates the hybrid Phillips curve of Gali and Gertler (1999) and others.

Others have merged partial information with the Calvo model. See, in particular, Morris and Shin (2006), Nimark (2008), and Angeletos and La'O (2009a). Nimark's (2008) results were similar to Dupor and Kitamura, while Morris and Shin (2006) and Angeletos and La'O (2009a) focused instead on the inertia of forward-looking expectations and the dynamics of higher order beliefs.

Another branch of work has merged imperfect information with fixed costs of changing prices as in state-dependent pricing models. Knotek (2006) did this for the sticky-information model. He found that the model fit well the micro facts from the price data, while keeping most of the predictions for aggregate supply as in the sticky-information model. Gorodnichenko (2008) examined a state-dependent model with partial information. He emphasized the positive externality from a price change: when a firm chooses to adjust its price, it releases some of its private information to other firms.

Another merger of these various models has been proposed by Woodford (2009). He assumed that firms can pay a fixed information cost at discrete times to perform a price review, and when they do so they obtain full information about the state of the economy at that moment, just as in delayed information models. At the same time, he assumed that between these adjustment dates, firms obtain signals as in partial information models. The cost of an information update is fixed similar to the theory of inattentiveness, while the informativeness of the signals is determined by a limited-capacity channel like the theory of rational inattention. Under the extra assumption that the calendar date is also a costly piece of information, so the price plan must consist of a single number, Woodford (2009) showed that this model generalizes the state-dependent pricing model. In the limit where the channel capacity is infinite, the model is exactly like a conventional state-dependent pricing model while when the channel capacity is zero the model becomes isomorphic to the Calvo model. For intermediate levels the model reproduces the generalized S_s model of Caballero and Engel (1999).

7.2 Heterogeneity in the frequency of information adjustment

Haltiwanger and Waldman (1989) studied the properties of equilibrium in models where some agents are informed, and so respond to shocks, while others are not. They showed that with strong strategic complementarity, the nonresponders have a disproportionate effect on the equilibrium. Intuitively, the firms that obtain information want

their prices to stay close to the those that are not adjusting, so the equilibrium ends up mimicking the lack of information of the uninformed. This may be clearer in the limit: as $\alpha \rightarrow 0$, firms want their price to equal the aggregate price level, so even if only a small fraction of firms do not have information on current shocks, the equilibrium will involve no firm responding to shocks at all.

Carvalho and Schwartzman (2008) proposed a sticky-information model with many sectors, where the frequency of information adjustment is different across sectors. Their important finding was that demand shocks are much more persistent in this economy than in an equivalent single-sector economy with the average frequency of information adjustment. Because of strategic complementarities, the sector that adjusts less often has a disproportionate effect on the aggregate dynamics since the other sectors want to keep their prices close to theirs.⁴⁴

7.3 Optimal policy with imperfect information

Ball, Mankiw, and Reis (2005) studied optimal monetary policy in a simple sticky-information economy. They show that price level targeting is better than inflation targeting. Because firms choose plans for prices and want to minimize their forecast errors well into the future, price level targeting dominates inflation targeting. That is, base drift is quite costly. The optimal policy is an elastic price standard: there is a deterministic target for the price level and the central bank deviates from it when output is expected to deviate from its full-information level.

Jinnai (2006) and Branch, Carlson, Evans, and McGough (2009) examined how policy choices affect the optimal frequency of information updating. The latter showed that if the central bank becomes more concerned with inflation relative to output, the firm's forecasting problem becomes easier. It therefore ends up lowering the variance of output together with that of inflation. This mechanism may partially explain the "Great Moderation," and suggests a fruitful avenue for future research to test models of inattentiveness using historical changes in the volatility of inflation and the business cycle

Reis (2009a) characterized optimal policy rules in an estimated medium-scale model with pervasive sticky information. Relative to models with rigidity in agents' actions such as habits by consumers, sticky prices by firms, sticky wages by workers, and adjustment costs by investors, sticky information leads to a larger focus on stabilizing real activity. This is true both in terms of the optimal variance of output relative to inflation as well as in terms of the optimal policy-rule coefficients.

Adam (2007) studied optimal monetary policy in a simple partial information economy similar to the one we presented in Section 4.2. He showed that in response to persistent shocks, policy should stabilize the output gap in the short run, focusing on

⁴⁴ Carvalho (2006) made the same point in Calvo sticky-price models, and Nakamura and Steinsson (forthcoming) discussed the interaction between heterogeneity and strategic complementarity in a menu cost model.

stabilizing the price level only in the medium run. Adam (2009) showed that with partial information, discretionary policy can be much more costly relative to commitment than with full information. He also confirmed the Branch et al. (2009) result described earlier in partial information economies: an increased focus on price stability may lower the variance of both inflation and output. Lorenzoni (2010) extends the analysis of optimal monetary policy to a setting where all price-setters have a common signal on productivity (similar to the policy announcement in Section 6.2).

Finally, Angeletos and Pavan (2007, 2009) provided a more general, but also more abstract, characterization of efficiency and optimal policy with incomplete information. They focused on the externalities that one agent's use of information imposes on others. Angeletos and La'O (2008) characterized optimal fiscal and monetary policy over the business cycle in a partial information economy.

7.4 Other choices with imperfect information

The resurgence of work on imperfect information models has not been constrained to the study of pricing decisions by firms. At the same time, an equally large literature has sprung up using very similar ideas and often the same authors, but applied to different questions in economics.⁴⁵

Mankiw and Reis (2003) and Koenig (2004) considered sticky information on the part of workers setting wages to explain unemployment. Carroll, Slacalek, and Sommer (2008) and Luo (2008) focused on consumption choices with sticky and partial information, respectively. Reis (2006b) investigated the inattentiveness model for consumers, while Tutino (2009) and Lewis (2007) extended the rational inattention model to deal with the dynamic decisions of consumers. Angeletos and Pavan (2004) considered physical investment decisions.

A fruitful line of work has applied the inattentiveness model to portfolio choice. Gabaix and Laibson (2001) emphasized the potential for delayed information to explain the equity premium. Building on Duffie and Sun (1990) and Reis (2006b), Abel, Eberly, and Panageas (2007) provided a micro-founded inattentiveness model of delayed adjustment and characterized its implications for portfolio choice and asset prices. Abel, Eberly, and Panageas (2009) combined delayed information with transaction costs and showed a remarkable result: the behavior of the consumer converges to time-dependent adjustments with constant intervals of inattentiveness, as if the transaction costs were not present. Huang and Liu (2007) studied portfolio choice with rational inattention.

In an important contribution, Lorenzoni (2009) shifted the focus of imperfect information from the demand to the supply shock. He showed that a common signal about productivity can generate business cycles that resemble those due to demand shocks.

⁴⁵ There is also an active parallel work in finance, surveyed in the book by Veldkamp (2009).

Angeletos and La'O (2009b) considered partial information about shocks on tastes, productivity, and desired markups. Finally, La'O (2009) applied the partial information model to financial contracting.

Finally, in the open economy literature, Bacchetta and van Wincoop (2006) considered a simple partial information model for traders in currency markets and showed this could explain some of the puzzling disconnect between exchange rates and fundamentals. Crucini, Shintani, and Tsuruga (2008) used instead a delayed information model and showed it can explain volatile and persistent real exchange rate movements both in the aggregate and at the sectoral level. Bacchetta and van Wincoop (2010) found that a delayed information model can explain the forward discount puzzle.

7.5 DSGE models with imperfect information

Dynamic stochastic general equilibrium modeling, surveyed by Christiano, Trabandt, and Walentin in Chapter 7 in this *Handbook*, has been an active area of intersection between academic and central-bank researchers. The first DSGE models with imperfect information have recently appeared, and this is likely an area of much future work.

In a series of papers, Mankiw and Reis (2006, 2007) and Reis (2009a,b) put forward a first DSGE model with sticky information in all markets.⁴⁶ In their model, firms when setting prices, households when choosing consumption, and workers when setting reservation wages are all allowed to be inattentive, and estimates using both Euro Area and U.S. data show that sticky information is pervasive across all of these markets. Their work also contributed algorithms to solve medium- to large-scale models with sticky information, and to evaluate likelihood functions.⁴⁷

Mackowiak and Wiederholt (2010) proposed a DSGE model with partial information. They showed that the utility and profit losses from inattentive behavior are small even though the aggregate dynamics are significantly different than the full-information alternative. Moreover, by allowing for different shocks and different signals, as explained in Section 4.3, they found that these individual losses are significantly smaller than those in standard sticky-price models.

The previous models still involve some simplifications to make the information heterogeneity manageable. In particular, it is often difficult to define equilibrium in markets where both sellers and buyers are inattentive. This is an active area of work.⁴⁸

⁴⁶ There had been some previous attempts, by Trabandt (2004), Andres, Nelson, and Lopez-Salido (2005), Kiley (2007), Laforge (2007), and Korenok and Swanson (2005, 2007) with sticky information only on the part of firms. Mankiw and Reis (2006) criticized this work and argued that information stickiness should be pervasive across all markets, both on grounds of methodological consistency and, more important, because such pervasive stickiness empirically helps to fit the U.S. data.

⁴⁷ Meyer-Gohde (2010) improved on these algorithms significantly, and his publicly available programs make the solution and estimation of sticky-information models as easy as conventional rational-expectations models.

⁴⁸ Reis (2009b) discussed the existing open questions on micro-founding sticky information in general equilibrium.

8. CONCLUSION

Since the birth of business cycle theory, economists have struggled with one overarching question: What is the nature of the market imperfection, if any, that causes the economy to deviate in the short run from full employment and the optimal allocation of resources? Or, to put the question more concretely and more prosaically in terms of undergraduate macroeconomics: What friction causes the short-run aggregate supply curve to be upward sloping rather than vertical, giving a role to aggregate demand in explaining economic fluctuations? The theme of the literature surveyed here is that the answer is to be found in the natural uncertainty of economic conditions coupled with peoples' inherent limitations in obtaining and processing information.

The models described here build on much of traditional macroeconomics. In his 1936 classic *The General Theory*, John Maynard Keynes emphasized vast uncertainty as a key fact of economic life; his famous “beauty contest” parable relates closely to the common-knowledge problem we described earlier. Similarly, in his 1968 AEA presidential address, Milton Friedman stressed the failure of some agents to correctly perceive monetary conditions as an explanation for the short-run Phillips curve — a theme that pervades the models surveyed in this chapter.

These models are also tied to more recent themes in macroeconomic research. The models examined here are all solved using mathematical tools that economists developed during the rational expectations revolution of the 1970s. But in contrast to early rational expectation theory, these models typically assume agents make decisions based on a much more limited set of information. This assumption of restricted information has been made more palatable in recent years by the growth of behavioral economics, which has stressed imperfections in human cognition.

Despite building on a long tradition, models on imperfect information and aggregate supply are still in their infancy. Without a doubt, much progress has been made in recent years, and we hope this chapter has given readers a taste of this research and some leads about where to learn more. This line of work still offers many attractive open questions concerning macro theory, empirics, and policy. We expect it to remain a fruitful area of research in the years to come.

REFERENCES

- Abel, A., Eberly, J., Panageas, S., 2007. Optimal inattention to the stock market. *Am. Econ. Rev.* 97 (2), 244–249.
- Abel, A., Eberly, J., Panageas, S., 2009. Optimal inattention to the stock market with information costs and transactions costs. NBER Working Paper 15010.
- Adam, K., 2007. Optimal monetary policy with imperfect common knowledge. *J. Monetary Econ.* 54 (2), 276–301.
- Adam, K., 2009. Monetary policy and aggregate volatility. *J. Monetary Econ.* 56 (S1), S1–S18.
- Akerlof, G.A., 2002. Behavioral macroeconomics and macroeconomic behavior. *Am. Econ. Rev.* 92 (3), 411–433.
- Akerlof, G.A., Yellen, J.L., 1985. A near-rational model of the business cycle, with wage and price inertia. *Quarterly Journal of Economics* 100 (5), 823–838.
- Amador, M., Weill, P.O., 2008. Learning from private and public observations of others' actions. Manuscript.
- Amato, J.D., Shin, H.S., 2006. Imperfect common knowledge and the information value of prices. *Econ. Theory* 27, 213–241.
- Andrés, J., Nelson, E., López-Salido, D., 2005. Sticky-price models and the natural rate hypothesis. *J. Monetary Econ.* 52 (5), 1025–1053.
- Ang, A., Bekaert, G., Wei, M., 2007. Do macro variables, asset markets or surveys forecast inflation better? *J. Monetary Econ.* 54, 1121–1163.
- Angeletos, G.M., La'O, J., 2008. Dispersed information over the business cycle: Optimal fiscal and monetary policy. Manuscript.
- Angeletos, G.M., La'O, J., 2009a. Incomplete information, higher-order beliefs, and price inertia. *J. Monetary Econ.* 56 (S1), S19–S37.
- Angeletos, G.M., La'O, J., 2009b. Noisy business cycles. NBER Macroeconomics Annual 24, 319–378.
- Angeletos, G.M., Pavan, A., 2004. Transparency of information and coordination in economies with investment complementarities. *Am. Econ. Rev.* 94 (2), 91–98.
- Angeletos, G.M., Pavan, A., 2007. Efficient use of information and social value of information. *Econometrica* 75 (4), 1103–1142.
- Angeletos, G.M., Pavan, A., 2009. Policy with dispersed information. *Journal of the European Economic Association* 7 (1), 11–60.
- Angeletos, G.M., Werning, I., 2006. Crises and prices: Information aggregation, multiplicity and volatility. *Am. Econ. Rev.* 96 (5), 1720–1736.
- Aumann, R.J., Serrano, R., 2008. An economic index of riskiness. *J. Polit. Econ.* 116 (5), 810–836.
- Bacchetta, P., van Wincoop, E., 2006. Can information heterogeneity explain the exchange rate determination puzzle? *Am. Econ. Rev.* 96 (3), 552–576.
- Bacchetta, P., van Wincoop, E., 2010. Infrequent portfolio decisions: A solution to the forward discount puzzle. *Am. Econ. Rev.* 100 (3), 870–904.
- Ball, L., 1994. Credible Disinflation with Staggered Price Setting. *Am. Econ. Rev.*, 84, 282–289.
- Ball, L., Mankiw, N.G., Reis, R., 2005. Monetary policy for inattentive economies. *J. Monetary Econ.* 52 (4), 703–725.

- Ball, L., Romer, D., 1989. The equilibrium and optimal timing of price changes. *Review of Economic Studies* 56 (2), 179–198.
- Ball, L., Romer, D., 1990. Real rigidities and the non-neutrality of money. *Review of Economic Studies* 57 (2), 183–203.
- Barro, R.J., 1977. Unanticipated money growth and unemployment in the United States. *Am. Econ. Rev.* 67 (2), 101–115.
- Basu, S., Fernald, J.G., 1997. Returns to scale in U.S. production: Estimates and implications. *J. Polit. Econ.* 105 (2), 249–283.
- Bergen, M., Chen, A., Ray, S., Levy, D., 2008. Asymmetric price adjustment in the small. *J. Monetary Econ.* 55, 728–737.
- Berkelmans, L., 2009. Imperfect information and monetary models: Multiple shocks and their consequences. Manuscript.
- Blanchard, O., Kiyotaki, N., 1987. Monopolistic competition and the effects of aggregate demand. *Am. Econ. Rev.* 77 (4), 647–666.
- Blinder, A.S., Canetti, E., Lebow, D., Rudd, J., 1998. *Asking about prices: a new approach to understanding price stickiness*. Russell Sage Foundation, New York.
- Bonomo, M., Carvalho, C., 2004. Endogenous time-dependent rules and inflation inertia. *Journal of Money, Credit and Banking* 36 (6), 1015–1041.
- Bonomo, M., Carvalho, C., Imperfectly-credible disinflation under endogenous time-dependent pricing. *Journal of Money, Credit and Banking*. (in press).
- Branch, W., 2007. Sticky information and model uncertainty in survey data on inflation expectations. *Journal of Economic Dynamics and Control* 31 (1), 245–276.
- Branch, W.A., Carlson, J., Evans, G., McGough, B., 2009. Monetary policy, endogenous inattention, and the volatility trade-off. *Econ. J.* 119, 123–157.
- Broda, C., Weinstein, D.E., 2006. Globalization and the gains from variety. *Quarterly Journal of Economics* 121 (4), 541–585.
- Caballero, R.J., Engel, E.M.R.A., 1999. Explaining investment dynamics in U.S. manufacturing: A generalized (S,s) approach. *Econometrica* 67 (4), 783–826.
- Caballero, R.J., Engel, E.M.R.A., 2007. Price stickiness in Ss models: New interpretations of old results. *J. Monetary Econ.* 54 (S1), 100–121.
- Calvo, G., 1983. Staggered prices in a utility-maximizing framework. *J. Monetary Econ.* 12, 383–398.
- Carroll, C.D., 2003. Macroeconomic expectations of households and professional forecasters. *Quarterly Journal of Economics* 118 (1), 269–298.
- Carroll, C.D., Slacalek, J., 2007. Sticky expectations and consumption dynamics. Johns Hopkins University Manuscript.
- Carroll, C.D., Slacalek, J., Sommer, M., 2008. International evidence on sticky consumption growth. NBER Working Paper 13876.
- Carvalho, C.V., 2006. Heterogeneity in price stickiness and the real effects of monetary shocks. *B.E. Journals: Frontiers of Macroeconomics* 2 (1).
- Carvalho, C.V., Schwartzman, F., 2008. Heterogeneous price-setting behavior and aggregate dynamics: Some general results. Manuscript.

- Chari, V.V., Kehoe, P.J., McGrattan, E.R., 2000. Sticky price models of the business cycle: Can the contract multiplier solve the persistence problem? *Econometrica* 68 (5), 1151–1180.
- Chetty, R., 2009. Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. NBER Working Paper 15616.
- Christiano, L., Trabandt, M., Walentin, K., 2010. DSGE Models for Monetary Policy Analysis. In: Friedman, B.M., Woodford, M. (Eds.), *Handbook of Monetary Economics* 3A, Elsevier/North-Holland, Amsterdam (Chapter 7).
- Coibion, O., 2006. Inflation inertia in sticky information models. *B.E. Journals: Contributions to Macroeconomics* 6 (1).
- Coibion, O., Gorodnichenko, Y., 2008. What can survey forecasts tell us about informational rigidities. NBER Working Paper 14586.
- Cooper, R., John, A., 1988. Coordinating coordination failures in Keynesian Models. *Quarterly Journal of Economics* 103 (3), 441–463.
- Cover, T., Thomas, J., 1991. *Elements of information theory*. John Wiley and Sons, New York.
- Crucini, M.J., Shintani, M., Tsuruga, T., 2008. Accounting for persistence and volatility of good-level real exchange rates: The role of sticky information. NBER Working Papers 14381.
- Curtin, R., 2009. Sticky information and inflation targeting: How people obtain accurate information about inflation. Manuscript.
- Dixon, H., Kara, E., 2006. How to compare Taylor and Calvo contracts: A comment on Michael Kiley. *Journal of Money, Credit and Banking* 38 (4), 1119–1126.
- Döpke, J., Dovern, J., Fritsche, U., Slacalek, J., 2008a. Sticky information Phillips Curves: European evidence. *Journal of Money, Credit and Banking* 40 (7), 1513–1520.
- Döpke, J., Dovern, J., Fritsche, U., Slacalek, J., 2008b. The dynamics of European inflation expectations. *B.E. Journals: Topics in Macroeconomics* 8 (1).
- Dotsey, M., King, R.G., 2005. Implications of state-dependent pricing for dynamic macroeconomic models. *J. Monetary Econ.* 52 (1), 213–242.
- Dovern, J., Fritsche, U., Slacalek, J., 2009. Disagreement among forecasters in G7 countries. Manuscript.
- Dupor, B., Tsuruga, T., 2005. Sticky information: The impact of different information updating decisions. *Journal of Money, Credit and Banking* 37 (6), 1143–1152.
- Dupor, B., Kitamura, T., Tsuruga, T., 2010. Integrating sticky information and sticky prices. *Rev. Econ. Stat.* 92 (3), 657–669.
- Duffie, D., Sun, T.S., 1990. Transaction costs and portfolio choice in a discrete-continuous-time setting. *Journal of Economic Dynamics and Control* 14 (1), 35–51.
- Eichenbaum, M., Jaimovich, N., Rebelo, S.T., 2008. Reference prices and nominal rigidities. NBER Working Paper 13829.
- Friedman, B., Kuttner, K., 2010. Implementation of monetary policy: How do central banks set interest rates? In: Friedman, B.M., Woodford, M. (Eds.), *Handbook of monetary economics*. 3B, Elsevier/North-Holland, Amsterdam (Chapter 24).
- Friedman, M., 1968. The role of monetary policy. *Am. Econ. Rev.* 58 (1), 1–17.
- Gabaix, X., Laibson, D., 2001. The 6D bias and the equity premium puzzle. *NBER Macroeconomics Annual* 16, 257–312.

- Gali, J., 2008. Monetary policy, inflation, and the business cycle: An introduction to the New Keynesian framework. Princeton University Press, Princeton, NJ.
- Gali, J., Gertler, M., 1999. Inflation dynamics: A structural econometric analysis. *J. Monetary Econ.* 44 (2), 195–222.
- Golosov, M., Lucas, R.E., 2007. Menu costs and Phillips Curves. *J. Polit. Econ.* 115 (2), 171–199.
- Gorodnichenko, Y., 2008. Endogenous information, menu costs and inflation persistence. NBER, Working Paper 14184.
- Hall, R.E., 1988. Intertemporal substitution in consumption. *J. Polit. Econ.* (2), 339–357.
- Haltiwanger, J.C., Waldman, M., 1989. Limited rationality and strategic complements: The implications for macroeconomics. *Quarterly Journal of Economics* 104 (3), 463–483.
- Hamilton, J.D., 1995. Time series analysis. Princeton University Press, Princeton, NJ.
- Heinemann, F., 2000. Unique equilibrium in a model of self-fulfilling currency attacks: Comment. *Am. Econ. Rev.* 90 (1), 316–318.
- Hellwig, C., 2006. Monetary business cycle models: Imperfect information. In: Durlauf, S.N., Blume, L.E. (Eds.), *New Palgrave Dictionary of Economics*. second ed. Palgrave-McMillan, London.
- Hellwig, C., 2008. Heterogeneous information and business cycle fluctuations. Manuscript.
- Hellwig, C., Veldkamp, L. 2009. Knowing what others know: Coordination motives in information acquisition. *Review of Economic Studies*. 76, 223–251.
- Hellwig, C., Venkateswaran, V., 2009. Setting the right prices for the wrong reasons. *J. Monetary Econ.* 56 (S1), S57–S77.
- Hirshleifer, J., 1971. The private and social value of information and the reward to inventive activity. *Am. Econ. Rev.* 61 (4), 561–574.
- Huang, L., Liu, H., 2007. Rational inattention and portfolio selection. *Journal of Finance* 62 (4), 1999–2040.
- Inoue, A., Kilian, L., Kiraz, F.B., 2009. Do actions speak louder than words? Household expectations of inflation based on micro consumption data. *Journal of Money, Credit and Banking* 41 (7), 1331–1363.
- Jinnai, R., 2006. Monetary policy with endogenous inattention. Manuscript.
- Kasa, K., 2000. Forecasting the forecast of others in the frequency domain. *Review of Economic Dynamics* 3, 726–756.
- Khan, H., Zhu, Z., 2006. Estimates of the sticky-information Phillips Curve for the United States. *Journal of Money, Credit and Banking* 38 (1), 195–207.
- Kiley, M.T., 2007. A quantitative comparison of sticky-price and sticky-information models of price setting. *Journal of Money, Credit and Banking* 39 (S1), 101–125.
- Kimball, M.S., Shapiro, M.D., 2008. Labor supply: Are income and substitution effects both large or both small?. NBER Working Paper 14208.
- Klenow, P.J., Malin, B.A., 2010. Microeconomic evidence on price-setting. In: Friedman, B.M., Woodford, M. (Eds.), *Handbook of monetary economics*. 3A, Elsevier/North-Holland, Amsterdam (Chapter 6).
- Klenow, P., Willis, J., 2007. Sticky information and sticky prices. *J. Monetary Econ.* 54 (S1), 79–99.
- Knotek, E.S., 2006. A tale of two rigidities: Sticky prices in a sticky-information environment. FRB, Kansas City, Working Paper 06–15.

- Knotek, E.S., 2008. Convenient prices, currency, and nominal rigidity: Theory with evidence from newspaper prices. *J. Monetary Econ.* 55 (7), 1303–1316.
- Koenig, E.F., 2004. Optimal monetary policy in economies with sticky-information wages. Federal Reserve Bank of Dallas, Dallas, TX Working Paper 04–05.
- Korenok, O., Swanson, N.R., 2005. The incremental predictive information associated with using theoretical New Keynesian DSGE models vs. simple linear econometric models. *Oxford Bull. Econ. Stat.* 67 (1), 905–930.
- Korenok, O., Swanson, N.R., 2007. How sticky is sticky enough? A distributional and impulse response analysis of New Keynesian DSGE models. *Journal of Money, Credit and Banking* 39 (6), 1481–1508.
- Laforte, J.P., 2007. Pricing models: A Bayesian DSGE approach for the US economy. *Journal of Money, Credit and Banking* 39 (S1), 127–154.
- La'O, J., 2009. Collateral constraints and noisy fluctuations. Manuscript.
- Lewis, K.F., 2007. The life-cycle effects of information-processing constraints. Manuscript.
- Lewis, K.F., 2009. The two-period rational inattention model: Accelerations and analyses. *Comput. Econ.* 33 (1), 79–97.
- Lorenzoni, G., 2009. A theory of demand shocks. *Am. Econ. Rev.* 99 (5), 2050–2084.
- Lorenzoni, G., 2010. Optimal monetary policy with uncertain fundamentals and dispersed information. *Rev. Econ. Stud.* 77 (1), 305–338.
- Lucas, R.E., 1972. Expectations and the neutrality of money. *J. Econ. Theory* 4 (2), 103–124.
- Lucas, R.E., 1973. Some international evidence on output-inflation trade-offs. *Am. Econ. Rev.* 63, 326–334.
- Luo, Y., 2008. Consumption dynamics under information processing constraints. *Review of Economic Dynamics* 11, 366–385.
- Mackowiak, B., Wiederholt, M., 2009. Optimal sticky prices under rational inattention. *Am. Econ. Rev.* 99 (3), 769–803.
- Mackowiak, B., Wiederholt, M., 2010. Business cycles dynamics under rational inattention. Manuscript.
- Mankiw, N.G., 1985. Small menu costs and large business cycles: A macroeconomic model of monopoly. *Quarterly Journal of Economics* 100 (2), 529–539.
- Mankiw, N.G., 2001. The inexorable and mysterious tradeoff between inflation and unemployment. *Econ. J.* 111, C45–C61.
- Mankiw, N.G., Reis, R., 2002. Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips Curve. *Quarterly Journal of Economics* 117 (4), 1295–1328.
- Mankiw, N.G., Reis, R., 2003. Sticky information: A model of monetary non-neutrality and structural slumps. In: Aghion, P., Frydman, R., Stiglitz, J., Woodford, M. (Eds.), *Knowledge, information, and expectations in modern macroeconomics: In honor of Edmund S. Phelps*. Princeton University Press, Princeton, NJ.
- Mankiw, N.G., Reis, R., 2006. Pervasive stickiness. *Am. Econ. Rev.* 96 (2), 164–169.
- Mankiw, N.G., Reis, R., 2007. Sticky information in general equilibrium. *Journal of the European Economic Association* 2 (2–3), 603–613.
- Mankiw, N.G., Reis, R., Wolfers, J., 2003. Disagreement about inflation expectations. *NBER Macroeconomics Annual* 18, 209–248.

- Matejka, F., 2008. Rationally inattentive seller: Sales and discrete pricing. Manuscript.
- Meyer-Gohde, A., 2010. Linear rational expectations models with lagged expectations: A synthetic method. *Journal of Economic Dynamics and Control*. 34(5), 984–1002..
- Mondria, J., forthcoming. Portfolio choice, attention allocation, and price comovement. *J. Econ. Theory*. (in press).
- Morris, S., Shin, H.S., 1998. Unique equilibrium in a model of self-fulfilling currency attacks. *Am. Econ. Rev.* 88 (3), 587–597.
- Morris, S., Shin, H.S., 2001. Rethinking multiple equilibria in macroeconomics. *NBER Macroeconomics Annual* 15, 139–161.
- Morris, S., Shin, H.S., 2002. The social value of public information. *Am. Econ. Rev.* 92 (5), 1521–1534.
- Morris, S., Shin, H.S., 2006. Inertia of forward-looking expectations. *Am. Econ. Rev.* 96 (2), 152–157.
- Moscarini, G., 2004. Limited information capacity as a source of inertia. *Journal of Economic Dynamics and Control* 28 (10), 2003–2035.
- Nakamura, E., Steinsson, J., Monetary non-neutrality in a multi-sector menu cost model. *Quarterly Journal of Economics*. (in press).
- Nimark, K., 2008. Dynamic pricing and imperfect common knowledge. *J. Monetary Econ.* 55 (2), 365–382.
- Phelps, E.S., 1968. Money-wage dynamics and labor market equilibrium. *J. Polit. Econ.* 76 (4), 678–711.
- Phillips, A.W., 1958. The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861–1957. *Economica* 25 (100), 283–299.
- Reis, R., 2006a. Inattentive producers. *Review of Economic Studies* 73 (3), 793–821.
- Reis, R., 2006b. Inattentive consumers. *J. Monetary Econ.* 53 (8), 1761–1800.
- Reis, R., 2009a. Optimal monetary policy rules in an estimated sticky-information model. *American Economic Journal: Macroeconomics* 1 (2), 1–28.
- Reis, R., 2009b. A sticky-information general-equilibrium model for policy analysis. In: Schmidt-Heubel, K., Walsh, C. (Eds.), *Monetary policy under uncertainty and learning*. Central Bank of Chile, Chile.
- Reis, R., 2010. When Should Policymakers Make Announcements? Manuscript.
- Roca, M., 2006. Transparency and monetary policy with imperfect common knowledge. Manuscript.
- Romer, D., 2008. Real rigidities. In: Durlauf, S.N., Blume, L. (Eds.), *New Palgrave Dictionary of Economics*. second ed. Palgrave-MacMillan, London.
- Rondina, G., 2008. Incomplete information and informative pricing. Manuscript.
- Rogerson, R., Wallenius, J., 2009. Micro and macro elasticities in a life cycle model with taxes. *J. Econ. Theory*. 144 (6), 2277–2292.
- Rudd, J., Whelan, K., 2007. Modeling inflation dynamics: A critical review of recent research. *Journal of Money, Credit and Banking* 39 (1), 155–170.
- Samuelson, P., Solow, R., 1960. Analytical aspects of anti-inflation policy. *Am. Econ. Rev.* 50 (2), 177–194.
- Simon, H.A., 1956. Dynamic programming under uncertainty with a quadratic criterion function. *Econometrica* 24 (1), 74–81.
- Sims, C.A., 2003. Implications of rational inattention. *J. Monetary Econ.* 50 (3), 665–690.

- Sims, C.A., 2006. Rational inattention: Beyond the linear-quadratic case. *Am. Econ. Rev.* 96 (2), 158–163.
- Sims, C.A., 2010. Rational inattention and monetary economics. In: Friedman, B.M., Woodford, M. (Eds.), *Handbook of monetary economics*. 3A, Elsevier/North-Holland, Amsterdam (Chapter 4).
- Svensson, L.E.O., 2006. Social value of public information: Morris and Shin (2002) is actually pro transparency, not con. *Am. Econ. Rev.* 96 (1), 448–452.
- Taylor, J.B., 1985. New econometric approaches to stabilization policy in stochastic models of macroeconomic fluctuations. In: Griliches, Z., Intriligator, M. (Eds.), *Handbook of econometrics*. 3, Elsevier/North-Holland, Amsterdam.
- Townsend, R., 1983. Forecasting the forecasts of others. *J. Polit. Econ.* 91 (4), 546–588.
- Trabandt, M., 2004. Sticky information vs. sticky prices: A horse race in a DSGE framework. Manuscript.
- Tutino, A., 2009. The rigidity of choice: Lifecycle savings with information-processing limits. Manuscript.
- van Nieuwerburgh, S., Veldkamp, L., Information acquisition and under-diversification. *Review of Economic Studies*. (in press).
- Veldkamp, L., 2009. Information choice in macroeconomics and finance. Manuscript.
- Woodford, M., 2002. Imperfect common knowledge and the effects of monetary policy. In: Aghion, P., Frydman, R., Stiglitz, J., Woodford, M. (Eds.), *Knowledge, information, and expectations in modern macroeconomics: In honor of Edmund S. Phelps*. Princeton University Press, Princeton, NJ.
- Woodford, M., 2003. *Interest and prices*. Princeton University Press, Princeton, NJ.
- Woodford, M., 2009. Information-constrained state-dependent pricing. *J. Monetary Econ.* 56 (S1), S100–S124.
- Zbaracki, M.J., Ritson, M., Levy, D., Dutta, S., Bergen, M., 2004. The managerial and customer costs of price adjustment: Direct evidence from industrial markets. *Rev. Econ. Stat.* 86 (2), 514–533.

This page intentionally left blank

Microeconomic Evidence on Price-Setting[☆]

Peter J. Klenow* and Benjamin A. Malin**

*Stanford University and NBER

**Federal Reserve Board

Contents

1. Introduction	232
2. Data Sources	234
3. Frequency Of Price Changes	238
3.1 Average frequency	238
3.2 Heterogeneity	242
3.3 Sales, product turnover, and reference prices	247
3.4 Determinants of frequency	254
4. Size of Price Changes	257
4.1 Average magnitude	257
4.2 Increases versus decreases	257
4.3 Higher moments of the size distribution	257
5. Dynamic Features of Price Changes	258
5.1 Synchronization	258
5.2 Sales, reference prices, and aggregate inflation	262
5.3 Hazard rates	266
5.4 Size versus age	266
5.5 Transitory relative price changes	267
5.6 Response to shocks	268
5.7 Higher moments of price changes and aggregate inflation	270
6. Ten Facts and Implications for Macro Models	271
6.1 Fact 1: Prices change at least once a year	271
6.2 Fact 2: Sales and product turnover are often important for micro price flexibility	272
6.3 Fact 3: Reference prices are stickier and more persistent than regular prices	273
6.4 Fact 4: There is substantial heterogeneity in the frequency of price change across goods	273
6.5 Fact 5: More cyclical goods change prices more frequently	274
6.6 Fact 6: Price changes are big on average, but many small changes occur	274
6.7 Fact 7: Relative price changes are transitory	275

[☆] This research was conducted with restricted access to U.S. Bureau of Labor Statistics (BLS) data. Rob McClelland provided us invaluable assistance and guidance in using BLS data. We thank Margaret Lay and Krishna Rao for excellent research assistance. We are grateful to Luis J. Álvarez, Mark Bils, Marty Eichenbaum, Etienne Gagnon, Emi Nakamura, Martin Schneider, Frank Smets, Jón Steinsson, and Michael Woodford for helpful suggestions. The views expressed here are those of the authors and do not necessarily reflect the views of the BLS or the Federal Reserve System.

6.8 Fact 8: Price changes are typically not synchronized over the business cycle	275
6.9 Fact 9: Neither frequency nor size is increasing in the age of a price	276
6.10 Fact 10: Price changes are linked to wage changes	276
6.11 Summary: Model features and the facts	277
7. Conclusion	278
References	279

Abstract

The last decade has seen a burst of micro price studies. Many studies analyze data underlying national CPIs and PPIs. Others focus on more granular subnational grocery store data. We review these studies with an eye toward the role of price setting in business cycles. We summarize with ten stylized facts: prices change at least once a year, with temporary price discounts and product turnover often playing an important role. After excluding many short-lived prices, prices change closer to once a year. The frequency of price changes differs widely across goods, however, with more cyclical goods exhibiting greater price flexibility. The timing of price changes is little synchronized across sellers. The hazard (and size) of price changes does not increase with the age of the price. The cross-sectional distribution of price changes is thick-tailed, but contains many small price changes too. Finally, strong linkages exist between price changes and wage changes.

JEL classification: E3, E31, E5

Keywords

Micro Price Data
Nominal Stickiness
Time-Dependent Pricing
State-Dependent Pricing
Contract Multiplier

1. INTRODUCTION

Recent years have seen a wealth of rich micro price data become available. Many studies have examined data underlying nationally representative consumer and producer price indices from national statistical agencies. A smaller set of studies have focused on finer scanner data for a subset of stores or products. The United States and Western European countries have received the most attention, but evidence on emerging markets has grown rapidly.

Such micro data offer many insights on the importance of price stickiness for business cycles. We review the literature by stating a series of ten facts modelers may want to know about price setting.

First, individual prices change at least once a year. The frequency is more like twice a year in the United States versus once a year in the Euro Area. Thus we need a “contract multiplier” to explain why real effects of nominal shocks appear to last several years.

Second, temporary price discounts (“sales”) and product turnover are important to micro price flexibility. This is particularly true in the United States, which plays a role

in its greater price flexibility than in the Euro Area. We provide evidence that such sale prices partially cancel out with cross-sectional and time aggregation, but appear to contain macro content.

Third, if one drops a broad set of short-lived prices (i.e., more than just temporary price discounts), a stickier “reference” price emerges that changes about once a year in the United States. This filtering conceals considerable novelty in nonreference prices, and these deviations could be responding to aggregate shocks as they do not seem to wash out with aggregation. Still, reference price inflation is considerably more persistent than overall inflation, perhaps suggesting some sort of sticky plan and/or sticky information.

Fourth, goods differ greatly in how frequently their prices change. At one extreme are goods that change prices at least once a month (fresh food, energy, airfares), and at the other extreme are services that change prices much less often than once a year. Such heterogeneity makes mean price durations much longer than median durations, and could help explain a big contract multiplier if combined with strategic complementarities.

Fifth, goods with more cyclical quantities (e.g., cars and apparel) exhibit greater micro price flexibility than goods with little business cycle (e.g., medical care). Durables, as a whole, change prices more frequently than nondurables and services. Including temporary price changes, nondurables change price more frequently than services. Such nonrandom heterogeneity in price stickiness may hold down the contract multiplier.

Sixth, micro price changes are, on average, much bigger than needed to keep up with aggregate inflation, suggesting the dominance of idiosyncratic forces (intertemporal price discrimination, inventory clearance, etc.). In state-dependent pricing models, price changers can be selected on their idiosyncratic shocks, thereby speeding price adjustment and depressing the contract multiplier. Micro evidence exists for such selection, but not as strong as predicted by models with a single menu cost. For example, many price changes are small, as with time-dependent or information-constrained pricing.

Seventh, relative price changes are transitory. Idiosyncratic shocks evidently do not persist as long as aggregate shocks do. Sellers may be implementing price changes for temporary, idiosyncratic reasons while failing to incorporate macro shocks (e.g., as in rational inattention models).

Eighth, the timing of price changes is little synchronized across products. Most movements in inflation (from month to month or quarter to quarter) are due to changes in the size rather than the frequency of price changes. This may be a by-product of the stable inflation rates in the past few decades in the United States and Euro Area. In countries with more volatile inflation, such as Mexico, the frequency of price changes has shown more meaningful variation. This lack of synchronization is consistent with the importance of idiosyncratic pricing considerations over macro ones. When combined with strategic complementarities, price staggering paves the way for coordination failure and a high contract multiplier. It is also consistent with rational inattention toward macro shocks. Perhaps related, consumer price changes (both increases and decreases) have increased noticeably in the recent U.S. recession.

Ninth, the hazard rate of price changes falls with the age of a price for the first few months (mostly due to sales and returning to regular prices), and is largely flat thereafter (other than a spike at one year for services). This finding holds in the United States and Euro Area, and for both consumer and producer prices. Such a pattern is consistent with a mix of Calvo and Taylor time-dependent pricing, but can also be generated under state-dependent pricing. Meanwhile, the size of price changes is largely unrelated to the time between price changes. This fact seems more discriminating, and favors state-dependent over time-dependent pricing. If price spell length is exogenous, more shocks should accumulate and make for bigger price changes after longer price spells. Under state-dependent pricing, longer price spells reflect stable desired prices rather than pent-up demand for price changes.

Tenth and finally, price changes are linked to wage changes. Firms in labor-intensive sectors adjust prices less frequently, potentially because wages adjust less frequently than other input prices. Furthermore, survey evidence suggests synchronization between wage and price adjustments over time. Thus, in addition to contributing directly to a higher contract multiplier, wage stickiness may be contributing indirectly by lowering the frequency of price changes.

We organize the rest of this chapter as follows. Section 2 briefly outlines the micro data sources commonly used in the recent literature. Section 3 discusses evidence on the frequency of price changes. Section 4 describes what we know about the size of price changes. Section 5 delves into price setting dynamics; for example, synchronization and what types of price changes cancel out with aggregation across products and time. Section 6 reviews, at greater length, the ten stylized facts we just discussed. Section 7 offers conclusions.

2. DATA SOURCES

The recent literature studies data underlying consumer (CPI) and producer (PPI) price indices, scanner and online data collected from retailers, and information gleaned from surveys of price setters. In this section we briefly describe these data sets.

Until recently, empirical evidence on price-setting at the microeconomic level was somewhat limited, consisting mostly of studies that focused on relatively narrow sets of products (e.g., Carlton, 1986; Cecchetti 1986; and Kashyap 1995).¹ This changed as data sets underlying official CPIs and PPIs became available to researchers. These data sets, compiled by national statistical agencies, contain a large number of monthly price quotes tracking individual items over several years or more. The samples aim to be broadly representative — in terms of products, outlets, and cities covered — of national consumer expenditure (or industrial production). For example, the CPI Research Database (CPI-RDB), maintained by the U.S. Bureau of Labor Statistics (BLS), contains prices

¹ Wolman (2007) provided a comprehensive survey of the older literature, while Mackowiak and Smets (2008) also surveyed the more recent literature.

for all categories of goods and services other than shelter, or about 70% of consumer expenditure. It begins in January 1988 and includes about 85,000 prices per month (Klenow & Kryvtsov, 2008; Nakamura & Steinsson, 2008a).

Although the CPI and PPI data sets are alike in many ways, Nakamura and Steinsson (2008a) pointed out that interpreting the PPI data is somewhat more complicated than interpreting evidence on consumer prices.² First, the BLS collects PPI data through a survey of firms rather than a sample of “on-the-shelf” prices. Second, the definition of a PPI good is meant to capture all “price-determining variables,” which often include the buyer of the good. Intermediate prices may be part of (explicit or implicit) long-term contracts, and thus observed prices might not reflect the true shadow prices faced by the buyer (Barro, 1977). Related, in wholesale markets the seller may choose to vary quality margins, such as delivery lags, rather than varying the price (Carlton, 1986). Mackowiak and Smets (2008) pointed out that repeated interactions (e.g., for legal services) and varying quality margins (e.g., waiting in order to purchase a good at the published price) are also present in some retail markets.

A critical open question for macroeconomists in interpreting prices is whether they conform to the Keynesian sticky-price paradigm of “call options with unlimited quantities.”³ On-the-shelf consumer prices may have this feature if they are available in inventory (see Bills, 2004, on stockouts in the CPI). Gopinath and Rigobon (2008) said import prices usually appear to be call options for buyers. Still, unlike for consumer prices, it is not clear whether new buyers of producer goods have the option to buy at prices prevailing for existing buyers.

Tables 1 and 2 list several studies that have made use of CPI and PPI data, respectively. These include studies for the United States, for countries in the Euro Area (Austria, Belgium, Finland, France, Germany, Italy, Luxembourg, The Netherlands, Portugal, and Spain), and for a handful of other developed (Denmark, Israel, Japan, Norway, South Africa) and developing economies (Brazil, Colombia, Chile, Hungary, Mexico, Sierra Leone, Slovakia). Although differences in methodology and coverage make cross-country comparisons challenging, the Inflation Persistence Network (IPN) has coordinated efforts of the many researchers in the Euro Area to allow for such comparisons (Dhyne et al., 2006; Vermeulen et al., 2007).

A related set of studies has made use of micro data the BLS collects to construct import and export price indices for the United States. These include Gopinath and Rigobon (2008), Gopinath, Itskhoki, and Rigobon (2010), Gopinath and Itskhoki (2010), and Nakamura and Steinsson (2009). The prices are collected from surveys of importing firms and thus represent wholesale markets. One benefit to using international data is the ability to analyze price-setting behavior in response to large, identified shocks (i.e., nominal exchange rate shocks).

² The challenges described are for United States PPI data, but Euro Area PPI data display similar features (Vermeulen et al., 2007).

³ We are grateful to Robert Hall for this phrase.

Table 1 Monthly mean frequency of CPI price changes

Country	Paper	Sample period	Frequency (in %)
Austria	Baumgartner et al. (2005)	1996:01–2003:12	15.1
Belgium	Aucremanne and Dhyne (2004)	1989:01–2001:01	16.9
Brazil	Barros et al. (2009)	1996:03–2008:12	37.2
	Gouvea (2007)	1996:01–2006:12	37.0
Chile	Medina et al. (2007)	1999:01–2005:07	46.1
Denmark	Hansen and Hansen (2006)	1997:01–2005:12	17.3
Euro Area	Dhyne et al. (2006)	1996:01–2001:01	15.1
Finland	Vilmunen and Laakkonen (2005)	1997:01–2003:12	16.5
France	Baudry et al. (2007)	1994:07–003:02	18.9
Germany	Hoffmann and Kurz-Kim (2006)	1998:02–2004:01	11.3
Hungary	Gabriel and Reiff (2008)	2001:12–2007:06	15.1
Israel	Baharad and Eden (2004)	1991:01–1992:12	24.5
Italy	Fabiani et al. (2006)	1996:01–2003:12	10.0
Japan	Saita et al. (2006)	1999:01–2003:12	23.1
Luxembourg	Lünnemann and Mathä (2005)	1999:01–2004:12	17.0
Mexico	Gagnon (2009)	1994:01–2004:12	29.4
Netherlands	Jonker et al. (2004)	1998:11–2003:04	16.5
Norway	Wulfsberg (2009)	1975:01–2004:12	21.3 (21.9)
Portugal	Dias et al. (2004)	1992:0–2001:01	22.2
Sierra Leone	Kovanen (2006)	1999:01–2003:04	51.5
Slovakia	Horvath and Coricelli (2006)	1997:01–2001:12	34.0
South Africa	Creamer and Rankin (2008)	2001:12–2006:02	16.0
Spain	Álvarez and Hernando (2006)	1993:01–2001:12	15.0
United Kingdom	Bunn and Ellis (2009)	1996:01–2006:01	15 (19)
United States	Bils and Klenow (2004)	1995:01–97:12	26.1
	Klenow and Kryvtsov (2008)	1988:02–2005:01	29.9 (36.2)
	Nakamura and Steinsson (2008a)	1988:01–2005:12	21.1 (26.5)

Notes: Source is Álvarez (2008) with three additional studies, Barros et al. (2009), Bunn and Ellis (2009), and Wulfsberg (2009), and updated versions of Gagnon (2009), Creamer and Rankin (2008), and Klenow and Kryvtsov (2008).

For studies that report frequencies of both regular (i.e., nonsale) and posted prices, the figures in parentheses correspond to posted prices. Frequencies for Nakamura and Steinsson (2008a) correspond to the 1998–2005 sample period (for contiguous observations, excluding substitutions). For Germany, frequencies refer to the sample considering item replacements and nonquality adjusted data. The Spanish sample excludes energy products, which lowers the aggregate frequency.

Table 2 Monthly mean frequency of PPI price changes

Country	Paper	Sample period	Frequency (in %)
Belgium	Cornille and Dossche (2008)	2001:01–2005:01	24
Colombia	Julio and Zárate (2008)	1999:06–2006:10	20.2
Euro Area	Vermuelen et al. (2007)	Various	21
France	Gautier (2008)	1994:01–2005:06	25
Germany	Stahl (2006)	1997:01–2003:09	22
Italy	Sabbatini et al. (2006)	1997:0–2002:12	15
Portugal	Dias et al. (2004)	1995:01–2001:01	23
South Africa	Creamer (2008)	2001:1–2006:02	19.5
Spain	Álvarez et al. (2008)	1991:01–1999:02	21
United Kingdom	Bunn and Ellis (2009)	2003:01–2007:12	26
United States	Nakamura–Steinsson (2008a)	1988:01–2005:12	24.7
	Goldberg–Hellerstein (2009)	1987:0–2008:10	33.3

Note: Source is Álvarez (2008), Bunn and Ellis (2009), Goldberg and Hellerstein (2009), and the published versions of Cornille and Dossche (2008) and Gautier (2008). Frequencies for Nakamura and Steinsson (2008a) correspond to finished goods. The Italian sample excludes energy products, while the French sample does not include business services.

Another source of microeconomic evidence on pricing comes from scanner (i.e., barcode) data collected from supermarkets, drugstores, and mass merchandisers. These data cover a narrower set of goods than the data underlying price indices, but they provide deeper information. Scanner data usually cover many more items per outlet, and often contain information on quantities sold (and sometimes wholesale cost). Data are usually collected at a weekly frequency and may come from one particular retailer (e.g., Eichenbaum, Jaimovich & Rebelo, 2009, or studies using Dominick’s data) or from multiple retailers (e.g., through AC Nielsen). A number of these studies are listed in Table 3.

Other researchers have begun collecting price information from retailers by “scraping” prices from Web sites. The ongoing “Billion Prices Project” of Cavallo and Rigobon (e.g., Cavallo, 2009) collects daily prices from numerous retailers in over 50 countries. Useful aspects of this data set include the daily frequency, comparability across many countries, and detailed information on each product including sales and price control indicators. Lünneemann and Wintr (2006) is another example.

A final source of microeconomic information comes from surveying firms about their price-setting practices, as opposed to collecting longitudinal information about individual prices. These surveys allow researchers to ask about aspects of pricing that cannot be captured from data sets of observed prices, such as the frequency with which price-setters review prices and the importance of particular theories of price stickiness

Table 3 Frequency of price change in scanner data sets

Data Source	Paper	Sample period	Frequency (in %)
Dominick's	Midrigan (2009)	1989–1997	45 (23)
	Burstein and Hellwig (2007)		41 (26)
Large U.S. Retailer	Eichenbaum et al. (2009)	2004–2006	43
AC Nielsen ERIM	Campbell and Eden (2005)	1985–1987	47
Database	Midrigan (2009)		36 (25)
AC Nielsen ScanTrak	Nakamura (2008)	2004	44 (19)
AC Nielsen Homescan	Broda and Weinstein (2007)	1994–2003	NA

Note: For most studies, weekly data are collected, but a monthly frequency of price change is reported. Eichenbaum et al. (2009) reported a weekly frequency of price change. Frequencies are for *posted* prices, and numbers in parentheses are for regular (i.e., nonsale) prices. Frequencies vary across studies using the same data set because of different sample choices and reported measures (e.g., for Dominick's, Midrigan (2009) reported mean frequencies for one store, while Burstein and Hellwig (2007) considered many stores and report the frequency of the median product category).

for explaining their pricing decisions. Blinder, Canetti, Lebow, and Rudd (1998) was a pioneering study for the United States, and subsequent surveys have been conducted in many countries, as shown in Table 4. The surveys typically ask firms to focus on their main product (or most important products).

3. FREQUENCY OF PRICE CHANGES

We begin our review of the substantive findings of the literature by looking at the frequency with which prices change. A theme that will arise throughout this chapter is the presence of a great deal of heterogeneity in price-setting behavior, and we therefore report results along several dimensions. These include measures of the “average” time between price changes, how these measures vary across different samples and types of goods, the importance of temporary sales and product turnover, and some discussion of the determinants of the frequency of price change.

3.1 Average frequency

Table 1, drawn primarily from a survey by Álvarez (2008), presents estimates of the *mean* frequency of price changes obtained from the data sets underlying national CPIs.⁴ Prices clearly exhibit nominal stickiness, as the (unweighted) median across these

⁴ For studies that contain information on price changes due to temporary sales, we report the frequency for both all prices (in parentheses) and nonsale prices. In many countries, the prices reported during sales periods are prices without rebates (i.e., posted prices *are* essentially nonsale prices), and we thus use the nonsale prices when we describe results across countries.

Table 4 Number of price changes per year (%) in survey data

Country	Paper	<1	1	2–3	≥4	Median	Mean (in months)
Austria	Kwapil et al. (2005)	24	51	15	11	1	12.7
Belgium	Aucremanne and Druant (2005)	18	55	18	8	1	11.9
Canada	Amirault et al. (2006)	8	27	23	44	2–3	6.8
Estonia	Dabusinskas and Randveer (2006)	14	43	25	18	1	10.0
Euro Area	Fabiani et al. (2005)	27	39	20	14	1	12.3
France	Loupias and Ricart (2004)	21	46	24	9	1	11.8
Germany	Stahl (2005)	44	14	21	21	1	13.5
Italy	Fabiani et al. (2007)	20	50	19	11	1	11.9
Japan	Nakagawa et al. (2000)	23	52	11	14	1	12.5
Luxembourg	Lünnemann and Mathä (2006)	15	31	27	27	2–3	9.0
Mexico	Castanon et al. (2008)	—	—	—	—	—	5.7
Netherlands	Hoeberichts and Stokman (2006)	10	60	19	11	1	10.7
Portugal	Martins (2005)	24	51	14	12	1	12.7
Romania	Copaciu et al. (2007)	—	—	—	—	—	4.1
Spain	Álvarez and Hernando (2007a)	14	57	15	14	1	11.1
Sweden	Apel et al. (2005)	29	43	6	20	1	12.7
Turkey	Sahinoz and Saracoglu (2008)	—	—	—	—	—	3.0
United Kingdom	Hall et al. (2000)	6	37	44	14	2–3	8.2
United States	et al. (1998)	10	39	29	22	1	8.8

Note: Source is Álvarez (2008), Table 3. Mean implicit durations obtained from the interval-grouped data using the following assumptions: for firms declaring “at least four price changes per year,” 8 price changes are considered (i.e., mean duration of 1.33 months); for those declaring “two or three price changes per year,” 2.5 price changes are considered (i.e., 4.8 months); for those declaring “one change per year” a duration of 12 months, and for those declaring “less than one price change per year,” a change every two years is considered (i.e., 24 months).

studies for the estimated mean frequency of price change is 19% (per month). The degree of stickiness varies considerably across countries, with prices in the Euro Area appearing to change less frequently than those in the United States, which in turn change less frequently than those in high-inflation developing countries (Brazil, Chile, Mexico, Sierra Leone, Slovakia). We will return to the question of what explains these cross-country differences after we take a closer look at individual country studies.

We will give particular attention to the multiple U.S. CPI studies (Bils & Klenow 2004; Klenow & Kryvtsov 2008; Nakamura & Steinsson 2008a) to shed light on a number of features of the data and provide understanding on how different methodologies impact results. Moreover, since we have access to the micro data from the BLS, we will be able to construct some new results as we proceed. We begin by describing the structure of the micro data. The BLS divides goods and services into 300 or so categories of consumption known as Entry Level Items (ELIs). Within these categories are prices for particular products sold at particular outlets (which we will refer to as quote-lines). The BLS collects prices monthly for all products in the three largest metropolitan areas (New York, Los Angeles, and Chicago) and for food and fuel products in all areas, and bimonthly for all other prices. The statistics we report from Klenow and Kryvtsov (2008) and Nakamura and Steinsson (2008a) are for prices collected monthly from the top three cities.⁵

To construct their average monthly frequency that we report in Table 1, Klenow and Kryvtsov (2008) first estimated frequencies for each ELI category and then took the weighted *mean* across categories to arrive at a figure of 36.2% (for posted prices) between 1988 and early 2005. Of course, this is not the only possible measure of the “average” frequency of price change. The weighted *median* frequency of price changes is 27.3%. The mean is higher than the median because the distribution of the frequency of price changes across ELIs, shown in Figure 1 for 1998–2009 data, is convex (Jensen’s inequality). Related, the mean implied duration (from the mean of the inverse frequencies across ELIs) of 6.8 months is higher than the median (the inverse of the median frequency) of 3.7 months.

Turning to producer prices, the median country in Table 2 has a mean frequency of price change of 23%. Nakamura and Steinsson (2008a) compared price flexibility with consumer goods by matching 153 ELI categories from the CPI with product codes from the PPI. In general, they found the frequency of price change for producer prices to be similar to that of consumer prices excluding sales. Goldberg and Hellerstein (2009), however, reported a higher frequency — closer to consumer prices including sales — and attributed the difference to weighting products by their use of BLS firm

⁵ Both studies weight ELIs by BLS estimates of their importance in consumer expenditures. Klenow and Kryvtsov (2008) also used some BLS weighting information for items within ELIs; Nakamura and Steinsson (2008a) did not, but it does not seem to affect statistics such as the median duration of prices across ELIs.

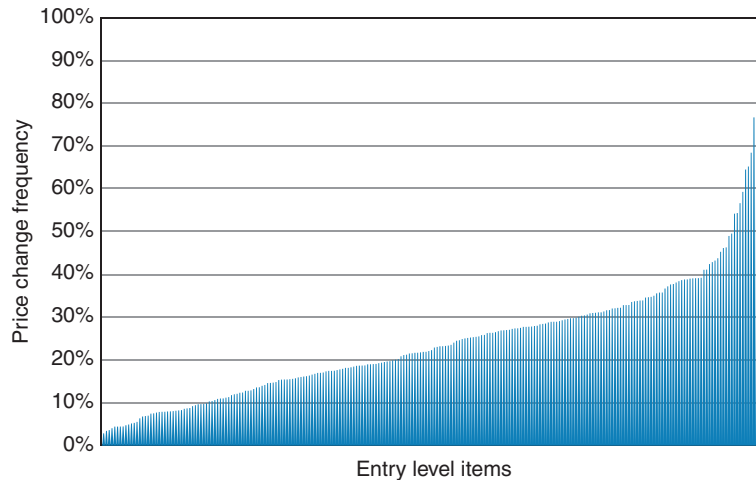


Figure 1 Price change frequency by product category. *Source:* CPI-RDB. Data are for the top three cities (New York, Los Angeles, and Chicago) January 1988 through October 2009. Each bar corresponds to an ELI product category (weighted by expenditure), and price change frequency is calculated as the weighted average across quotes within the ELI. Prices include sales and regular prices.

and industry weights. Using these weights makes a large difference because larger firms change prices more frequently than small firms, they found. Vermeulen et al. (2007) documented producer price setting in six European countries. Not controlling for the composition of the CPI and PPI baskets, they found producer prices changed somewhat more frequently than what Dhyne et al. (2006) reported for consumer prices. This pattern persisted when they focused on similar products in the “processed food” and “nonfood nonenergy consumer goods” sectors.

Gopinath and Rigobon (2008) found that U.S. export and import (wholesale) prices are sticky in the currency in which they are reported. The median implied duration in the currency of pricing is 10.6 (12.8) months for imports (exports) during their 1994–2005 sample.⁶ They go on to compare the duration of their cross-border transactions to domestic transactions by using product category descriptions to match international price (IPP) categories with PPI categories. Restricting their sample to these matched categories (69 of them), they found a mean duration of 10.3 months for the IPP and 10.6 months for the PPI.

Table 3 provides frequencies from a number of scanner data studies. Although the underlying data is weekly, the numbers in the table are monthly frequencies of price change (with the exception of Eichenbaum et al., 2009, who reported a weekly

⁶ These numbers correspond to their benchmark specification, in which price changes across nonadjacent prices and product substitutions are included and the frequencies of goods whose price never changes are adjusted by the probability of discontinuation.

frequency). Frequencies vary, even across studies using the same data set, because of different sample choices and reported measures. For example, for Dominick's, Midrigan (2009) reported a mean frequency for one store, while Burstein and Hellwig (2007) considered many stores and reported the frequency of the median product category. Despite these differences, the studies have very similar results: average posted (nonsale) prices change at least every three (five) months.

Table 4 presents information on price flexibility that comes from asking firms how frequently they changed their prices in the past year (or on average in recent years). The firms surveyed tend to sell their main product to other firms, and thus, the survey data pertain primarily to producer prices. The median frequency of price change, about once a year in most countries, exhibits more stickiness than the PPI micro data, although the results are not directly comparable due to different time periods, samples of firms, and so forth.

To recap, prices do not change continuously but do change “on average” *at least* once a year. We use “on average” in a loose sense, as we have already seen that the complexity of the micro data makes it difficult to summarize with one statistic (such as the mean or median). We now explore this complexity in more detail, first investigating heterogeneity in the frequency with which prices change across different types of goods and then discussing the treatment of sales and product substitution.

3.2 Heterogeneity

Figure 1 shows a tremendous amount of heterogeneity across ELI categories, as the price change frequency ranges from 2.7% for “Intracity Mass Transit” to 91% for “Regular Unleaded Gasoline.” Indeed, while half of all prices have an implied duration less than 3.4 months, almost a fifth last longer than a year. This heterogeneity also helps to explain the finding that “average” consumer prices adjust more frequently than in the narrower investigations predating the latest generation of micro studies. For example, Cecchetti (1986) found the length of time between changes in the newsstand prices of U.S. magazines ranged from 1.8 to 14 years, but Nakamura and Steinsson (2008a) found that “Single-Copy Newspapers and Magazines,” with a duration of 17.2 months in their sample, change prices less frequently than 84% of nonhousing consumption.

Table 5 illustrates the heterogeneity in the frequency of price changes in additional ways. It reports the weighted median and mean implied price durations in the U.S. CPI from January 1988 through October 2009 separately for posted and regular (i.e., nonsale) prices covering: (a) All Items; (b) Durables, Nondurables, and Services; (c) Raw and Processed goods; and (d) eight Major Groups. For conciseness, consider the mean durations of posted prices. For Durables the mean price duration is 3.0 months, whereas for Nondurables it is 5.8 months and for Services it is 9.4 months. For raw goods (energy and food commodities) prices last about 1.1 months, whereas

Table 5 Price durations by category in the U.S. CPI

Durations in Months	Posted		Regular		% of CPI
	Median	Mean	Median	Mean	
All items	3.4	6.2	6.9	8.0	100.0%
Durable Goods	1.8	3.0	1.8	5.0	21.7
Nondurable Goods	3.4	5.8	7.3	8.3	48.6
Services	7.6	9.4	7.6	9.6	29.7
Raw Goods	1.0	1.1	1.0	1.2	12.0
Processed Goods	4.4	6.9	7.7	8.9	88.0
Apparel	2.8	2.9	9.2	10.1	7.0
Education and Communication	5.4	6.2	6.7	6.3	7.3
Food	3.4	6.9	8.5	9.3	22.4
Home Furnishings	1.9	3.5	2.0	5.4	17.0
Medical Care	10.0	14.2	12.6	14.7	7.8
Recreation	6.3	7.5	9.4	9.8	8.5
Transportation	1.8	3.7	1.8	3.8	24.5
Other Goods and Services	8.6	14.7	12.1	16.7	5.5

Source: CPI-RDB. Data are for the top three cities (New York, Los Angeles, and Chicago) from February 1988 through October 2009. Durations are weighted medians or means of implied durations from weighted average frequencies within ELIs. Durables, Nondurables, and Services coincide with U.S. National Income and Product Account classifications. Raw goods include gasoline, motor oil and coolants, fuel oil and other fuels, electricity, natural gas, meats, fish, eggs, fresh fruits, fresh vegetables, and fresh milk and cream. Apparel, etc., are Major Groups in the CPI (1998-onward definition).

for processed goods and services it is 6.9 months. Among Major Groups, price durations range from 2.9 months in Apparel to 14.7 months in Other Goods and Services.

We further explore the connection between durability and price change frequency at a more disaggregated level in U.S. consumer prices. For each of 65 Expenditure Classes of the CPI (more aggregated than the 300+ ELIs, less aggregated than the Major Groups), we were able to estimate durability from the data in Bils and Klenow (1998). For interpretability, in Figure 2 we aggregate back up to the Major Groups. The figure plots the average frequency of posted price changes against average durability in years, with each dot proportional to the group's average expenditure weight. Transportation stands out as durable, flexible, and important. For example, Motor Vehicles have a high weight (16% of the nonshelter sample weight), high durability (9 years), and high frequency (38% per month). Food, on the other hand, is *nondurable*, flexible, and important. As a result, Table 6 reports no significant relationship between

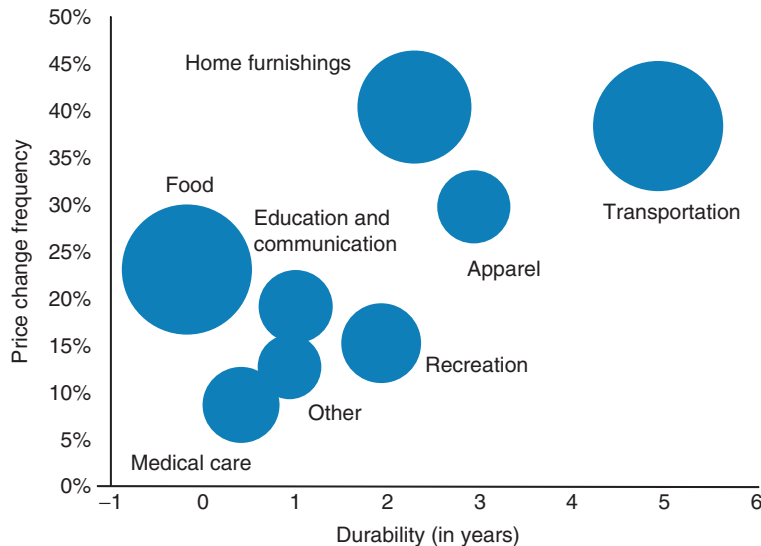


Figure 2 Frequency versus durability in the U.S. CPI. *Source:* CPI-RDB. Data are for the top three cities (New York, Los Angeles, and Chicago) January 1988 through October 2009. Each circle is one of the eight Major Groups in the CPI (1998-onward definition), with the area proportional to the average expenditure weight over the sample. Prices include sales and regular prices. Frequency is calculated as the weighted mean across ELIs (with each ELI mean itself a weighted average across quotes within the ELI). Durability is based on estimates reported in Bils and Klenow (1998).

Table 6 Frequency vs. durability and cyclicity across U.S. CPI categories

WLS of Frequency on →	Durability	Cyclicity
All items		
Posted prices	0.60 (0.69)	3.23 (1.05)
Regular prices	0.49 (0.71)	3.77 (1.07)
Processed items		
Posted prices	1.47 (0.41)	3.29 (0.58)
Regular prices	1.42 (0.39)	3.79 (0.48)

Note: Source is CPI-RDB. Data are for the top three cities (New York, Los Angeles, and Chicago) from January 1988 through October 2009. Entries are Weighted Least Square (WLS) coefficients across Expenditure Classes, where the weights are based on average shares of consumer expenditure. Regressions include a constant. Coefficient standard errors are in parentheses. We were able to match estimates of durability (cyclicity) for 65 (64) BLS Expenditure Classes (1998-onward definition) for All Items, and 59 (58) for Processed Items. Frequency is calculated as the weighted mean across ELIs (with each ELI mean itself a weighted average across quotes within the ELI). Durability is based on estimates reported in Bils and Klenow (1998). Cyclicity is based on an OLS regression coefficient for each Expenditure Class of *real* monthly consumption growth on aggregate real monthly consumption growth, using seasonally adjusted Detailed Expenditure data from the U.S. National Income and Product Accounts, January 1990 through May 2009.

frequency and durability across the 65 Expenditure Classes: the weighted least squares estimate is 0.60 percentage points (standard error 0.69). Excluding the six Expenditure Classes with raw goods (fresh food and energy) — which are nondurable, flexibly priced, and often dropped from the data for business cycle analysis — the relationship becomes significantly positive. For processed goods, each year of durability goes along with 1.47 percentage points higher frequency (standard error 0.41), so that a good lasting 10 years tends to have 14 percentage points higher price change frequency than a nondurable. The connection is similar for regular price changes among processed goods.

The positive correlation between durability and price flexibility (at least for processed goods and services) could have important implications for business cycles. The more durable the good, the more cyclical expenditures and production tend to be. This is true in both theory and practice (see Bils & Klenow, 1998, for one example). Barsky, House, and Kimball (2007) presented a model in which monetary non-neutrality is closely connected to price stickiness for durables, with the stickiness of nondurables of no importance. That said, we hasten to reiterate that the relationship in the data is not significant if raw good categories are included.

Durability and cyclicalities are not synonymous. We therefore gauged cyclicalities directly for each of 64 BLS Expenditure Classes based on the coefficient from regressing its NIPA real consumption expenditure growth on NIPA aggregate real consumption growth. For each Expenditure Class we ran a single OLS regression (with a constant) for the available NIPA sample from February 1990 through March 2009, which is close to our CPI-RDB 1988–2009 monthly sample. Figure 3 shows a clear positive association between price change frequencies and cyclicalities across the Major Groups. Inside Transportation, for example, Motor Vehicles stands out in terms of its combination of price flexibility (38% frequency), cyclicalities (5.7% higher expenditure growth for every 1% higher aggregate consumption growth), and sampling weight (16% of the nonshelter sample). Apparel is also fairly flexible (30% frequency if one includes sale prices) and fairly cyclical (cyclicalities coefficient 1.75). In Table 6, the WLS regression coefficient of price frequency on cyclicalities across the 64 Expenditure Classes is 3.23 percentage points (standard error 1.05). If we look only at the 58 Expenditure Classes for processed goods, the WLS coefficient is similar at 3.29 (but more precisely estimated, with a standard error of 0.59). The relationship is stronger still for regular price changes across processed goods at 3.79 (standard error 0.48).⁷

The frequency–cyclicalities nexus could arise because cyclicalities induces price changes, the other way around, or because they share driving forces. If price flexibility

⁷ Frequency is positively correlated with cyclicalities across categories even when controlling for durability. When we regress the frequency of regular price changes for processed goods on durability and cyclicalities across 64 ECs, the coefficient on durability is 0.34 (SE 0.05) and the coefficient on cyclicalities is 9.24 (SE 1.39).

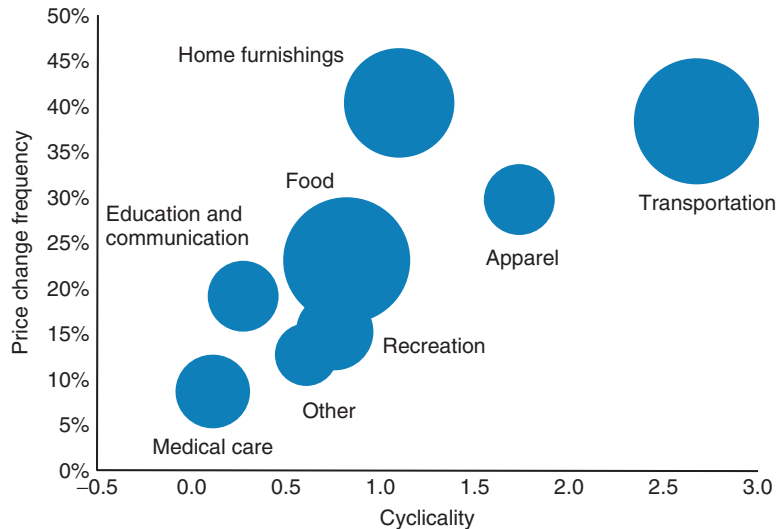


Figure 3 Frequency versus cyclical frequency in the U.S. CPI. *Source:* CPI-RDB. Data are for the top three cities (New York, Los Angeles, and Chicago) January 1988 through October 2009. Each circle is one of the eight Major Groups in the CPI (1998-onward definition), with the area proportional to the average expenditure weight over the sample. Prices include sales and regular prices. Frequency is calculated as the weighted mean across ELIs (with each ELI mean itself a weighted average across quotes within the ELI). Cyclical frequency is based on an OLS regression coefficient for each Major Group of *real* monthly consumption growth on aggregate real monthly consumption growth, using seasonally adjusted Detailed Expenditure data from the U.S. National Income and Product Accounts, January 1990 through May 2009.

is responding to cyclical shocks, then the pattern may suggest more macro price flexibility than in models where the variation in price change frequency across goods is exogenous, as in Carvalho (2006), or reflects variation in noncyclical factors, as in Nakamura and Steinsson (2008b). Causality running from frequency to cyclical frequency would presumably work for “supply” shocks (for which price flexibility should amplify the response of real expenditure growth) but not for “demand” shocks (for which price flexibility would dampen the response of real expenditure growth; see Bills, Klenow, & Kryvtsov, 2003).⁸

Heterogeneity is also evident for producer prices. Nakamura and Steinsson (2008a) reported a median implied duration of 8.7 months for finished producer goods, 7.0 months for intermediate goods, and 0.2 months for crude materials from 1998 to 2005. Within finished producer goods, the median frequency of price change ranges

⁸ We do not know how the analysis would change with shelter. On the one hand, rents and owner equivalent rents may be sticky. On the other hand, shelter *quantities* may be just as sticky — contrary to the Keynesian paradigm of flexible quantities relative to prices. And housing services are presumably less cyclical than housing construction, for which prices may be more flexible.

from 1.3% for “Lumber and Wood Products” to 87.5% for “Food Products”. Vermeulen et al. (2007) also documented significant heterogeneity across sectors and investigated the causes. Firms with a higher labor share in total costs tend to change price less frequently, whereas firms with higher shares of energy and nonenergy intermediate goods change price more frequently. Moreover, they find that the higher the degree of competition, the higher the frequency of price changes, particularly price decreases.

Another dimension of heterogeneity that apparently affects the frequency of price changes is the type of establishment at which goods are sold. In Europe, consumer prices are more flexible in large outlets, such as supermarkets and department stores, than in smaller retail outlets (e.g., Dias, Dias, & Neves, 2004; Fabiani, Gattulli, Sabbatini, & Veronese, 2006; Jonker, Folkertsma, & Blijenberg, 2004). For U.S. producer prices, Goldberg and Hellerstein (2009) found that large firms change prices two to three times more frequently than small firms. Survey studies have also reported similar patterns (e.g., Amirault, Kwan & Wilkinson, 2006; Buckle & Carlson 2000; Fabiani et al., 2005).

3.3 Sales, product turnover, and reference prices

One lesson from the theoretical price-setting literature is that different types of price adjustments (e.g., transitory or permanent, selected or random) can have substantially different macroeconomic implications. Researchers have thus investigated how measures of the frequency of price change are altered when the data are filtered in different ways, such as excluding temporary sales and product turnover. It turns out that the answer can vary considerably depending on how exactly this is done.

In the U.S. CPI data, a “sale” price is (a) temporarily lower than the “regular” price, (b) available to all consumers, and (c) usually identified by a sign or statement on the price tag. Klenow and Kryvtsov (2008) reported that roughly 11% of the prices in their sample are identified as sale prices by BLS price collectors. Another approach is to use a “sales filter” to identify “V-shaped” patterns in the data as sales. Nakamura and Steinsson (2008a) reported results for a variety of sales filters, allowing for asymmetric and multi-period Vs. Concerning product turnover, “forced item substitutions” occur when an item in the sample has been discontinued from its outlet and the price collector identifies a similar replacement item in the outlet to price going forward, often taking the form of a product upgrade or model changeover. The monthly rate of forced item substitutions is about 3% in the BLS sample.⁹

Table 7 demonstrates the impact on the implied duration of prices of applying various filters to the data. We follow Klenow and Kryvtsov (2008) but with a U.S. CPI sample that extends through October 2009 (rather than January 2005). Depending on the treatment of sales, the median duration of prices ranges from 3.4 months (all

⁹ See Broda and Weinstein (2007) for the importance of product turnover in AC Nielsen Homescan data.

Table 7 U.S. CPI price durations under various exclusions

Case	Implied median duration	Implied mean duration
Posted prices	<i>3.4 months</i>	6.2 months
No V shapes	<i>5.0</i>	7.1
Like prices	<i>5.9</i>	7.2
Regular prices	<i>6.9</i>	8.0
No substitutions	<i>8.3</i>	10.1
Adjacent prices	<i>9.0</i>	11.2
1988–1997 posted	<i>3.6</i>	<i>6.3</i>
1998–2009 posted	<i>3.4</i>	<i>6.6</i>

Notes: Source is CPI-RDB. Data are for the top three cities (New York, Los Angeles, and Chicago) from January 1988 through October 2009. The implied durations are inverses of the monthly frequencies.

Means and medians use ELI weights based on the BLS consumer expenditure surveys and unpublished weights for each quote based on BLS point-of-purchase surveys.

No V shapes: lower middle prices are replaced with identical neighbors.

Like prices: regular (sale) price is compared only to the previous regular (sale) price.

Regular prices: posted prices excluding sale prices.

No substitutions: only regular prices in between item substitutions are compared.

Adjacent prices: only consecutive monthly regular prices in between substitutions.

1988–1997 only: sample is confined to 1988–1997 (posted prices).

1998–2009 only: sample is confined to 1998–2009 (posted prices).

prices included) to 6.9 months (excluding BLS-flagged sales). The one-period V-shape filter — in which every time a middle price is lower than its identical neighbors it is replaced by its neighbors — produces an intermediate duration of 5.0 months, reflecting that many sales, such as clearance sales, are not V-shaped.¹⁰ “Like” prices compare a regular price only to the previous regular price and a sale price only to the previous sale price, thus allowing for the possibility that sale prices are sticky even if they do not return to the previous regular price. This raises the implied median duration to 5.9 months. The fact that “like” prices change more frequently than regular prices (every 5.9 months vs. every 6.9 months) indicates a sale price is more likely to differ from a previous sale price than a regular price is to differ from a previous regular price.

Table 7 also shows that removing all forced item substitutions from the data increases the median implied duration to 8.3 months from 6.9 months. Klenow and Kryvtsov (2008) reported that item substitutions display price changes about 80% of the time, much more frequently than the average over a product’s life cycle. Purging substitution-related price changes can imply prices have a longer duration than the products themselves: for apparel, regular prices excluding substitutions change about every

¹⁰ Nakamura and Steinsson (2008a) found that removing BLS-flagged sales generated a higher estimate of price duration than did any of their V-shaped sales filters.

27 months, whereas the average item lasts only 10 months. Finally, comparing only *consecutive* regular monthly prices between substitutions pushes the duration up from 8.3 months to 9.0 months. Price changes are more frequent after items return to stock, come back into season, or return from sales.

Nakamura and Steinsson (2008a) underscored that sales and forced item substitutions are much more important in some categories than others in the U.S. CPI. For example, 87% of price changes in Apparel and 67% in Home Furnishings are sale-related price changes, while Utilities, Vehicle Fuel, and Services have virtually no sale-related changes. The monthly rate of forced item substitutions is about 10% in Apparel and in Transportation, compared to 3% for all goods. This uneven distribution of sales and substitutions is important for explaining why excluding them has a sizable impact on the implied median duration of U.S. consumer prices: the sectors in which sales and substitutions are concentrated are those with a frequency of price change that is relatively close to the median.

Sales have become more important over time for U.S. consumer prices but continue to play a small role in other countries. Nakamura and Steinsson (2008a) documented a strong increase in the frequency of U.S. sales from 1988 to 2005, especially in processed food and apparel where the frequency of sales doubled. Available evidence from European countries suggests that sales are a less important source of price flexibility. Wulfsberg (2009) reported that sale prices account for only 3% of price observations in Norwegian CPI data, and removing these observations increases the mean duration by only 0.3 months. Dhyne et al. (2006) similarly reported that sales have small effects on the estimated frequency of price change in France and Austria.

As emphasized by Mankiw and Reis (2002), Burstein (2006), Woodford (2009) and others, price changes may be part of a sticky plan and hence fail to incorporate current macro information. This can be true of regular price changes, not just movements between regular and sale prices. One sign of such a plan might be the existence of only a few prices over the life of an item. In the top three cities of the U.S. CPI, the weighted median (mean) length of a quote-line is 50 (53) monthly prices. In Table 8 we report the cumulative share of price quotes represented by the “top” (i.e., most common) 1, 2, 3, and 4 prices over the quote-line. The median (mean) share of the most common price is 31% (38%). The top four prices over the typical 4.3 year quote-line together represent a median (mean) of 70% (66%) of all prices. Table 8 reports the figures for Major Groups as well. There is a natural tendency for higher shares where there are shorter quote-lines (apparel) or less frequent nominal price changes (medical care, recreation). Relative to its moderate frequency of price changes, Food stands out in having a 42% median for the top price and 86% median for the top 4 prices. Even Transportation — which is highly cyclical and exhibits frequent price changes — has a 16% median share of the top price and 46% median share of the top 4 prices.

Table 8 Share of “top” prices in each U.S. CPI quote-line

Median % (Mean %)	Top price	Top 2 prices	Top 3 prices	Top 4 prices	# of quotes
All items	31.4 (37.6)	50.9 (53.2)	62.7 (61.3)	70.1 (66.2)	50 (53.1)
By major group					
Apparel	36.7 (41.2)	58.3 (59.9)	71.4 (70.5)	79.0 (77.0)	38 (41.2)
Education and communication	29.2 (33.1)	47.4 (50.1)	61.3 (61.7)	73.0 (69.3)	42 (52.3)
Food	42.4 (47.0)	66.7 (66.7)	79.2 (76.3)	85.5 (81.3)	51 (51.7)
Home furnishings	24.0 (30.5)	40.0 (43.5)	50.0 (50.4)	56.9 (54.7)	56 (63.4)
Medical care	48.9 (53.3)	72.1 (71.1)	82.8 (79.0)	88.0 (84.1)	50 (49.5)
Recreation	40.4 (46.2)	64.9 (65.4)	77.6 (75.4)	85.1 (81.2)	49 (48.9)
Transportation	16.3 (22.6)	28.6 (34.8)	38.0 (43.1)	45.9 (49.8)	49 (53.1)
Other goods and services	50.0 (54.1)	70.6 (68.5)	79.3 (74.6)	84.0 (78.2)	51 (52.6)

Note: Source is CPI-RDB. Data are for the top three cities (New York, Los Angeles, and Chicago) from January 1988 through October 2009. Both regular and sale prices are included. Entries are weighted medians (means) across quote-lines of the top (most common) n prices in each quote-line as a share of all prices in the quote-line.

Given that nominal prices change every four months or so, there are on average around 13 prices per quote-line. A Taylor model (and perhaps Calvo model as well) would therefore imply a *median* top 2 price share of less than 20%, whereas the actual share in the data is over 50%.¹¹ The disproportionate importance of a few prices appears supportive of downward-sloping hazards and/or sticky nominal plans.¹² In favor of the former, two-thirds of top price spells are uninterrupted by other prices. Before discarding less common prices, however, research could explore how aggregate quantities produced and sold relate to changes in common versus rare prices. It is conceivable that cyclical quantities are sensitive to the rare prices (e.g., clearance sales in apparel).

Eichenbaum et al. (2009) usefully proposed a way of measuring sticky “reference” prices amidst shorter lived new prices. Using weekly price data covering 2004–2006 from a large U.S. supermarket chain, they defined the reference price for each UPC as the modal price in each quarter. They found such reference prices are responsible

¹¹ The prediction of a menu-cost model for the top price shares would be more sensitive to the distribution and timing of large idiosyncratic changes in the desired price. In Golosov and Lucas (2007) high variance shocks are realized every period, whereas in Gertler and Leahy (2008) they followed a Poisson process.

¹² In addition to costs of collecting and processing information and formulating and implementing new plans, the use of a few prices may reflect “price points.” Levy et al. (2007) found prices ending in 9 are most common (whether in cents, dollars, or tens of dollars), less likely to change, and change by bigger amounts.

for 62% of all weekly prices and 50% of quantities sold. Importantly, they reported that reference prices only change every 11.1 months. This is considerably stickier than regular (nonsale) prices in the same supermarket, which change about once per quarter. They went on to present a simple model in which the frequency of reference price changes is the key to monetary non-neutrality, as deviations from reference prices largely reflect idiosyncratic considerations.

Do sticky reference prices exist in the U.S. CPI more broadly? The CPI data is monthly, so it is not possible to implement the exact Eichenbaum et al. (2009) methodology on the CPI. We instead defined the reference price in each month (for an item) as the most common price in the 13-month window centered on the current month. We broke ties in favor of the current price. An advantage of a rolling window is that it allows the reference price to change every month, whereas the calendar definition imposes at most one reference price change per quarter. Using this 13-month window, we find that the posted price equals the reference price 78.5% of the time on average (84.2% median) when looking across weighted quote-lines. Table 9 provides reference price statistics for all items and Major Groups. The share of reference prices is modestly higher in Food (88.1% median). The only Major Group with a reference price share below 70% is Transportation (62.5% median); albeit an important exception given its cyclicity. We find that the weighted median duration of reference prices is 11.0 months across ELIs in the CPI. Note that reference prices change less frequently than regular prices (median duration 6.9 months), so that some regular price changes must be temporary deviations from reference prices. The median duration is higher for Food at 13.5 months, and lower for Transportation at 6.1 months. We conclude that the Eichenbaum et al. (2009) reference price phenomenon extends not only to most food items, but to most items in the nonshelter CPI more generally.

We add several caveats to our reference price results for the U.S. CPI. First, our definition of reference prices is not strictly comparable to that of Eichenbaum et al. (2009). With our definition, a combination of a high share and high duration of references — which we do observe — is more suggestive of stickiness than either of these without the other. Second, there is considerable variation in the share of reference prices across quote-lines (weighted standard deviation 41%), even within product categories (see Table 9). Third, it is possible that cyclical quantities are sensitive to nonreference prices along with reference prices.

A final caveat has to do with the modeling implications that can be drawn from the reference-price findings. Although reference prices constitute a large share of total prices and have long durations, our statistics do not imply that sellers choose prices from a small set or that prices display “memory”. Indeed, in the U.S. CPI, we found that only 30% of deviations from reference prices *ever* return to the previous reference price. We next consider two statistics that may be more directly revealing about these issues.

Table 9 Share of “reference” prices in each U.S. CPI quote-line

	Median %	Mean %	S.D. %	Median duration (months)
All items	84.2	78.5	41.0	11.0
By major group				
Apparel	77.8	77.1	35.1	11.0
Education and communication	93.1	83.0	57.4	7.3
Food	88.1	82.2	27.2	13.5
Home furnishings	82.1	77.9	46.8	6.9
Medical care	100.0	93.7	27.1	18.8
Recreation	92.3	87.4	36.3	14.6
Transportation	62.5	64.2	70.4	6.1
Other goods and services	96.3	89.4	36.9	19.3

Note: Source is CPI-RDB. Data are for the top three cities (New York, Los Angeles, and Chicago) from January 1988 through October 2009. Both regular and sale prices are included. “Reference” prices are the most common prices in the 13-month window centered on the current month within each quote-line. The share of reference prices in all prices was calculated for each quote-line, and then the weighted median (mean) and standard deviation of this % was calculated across quote-lines. We calculated the weighted median duration of reference prices across Major Groups using MLE as in Klenow and Kryvtsov (2008).

The first statistic is the fraction of prices that are “novel,” which we define as prices that do not appear in any of the previous 12 months for the same item (quote-line). For the top three cities of the U.S. CPI, we find the weighted median (mean) share of prices that are novel to be 16.1% (25.0%). These fractions are consistent with genuinely new prices every four to seven months. We take this to mean that deviations from the most common prices exhibit considerably novelty. Table 10 also breaks the statistic down by Major Group. Novelty is naturally correlated with the frequency of price changes. Food exhibits less than average novelty (10.3% median vs. the overall item median of 16.1%, 13.8% average vs. the overall average of 25.0%) despite having average frequency of price changes. Thus caution may be warranted in drawing lessons from grocery store scanner data sets. Prices appear more novel in more cyclical categories (Transportation, Apparel). If cyclical quantities are linked to these novel prices, they could well contribute to macro price flexibility.

We also compute the fraction of prices that are “comeback” prices. We define the current price as a comeback price if the same price appeared any time in the previous 12 months with a *different* price occurring at least once in between. As a hypothetical example, we would label a current price of \$10 as a comeback price if the price was stuck at \$10 for the previous six months, was at \$11 seven months ago, but was also at \$10 eight months ago. For the top three cities of the U.S. CPI, Table 11 reports

Table 10 Share of prices that are “novel” in each U.S. CPI quote-line

	Median	Mean
All items	16.1	25.0
By major group		
Apparel	21.2	26.3
Education and communication	23.3	28.3
Food	10.3	13.8
Home furnishings	23.6	35.9
Medical care	6.7	8.2
Recreation	9.9	12.9
Transportation	35.0	38.3
Other goods and services	7.1	10.7

Note: Source is CPI-RDB. Data are for the top three cities (New York, Los Angeles, and Chicago) from January 1988 through October 2009. Both regular and sale prices are included. “Novel” prices are prices that did not appear in any of the previous 12 months for the same quote-line.

Table 11 Share of “comeback” prices in each U.S. CPI quote-line

	Median	Mean
All items	0.8	14.0
By major group		
Apparel	19.5	25.2
Education and communication	0	4.8
Food	14.9	24.7
Home furnishings	6.3	16.4
Medical care	0	3.7
Recreation	0	11.0
Transportation	0	6.2
Other goods and services	0	8.8

Note: Source is CPI-RDB. Data are for the top three cities (New York, Los Angeles, and Chicago) from January 1989 through October 2009. Both regular and sale prices are included. “Comeback” prices are current prices that both occurred and were interrupted in the previous 12 months for the same quote-line.

the weighted median (mean) share of comeback prices to be 0.8% (14.0%). The Major Groups with the highest share of comeback prices are Apparel and Food, which have means of around 25%. The share of comeback prices for the median quote-line is zero, however, in five of the eight Major Groups.

The upshot is that, outside of Apparel and Food, there appears to be little memory in monthly U.S. CPI prices. A corollary is that most reference prices are not comeback prices. The typical nonreference price must be a short-lived (transition) price in between successive reference prices, at least as far as we can tell. It is possible that monthly observations are obscuring many temporary price changes we would see if we had weekly or even daily data. This is most plausible for Apparel and Food, where a nontrivial share of comeback prices occur after or during sales (roughly three-fourths of comeback prices in apparel, and roughly one-half of comeback prices in food). It would appear less likely for services, such as medical care, where monthly prices largely go from one reference price to another.

3.4 Determinants of frequency

Researchers have also investigated factors *affecting* the frequency of price change. Some studies have made use of the substantial variation in frequencies along different dimensions to identify important determinants, while others have directly asked price-setters to assess the importance of various theories of price stickiness. A (nonexhaustive) list of determinants includes (a) the level and variability of inflation, (b) the frequency and magnitude of cost and demand shocks (broadly construed to include price discrimination), (c) the structure and degree of market competition (including regulation of discounts), and (d) the price-collecting methods of statistical agencies (e.g., do they report temporary sales?).

We begin by looking at the cross-country evidence reported in Table 1. Following Golosov and Lucas (2007) and Mackowiak and Smets (2008), Figure 4 simply plots the mean frequency of price change against the average inflation rate for these studies. The OLS regression coefficient of price frequency on inflation is 17.0 (standard error 6.8). Of course, this exercise comes with a few well-known caveats. First, the studies differ in many details, such as sample composition and different price-collecting methodologies. Second, periods of high inflation are often periods of volatile inflation, so it is unclear whether the relationship in Figure 4 reflects the importance of the level or the variability of inflation (probably both). Still, the relationship is provocative.

Dhyne et al. (2006) investigated the impact of inflation (and other factors) on the frequency of price change by running regressions on European data. They regressed the frequency of price change across 50 product categories in 9 countries on dummy variables for product type (unprocessed and processed food, energy, nonenergy industrial goods, and services), country dummies, the mean and standard deviation of inflation at the product category level, an indicator for whether sale prices occur and are reported, the share of prices set at attractive levels (“price points”), and an indicator for whether the price is typically regulated. They found that mean inflation is *not* significantly correlated with the overall frequency, but is correlated with the frequency of increases and decreases separately. The overall frequency (and increases and decreases,

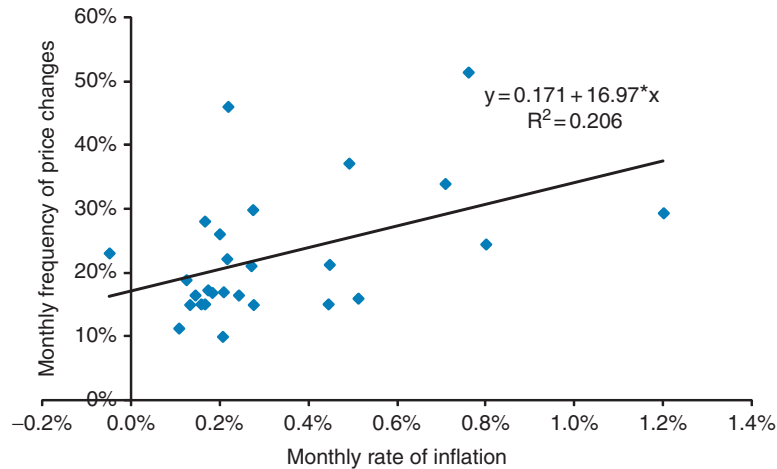


Figure 4 Cross-country frequency versus inflation. Notes: Each data point represents one study from Table 1, with the Gouvea (2007) study excluded because of its overlap with the Barros et al. (2009) study. The monthly frequency of price change is as reported in Table 1, while the monthly rate of inflation is based on the authors' calculations.

respectively) are significantly higher in sectors in which the variability of inflation is higher. A higher frequency of price change is also found when sales and temporary price cuts are included, when the share of attractive prices is lower, and when prices are not subject to regulatory control.

For U.S. consumer prices, Bils and Klenow (2004) considered regressions relating the frequency of price change in different product categories to measures of market structure: the concentration ratio, wholesale markup, and rate of noncomparable substitutions in those categories. After controlling for whether a good is raw or processed, they found that the first two measures do not have statistically significant effects, while the rate of product turnover remains a robust predictor of the frequency of price changes. They interpreted the role of product turnover and raw materials in explaining the frequency of price changes as reflecting the importance of the volatility of shocks to the supply of and demand for goods.

Boivin, Giannoni, and Mihov (2009) found a relationship between the volatility of sectoral shocks and the frequency of price change. Using disaggregated PCE inflation series, they disentangled inflation fluctuations due to sector-specific conditions from those due to macroeconomic factors. They found that sectors with more volatile sector-specific shocks have a higher frequency of price change.

Other studies provide evidence consistent with the importance of cost shocks, as goods with more sticky input prices tend to change price less frequently. For example, Nakamura and Steinsson (2008a) reported a high correlation between the frequency of price changes upstream (PPI) and downstream (CPI), and Eichenbaum et al. (2009)

found a similar pattern using detailed cost and price data for one large U.S. retailer. Looking at Euro Area producer prices and noting that wages tend to be stickier than goods prices, Vermeulen et al. (2007) found that goods with a higher labor share in total costs tend to change price less frequently, whereas firms with a higher share of intermediate goods change prices more frequently. Peneva (2009) matched categories of U.S. consumer goods to the manufacturing industries in which they are produced and also found that higher labor intensity is associated with less frequent price changes.

Survey data also provides useful insights on the determinants of price change. Fabiani et al. (2005) reported the top four reasons Euro Area firms refrain from changing prices: (1–2) implicit and explicit contracts with customers, (3) cost-based pricing (i.e., input costs are slow to change), and (4) coordination failure (not wanting to raise one's price out of fear of losing market share to competitors who do not follow suit). These reasons were also ranked in the top five by firms in the United States (Blinder et al., 1998), UK (Hall, Walsh, & Yates, 2000), Sweden (Apel, Friberg, & Hallsten, 2005), and Canada (Amirault et al., 2006). On the other hand, physical costs (menu costs) and costly information are among the reasons least favored by firms. Finally, the main impediments to more frequent price adjustment appear to be associated with price *changes* rather than price *reviews*; that is, surveyed firms report reviewing prices much more often than changing prices.

Drilling down further into the survey evidence for cost-based pricing and, in particular, the role of labor costs, Álvarez and Hernando (2007b) reported that Spanish sectors with relatively high labor costs tend to contain a small number of firms that change prices often. Druant et al. (2009) provided insight into the relationship between wage and price rigidity based on a survey on wage and pricing policies of Euro Area firms. They found that 40% of firms indicate a relationship (formal or informal) between the timing of their wage and price adjustment decisions. Moreover, firms with a higher labor cost share report a tighter link between wage and price changes and a lower frequency of price adjustment (as wages change less frequently than prices). Finally, even accounting for the likely simultaneity between price and wage changes, a statistically significant relationship is found, running from the frequency of wage changes to that of prices.

Dhyne et al. (2006) tried to account for the higher frequency of price changes in the United States than in the Euro Area. The United States had somewhat higher level and volatility of inflation, but to arguably little effect. Differences in consumption patterns do not help at all, as the expenditure share of more flexible components of the CPI is actually larger in the Euro Area. Heterogeneity of outlets may play a role: smaller shops, which change prices less frequently, have a higher market share in Euro Area. Differences in occurrence and treatment of temporary sales are important. For example, 1 in 5 price changes is related to sales in the United States, compared to less than 1 in 8 in France. Many other Euro Area countries do not record sale prices. Finally, a higher variability of wages (and other input prices) and less anti-competitive regulation may help explain the higher frequency of price changes in the United States.

4. SIZE OF PRICE CHANGES

We now move from the extensive margin (how often prices change) to the intensive margin (how large are the price changes). Again, there is substantial heterogeneity in the micro data, and it is thus useful to characterize the distribution of the size of price changes.

4.1 Average magnitude

A common finding across studies is that price changes are large on average. For example, in the U.S. CPI Klenow and Kryvtsov (2008) reported a mean (median) absolute change in posted prices of 14% (11.5%), while regular price changes are smaller but still large with a mean (median) of 11% (10%). The average consumer price decrease (increase) is 10% (8%) in the Euro Area (Dhyne et al., 2006), and emerging markets also display large changes (e.g., Barros, Bonomo, Carvalho, & Matos, 2009; Konieczny & Skrzypacz 2005). For U.S. finished goods producer prices, Nakamura and Steinsson (2008a) reported a median magnitude of 7.7%.

Given the low level and volatility of aggregate inflation in the United States and Europe, most price changes are not simply keeping up with overall inflation (i.e., indexed). But perhaps many micro price changes incorporate idiosyncratic or sectoral considerations, but not aggregate shocks. See Mackowiak and Smets (2008) for further discussion of this rational inattention hypothesis in the context of large micro price changes.

4.2 Increases versus decreases

A second feature of the size distribution is that price declines are very common. Nakamura and Steinsson (2008a) reported that around 40% of both CPI and PPI monthly price changes in the United States are decreases, and Dhyne et al. (2006) reported similar numbers for the Euro Area. These facts help reconcile the finding of large average *absolute* price changes with small average price changes (14% vs. 0.8% according to Klenow & Kryvtsov, 2008) and suggest an important role for idiosyncratic shocks (or price discrimination) in driving price changes. The prevalence of price declines also varies across sectors; in particular, they are relatively uncommon in the services sector (Dhyne et al., 2006; Nakamura & Steinsson 2008a).

4.3 Higher moments of the size distribution

Using scanner data, Midrigan (2009) emphasized that the distribution of nonzero price changes has more weight in the vicinity of zero than predicted by a normal distribution, while the tails are somewhat fatter. Formally, the distribution of price changes is leptokurtic (i.e., has positive excess kurtosis). We have confirmed this pattern in the U.S. CPI data, where the kurtosis of the price change distribution is 10.0 for posted prices and 17.4 for regular prices (vs. 3 for a normal distribution). As maintained by Midrigan, fat tails suggest weaker selection, with more price changes large and

therefore inframarginal. For this reason, the fraction of price increases versus decreases can be less sensitive to macro shocks (see also Gertler & Leahy, 2008).

Other studies have documented the prevalence of small price changes more directly. In the U.S. CPI, around 44% of regular price changes are smaller than 5% in absolute value, 25% are smaller than 2.5%, and 12% are smaller than 1% (Klenow & Kryvtsov, 2008). Note that this is not just due to frequent shopper cards (as in scanner data that report average weekly prices inclusive of coupons and frequent shopper discounts). Wulfsberg (2009) reported that 45% of price changes are smaller than 5% in Norway, while in Brazil, where the mean absolute size is 13%, over one-third of price changes are smaller than 5% (Barros et al., 2009). For producer prices, Vermeulen et al. (2007) found that one-quarter of both increases and decreases are smaller than 1%, compared to a mean price change of 4% in the Euro Area.

5. DYNAMIC FEATURES OF PRICE CHANGES

In addition to the unconditional statistics that we have highlighted so far, researchers have documented a number of features concerning how prices change over time. These include the synchronization of price changes, how the frequency and size of price changes correlate with the duration of the existing price, and the response of prices to shocks that would be expected to alter a firm's desired price. These features are of interest because in the presence of nominal stickiness (like we see in the data), price-setters have dynamic decision problems, and thus, dynamic features of the data are particularly helpful in distinguishing between the various theories of price-setting.

5.1 Synchronization

Since at least Taylor (1980), staggered price-setting has played an important role in modeling persistent real effects of monetary shocks. The staggering of price adjustments is readily apparent from the observation that not all prices change in any given period. Moreover, when price-setters do change their price, they often take the prevailing price of their competition into consideration (Levy, Dutta, Bergen, & Venable, 1998), which only makes sense if (at least some) competitors' prices are expected to remain active for a period of time. Recent studies have looked at time variation in the frequency of price changes as a measure of the extent of synchronization (or, conversely, uniform staggering) in price-setting. Time variation also provides relevant evidence for distinguishing between some time-dependent pricing models (TDP) versus state-dependent pricing models (SDP).

Klenow and Kryvtsov (2008) decomposed monthly inflation (π_t) into the fraction (f_{r_t}) of items with price changes and the average size (sz_t) of those price changes: $\pi_t = f_{r_t} \times sz_t$. In their sample (U.S., 1988–2004), they found the fraction to be relatively stable and not so correlated with inflation (correlation 0.25), while the average

size was more volatile and had comoved almost perfectly with inflation (correlation 0.99). In addition, they decompose the variance of inflation over time into an “intensive margin” (IM) and “extensive margin” (EM) as follows:

$$\text{var}(\pi_t) = \underbrace{\text{var}(sz_t) \cdot \bar{fr}^2}_{IM} + \underbrace{\text{var}(fr_t) \cdot \bar{sz}^2 + 2 \cdot \bar{fr} \cdot \bar{sz} \cdot \text{cov}(fr_t, sz_t)}_{EM} + O_t$$

This decomposition is interesting because different models of price-setting have distinct implications for this decomposition. For example, in staggered TDP models, the intensive margin will account for all of inflation’s variance, whereas the fraction of items changing price plays a substantial role in some SDP models, such as Dotsey, King, and Wolman (1999). Klenow and Kryvtsov (2008) found that the IM term accounts for between 86% and 113% of the variance of inflation, implying that the fraction of items changing price are a relatively unimportant source of fluctuations in inflation, at least for their sample period. Note that other SDP models (such as Golosov & Lucas, 2007) can fit this decomposition; the key is to include (realistically) large idiosyncratic price changes so that aggregate shocks have offsetting effects on the frequency of increases versus decreases.

Gagnon (2009) first emphasized the usefulness of further decomposing inflation into terms due to price increases and decreases. Note that $fr = fr^+ + fr^-$ and $sz = fr^+ sz^+ - fr^- sz^-$, where fr^+ and fr^- (sz^+ and sz^-) denote the frequency (absolute size) of price increases and decreases, respectively. So, even if the average size of price increases and decreases remain constant over time, significant variation in the average size of price changes could result due to offsetting movements in the frequency of price increases and decreases, which would also imply little variation in the overall frequency of price changes. This is the type of pattern we see in the United States. Klenow and Kryvtsov (2008) reported that a 1 percentage point increase in inflation is associated with a 5.5 (–3.1) percentage point change in the fraction of price increases (decreases), and a 0.6 (–1) percentage point change in the size of price increases (decreases). When they additively decompose the variance of inflation (splitting the covariance term), they find fluctuations in increases and decreases equally important to inflation volatility.¹³

While inflation was relatively low and stable in the United States during the sample period considered by Klenow and Kryvtsov (2008), Gagnon (2009) noted that Mexico experienced episodes of both high and low inflation from 1994 to 2002. He found that when the annual rate of inflation is below 10–15%, the average frequency (size) of price changes comoves weakly (strongly) with inflation due to offsetting movements in the frequency of price increases and decreases. When inflation rises beyond

¹³ Nakamura and Steinsson (2008a), on the other hand, found that the frequency of price increases, and not decreases, are important for driving inflation movements. They look at the median frequency of price changes across sectors rather than cross-sector means.

10–15%, however, few price decreases are observed and both the frequency and average size are important determinants of inflation. Thus, over the entire sample period, the extensive margin accounts for more than half of the variance of inflation.

Wulfsberg (2009) studied Norwegian consumer price data over a 30-year period in which inflation was at first high and volatile (1975–1989) and then low and stable (1990–2004). He used a different metric to assess the importance of the extensive and intensive margins. Specifically, he constructed the average monthly inflation in year t as the weighed product-sum of item-specific average frequencies and magnitudes of price changes, $\hat{\pi}_t = \sum \omega_{i,t} [f_{i,t}^+ \cdot s_{i,t}^+ + f_{i,t}^- \cdot s_{i,t}^-]$, where $f_{i,t}^+$ is the average frequency of price increases, and so on. To assess the importance of the extensive (intensive) margin, he computed the conditional inflation rate where the size (frequency) of price changes are kept constant at their means while allowing the frequency (size) of price changes to vary. He found that the EM is strongly correlated with CPI inflation (0.91), while the IM is negatively correlated with CPI inflation (−0.12) and interpreted this as evidence of strong state dependence in price-setting. Restricting his sample to the low and stable inflation period of 1990–2004, he found that the IM has a higher correlation with CPI inflation: 0.51 versus 0.36.

We next consider evidence of price-change synchronization in the United States during the recent recession. Figure 5 plots the average frequency of price changes in the top three cities, based on regular prices for processed items. We first calculated the monthly frequency of price changes (the weighted mean across ELIs) from April 1988 through September 2009. We then took out seasonal (monthly) dummies, and calculated deviations from them. Finally, we averaged the monthly deviations in each quarter, and added back the mean across all quarters to produce quarterly data from 1988:2 through 2009:3. The *monthly* frequency of price changes increased a couple of percentage points from the end of 2007 onward, from about 18.5% to about 21%. We do not have a good metric for whether this is an economically large shift, but the issue deserves deeper investigation. For example, did the frequency increase more rapidly for cyclical goods? Did the increase reflect endogenously more attention given the magnitude of the recession? We do know that the increase holds for non-raw posted prices as well as all posted prices. Also, the size of increases and decreases for regular processed items were little changed (if we include raw items, there were some unusually large energy price declines near the end of the sample).

Another form of synchronization is seasonality. Nakamura and Steinsson (2008a) found that the weighted median frequency of (regular) consumer price changes declined monotonically over the four quarters of the year, with local spikes in the first month of each quarter. The quarterly seasonal pattern in producer prices mirrors the seasonal pattern in consumer prices qualitatively, but is substantially larger, with the frequency of price change in January more than twice the average for the rest of the year. For the Euro Area, Dhyne et al. (2006) emphasized that various goods are more likely

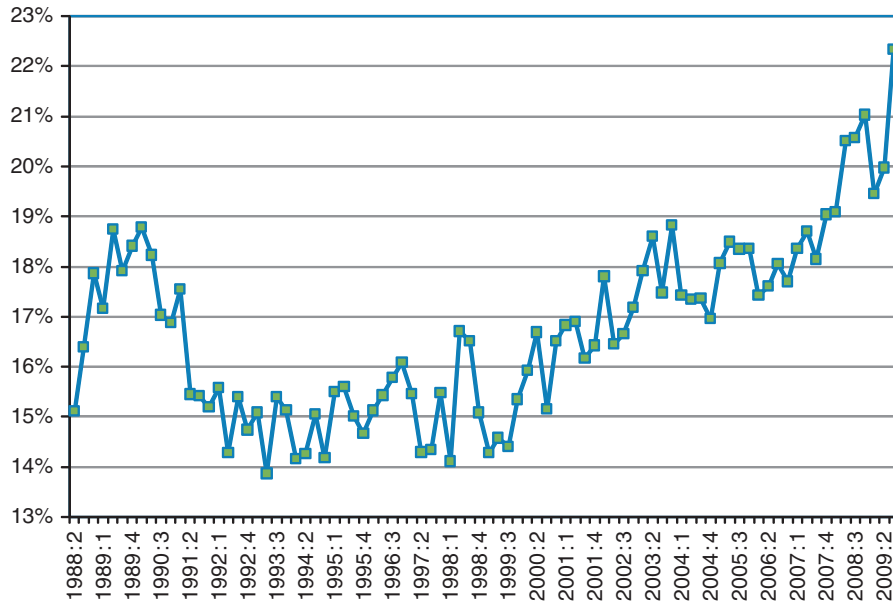


Figure 5 U. S. Monthly CPI frequency over time. *Source:* CPI-RDB. Data are for the top three cities (New York, Los Angeles, and Chicago), based on regular prices for processed items. We first calculated the monthly frequency of price changes (weighted mean across ELIs) from April 1988 through September 2009. We then took out seasonal (monthly) dummies, and calculated deviations from them. Finally, we averaged the monthly deviations in each quarter, and added back the mean across all quarters, to produce quarterly data from 1988:2 through 2009:3.

to exhibit seasonal patterns: unprocessed foods due to seasonality in agricultural producer prices, certain industrial goods due to “end-of-season” sales, and services because they show an inclination to change prices at the beginning of the year and refrain from changing price at the end of the year.

The Euro Area consumer price studies (Dhyne et al., 2006) also investigated synchronization at the product level using the Fisher–Konieczny (2000) measure, which takes a value of 1 in the case of perfect synchronization and a value of 0 in the case of uniform staggering of price changes. They calculated the measure for each of the 50 product categories in their common sample. The degree of synchronization was, in general, rather low except for energy prices, with the median synchronization ratio across products ranging between 0.13 in Germany and 0.48 in Luxembourg. The higher ratio observed in Luxembourg compared to Germany likely reflects the difference in the size of the market upon which the ratio is computed and the relatively small number of outlets in Luxembourg. Related, using Belgian CPI data, Dhyne and Konieczny (2010) find that the more aggregated the data, both in product and geographic space, the more staggering of price changes.

Other studies have focused on the synchronization of prices at the outlet level. Using scanner data for a number of retailers in the United States, Nakamura (2008) decomposed price variation for individual products into variation that is common across all stores (16%), variation that is common only to stores within the same retail chain (65%), and variation that is completely idiosyncratic to particular stores (17%). These findings suggest that retail-level shocks, rather than manufacturer-level shocks, may be quite important for understanding fluctuations in retail prices. Midrigan (2009) found that the probability a particular product experiences a price change depends on the fraction of other prices within its store that change, especially those in its own product category. He also found some evidence, albeit weaker, of synchronization across stores in a particular city. Related, Lach and Tsiddon (1992, 1996) analyzed prices of different meat products and wines at retail stores in Israel. They found that, when stores change price, they seem to change price for most of their products at the same time, but these adjustments occur at different times for different stores.

5.2 Sales, reference prices, and aggregate inflation

A key question for macroeconomists when thinking about sale-related price changes is whether they respond to macro shocks or instead reflect entirely idiosyncratic forces. Midrigan (2009) and Nakamura and Steinsson (2008a) implicitly took the latter view when they replaced sale prices with previous regular prices in their empirical analysis. Guimaraes and Sheedy (2008) rationalized the latter in a specific model, generating oscillation between sticky regular and sale prices as an optimal form of price discrimination. Similarly, Eichenbaum et al., (2009) described a model in which sellers choose a “sticky pair” of prices that they can freely bounce between, but a menu cost applies whenever the pair is changed. Kehoe and Midrigan (2008) took an intermediate position, modeling sale-related price changes as subject to (smaller) menu costs and responding to macro as well as idiosyncratic shocks. Because sale prices tend to revert to previous regular prices (both in their theory and in the data), their theoretical sale prices contribute notably less to macro price flexibility than do regular price changes (which can be arbitrarily persistent).

The answer to whether sale-related price changes contribute to macro price flexibility is ultimately empirical. According to Klenow and Kryvtsov (2008), more than 40% of sale price episodes do *not* return to the previous regular price, opening the door to more macro price flexibility. Klenow and Willis (2007), using bi-monthly data for all cities in the CPI, found that the magnitude of price discounts indeed correlates with cumulative inflation since the item last changed price. Another possibility, which has not yet been investigated, is that clearance sales (more common for apparel and appliances, less common for food) react to unwanted inventory build-up at the macro level.

To provide some evidence on the macro content of sale prices, we calculated inflation for posted prices and regular prices separately in the U.S. CPI from February 1988 through October 2009 for the top three cities. We took out separate monthly dummies for each to remove seasonal effects. Table 12 summarizes some of the resulting moments for posted versus regular price inflation in the U.S. CPI. For the residuals, the variance of regular price inflation equaled 87.5% of the variance of posted price inflation, so that sale prices “accounted for” 12.5% of the aggregate variance. By comparison, sale prices represent about 19% of all price changes (see Klenow & Kryvtsov, 2008). When we aggregated up to the quarterly level, however, sale prices contributed only 7.5% of the variance of quarterly posted price inflation. Thus sale-related price changes do not fully wash out with cross-sectional aggregation, but do significantly cancel out with time aggregation.

The serial correlation of posted price inflation in the U.S. CPI (0.394, standard error 0.053) is similar to that of regular price inflation (0.420, SE. 0.051) — see also Bils, Klenow, and Malin (2009). The same is true of quarterly inflation rates (0.174 serial correlation for posted price inflation, 0.142 for regular price inflation). To get at whether price changes build or fade, we regressed cumulative inflation from month t to $t + 12$ on inflation from month t to $t + 1$. As given in Table 12, following a 1% price increase in the first month, regular prices are 1.48% (SE 0.27) higher after 12 months. The aggregate component of sale-related price changes is surprisingly

Table 12 Posted versus regular price inflation in the U.S. CPI

	Posted price π	Regular price π	Sale-related π
Monthly			
Variance relative to that of posted price π	1	0.875	0.125
Correlation with posted price π	1	0.963 (0.005)	0.365 (0.054)
Serial correlation	0.394 (0.053)	0.420 (0.051)	-0.076 (0.062)
Cumulative inflation after 1 year	1.507 (0.270)	1.480 (0.267)	1.089 (0.168)
Quarterly			
Variance relative to that of posted price π	1	0.925	0.075
Correlation with posted price π	1	0.979 (0.005)	0.285 (0.100)
Serial correlation	0.174 (0.106)	0.142 (0.108)	-0.017 (0.110)

Note: Source is CPI-RDB. Inflation rates are for the top three cities (New York, Los Angeles, and Chicago). Monthly series go from February 1988 through October 2009, and quarterly series from 1988:2 through 2009:3. Posted Prices include sale prices, whereas Regular Prices replace sale prices with the previous regular price. Monthly (seasonal) dummies are removed separately from Posted Price Inflation and Regular Price Inflation before moments are calculated. “Sale-related inflation” refers to the series obtained by subtracting Regular Price Inflation from Posted Price Inflation. The “Cumulative Inflation After 1 Year” is obtained from regressing $\ln P_{t+12} - \ln P_t$ on $\ln P_{t+1} - \ln P_t$.

Table 13 Posted versus reference price inflation in the U.S. CPI

	Posted price π	Reference price π	Non-reference π
Monthly			
Variance relative to that of posted price π	1	0.462	0.538
Correlation with posted price π	1	0.689 (0.033)	0.734 (0.029)
Serial correlation	0.398 (0.054)	0.461 (0.050)	0.119 (0.063)
Cumulative inflation after 1 year	1.569 (0.281)	3.067 (0.395)	0.906 (0.162)
Quarterly			
Variance relative to that of posted price π	1	0.592	0.408
Correlation with posted price π	1	0.822 (0.037)	0.642 (0.066)
Serial correlation	0.175 (0.109)	0.441 (0.091)	-0.217 (0.107)

Note: Source is CPI-RDB. Inflation rates are for the top three cities (New York, Los Angeles, and Chicago). Monthly series go from August 1988 through April 2009, and quarterly series from 1988:3 through 2009:1. Reference Prices represent the most common price in the 13-month window centered on the current month for each item. Monthly (seasonal) dummies are removed separately from Posted Price Inflation and Reference Price Inflation before moments are calculated. “Non-reference Inflation” refers to the series obtained by subtracting Reference Price Inflation from Posted Price Inflation. The “Cumulative Inflation After 1 Year” is obtained from regressing $\ln P_{t+12} - \ln P_t$ on $\ln P_{t+1} - \ln P_t$.

persistent: the difference between posted and regular prices is 1.09% (SE 0.17) higher after 12 months (following a 1% increase in the first month in the difference between posted and regular price inflation).

Table 13 provides statistics on aggregate *reference* price inflation, again obtained as residuals from separate monthly dummies for the top three cities in the U.S. CPI. Recall that, in the spirit of Eichenbaum et al. (2009), we defined the reference price as the most common price in the 13-month window centered on the current month for each item. As discussed, reference prices represent about 80% of all prices in the CPI by this definition, and change every 11 months. By comparison, regular prices represent about 90% of all prices and change roughly every 7 months.¹⁴ Thus, in practice, reference prices more aggressively filter out short-lived prices (although this is not the case in principle, as our definition allows the reference price to conceivably change every month). The variance of aggregate *monthly* reference price inflation is 46% of the variance of posted price inflation in the U.S. CPI. The monthly posted price inflation rate is similarly correlated with reference price inflation (0.69) and with the deviations

¹⁴ The correlation between *monthly* regular and reference price inflation rates is 0.69 (SE 0.03), whereas that between *quarterly* regular and reference price inflation rates is 0.79 (0.04).

from reference price inflation (0.73). Monthly reference price inflation is more serially correlated (0.46, SE 0.05), than are the deviations from reference price inflation (0.12, SE 0.06). This is not by construction, as the deviations from reference prices could have exhibited more persistent changes. Reference price changes build to 3.07% (SE 0.40) after one year, whereas deviations from reference prices neither build nor fade (coefficient 0.91%, SE 0.16).

As Table 13 shows, reference price inflation becomes more volatile and persistent relative to posted price inflation when we aggregate up to the quarterly level. The variance of reference price inflation rises to about 59% of overall quarterly inflation (from about 46% at the monthly frequency). At the quarterly level, posted price inflation is more correlated with reference price inflation (0.82) than with deviations from it (0.64). The serial correlation of quarterly reference price inflation is 0.44 (SE 0.09), versus -0.21 (0.11) for deviations from quarterly reference price inflation. Figure 6 plots these quarterly series. Both rates plunge at the end of 2008 due to energy price declines, but these two quarters do not drive any of the statistics. We conclude that

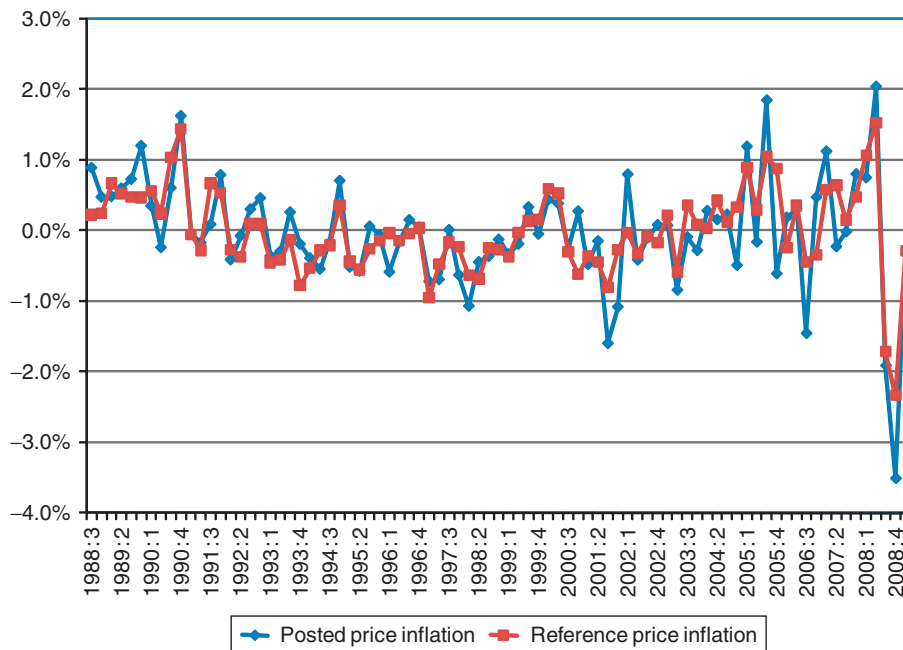


Figure 6 Posted versus reference price inflation, U.S. CPI. *Source:* CPI-RDB. Data are for the top three cities (New York, Los Angeles, and Chicago) for the quarters 1988:3 through 2009:1. Posted price inflation includes sale and regular prices. Reference price inflation is based on replacing the current posted price with the modal price in the 13-month window centered on current month (for each micro data point). Inflation rates were calculated at the monthly level, and monthly (seasonal) dummies were removed for each series separately. The monthly residuals were summed to create quarterly data.

reference price inflation picks up responses to more persistent shocks to inflation, whereas deviations from reference price inflation reflect more fleeting disturbances.

5.3 Hazard rates

Another “dynamic” feature that has been documented in many studies is the shape of the hazard function of price change; that is, how the probability of changing price varies depending upon the age of the price. One reason for this is that price-setting models often have fairly stark predictions for the shape of the hazard function. The original Calvo (1983) model assumed a flat hazard function, while the deterministic timing of price adjustment in the Taylor (1980) model predicted a zero hazard except at a single age, where the hazard is one. Menu-cost models can generate a variety of shapes, depending on, among other things, the relative importance of transitory and permanent shocks to marginal costs. Permanent shocks, which accumulate over time, tend to yield an upward-sloping hazard function, while transitory shocks tend to flatten or even produce a downward-sloping hazard (e.g., sellers may be more attentive to getting prices right when revenue is temporarily high for a product due to idiosyncratic supply or demand considerations).

The general finding in the literature is that hazard rates for individual products are *not* upward-sloping.¹⁵ Klenow and Kryvtsov (2008) found the frequency of (regular) price changes conditional on reaching a given age is downward-sloping if all goods are considered together, but note that this could simply reflect a mix of heterogeneous flat hazards and survivor bias. Once they took out decile fixed effects (i.e., average frequencies of price change in each decile of price change frequency), they characterized the hazard rate as flat (other than a spike at 12 months). Álvarez (2008) reported a similar pattern in Euro Area consumer price data. Nakamura and Steinsson (2008a) estimated separate hazard functions for each Major Group, however, and found hazards are downward-sloping for the first few months and generally flat after that. They found similar patterns for producer prices. Finally, Cavallo (2009) found initially downward-sloping hazards for daily online prices from a large supermarket in each of several Latin American countries.

5.4 Size versus age

Klenow and Kryvtsov (2008) undertook an analogous exercise for the (absolute) size of price changes. They found that the size of price changes rises with age, but once they controlled for heterogeneity across goods, the size of price changes is unrelated to the age of the price for a given item. Likewise, Eden (2001), used data from Israel and several different approaches, and found no strong correlation between the size and age.

¹⁵ One exception is Ikeda and Nishioka (2007), who used a finite mixture model (to allow for heterogeneity across price-setters) and assumed price changes occur according to a Weibull distribution (to allow for increasing, flat, or decreasing hazards). They estimated increasing hazard functions for some Japanese products, and flat or Taylor-type spiked hazards for others.

Álvarez and Hernando (2004), using Spanish data and Heckman's sample selection correction, reported that age does not significantly contribute to the size of price changes.

In contrast, many models predict a positive relationship between size and age. For example, models in which shocks accumulate over time and the timing of price changes are exogenous (Calvo 1983) or driven primarily by having a low menu-cost draw (Dotsey et al., 1999) fall into this category.

5.5 Transitory relative price changes

The persistence of relative price changes can shed light on the relative importance of idiosyncratic versus aggregate shocks and also help to distinguish between different price-setting models. Indirect evidence on this aspect of the data comes from investigating the behavior of sales. Nakamura and Steinsson (2008a) reported that (a) sales spells are short — the average length is just 1.8–2.3 months, (b) sales price changes are twice as large as other price changes on average, and (c) many prices return to their original price following a sale. Taking these facts together suggests highly transient *relative* price changes for both sales prices and all posted prices, as sales are quite common in the U.S. CPI data (representing about 20% of price changes according to Klenow & Kryvtsov, 2008). Scanner data studies uncover similar patterns. For example, Kehoe and Midrigan (2008) found that price changes are large and dispersed and most are of a “temporary” nature, again suggesting transitory relative price changes.

The transitory nature of relative price changes is not just a sales-related phenomenon, however, as it has been found after filtering out sales as well. Campbell and Eden (2005) documented that grocers choose, and then quickly abandon, extreme relative prices (rather than arriving at those prices through the gradual erosion of their fixed nominal price by other sellers' price adjustments). Midrigan (2009) reported that the probability a firm's next price change will have the same sign as the current one is fairly low (between 32% and 41%). Klenow and Willis (2006) constructed relative prices as the ratio of individual prices to the weighted average of prices in a sector, and reported an across-sector weighted mean serial correlation of *new* relative prices (i.e., across months with price changes) of 0.32.

To provide a metric for how transitory these relative price changes are, it is useful to consider what they imply for the persistence of shocks in structural models. The general finding is that transitory, idiosyncratic shocks are an important input that enables structural models to replicate pricing patterns in the data. Klenow and Willis (2006) estimated a monthly persistence of 0.7 for the idiosyncratic technology shock in their model in order to match the serial correlation of *new* relative prices. The shocks that trigger price changes in Midrigan (2009) are similarly transitory and tend to be reversed. Nakamura and Steinsson (2008a) calibrated the size of the menu cost and the idiosyncratic-shock persistence and volatility to match the frequency of (regular) price changes, the fraction of changes that are increases, and the size of changes.

The idiosyncratic shock is large and has a serial correlation of 0.66. Other models that produce transitory relative price changes include Burstein and Hellwig (2007) who incorporated transitory demand shocks in addition to marginal cost shocks, Kehoe and Midrigan (2008) who gave firms the option of a “one-period markdown” at a lower menu cost than a permanent price change, and Eichenbaum et al. (2009) whose sticky-plan firms could switch between a small number of prices at no cost.

5.6 Response to shocks

Evidence on the response of prices to identified shocks is particularly useful for shedding light on the nature of firms’ price-setting decisions because most (if not all) theories of price-setting posit a tight link between (desired) prices and the underlying cost and demand conditions facing the firm. Precisely identifying shocks in the data is not an easy task, but recent studies have made use of novel data sets to do just this. These include studies of the response of import/export prices or of prices on either side of a border to exchange rate shocks, evidence of price-setting around the time of a significant event like the Euro changeover or an increase in VAT rates, and studies of scanner data that contain both wholesale costs and retail prices.

Researchers have made use of well-identified and sizable movements in exchange rates to evaluate how prices respond to shocks. Gopinath and Rigobon (2008) measured *medium-run* exchange rate pass-through as the change in a price in response to the cumulative change in the exchange rate since the price was last changed and found that the trade weighted exchange rate pass-through into U.S. import prices is low at 22%. Gopinath, Itskhoki, and Rigobon (2010) documented that non-dollar priced goods display much higher medium-run pass-through than do dollar priced goods. They presented a sticky-price model in which firms chose their currency based on their average desired pass-through over the period of price nonadjustment. Incomplete desired pass-through — whether it is driven by variable markups, imported inputs, or decreasing returns to scale in production — is thus seen as the reason that the majority of U.S. import goods are priced in dollars. Gopinath and Itskhoki (2010) found further evidence consistent with variable markups or variable marginal costs in the form of a positive correlation between the frequency of price adjustment and *long-run* pass-through, the life-long change in the price of a good (relative to U.S. inflation) on the (real) exchange rate over the same period. Fitzgerald and Haller (2009) used micro data on domestic and export prices set by Irish producers. Because they had matched price quotes for the same product produced in the same plant but sold in multiple markets, they were able to control for changes in marginal costs in addition to exchange rate movements. They found strong evidence of pricing-to-market for products whose export prices are invoiced in the destination currency: conditioning on price changes in both markets, relative prices move one-for-one with exchange rate changes (i.e., zero pass-through of exchange rate changes).

Other research has found evidence that borders effectively segment markets. Burstein and Jaimovich (2009) and Gopinath, Gourinchas, Hsieh, and Li (2009) evaluated scanner data from a major retailer that has multiple locations in Canada and the United States. Gopinath et al. (2009) found large discontinuities of retail prices and wholesale costs between adjacent stores on either side of the border, and noted that these border gaps move almost one-for-one with changes in the nominal exchange rate. Burstein and Jaimovich (2009) documented substantial pricing-to-market for traded goods. Using a data set of online book prices (and some information on quantities) for a number of U.S. and Canadian retailers, Boivin, Clark, and Vincent (2009) concluded that market segmentation is probably behind the lack of exchange rate pass-through in their sample.

Other identifiable aggregate shocks include changes in tax rates or in the nation's currency. A number of European studies, summarized in Dhyne et al. (2005), have consistently found that changes in the value-added tax rate lead to temporary increases in the frequency of CPI price changes. The frequency of price change also increased noticeably after countries converted to the euro in January 2002 (Dhyne et al. 2005). Although the response of aggregate inflation to the introduction of the euro was not particularly large, Hobijn, Ravenna, and Tambalotti (2006) emphasized the dramatic increase in Euro Area restaurant prices and explain this increase with a menu-cost model in which firms decide when to adopt the new currency.

Eichenbaum et al. (2009) were able to provide information on the price response to (potentially disaggregated) shocks. They investigated the relationship of prices to costs using scanner data for one large U.S. retailer. Prices typically do not change in the absence of a change in cost, but a cost change is not sufficient to induce a change in price. This allows markups (both actual and reference) to display substantial variation. When the retailer does decide to change its reference price, it almost always reestablishes the unconditional markup over the reference cost. That is, the retailer passes through 100% of the cumulative change in reference cost that occurred since the last reference price change.

Nakamura and Zerom (2010) documented delayed and incomplete pass-through of (observable) commodity cost shocks to wholesale/retail prices in the coffee industry. They estimated that, relative to a CES benchmark, local costs reduce long-run pass-through (after 6 quarters) by 59% and markup adjustment reduces it by an additional 33%. Barriers to price adjustment are important for the delayed response of prices to cost but have a negligible role in incomplete long-run pass-through.

Bils et al. (2009) did not condition upon identified shocks but instead documented how "reset prices" evolve over time in response to reduced-form impulses. A reset price for an individual firm is the price it would choose if it implemented a price change in the current period. Models with sizeable strategic complementarities predict that the impact of a shock to reset prices will build over time, as price-setters wait for

the average price to respond. In the U.S. CPI data, however, the reset price actually overshoots — it changes more on impact than it does in the long-run. At the same time, Klenow and Willis (2007) found some evidence that micro price changes respond to “old” inflation (i.e., since before the item’s previous price change), as in sticky information models.

5.7 Higher moments of price changes and aggregate inflation

Various authors have investigated the relationship between higher moments of the distribution of price changes and inflation (i.e., the first moment). Relative price variability is generally found to be positively correlated with inflation, as predicted by menu cost or incomplete information models. Lach and Tsiddon (1992) pointed out that menu-cost models imply that the variability is affected by *expected* inflation, while incomplete information models imply a relationship with *unexpected* inflation. They found that expected inflation had a stronger effect on variability than unexpected inflation for a sample of foodstuffs in Israel. Konieczny and Skrzypacz (2005) reached similar conclusions using data on 52 goods in Poland from 1990 to 1996.

Ball and Mankiw (1995) focused on the third moment and found that innovations in inflation are positively associated with the skewness in relative-price changes for four-digit U.S. PPI data from 1949 to 1989. Indeed, the inflation-skewness relationship is stronger than the inflation-variance relationship. They showed these findings are consistent with a menu-cost model: when price adjustment is costly, firms adjust to large shocks but not to small shocks, and therefore large shocks have disproportionate effects on the price level. They also explored the idea that asymmetric relative-price changes represent aggregate supply shocks.

Domberger and Fiebig (1993) looked at individual price changes in 80 disaggregated industry groups in the UK. They also found that an increasing (decreasing) average price is associated with a rightward (leftward) skew, but noted that when the absolute magnitude of inflation (deflation) increases, the distribution of individual price changes becomes less skewed. They suggested that the higher the skewness of the price-change distribution, the more staggered (less synchronized) the changes.

We complement these studies by calculating monthly measures of higher moments of the distribution of price changes for the top three cities in the U.S. CPI from 1988 to 2009. We then construct the correlation of the variance, skewness, and kurtosis of (nonzero) price changes with inflation. We find few robust significant relationships. The variance is *negatively* correlated with inflation, both for posted prices (-0.16, SE 0.06) and for regular prices (-0.29, SE 0.06). These results are fragile, however, as they are highly influenced by deflation during the last few months of 2008. Omitting the last 15 months of the sample (August 2008–October 2009), the inflation-variance correlation is still statistically significant for regular prices (-0.15, SE 0.06) but not for

posted prices (-0.07, SE 0.06). Moreover, the correlation between skewness and inflation becomes significant: correlation of 0.15 (SE 0.06) for regular prices and 0.17 (SE 0.06) for posted prices. That is, the late-2008 plunge in inflation was accompanied by a spike in the variance of price changes, but no fall in skewness. We find that skewness is more robustly (if modestly) related to *absolute* inflation (0.15, SE 0.06 for both posted prices and regular prices).

More study of the relationship between inflation and higher moments thus seems warranted. For example, Ball and Mankiw's (1995) interpretation of skewness as representing an aggregate supply shock suggests a more nuanced inflation-skewness relationship than might be picked up by a simple correlation. Indeed, they argued that the absence of asymmetry when inflation fell in the 1975 and 1982 recessions can be viewed as evidence in favor of their hypothesis, as it suggests that causation does not run from inflation to skewness.

6. TEN FACTS AND IMPLICATIONS FOR MACRO MODELS

We now summarize the findings from the empirical literature on individual price data with ten stylized facts and their potential lessons for models. These touch on two broad themes. The first is the frequency and nature (e.g., state-dependent vs. time-dependent pricing) of micro price changes. The second is the potential magnitude and source of the “contract multiplier”; that is, how long it takes for the macro effects of price stickiness to fade after a permanent shock relative to the duration of individual prices.

6.1 Fact 1: Prices change at least once a year

In the U.S. CPI, prices change every 4 months or so. Prices appear stickier in the U.S. PPI (around 6–8 months for the median), and stickier still in the Euro Area (around once a year for the median). But median price durations in emerging markets (tilted toward food, as actual household budgets are) are typically closer to the United States than the Euro Area.

In contrast, a large literature has estimated that permanent monetary shocks have real effects lasting several years.¹⁶ Real effects of nominal shocks therefore last three to five times longer than individual prices. Nominal stickiness appears insufficient to explain why aggregate prices respond so sluggishly to monetary policy shocks. For this reason, nominal price stickiness is usually combined with a “contract multiplier” (in Taylor's 1980 phrase).

Ball and Romer (1990) and Kimball (1995) are early examples, whereas Christiano, Eichenbaum, and Evans (2005) are a more recent one. Possible sources of big contract

¹⁶ Christiano, Eichenbaum, and Evans (1999); Romer and Romer (2004); and Bernanke, Boivin, and Elias (2005) are a few examples based on U.S. data. Peersman and Smets (2003) reported similar Euro Area results.

multipliers include strategic complementarities, countercyclical markups, sticky plans, and sticky information (the latter including rational inattention, e.g., Mackowiak & Wiederholt, 2008).

Prices do not change continuously either. Nominal stickiness is a pervasive feature of the data in all country-years with moderate inflation rates. Models that generate persistence in aggregate prices without any nominal stickiness, for example via indexation (Christiano et al., 2005) or convex costs of price adjustment (Rotemberg, 1982), are at odds with this fact.

6.2 Fact 2: Sales and product turnover are often important for micro price flexibility

Large price discounts and product turnover account for about one-third of consumer price changes in the United States. Sales are much more frequent for some types of goods (food and apparel), and usually revert to the previous regular price afterward (more so for food, less so for apparel). Sales are much less frequent in the Euro Area than in the United States. Within the United States, sales are less prevalent for producer prices than consumer prices. Product turnover ushers in nominal price changes, and is much more common for durables (and apparel) than food or services. Both sales and product turnover exhibit clear seasonal patterns, suggesting a time-dependent component.

Given their temporary and seasonal nature, sale-related price changes may contribute more to micro than macro price flexibility. Put differently, the contract multiplier may be larger for price changes related to sales and product turnover. That said, sale- and turnover-related price changes do not wash out with aggregation in the United States, consistent with macro content. It remains an open question whether such prices respond to aggregate shocks.

In addition to interesting empirical questions concerning the macroeconomic importance of sales and substitutions, more theoretical research is necessary for us to better understand how these types of price changes affect the monetary transmission mechanism. It is awkward to simply apply menu-cost models to data in which some price changes have been filtered out. Alternative approaches have been pursued in the literature. Kehoe and Midrigan (2008) modeled one-period price discounts as incurring a lower menu cost. They found sales-related price changes can contribute to macro price flexibility, but not as much as regular price changes. Guimaraes and Sheedy (2008) built a model in which firms face consumers with different price sensitivities and therefore choose a pricing strategy with a high price at some moments and a low price at other moments. Even though “sales” decisions are completely flexible at the micro level, this does not translate into flexibility at the macro level because sales are strategic substitutes: a firm gains little from a sale if its competitors are having sales.

Thus, different models of sales have different implications for aggregate flexibility. Empirical work could help distinguish between and refine this theorizing.

6.3 Fact 3: Reference prices are stickier and more persistent than regular prices

Eichenbaum et al. (2009) documented many temporary regular price changes in a major U.S. grocery store chain. We find a similar pattern of sticky “reference” prices in most U.S. CPI items. Whereas posted prices change every 4 months and regular prices every 7 months, reference prices change every 10–11 months.

Deviations from reference prices do not cancel out across items, thereby affecting aggregate inflation. Perhaps not surprising, reference price inflation is considerably more persistent than either posted or regular price inflation. Changes in reference prices build over time, whereas other price changes tend to be much more transitory. Reference pricing behavior may suggest some form of sticky plan and/or sticky information, rather than menu costs per price change, and hence could explain the sizable contract multiplier. That said, a novel price (relative to prices in the previous 12 months) tends to appear every six months or so for a given item. And the vast majority of prices are not “comeback” prices in the sense of having existed in an earlier price spell in the prior 12 months.

A related fact is that U.S. CPI quote-lines, which typically span over four years, are dominated by a few prices. The top price alone captures one-third of prices, and the top two prices more than one-half of prices. This fact may reflect downward-sloping hazards rather than sticky information or plans, given that two-thirds of top prices occur without interruption by other prices.

Future research could explore whether the reference price phenomenon extends to U.S. producer prices and to prices in other countries. There might be more frequent reference price changes in countries and time frames with more persistent inflation than in recent U.S. decades. Moreover, a vital question remains: how do reference versus nonreference prices respond to aggregate shocks?

6.4 Fact 4: There is substantial heterogeneity in the frequency of price change across goods

Every micro data set to date has displayed large, persistent heterogeneity in the frequency of price changes across types of goods. Universally, service prices are stickier than those of goods. Among goods, “raw” goods (energy, fresh produce) are more flexible than “processed” goods (goods with low intermediate shares and high labor shares).

As stressed by Carvalho (2006) and Nakamura and Steinsson (2008b), such heterogeneity can combine powerfully with strategic complementarities to boost the contract multiplier. But evidence for strategic complementarities in micro price data is mixed. In the supportive column, Gopinath (2010) and collaborators consistently found that exchange rate changes gradually pass-through into import prices. In the skeptical

column, Klenow and Willis (2006) stressed that sellers routinely change prices relative to close substitutes (e.g., temporary sales are not synchronized among competing brands of a narrow product) and argued that large relative price movements are inconsistent with strong “micro” real rigidities. Using evidence on product-level prices and market shares, Burstein and Hellwig (2007) also concluded that “micro” pricing complementarities are too weak to generate large aggregate real effects. Kryvtsov and Midrigan (2009) noted that strong “macro” complementarities imply stable markups, and hence are hard to reconcile with countercyclical inventory-sales ratios. And Bils et al. (2009) found that consumer price changers tend to overshoot, rather than undershoot as predicted by slow pass-through.

6.5 Fact 5: More cyclical goods change prices more frequently

In the U.S. CPI, prices change more frequently in categories with more procyclical real consumption growth — in particular transportation (cars, airfares) and apparel. This fact probably extends to U.S. producer prices, given the high correlation between upstream and downstream price change frequency (Nakamura & Steinsson, 2008a). In other countries, we also note the strong tendency for durables to change prices more frequently than services.

This form of price change heterogeneity could have the opposite effect of idiosyncratic sources of heterogeneity (such as different-sized menu costs or idiosyncratic shocks); that is, higher frequency price changes among cyclical goods could reduce the contract multiplier. See, in particular, Barsky, House, and Kimball (2007). Natural topics for further investigation include the source of this connection (does cyclicality cause flexibility or the other way around?) and its quantitative importance for the contract multiplier in various models.

6.6 Fact 6: Price changes are big on average, but many small changes occur

Micro price changes are an order of magnitude larger than needed to keep up with aggregate inflation in the United States and Euro Area. Thus idiosyncratic forces dominate macro shocks. Golosov and Lucas (2007) showed that the former can limit the contract multiplier. In their SDP model, a positive aggregate monetary shock results in more idiosyncratic price increases (and fewer idiosyncratic price decreases), thereby speeding the response of the aggregate price level. Bils et al. (2009) found that “reset price inflation” behaves as if it strongly affected by such selection.

Many price changes *are* small, too, so that there is a “missing middle” in the distribution of price changes emerging from SDP models with a single, large menu cost. This need not be true if menu costs are variable (as in Dotsey et al., 1999) or are small but shocks arrive infrequently (as in Gertler and Leahy, 2008).

Meanwhile, Midrigan (2009) made the point that TDP models can easily fit the whole size distribution. He argued that the missing middle in SDP models is a by-product of overstated selection. Alternatively, the small price changes generated by his SDP model with multiproduct firms reflect weakened selection and hence greater departures from monetary neutrality, akin to TDP models. Woodford (2009) laid out a model in which information constraints can lead to small price changes as well (costly information updating results in little change ex post sometimes, although not on average). As in TDP, the result of information constraints can be a bigger contract multiplier than under SDP.

6.7 Fact 7: Relative price changes are transitory

A corollary to large idiosyncratic price changes is big relative price movements. Even within narrower categories of the CPI (Expenditure Classes, ELIs) and U.S. grocery scanner data, these relative price changes are large. Such relative price movements tend to fade over time — they are far less persistent than a random walk. This is true across stores of different chains, but also across competing brands within chains. It is also true even for regular prices; that is, excluding the (often temporary) sale price discounts from regular prices. This fact is a corollary to many regular price changes being temporary in nominal terms.

The persistence of relative prices can matter for macroeconomic questions. First, in menu-cost models the S_s band for nonadjustment is narrower when firms are selling more and wider when firms are selling less. Thus the selection effect of macro shocks can change with the persistence of idiosyncratic shocks. Second, big transitory movements in relative prices require either that micro complementarities be weak, or that idiosyncratic shocks be large. Weaker complementarities imply a smaller contract multiplier in response to monetary shocks. Third, transitory movements may reflect micro flexibility amid a sticky macro plan/information (i.e., firms frequently respond to big, temporary idiosyncratic shocks within the plan, but less frequently change the macro plan/information given that persistent macro shocks are smaller).

6.8 Fact 8: Price changes are typically not synchronized over the business cycle

The contract multiplier can be lower if sellers accelerate price increases in response to positive monetary shocks and postpone them in response to negative monetary shocks. For moderate inflation episodes such as the last two decades in the United States, sellers do not seem to synchronize their timing in this way. This is consistent with time-dependent pricing, but does not nullify the selection effect of state-dependent pricing completely. As predicted by SDP models and TDP models alike, the composition of price increases versus decreases correlates positively with inflation movements. But lack of synchronization may mean weaker selection, or even a preoccupation with idiosyncratic over aggregate shocks (e.g., rational inattention).

Staggered nominal price stickiness is also critical for how strategic complementarities can help produce a big contract multiplier. Synchronized price changes are not subject to coordination failure in magnitude. The less price changes are synchronized, the more sellers may slow their adjustment as they wait for other prices (of competitors and input suppliers) to adjust. That said, the evidence for strategic complementarities appears mixed. The lack of synchronization even among closely competing brands, is hard to understand if micro real rigidities are strong.

Periods of greater macro volatility may exhibit more synchronization. In Mexico in recent decades, the frequency of price changes moved up and down with aggregate inflation movements. In the recent U.S. recession, the frequency of consumer price changes appears to have surged. Intriguingly, both price increases and decreases became more common, perhaps because of more frequent updating of sticky plans or sticky information.

6.9 Fact 9: Neither frequency nor size is increasing in the age of a price

In both the United States and Euro Area, and for both consumer and producer prices, the hazard rate of price changes is falling over the first few months and largely flat afterward. The exception is a spike in price change frequency for services, suggestive of annual updating for the stickiest category of products. The downward slope is much less pronounced if one looks at regular prices (i.e., excludes sale-related price changes), and may be flat if one fully controls for survivor bias.

For a state-dependent pricing model, standard intuition would have the hazard of price changes increasing in the time since a price change as shocks accumulate and the desired price drifts farther away from the current price. This force can be weakened by thicker tailed or Poisson shocks, and by periods with wider versus narrower S s bands (which create an intertemporal form of survivor bias). Scanner studies find clear evidence of state-dependence in rising hazards in distance from the average markup. Again, state-dependence can imply a smaller contract multiplier than time-dependent pricing.

For a time-dependent model, analogous intuition would imply that the *size* of price changes is *increasing* in the duration of stickiness, as more shocks accumulate the longer the spell between price changes. Evidence on this question is somewhat more limited, but there seems to be little connection between the size of price changes and the duration of price spells. This pattern is consistent with state-dependent pricing, under which spell length is endogenous to the shocks accumulated. A long spell without a price change may signal that the desired price did not move much, so that it does not change unduly when the spell ends. This question could use more study in the United States and elsewhere.

6.10 Fact 10: Price changes are linked to wage changes

Recent research has revealed a noticeable link between price and wage rigidity. In the cross-section, firms (or categories of goods) with a higher share of labor costs in total

costs make less frequent price adjustments, potentially resulting from the fact that wages adjust less frequently than other input prices. Survey evidence also suggests synchronization between wage and price adjustments over time, as well as a cross-sectional correlation between wage flexibility and price flexibility.

Wage stickiness can contribute directly to a high contract multiplier. But the aforementioned evidence suggests it may be contributing indirectly, as well, by lowering the frequency of price changes. This relationship is all the stronger where production is labor-intensive. Moreover, wage adjustment exhibits a substantial degree of time-dependence. Firms tend to concentrate wage changes in a specific month, mostly January in a majority of European countries (Druant et al., 2009). This could lend a degree of time-dependence into price-setting, further contributing to a higher contract multiplier.

6.11 Summary: Model features and the facts

Table 14 provides a quick summary of how some common features of macro models of price-setting stack up against the stylized facts about micro price setting. Most of the entries should be self-explanatory given the preceding discussion, but a few warrant elaboration. Take the fact that many consumer price changes are reversed (i.e., prices

Table 14 Model features and the facts

	Consistent features	Inconsistent features
Half of prices change only a few times a year	Menu costs; sticky input prices	Price indexation; convex adjustment costs
Temporary price changes are common	Sticky set; price discrimination	Menu costs with flexible marginal cost
Frequency differs persistently across goods	Menu costs; sticky information	Exogenous frequency of price changes
Price changes are large on average	Big menu costs; rational inattention	Small idiosyncratic shocks with strong complementarities
Many price changes are small	Sticky information; sticky path	Large menu costs for changing each individual price
Relative price changes are transitory	Transitory idiosyncratic shocks	Strong micro complementarities
Price changes are not well synchronized	Big idiosyncratic shocks	Small idiosyncratic shocks with strong complementarities
Old prices do not change by bigger amounts	Menu costs; lumpy shocks	Time-dependent pricing with persistent shocks

sometimes exhibit “memory,” particularly after sales in food and apparel in the United States). This pattern is consistent with sellers periodically choosing a “sticky set” of a few prices that they bounce between, a form of sticky plan advocated by Eichenbaum et al. (2009). A fuller explanation might be price discrimination as in Guimaraes and Sheedy (2008).

Probably the most robust fact across all micro pricing studies is that goods persistently differ in their frequency of price changes. The variation is far from random, as raw goods (fresh produce, energy) change price more frequently than do services in country after country. Thus price flexibility appears to respond to economic fundamentals (e.g., the average size of sectoral shocks and the trend rate of inflation). And, in the United States at least, more cyclical categories exhibit more consumer price flexibility.

Another consistent finding is that most micro price changes are much larger than needed to keep up with average inflation. This fact would seemingly point to a combination of big idiosyncratic shocks and big menu costs. But rational inattention could coexist with these ingredients. At the same time, the number of small price changes in the data is far from trivial. This may be because smaller menu costs are combined with sticky information (sellers sometimes find only a small price change is needed after a periodic information update) or with sticky plans (sellers periodically update a path of prices, as in Burstein, 2006).

7. CONCLUSION

We reviewed the recent empirical literature on individual price data and distill ten salient facts for macro models. Prices change quite frequently, although much of this flexibility is associated with price movements that are temporary in nature. Even if all short-lived prices are excluded, however, the resulting nominal stickiness, by itself, appears insufficient to account for the sluggish movement of aggregate prices. These findings point to the need for a large contract multiplier to bridge the gap between micro flexibility and macro inertia.

Other micro price facts provide evidence on the plausibility of various mechanisms for generating a large contract multiplier. The lack of synchronization in the timing of price changes provides scope for strategic complementarities to amplify the real effects of nominal stickiness. The presence of substantial heterogeneity across goods in the average frequency of price change can bolster this channel, although the fact that more cyclical goods exhibit greater price flexibility may work in the opposite direction. The presence of many large, transitory price changes raises concerns about the relevance of real rigidities at the individual-item level, but is consistent with macro rigidities or with rationally inattentive sellers who respond to sizable idiosyncratic shocks but not to smaller aggregate impulses. Finally, the fact that the size of price changes does not increase with the age of a price

provides evidence for state-dependent pricing and thus selection. The stronger the selection effect, the smaller the contract multiplier, although the presence of many small price changes suggests the selection effect may be muted.

A number of open empirical (and related theoretical) questions remain. How do temporary price changes (related to sales, product turnover, or a movement to a non-reference price) respond to aggregate shocks? How should these temporary movements be modeled, and what impact do they have on the contract multiplier? What drives the relationship between the cyclicity of goods and the frequency with which prices change? Related, what aspects of micro heterogeneity in price-setting are important to consider in macro models? Finally, what evidence can be used to distinguish between different sources of the contract multiplier, such as rational inattention, sticky plans, or strategic complementarities?

REFERENCES

- Álvarez, L.J., 2008. What do micro price data tell us on the validity of the New Keynesian Phillips Curve? *Economics: The Open-Access, Open-Assessment E-Journal* 2 (19), 1–36.
- Álvarez, L.J., Burriel, P., Hernando, I., 2008. Price setting behaviour in Spain: Evidence from micro PPI data. *Managerial and Decision Economics* in press.
- Álvarez, L.J., Hernando, I., 2004. Price setting behaviour in Spain: Stylized facts using consumer price micro-data. Banco de España, Spain Working Paper 0422.
- Álvarez, L.J., Hernando, I., 2006. Price setting behaviour in Spain. Evidence from consumer price micro-data. *Economic Modelling* 23, 699–716.
- Álvarez, L.J., Hernando, I., 2007a. The price setting behaviour of Spanish firms: Evidence from survey data. In: Fabiani, S., Loupias, C., Martins, F., Sabbatini, R. (Eds.), *Pricing decisions in the Euro Area: How firms set prices and why*. Oxford University Press, Oxford, UK.
- Álvarez, L.J., Hernando, I., 2007b. Competition and price adjustment in the Euro Area. In: Fabiani, S., Loupias, C., Martins, F., Sabbatini, R. (Eds.), *Pricing decisions in the Euro Area: How firms set prices and why*. Oxford University Press, Oxford, UK.
- Amirault, D., Kwan, C., Wilkinson, G., 2006. Survey of price-setting behaviour of Canadian companies. Bank of Canada, Canada Working Paper 2006–35.
- Apel, M., Friberg, R., Hallsten, K., 2005. Microfoundations of macroeconomic price adjustment: Survey evidence from Swedish firms. *Journal of Money, Credit and Banking* 37 (April), 313–338.
- Aucremanne, L., Dhyne, E., 2004. How frequently do prices change? Evidence based on the micro data underlying the Belgian CPI. ECB Working Paper 331.
- Aucremanne, L., Druant, M., 2005. Price-setting behaviour in Belgium: What can be learned from an ad hoc survey. ECB Working Paper 448.
- Baharad, E., Eden, B., 2004. Price rigidity and price dispersion: Evidence from micro data. *Review of Economic Dynamics* 7 (July), 613–641.
- Ball, L., Mankiw, N.G., 1995. Relative-price changes as aggregate supply shocks. *Quarterly Journal of Economics* 110 (February), 161–193.
- Ball, L., Romer, D., 1990. Real rigidities and the non-neutrality of money. *Review of Economic Studies* 57 (April), 183–203.
- Barro, R.J., 1977. Long-term contracting, sticky prices and monetary policy. *J. Monetary Econ.* 3 (July), 305–316.
- Barros, R., Bonomo, M., Carvalho, C., Matos, S., 2009. Price setting in a variable macroeconomic environment: Evidence from Brazilian CPI. Getulio Vargas Foundation and Federal Reserve Bank of New York Unpublished paper.

- Barsky, R., House, C.L., Kimball, M., 2007. Sticky-price models and durable goods. *Am. Econ. Rev.* 97 (June), 984–998.
- Baudry, L., Le Bihan, H., Sevestre, P., Tarrieu, S., 2007. What do thirteen million price records have to say about consumer price rigidity? *Oxford Bull. Econ. Stat.* 69 (2), 139–183.
- Baumgartner, J., Glatzer, E., Ruml, F., Stiglitz, A., 2005. How frequently do consumer prices change in Austria? Evidence from micro CPI Data. ECB Working Paper 523.
- Bernanke, B.S., Boivin, J., Elias, P., 2005. Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *Quarterly Journal of Economics* 120 (February), 387–422.
- Bils, M., 2004. Studying price markups from stockout behavior. University of Rochester, New York Unpublished paper (December).
- Bils, M., Klenow, P.J., 1998. Using consumer theory to test competing business cycle models. *J. Polit. Econ.* 103 (April), 233–261.
- Bils, M., Klenow, P.J., 2004. Some evidence on the importance of sticky prices. *J. Polit. Econ.* 112 (October), 947–985.
- Bils, M., Klenow, P.J., Kryvtsov, O., 2003. Sticky prices and monetary policy shocks. Federal Reserve Bank of Minneapolis Quarterly Review (Winter), 2–9.
- Bils, M., Klenow, P.J., Malin, B.A., 2009. Reset price inflation and the impact of monetary policy shocks. NBER Working Paper 14787.
- Blinder, A.S., Canetti, E., Lebow, D., Rudd, J., 1998. Asking about prices: A new approach to understanding price stickiness. Russell Sage Foundation, New York.
- Boivin, J., Clark, R., Vincent, N., 2009. Virtual borders: Online nominal rigidities and international market segmentation. HEC Montreal Unpublished paper.
- Boivin, J., Giannoni, M., Mihov, I., 2009. Sticky prices and monetary policy: Evidence from disaggregated U.S. data. *Am. Econ. Rev.* 99 (March), 350–384.
- Broda, C., Weinstein, D., 2007. Product creation and destruction: Evidence and price implications. NBER Working Paper 13041.
- Buckle, R.A., Carlson, J.A., 2000. Menu costs, firm size and price rigidities. *Economics Letters* 66 (January), 59–63.
- Bunn, P., Ellis, C., 2009. Price-setting behaviour in the United Kingdom: A microdata approach. Bank of England Quarterly Bulletin 2009 Q1.
- Burstein, A., 2006. Inflation and output dynamics with state dependent pricing decisions. *J. Monetary Econ.* 53 (7/October), 1235–1257.
- Burstein, A., Hellwig, C., 2007. Prices and market shares in a menu cost model. UCLA, California Unpublished paper.
- Burstein, A., Jaimovich, N., 2009. Understanding movements in aggregate and product-level real-exchange rates. UCLA and Stanford University, California Unpublished paper.
- Calvo, G.A., 1983. Staggered prices in a utility-maximizing framework. *J. Monetary Econ.* 12 (September), 383–398.
- Campbell, J.R., Eden, B., 2005. Rigid prices: Evidence from U.S. scanner data. Federal Reserve Bank of Chicago and Vanderbilt University Unpublished paper.
- Carlton, D.W., 1986. The rigidity of prices. *Am. Econ. Rev.* 76 (September), 637–658.
- Carvalho, C., 2006. Heterogeneity in price stickiness and the real effects of monetary shocks. *Frontiers of Macroeconomics* 2 (1) Article 1.
- Castanon, V., Murillo, J.A., Salas, J., 2008. Formación de precios en la industria manufacturera de México. *El Trimestre Económico* 75 (1), 143–181.
- Cavallo, A., 2009. Scraped online data and sticky prices: Frequency, hazards and synchronization. Harvard University Unpublished paper.
- Cecchetti, S.G., 1986. The frequency of price adjustment: A study of newsstand prices of magazines. *Journal of Econometrics* 31 (April), 255–274.
- Christiano, L.J., Eichenbaum, M., Evans, C., 1999. Monetary policy shocks: What have we learned and to what end? In: Taylor, J.B., Woodford, M. (Eds.), *Handbook of macroeconomics*. 1A, Elsevier, New York.
- Christiano, L.J., Eichenbaum, M., Evans, C., 2005. Nominal rigidities and the dynamic effects of shocks to monetary policy. *J. Polit. Econ.* 113 (February), 1–45.

- Copaciu, M., Florian, N., Horia, B.E., 2007. Survey evidence on price setting patterns of Romanian firms. National Bank of Romania Unpublished paper.
- Cornille, D., Dossche, M., 2008. Some evidence on the adjustment of producer prices. *Scandinavian J. Econ.* 110 (September), 489–518.
- Cremer, K., 2008. Price setting behaviour in South Africa Stylised facts using producer price microdata. University of the Witwatersrand Unpublished paper.
- Cremer, K., Rankin, N.A., 2008. Price setting in South Africa 2001–2007 — stylised facts using consumer price micro data. *Journal of Development Perspectives* 1 (4), 93–118.
- Dabusinskas, A., Randveer, M., 2006. Comparison of pricing behaviour of firms in the Euro Area and Estonia. Bank of Estonia Working Paper 2006-08.
- Dhyne, E., Álvarez, L.J., Le Bihan, H., Veronese, G., Dias, D., Hoffmann, J., et al., 2005. Price setting in the Euro Area: Some stylized facts from individual consumer price data. ECB Working Paper 524.
- Dhyne, E., Alvarez, L.J., Le Bihan, H., Veronese, G., Dias, D., Hoffmann, J., et al., 2006. Price changes in the Euro Area and the United States: Some facts from individual consumer price data. *J. Econ. Perspect.* 20 (Spring), 171–192.
- Dhyne, E., Konieczny, J., 2010. Aggregation and the Staggering of Price Changes. Wilfrid Laurier University Unpublished paper.
- Dias, M., Dias, D., Neves, P.D., 2004. Stylised features of price setting behaviour in Portugal: 19922001. ECB Working Paper 332.
- Domberger, S., Fiebig, D.G., 1993. The distribution of price changes in oligopoly. *The Journal of Industrial Economics* 41 (September), 295–313.
- Dotsey, M., King, R., Wolman, A., 1999. State-dependent pricing and the general equilibrium dynamics of money and output. *Quarterly Journal of Economics* 114 (May), 655–690.
- Druant, M., Fabiani, S., Kezdi, G., Lamo, A., Martins, F., Sabbatini, R., 2009. How are firms' wages and prices linked: Survey evidence in Europe?. ECB Working Paper 1084 (August).
- Eden, B., 2001. Inflation and price adjustment: An analysis of microdata. *Review of Economic Dynamics* 4 (October), 607–636.
- Eichenbaum, M., Jaimovich, N., Rebelo, S., 2009. Reference prices and nominal rigidities. Northwestern University and Stanford University Unpublished paper.
- Fabiani, S., Druant, M., Hernando, I., Kwapil, C., Landau, B., Loupias, C., et al., 2005. The pricing behavior of firms in the Euro Area: New survey evidence. ECB Working Paper 535.
- Fabiani, S., Gattulli, A., Sabbatini, R., 2007. The pricing behavior of Italian firms. New survey evidence on price stickiness. In: Fabiani, S., Loupias, C., Martins, F., Sabbatini, R. (Eds.), *Pricing decisions in the Euro Area: How Firms Set Prices and Why*. Oxford University Press, Oxford, UK.
- Fabiani, S., Gattulli, A., Sabbatini, R., Veronese, G., 2006. Consumer price setting in Italy. *Giornale degli Economisti e Annali di Economia* 65 (1), 31–74.
- Fisher, T.C.G., Konieczny, J.D., 2000. Synchronization of price changes by multiproduct firms: Evidence from Canadian newspaper prices. *Economics Letters* 68, 271–277.
- Fitzgerald, D., Haller, S., 2009. Pricing-to-market: Evidence from producer prices. Unpublished paper.
- Gabriel, P., Reiff, A., 2008. Price setting in Hungary — A store-level analysis. Magyar Nemzeti Bank, Hungary Unpublished paper.
- Gagnon, E., 2009. Price setting during low and high inflation: Evidence from Mexico. *Quarterly Journal of Economics* 124 (August), 1221–1263.
- Gautier, E., 2008. The behaviour of producer prices: Some evidence from the French PPI micro data. *Empirical Economics* 35 (September), 301–332.
- Gertler, M., Leahy, J., 2008. A Phillips Curve with an Ss foundation. *J. Polit. Econ.* 116 (June), 533–572.
- Goldberg, P.K., Hellerstein, R., 2009. How rigid are producer prices. FRBNY Staff Report 407.
- Golosov, M., Lucas, R.E., 2007. Menu costs and Phillips curves. *J. Polit. Econ.* 115 (April), 171–199.
- Gopinath, G., Gourinchas, P.O., Hsieh, C.T., Li, N., 2009. Estimating the border effect: Some new evidence. Harvard University, University of California at Berkeley, and University of Chicago Unpublished paper.
- Gopinath, G., Itkhoki, O., 2010. Frequency of price-adjustment and pass-through. *Q. J. Econ.* 125 (May), 675–727.

- Gopinath, G., Itskhoki, O., Rigobon, R., 2010. Currency choice and exchange rate pass-through. *Am. Econ. Rev.* 100 (March), 304–336.
- Gopinath, G., Rigobon, R., 2008. Sticky borders. *Quarterly Journal of Economics* 123 (May), 531–575.
- Gouvea, S., 2007. Price rigidity in Brazil: Evidence from CPI micro data. Central Bank of Brazil, Brazil Working Paper 143.
- Guimaraes, B., Sheedy, K.D., 2008. Sales and monetary policy. CEPR Discussion Paper 6940.
- Hall, S., Walsh, M., Yates, A., 2000. Are UK companies' prices sticky? *Oxford Economic Papers* 52 (3), 425–446.
- Hansen, B.W., Hansen, N.L., 2006. Price setting behavior in Denmark: A study of CPI micro data 1997–2005. Danmarks Nationalbank, Denmark Working Paper 39.
- Hobijn, B., Ravenna, F., Tambalotti, A., 2006. Menu costs at work: Restaurant prices and the introduction of the Euro. *Quarterly Journal of Economics* 121 (August), 1103–1131.
- Hoerberichts, M., Stokman, A., 2006. Pricing behaviour of Dutch companies: Results of a survey. ECP Working Paper 607.
- Hoffmann, J., Kurz-Kim, J.R., 2006. Consumer price adjustment under the microscope: Germany in a period of low inflation. ECB Working Paper 652.
- Horvath, R., Coricelli, F., 2006. Price setting behaviour: Micro evidence on Slovakia. CEPR Discussion Papers 5445.
- Ikeda, D., Nishioka, S., 2007. Price setting behavior and hazard functions: Evidence from Japanese CPI micro data. Bank of Japan, Japan Working Paper 07-E-19.
- Jonker, N., Folkertsma, K., Blijenberg, H., 2004. Empirical analysis of price setting behaviour in the Netherlands in the period 19982003 using micro data. ECB Working Paper 413.
- Julio, J.M., Zárate, H.M., 2008. The price setting behaviour in Colombia: Evidence from PPI micro data. Banco de la República, Colombia Borradores de Economía.
- Kashyap, A.K., 1995. Sticky prices: New evidence from retail catalogues. *Quarterly Journal of Economics* 110 (February), 245–274.
- Kehoe, P.J., Midrigan, V., 2008. Temporary price changes and the real effects of monetary policy. Federal Reserve Bank of Minneapolis Research Department Staff Report 413 (September).
- Kimball, M.S., 1995. The quantitative analytics of the basic neomonetarist model. *Journal of Money, Credit and Banking* 27 (November), 1241–1277.
- Klenow, P.J., Kryvtsov, O., 2008. State-dependent or time-dependent pricing: Does it matter for recent U.S. inflation? *Quarterly Journal of Economics* 123 (August), 863–904.
- Klenow, P.J., Willis, J.L., 2006. Real rigidities and nominal price changes. Stanford University and Federal Reserve Bank of Kansas City Unpublished paper.
- Klenow, P.J., Willis, J.L., 2007. Sticky information and sticky prices. *J. Monetary Econ.* 54 (September), 79–99.
- Konieczny, J.D., Skrzypacz, A., 2005. Inflation and price setting in a natural experiment. *J. Monetary Econ.* 52 (April), 621–632.
- Kovanen, A., 2006. Why do prices in Sierra Leone change so often? A case study using micro-level price data. International Monetary Fund Working Paper 06/53.
- Kryvtsov, O., Midrigan, V., 2009. Inventories, markups, and real rigidities in menu cost models. Bank of Canada Unpublished paper.
- Kwapil, C., Baumgartner, J., Scharler, J., 2005. The price-setting behaviour of Austrian Firms: Some survey evidence. ECB Working Paper 464.
- Lach, S., Tsiddon, D., 1992. The behavior of prices and inflation: An empirical analysis of disaggregated price data. *J. Polit. Econ.* 100 (April), 349–389.
- Lach, S., Tsiddon, D., 1996. Staggering and synchronization in price-setting: Evidence from multiproduct firms. *Am. Econ. Rev.* 86 (December), 1175–1196.
- Levy, D., Dutta, S., Bergen, M., Venable, R., 1998. Price adjustment and multiproduct retailers. *Managerial and Decision Economics* 19, 81–120.
- Levy, D., Lee, D., Chen, A., Kauffman, R.J., Bergen, M., 2007. Price points and price rigidity. The Rimini Center for Economic Analysis Working Paper 04–07.

- Loupas, C., Ricart, R., 2004. Price setting in France: New evidence from survey data. ECB Working Paper 423.
- Lünnemann, P., Mathä, T.Y., 2005. Consumer price behaviour in Luxembourg: Evidence from micro CPI Data. ECB Working Paper 541.
- Lünnemann, P., Mathä, T.Y., 2006. New survey evidence on the pricing behavior of Luxembourg firms. ECB Working Paper 617.
- Lünnemann, P., Wintr, L., 2006. Are Internet prices sticky?. ECB Working Paper 645.
- Mackowiak, B., Smets, F., 2008. Implications of micro price data for macroeconomic models. CEPR Discussion Paper 6961.
- Mackowiak, B., Wiederholt, M., 2008. Business cycle dynamics under rational inattention. European Central Bank and Northwestern University Unpublished paper.
- Mankiw, N.G., Reis, R., 2002. Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve. *Quarterly Journal of Economics* 117 (November), 1295–1328.
- Martins, F., 2005. The price setting behavior of Portuguese Firms. Evidence from survey data. ECP Working Paper 562.
- Medina, J.P., Rappoport, D., Soto, C., 2007. Dynamics of price adjustment: Evidence from micro level data for Chile. Central Bank of Chile, Chile Working Paper 432.
- Midrigan, V., 2009. Menu costs, multiproduct firms, and aggregate fluctuations. New York University Unpublished paper.
- Nakagawa, S., Hattori, R., Takagawa, I., 2000. Price setting behavior of Japanese companies. Bank of Japan Research paper.
- Nakamura, E., 2008. Pass-through in retail and wholesale. *Am. Econ. Rev.* 98 (May), 430–437.
- Nakamura, E., Steinsson, J., 2008a. Five facts about prices: A reevaluation of menu cost models. *Q. J. Econ.* 123 (November), 1415–1464.
- Nakamura, E., Steinsson, J., 2008b. Monetary non-neutrality in a multi-sector menu cost model. NBER Working Paper 14001 (May).
- Nakamura, E., Steinsson, J., 2009. Lost in transit: Product replacement bias and pricing to market. Columbia University Unpublished paper.
- Nakamura, E., Zerom, D., 2010. Accounting for incomplete pass-through. *Rev. Econ. Stud.* 77 (July), 1192–1230.
- Peersman, G., Smets, F., 2003. The monetary transmission mechanism in the Euro Area: More evidence from VAR analysis. In: Angeloni, I., Kashyap, A., Mojon, B. (Eds.), *Monetary policy transmission in the Euro Area*. University Press, Cambridge, UK.
- Peneva, E., 2009. Factor intensity and price rigidity: Evidence and theory. FEDS Working Paper No. 2009-07 (January).
- Romer, D.H., Romer, C.D., 2004. A new measure of monetary shocks: Derivation and implications. *Am. Econ. Rev.* 94 (September), 1055–1084.
- Rotemberg, J.J., 1982. Monopolistic price adjustment and aggregate output. *Review of Economic Studies* 49 (October), 517–531.
- Sabbatini, R., Fabiani, S., Gatulli, A., Veronese, G., 2006. Producer price behaviour in Italy: Evidence from micro PPI data. Banca d'Italia, Italy Unpublished paper.
- Sahinoz, S., Saracoglu, B., 2008. Price setting behaviour in Turkish industries: Evidence from survey data. Turkish Economic Association Discussion Paper 2008/3.
- Saita, Y., Takagawa, I., Nishizaki, K., Higo, M., 2006. Price setting in Japan: Evidence from individual retail price data. Bank of Japan Working Paper Series, No. 06-J-02 (in Japanese).
- Stahl, H., 2005. Price setting in German manufacturing: New evidence from new survey data. Deutsche Bundesbank, Germany Discussion Paper 43/2005.
- Stahl, H., 2006. Producer price adjustment at the micro level: Evidence from individual price records underlying the German PPI. Deutsche Bundesbank, Germany Unpublished paper.
- Taylor, J.B., 1980. Aggregated dynamics and staggered contracts. *J. Polit. Econ.* 88 (February), 1–24.
- Vermeulen, P., Dias, D., Dossche, M., Gautier, E., Hernando, I., Sabbatini, R., et al., 2007. Price setting in the Euro Area: Some stylized facts from individual producer price data. ECB Working Paper 727.

- Vilmunen, J., Laakkonen, H., 2005. How often do prices change in Finland? Micro-level evidence from the CPI. Bank of Finland Unpublished paper.
- Wolman, A., 2007. The frequency and costs of individual price adjustment. *Managerial and Decision Economics* 28 (6), 531–552.
- Woodford, M., 2009. Information-constrained state-dependent pricing. *J. Monetary Econ.* 56 (October), S100–S124.
- Wulfsberg, F., 2009. Price adjustments and inflation: Evidence from consumer price data in Norway 1975–2004. Norges Bank WP 2009/11.

PART *Three*

Models of the Monetary Transmission Mechanism

This page intentionally left blank

DSGE Models for Monetary Policy Analysis[☆]

Lawrence J. Christiano,^{*} Mathias Trabandt,^{**} and Karl Walentin[†]

^{*}Department of Economics, Northwestern University

^{**}European Central Bank, Germany and Sveriges Riksbank, Sweden

[†]Research Division, Sveriges Riksbank, Sweden

Contents

1. Introduction	286
2. Simple Model	289
2.1 Private economy	290
2.1.1 Households	290
2.1.2 Firms	290
2.1.3 Aggregate resources and the private sector equilibrium conditions	294
2.2 Log-linearized equilibrium with Taylor rule	296
2.3 Frisch labor supply elasticity	299
3. Simple Model: Some Implications for Monetary Policy	302
3.1 Taylor principle	303
3.2 Monetary policy and inefficient booms	309
3.3 Using unemployment to estimate the output gap	311
3.3.1 A measure of the information content of unemployment	311
3.3.2 The CTW model of unemployment	312
3.3.3 Limited information Bayesian inference	315
3.3.4 Estimating the output gap using the CTW model	319
3.4 Using HP-filtered output to estimate the output gap	326
4. Medium-Sized DSGE Model	331
4.1 Goods production	331
4.2 Households	334
4.2.1 Households and the labor market	335
4.2.2 Wages, employment and monopoly unions	338
4.2.3 Capital accumulation	340
4.2.4 Household optimization problem	343
4.3 Fiscal and monetary authorities and equilibrium	344
4.4 Adjustment cost functions	344
5. Estimation Strategy	345
5.1 VAR step	345

[☆] We are grateful for advice from Michael Woodford and for comments from Volker Wieland. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the European Central Bank or of Sveriges Riksbank. We are grateful for assistance from Daisuke Ikeda and Matthias Kehrig.

5.2 Impulse response matching step	347
5.3 Computation of V	348
5.4 Laplace approximation of the posterior distribution	350
6. Medium-Sized DSGE Model: Results	351
6.1 VAR results	351
6.1.1 <i>Monetary policy shocks</i>	351
6.1.2 <i>Technology shocks</i>	355
6.2 Model results	355
6.2.1 <i>Parameters</i>	355
6.2.2 <i>Impulse responses</i>	358
6.3 Assessing VAR robustness and accuracy of the Laplace approximation	360
7. Conclusion	362
References	364

Abstract

Monetary DSGE models are widely used because they fit the data well and they can be used to address important monetary policy questions. We provide a selective review of these developments. Policy analysis with DSGE models requires using data to assign numerical values to model parameters. The chapter describes and implements Bayesian moment matching and impulse response matching procedures for this purpose.

JEL Classification: E2, E3, E5, J6

Keywords

Frisch Labor Supply Elasticity
 HP Filter
 Impulse Response Function
 Limited Information Bayesian Estimation
 Materials Input for Production
 New Keynesian DSGE Models
 Output Gap
 Potential Output
 Taylor Principle
 Unemployment
 Vector Autoregression
 Working Capital Channel

1. INTRODUCTION

There has been enormous progress in recent years in the development of dynamic, stochastic general equilibrium (DSGE) models for the purpose of monetary policy analysis. These models have been shown to fit aggregate data well by conventional econometric measures. For example, they have been shown to do as well or better than simple atheoretical statistical models at forecasting outside the sample of data on which they were estimated. In part because of these successes, a consensus has formed around a particular model structure, the New Keynesian model.

Our objective is to present a selective review of these developments. We present several examples to illustrate the kind of policy questions the models can be used to address. We also convey a sense of how well the models fit the data. In all cases, our discussion takes place in the simplest version of the model required to make our point. As a result, we do not develop one single model. Instead, we work with several models.

We begin by presenting a detailed derivation of a version of the standard New Keynesian model with price-setting frictions and no capital or other complications. We then use versions of this simple model to address several important policy issues. For example, the past few decades have witnessed the emergence of a consensus that monetary policy ought to respond aggressively to changes in actual or expected inflation. This prescription for monetary policy is known as the “Taylor principle.” The standard version of the simple model is used to articulate why this prescription is a good one. However, alternative versions of the model can be used to identify potential pitfalls for the Taylor principle. In particular, a policy-induced rise in the nominal interest rate may destabilize the economy by perversely giving a direct boost to inflation. This can happen if the standard model is modified to incorporate a so-called working capital channel, which corresponds to the assumption that firms must borrow to finance their variable inputs.

We then turn to the much discussed issue of the interaction between monetary policy and volatility in asset prices and other aggregate economic variables. We explain how vigorous application of the Taylor principle could inadvertently trigger an inefficient boom in output and asset prices.

Finally, we discuss the use of DSGE models for addressing a key policy question: How big is the gap between the level of economic activity and the best level that is achievable by policy? An estimate of the output gap not only provides an indication about how efficiently resources are being used, but in the New Keynesian framework, the output gap is also a signal of inflation pressure. Informally, the unemployment rate is thought to provide a direct observation on the efficiency of resource allocation. For example, a large increase in the number of people reporting to be “ready and willing to work” but not employed suggests, at least at a casual level, that resources are being wasted and that the output gap is negative. DSGE models can be used to formalize and assess these informal hunches. We do this by introducing unemployment into the standard New Keynesian model along the lines recently proposed in Christiano, Trabandt, and Walentin (2010a; CTW). We use the model to describe circumstances in which we can expect the unemployment rate to provide useful information about the output gap. We also report evidence suggesting that these conditions may be satisfied in the U.S. data.

Although the creators of the Hodrick and Prescott (1997; HP) filter never intended it to be used to estimate the New Keynesian output gap concept, it is often used for this purpose. We show that whether the HP filter is a good estimator of the gap depends sensitively on the details of the underlying model economy. This discussion involves a careful review of the intuition of how the New Keynesian model responds to shocks. Interestingly, a New Keynesian model fit to U.S. data suggests the conditions are satisfied for the HP filter to be a good estimator of the output gap. In our

discussion, we explain that there are several caveats that must be taken into account before concluding that the HP filter is a good estimator of the output gap.

Policy analysis with DSGE models, even the simple analyses summarized earlier, require assigning values to model parameters. In recent years, the Bayesian approach to econometrics has taken over as the dominant one for this purpose. In conventional applications, the Bayesian approach is a so-called full information procedure because the analyst specifies the joint likelihood of the available observations in complete detail. As a result, many of the limited information tools in macroeconomists' econometric toolbox have been deemphasized in recent times. These tools include methods that match model and data second moments and that match model and empirical impulse response functions. Following the work of Chernozhukov and Hong (2003), Kim (2002), Kwan (1999) and others, we show how the Bayesian approach can be applied in limited information contexts. We apply a Bayesian moment matching approach in Section 3.3.3 and a Bayesian impulse response function matching approach in Section 5.2.

The new monetary DSGE models are of interest not just because they represent laboratories for the analysis of important monetary policy questions. They are also of interest because they appear to resolve a classic empirical puzzle about the effects of monetary policy. It has long been thought that it is virtually impossible to explain the very slow response of inflation to a monetary disturbance without appealing to completely implausible assumptions about price frictions (see, e.g., Mankiw, 2000). However, it turns out that modern DSGE models do provide an account of the inertia in inflation and the strong response of real variables to monetary policy disturbances, without appealing to seemingly implausible parameter values. Moreover, the models simultaneously explain the dynamic response of the economy to other shocks. We review these important findings. We explain in detail the contribution of each feature of the consensus medium-sized New Keynesian model in achieving this result. This discussion closely follows the analyses in Christiano, Eichenbaum, and Evans (2005; CEE) and Altig, Christiano, Eichenbaum, and Lindé (2005; ACEL).

There is an econometric technique that is particularly well-suited to the shock-based analysis described in the previous paragraph. It is the one that matches impulse response functions estimated by vector autoregressions (VARs) with the corresponding objects in a model. Using U.S. macroeconomic data, we show how the parameters of the consensus DSGE model are estimated by this impulse response matching procedure. The advantage of this econometric approach is transparency and focus. The transparency reflects that the estimation strategy has a simple graphical representation, involving objects — impulse response functions — about which economists have strong intuition. The advantage of focus comes from the possibility of studying the empirical properties of a model without having to specify a full set of shocks. As noted previously, we show how to implement the impulse response matching strategy using Bayesian methods. In particular, we are able to implement all the machinery of priors and posteriors, as well as the marginal likelihood as a measure of model fit in our impulse response function matching exercise.

This chapter is organized as follows. Section 2 describes the simple New Keynesian model without capital. The following section reviews some policy implications of that model. The medium-sized version of the model, designed to econometrically address a rich set of macroeconomic data, is described in Section 4. Section 5 reviews our Bayesian impulse response matching strategy. Section 6 reviews the results, and conclusions are offered in Section 7. Many algebraic derivations are relegated to a separate technical appendix.¹

2. SIMPLE MODEL

This section analyzes versions of the standard Calvo-sticky price New Keynesian model without capital. In practice, the analysis of the standard New Keynesian model often begins with the familiar three equations: the linearized “Phillips curve,” “IS curve,” and monetary policy rule. We cannot simply begin with these three equations here because we also study departures from the standard model. For this reason, we must derive the equilibrium conditions from their foundations.

The version of the New Keynesian model studied in this section is the one considered in Clarida, Gali, and Gertler (1999) and Woodford (2003), modified in two ways. First, we introduce the working capital channel emphasized by CEE and Barth and Ramey (2002).² The working capital channel results from the assumption that firms’ variable inputs must be financed by short-term loans. With this assumption, changes in the interest rate affect the economy by changing firms’ variable production costs, in addition to operating through the usual spending mechanism. There are several reasons to take the working capital channel seriously. Using U.S. Flow of Funds data, Barth and Ramey (2002) argued that a substantial fraction of firms’ variable input costs are borrowed in advance. Christiano, Eichenbaum, and Evans (1996) provided VAR evidence suggesting the presence of a working capital channel. Chowdhury, Hoffmann, and Schabert (2006) and Ravenna and Walsh (2006) provided additional evidence supporting the working capital channel, based on instrumental variables estimates of a suitably modified Phillips curve. Finally, Section 4 shows that incorporating the working capital channel helps to explain the “price puzzle” in the VAR literature and provides a response to Ball’s (1994) “dis-inflationary boom” critique of sticky price models.

We explore a second modification to the classic New Keynesian model by incorporating the assumption about materials inputs proposed in Basu (1995). Basu argued that a large part — as much as half — of a firm’s output is used as inputs by other firms. The working capital channel introduces the interest rate into costs while the materials assumption makes those costs big. In the next section we show that these two factors have potentially far-reaching consequences for monetary policy.

¹ The technical appendix can be found at http://www.faculty.econ.northwestern.edu/faculty/christiano/research/Handbook/technical_appendix.pdf.

² The first monetary DSGE model we are aware of that incorporates a working capital channel is Fuerst (1992). Other early examples include Christiano (1991) and Christiano and Eichenbaum (1992b).

This section is organized as follows. We begin in subsection 2.1 by describing the private sector of the economy, and deriving equilibrium conditions associated with optimization and market clearing. In subsection 2.2, we specify the monetary policy rule and define the Taylor rule equilibrium. Subsection 2.3 discusses the interpretation of a key parameter in our utility function. The parameter controls the elasticity with which the labor input in our model economy adjusts in response to a change in the real wage. Traditionally, this parameter has been viewed as being restricted by microeconomic evidence on the Frisch labor supply elasticity. We summarize recent thinking stimulated by the seminal work of Rogerson (1988) and Hansen (1985), according to which this parameter is not restricted by evidence on the Frisch elasticity.

2.1 Private economy

2.1.1 Households

We suppose there is a large number of identical households. The representative household solves the following problem:

$$\max_{\{C_t, H_t, B_{t+1}\}} E_0 \sum_{t=0}^{\infty} \beta^t \left(\log C_t - \frac{H_t^{1+\phi}}{1+\phi} \right), 0 < \beta < 1, \phi \geq 0, \quad (1)$$

subject to

$$P_t C_t + B_{t+1} \leq B_t R_{t-1} + W_t H_t + \text{Transfers and profits}_t. \quad (2)$$

Here, C_t and H_t denote household consumption and market work, respectively. In Eq. (2), B_{t+1} denotes the quantity of a nominal bond purchased by the household in period t and R_t denotes the one-period gross nominal rate of interest on a bond purchased in period t . Finally, W_t denotes the competitively determined nominal wage rate. The parameter, ϕ , is discussed in Section 2.3.

The representative household equates the marginal cost of working, in consumption units, with the marginal benefit, the real wage:

$$C_t H_t^\phi = \frac{W_t}{P_t}. \quad (3)$$

The representative household also equates the utility cost of the consumption foregone in acquiring a bond with the corresponding benefit:

$$\frac{1}{C_t} = \beta E_t \frac{1}{C_{t+1}} \frac{R_t}{\pi_{t+1}}. \quad (4)$$

Here, π_{t+1} denotes the gross rate of inflation from t to $t + 1$.

2.1.2 Firms

A key feature of the New Keynesian model is its assumption that there are price-setting frictions. These frictions are introduced to accommodate the evidence of inertia in

aggregate inflation. Obviously, the presence of price-setting frictions requires that firms have the power to set prices, and this in turn requires the presence of monopoly power. A challenge is to create an environment in which there is monopoly power, without contradicting the obvious fact that actual economies have a very large number of firms. The Dixit-Stiglitz framework of production handles this challenge very nicely, because it has a very large number of price-setting monopolist firms. In particular, gross output is produced using a representative, competitive firm using the following technology:

$$Y_t = \left(\int_0^1 Y_{i,t}^{\frac{1}{\lambda_f}} di \right)^{\lambda_f}, \lambda_f > 1, \quad (5)$$

where λ_f governs the degree of substitution between the different inputs. The representative firm takes the price of gross output, P_t , and the price of intermediate inputs, $P_{i,t}$, as given. Profit maximization leads to the following first-order condition:

$$Y_{i,t} = Y_t \left(\frac{P_{i,t}}{P_t} \right)^{-\frac{\lambda_f}{\lambda_f-1}}. \quad (6)$$

Substituting Eq. (6) into Eq. (5) yields the following relation between the aggregate price level and the prices of intermediate goods:

$$P_t = \left(\int_0^1 P_{i,t}^{\frac{1}{\lambda_f-1}} di \right)^{-(\lambda_f-1)}. \quad (7)$$

The i^{th} intermediate good is produced by a single monopolist, who takes Eq. (6) as its demand curve. The value of λ_f determines how much monopoly power the i^{th} producer has. If λ_f is large, then intermediate goods are poor substitutes for each other, and the monopoly supplier of good i has a lot of market power. Consistent with this, note that if λ_f is large, then the demand for $Y_{i,t}$ is relatively price inelastic (see Eq. 6). If λ_f is close to unity, so that each $Y_{i,t}$ is almost a perfect substitute for $Y_{j,t}$, $j \neq i$, then the i^{th} firm faces a demand curve that is almost perfectly elastic. In this case, the firm has virtually no market power.

The production function of the i^{th} monopolist is:

$$Y_{i,t} = z_t H_{i,t}^\gamma I_{i,t}^{1-\gamma}, 0 < \gamma \leq 1, \quad (8)$$

where z_t is a technology shock whose stochastic properties are specified below. Here, $H_{i,t}$, denotes the level of employment by the i^{th} monopolist. We follow Basu (1995) in supposing that the i^{th} monopolist uses the quantity of materials, $I_{i,t}$, as inputs to production. The materials, $I_{i,t}$, are converted one-for-one from Y_t in Eq. (5). For $\gamma < 1$, each intermediate good producer in effect uses the output of all the other intermediate producers as input. When $\gamma = 1$, then materials inputs are not used in production.

The nominal marginal cost of the intermediate good producer is the following Cobb-Douglas function of the price of its two inputs:

$$\text{marginal cost}_t = \left(\frac{\bar{P}_t}{1 - \gamma} \right)^{1-\gamma} \left(\frac{\bar{W}_t}{\gamma} \right)^\gamma \frac{1}{z_t}.$$

Here, $\bar{W}$ and $\bar{P}$ are the effective prices of $H_{i,t}$ and $I_{i,t}$, respectively:

$$\begin{aligned} \bar{W}_t &= (1 - \nu_t)(1 - \psi + \psi R_t) W_t \\ \bar{P}_t &= (1 - \nu_t)(1 - \psi + \psi R_t) P_t. \end{aligned} \quad (9)$$

In this expression, ν_t denotes a subsidy to intermediate good firms and the term involving the interest rate reflects the presence of a “working capital channel.” For example, $\psi = 1$ corresponds to the case where the full amount of the cost of labor and materials must be financed at the beginning of the period. When $\psi = 0$, no advanced financing is required. A key variable in the model is the ratio of nominal marginal cost to the price of gross output, P_t :

$$s_t = (1 - \nu_t) \left(\frac{1}{1 - \gamma} \right)^{1-\gamma} \left(\frac{\bar{w}_t}{\gamma} \right)^\gamma (1 - \psi + \psi R_t), \quad (10)$$

where $\bar{w}_t$ denotes the scaled real wage rate:

$$\bar{w}_t \equiv \frac{W_t}{z_t^\frac{1}{1-\gamma} P_t}. \quad (11)$$

If intermediate good firms faced no price-setting frictions, they would all set their price as a fixed markup over nominal marginal cost:

$$\lambda_f P_t s_t. \quad (12)$$

In fact, we assume there are price-setting frictions along the lines proposed by Calvo (1983). An intermediate firm can set its price optimally with probability $1 - \xi_p$, and with probability ξ_p it must keep its price unchanged relative to what it was in the previous period:

$$P_{i,t} = P_{i,t-1}.$$

Consider the $1 - \xi_p$ intermediate good firms that are able to set their prices optimally in period t . There are no state variables in the intermediate good firm problem and all the firms face the same demand curve. As a result, all firms able to optimize their prices in period t choose the same price, which we denote by $\tilde{P}_t$. It is clear that optimizing firms do not set $\tilde{P}_t$ equal to Eq. (12). Setting $\tilde{P}_t$ to Eq. (12) would be optimal from the perspective of the current period, but it does not take into account the possibility that the firm may be stuck with $\tilde{P}_t$ for several periods into the future. Instead, the intermediate good firms that have an opportunity to reoptimize their price in the current period, do so to solve:

$$\max_{\tilde{P}_t} E_t \sum_{j=0}^{\infty} (\xi_p \beta)^j \nu_{t+j} (\tilde{P}_t Y_{i,t+j} - P_{t+j} s_{t+j} Y_{i,t+j}), \quad (13)$$

subject to the demand curve, Eq. (6), and the definition of marginal cost, Eq. (10). In Eq. (13), $\beta^j v_{t+j}$ is the multiplier on the household's nominal period $t + j$ budget constraint. Because they are the owners of the intermediate good firms, households are the recipients of firm profits. In this way, it is natural that the firm should weigh profits in different dates and states of nature using $\beta^j v_{t+j}$. Intermediate good firms take v_{t+j} as given. The nature of the family's preferences, Eq. (1), implies:

$$v_{t+j} = \frac{1}{P_{t+j} C_{t+j}}.$$

In Eq. (13) the presence of ξ_p reflects that intermediate good firms are only concerned with future scenarios in which they are not able to reoptimize the price chosen in period t .

The first-order condition associated with Eq. (13) is

$$\tilde{p}_t = \frac{E_t \sum_{j=0}^{\infty} (\beta \xi_p)^j (X_{t,j})^{-\frac{\lambda_f}{\lambda_f-1}} \lambda_f s_{t+j}}{E_t \sum_{j=0}^{\infty} (\beta \xi_p)^j (X_{t,j})^{-\frac{1}{\lambda_f-1}}} = \frac{K_t^f}{F_t^f}, \quad (14)$$

where K_t^f and F_t^f denote the numerator and denominator of the ratio after the first equality, respectively. Also,

$$\tilde{p}_t \equiv \frac{\tilde{P}_t}{P_t}, X_{t,j} \equiv \begin{cases} \frac{1}{\pi_{t+j} \cdots \pi_{t+j}} & j > 0 \\ 1 & j = 0 \end{cases}.$$

Not surprisingly, Eq. (14) implies $\tilde{P}_t$ is set to Eq. (12) when $\xi_p = 0$. When $\xi_p > 0$, optimizing firms set their prices so that Eq. (12) is satisfied on average. It is useful to write the numerator and denominator in Eq. (14) in recursive form. Thus,

$$K_t^f = \lambda_f s_t + \beta \xi_p E_t \pi_{t+1}^{\frac{\lambda_f}{\lambda_f-1}} K_{t+1}^f, \quad (15)$$

$$F_t^f = 1 + \beta \xi_p E_t \pi_{t+1}^{\frac{1}{\lambda_f-1}} F_{t+1}^f. \quad (16)$$

Expression (7) simplifies when we take into account that (i) the $1 - \xi_p$ intermediate good firms that set their price optimally all set it to $\tilde{P}_t$ and (ii) the ξ_p firms that cannot reset their price are selected at random from the set of all firms. Doing so,

$$\tilde{p}_t = \left[\frac{1 - \xi_p \pi_t^{\frac{1}{\lambda_f-1}}}{1 - \xi_p} \right]^{-(\lambda_f-1)}. \quad (17)$$

It is convenient to use Eq. (17) to eliminate $\tilde{p}_t$ in Eq. (14):

$$K_t^f = F_t^f \left(\frac{1 - \xi_p \pi_t^{\frac{1}{\lambda_f - 1}}}{1 - \xi_p} \right)^{-(\lambda_f - 1)}. \quad (18)$$

When $\gamma < 1$, cost minimization by the i^{th} intermediate good producer leads it to equate the relative price of its labor and materials inputs to the corresponding relative marginal productivities:

$$\frac{\bar{W}_t}{\bar{P}_t} = \frac{W_t}{P_t} = \frac{\gamma}{1 - \gamma} \frac{I_{i,t}}{H_{i,t}} = \frac{\gamma}{1 - \gamma} \frac{I_t}{H_t}. \quad (19)$$

Evidently, each firm uses the same ratio of inputs, regardless of its output price, $P_{i,t}$.

2.1.3 Aggregate resources and the private sector equilibrium conditions

A notable feature of the New Keynesian model is the absence of an aggregate production function. That is, given information about aggregate inputs and technology, it is not possible to say what aggregate output, Y_t , is. This is because Y_t also depends on how inputs are distributed among the various intermediate good producers. For a given amount of aggregate inputs, Y_t is maximized by distributing the inputs equally across producers. An unequal distribution of inputs results in a lower level of Y_t . In the New Keynesian model with Calvo price frictions, resources are unequally allocated across intermediate good firms if, and only if, $P_{i,t}$ differs across i . Price dispersion in the model is caused by the interaction of inflation with price-setting frictions. With price dispersion, the price mechanism ceases to allocate resources efficiently, as too much production is done in firms with low prices and too little in the firms with high prices. Yun (1996) derived a very simple formula that characterizes the loss of output due to price dispersion. We re-derive the analog of Yun's (1996) formula that is relevant for our setting.

Let Y_t^* denote the unweighted integral of gross output across intermediate good producers:

$$Y_t^* \equiv \int_0^1 Y_{i,t} di = \int_0^1 z_t \left(\frac{H_{i,t}}{I_{i,t}} \right)^\gamma I_{i,t} di = z_t \left(\frac{H_t}{I_t} \right)^\gamma I_t = z_t H_t^\gamma I_t^{1-\gamma}.$$

Here, we have used linear homogeneity of the production function, as well as the result in Eq. (19), that all intermediate good producers use the same labor to materials ratio. An alternative representation of Y_t^* makes use of the demand curve, Eq. (6):

$$Y_t^* = Y_t \int_0^1 \left(\frac{P_{i,t}}{P_t} \right)^{-\frac{\lambda_f}{\lambda_f - 1}} di = Y_t P_t^{\frac{\lambda_f}{\lambda_f - 1}} \int_0^1 (P_{i,t})^{-\frac{\lambda_f}{\lambda_f - 1}} di = Y_t P_t^{\frac{\lambda_f}{\lambda_f - 1}} (P_t^*)^{-\frac{\lambda_f}{\lambda_f - 1}}. \quad (20)$$

Thus,

$$Y_t = p_t^* z_t H_t^\gamma I_t^{1-\gamma},$$

where

$$p_t^* \equiv \left(\frac{P_t^*}{P_t} \right)^{\frac{\lambda_f}{\lambda_f - 1}}. \quad (21)$$

Here, $P_t^* \leq 1$ denotes Yun's (1996) measure of the output lost due to price dispersion. From Eq. (20),

$$P_t^* = \left[\int_0^1 (P_{i,t})^{-\frac{\lambda_f}{\lambda_f - 1}} di \right]^{-\frac{\lambda_f - 1}{\lambda_f}}. \quad (22)$$

According to Eq. (21), P_t^* is a monotone function of the ratio of two different weighted averages of intermediate good prices. The ratio of these two weighted averages can only be at its maximum of unity if all prices are the same.³

Taking into account observations (i) and (ii) after Eq. (16), Eq. (22) reduces (after dividing by P_t and taking into account Eq. 21) to:

$$p_t^* = \left[(1 - \zeta_p) \left(\frac{1 - \zeta_p \pi_t^{\frac{1}{\lambda_f - 1}}}{1 - \zeta_p} \right)^{\lambda_f} + \zeta_p \frac{\pi_t^{\frac{\lambda_f}{\lambda_f - 1}}}{p_{t-1}^*} \right]^{-1}. \quad (23)$$

According to Eq. (23), there is price dispersion in the current period if there was dispersion in the previous period and/or if there is a current shock to dispersion. Such a shock must operate through the aggregate rate of inflation.

We conclude that the relation between aggregate inputs and gross output is given by:

$$C_t + I_t = p_t^* z_t H_t^\gamma I_t^{1-\gamma}. \quad (24)$$

Here, $C_t + I_t$ represents total gross output, while C_t represents value added.

The private sector equilibrium conditions of the model are Eqs. (3), (4), (10), (15), (16), (18), (19), (23), and (24). This represents 9 equations in the following 11 unknowns:

$$C_t, H_t, I_t, R_t, \pi_t, p_t^*, K_t^f, F_t^f, \frac{W_t}{P_t}, s_t, v_t. \quad (25)$$

³ The distortion, p_t^* , is of interest in its own right. It is a sort of "endogenous Solow residual" of the kind called for by Prescott (1998). Whether the magnitude of fluctuations in p_t^* are quantitatively important given the actual price dispersion in data is something that deserves exploration. A difficulty that must be overcome, in such an exploration, is determining what the benchmark efficient dispersion of prices is in the data. In the model it is efficient for all prices to be exactly the same, but that is obviously only a convenient normalization.

As it stands, the system is underdetermined. This is not surprising, since we have said nothing about monetary policy or how v_t is determined. We turn to this in the following section.

2.2 Log-linearized equilibrium with Taylor rule

We log-linearize the equilibrium conditions of the model about its nonstochastic steady state. We assume that monetary policy is governed by a Taylor rule, which responds to the deviation between actual inflation and a zero inflation target. As a result, inflation is zero in the nonstochastic steady state. In addition, we suppose that the intermediate good subsidy, v_t , is set to the constant value that causes the price of goods to equal the social marginal cost of production in steady state. To see what this implies for v_t , recall that in steady state firms set their price as a markup, λ_f , over marginal cost. That is, they equate the object in Eq. (12) to P_t , so that:

$$\lambda_f s = 1.$$

Using Eq. (10) to substitute out for the steady state value of s , the latter expression reduces, in steady state, to:

$$\lambda_t(1-v)(1-\psi+\psi R) \left[\left(\frac{1}{1-\gamma} \right)^{1-\gamma} \left(\frac{\bar{w}}{\gamma} \right)^\gamma \right] = 1.$$

Because we assume competitive labor markets, the object in square brackets is the ratio of social marginal cost to price. As a result, it is socially efficient for this expression to equal unity. This is accomplished in the steady state by setting v as follows:

$$1-v = \frac{1}{\lambda_f(1-\psi+\psi R)}. \quad (26)$$

Our treatment of policy implies that the steady-state allocations of our model economy are efficient in the sense that they coincide with the solution to a particular planning problem. To define this problem, it is convenient to adopt the following scaling of variables:

$$c_t \equiv \frac{C_t}{z_t^{1/\gamma}}, i_t \equiv \frac{I_t}{z_t^{1/\gamma}}. \quad (27)$$

The planning problem is:

$$\max_{\{c_t, H_t, i_t\}} E_0 \sum_{t=0}^{\infty} \beta^t \left[\log c_t - \frac{H_t^{1+\phi}}{1+\phi} \right], \text{ subject to } c_t + i_t = H_t^\gamma i_t^{1-\gamma}. \quad (28)$$

The problem, (28), is that of a planner who allocates resources efficiently across intermediate goods and who does not permit monopoly power distortions. Because there is no

state variable in the problem, it is obvious that the choice variables that solve Eq. (28) are constant over time. This implies that the C_t and I_t that solve the planning problem are a fixed proportion of $z_t^{1/\gamma}$ over time. It turns out that the allocations that solve Eq. (28) also solve the Ramsey optimal policy problem of maximizing Eq. (1) with respect to the 11 variables listed in Eq. (25) subject to the 9 equations listed before Eq. (25).⁴

Because inflation, π_t , fluctuates in equilibrium, Eq. (23) suggests that p_t^* fluctuates too. It turns out, however, that p_t^* is constant to a first-order approximation. To see this, note that the absence of inflation in the steady state also guarantees there is no price dispersion in steady state in the sense that p_t^* is at its maximal value of unity (see Eq. 23). With p_t^* at its maximum in steady state, small perturbations have a zero first-order impact on p_t^* . This can be seen by noting that π_t is absent from the log-linear expansion of Eq. (23) about $p_t^* = 1$:

$$\hat{p}_t^* = \xi_p \hat{p}_{t-1}^*. \quad (29)$$

Here, a hat over a variable indicates:

$$\hat{\varrho}_t = \frac{d\varrho_t}{\varrho},$$

where ϱ denotes the steady state of the variable, ϱ_t , and $d\varrho_t = \varrho_t - \varrho$ denotes a small perturbation in ϱ_t from steady state. We suppose that in the initial period, $\hat{p}_{t-1}^* = 0$, so that, to a first-order approximation, $\hat{p}_t^* = 0$ for all t .

Log-linearizing Eqs. (15), (16), and (18) we obtain the usual representation of the Phillips curve:

$$\hat{\pi}_t = \frac{(1 - \beta\xi_p)(1 - \xi_p)}{\xi_p} \hat{s}_t + \beta E_t \hat{\pi}_{t+1}. \quad (30)$$

Combining Eq. (3) with Eq. (10), taking into account Eq. (27) and the setting of v in Eq. (26), real marginal cost is:

$$s_t = \frac{1}{\lambda_f} \frac{1 - \psi + \psi R_t}{1 - \psi + \psi R} \left(\frac{1}{1 - \gamma} \right)^{1-\gamma} \left(\frac{c_t H_t^\phi}{\gamma} \right)^\gamma.$$

Then,

$$\hat{s}_t = \gamma(\phi \hat{H}_t + \hat{c}_t) + \frac{\psi}{(1 - \psi)\beta + \psi} \hat{R}_t. \quad (31)$$

Substituting out for the real wage in Eq. (19) using Eq. (3) and applying Eq. (27),

⁴ The statement in the text is strictly true only in the case where the initial distortion in prices is zero, that is $p_{t-1}^* = 1$. If this condition does not hold, then the statement still holds asymptotically and may even hold as an approximation after a small number of periods.

$$H_t^{\phi+1} c_t = \frac{\gamma}{1-\gamma} i_t. \quad (32)$$

Similarly, scaling Eq.(24):

$$c_t + i_t = H_t^\gamma i_t^{1-\gamma}.$$

Using Eq. (32) to substitute out for i_t in the above expression, we obtain:

$$c_t + \frac{1-\gamma}{\gamma} H_t^{\phi+1} c_t = H_t^\gamma \left[\frac{1-\gamma}{\gamma} H_t^{\phi+1} c_t \right]^{1-\gamma}.$$

Log-linearizing this expression around the steady state implies, after some algebra,

$$\hat{c}_t = \hat{H}_t. \quad (33)$$

Substituting the latter into Eq. (31), we obtain:

$$\hat{s}_t = \gamma(1+\phi)\hat{c}_t + \frac{\psi}{(1-\psi)\beta + \psi} \hat{R}_t. \quad (34)$$

In Eq. (34), $\hat{c}_t$ is the percent deviation of c_t from its steady-state value. Since this steady-state value coincides with the constant c_t that solves Eq. (28) for each t , $\hat{c}_t$ also corresponds to the output gap. The notation we use to denote the output gap is x_t . Using this notation for the output gap and substituting out $\hat{s}_t$ in the Phillips curve, we obtain:

$$\hat{\pi}_t = \kappa_p \left[\gamma(1+\phi)x_t + \frac{\psi}{(1-\psi)\beta + \psi} \hat{R}_t \right] + \beta E_t \hat{\pi}_{t+1}, \quad (35)$$

where

$$\kappa_p \equiv \frac{(1 - \beta\xi_p)(1 - \xi_p)}{\xi_p}.$$

When $\gamma = 1$ and $\psi = 0$, Eq. (35) reduces to the “Phillips curve” in the classic New Keynesian model. When materials are an important factor of production, so that γ is small, then a given jump in the output gap, x_t , has a smaller impact on inflation. The reason is that in this case the aggregate price index is part of the input cost for intermediate good producers. So, a small price response to a given output gap is an equilibrium because individual intermediate good firms have less of an incentive to raise their prices in this case. With $\psi > 0$, Eq. (35) indicates that a jump in the interest rate drives up prices. This is because with an active working capital channel a rise in the interest rate drives up marginal cost.⁵

⁵ Equation (35) resembles equation (13) in Ravenna and Walsh (2006), except that we also allow for materials inputs, i.e., $\gamma < 1$.

Now consider the intertemporal Euler equation. Expressing (4) in terms of scaled variables,

$$1 = E_t \frac{\beta \hat{c}_t}{c_{t+1} \mu_{z,t+1}^{\frac{1}{\gamma}}} \frac{R_t}{\pi_{t+1}}, \mu_{z,t+1} \equiv \frac{z_{t+1}}{z_t}.$$

Log-linearly expanding about steady state and recalling that $\hat{c}_t$ corresponds to the output gap:

$$0 = E_t \left[x_t - x_{t+1} - \frac{1}{\gamma} \hat{\mu}_{z,t+1} + \hat{R}_t - \hat{\pi}_{t+1} \right],$$

or,

$$x_t = E_t [x_{t+1} - (\hat{R}_t - \hat{\pi}_{t+1} - \hat{R}_t^*)], \quad (36)$$

Where

$$\hat{R}_t^* \equiv \frac{1}{\gamma} E_t \hat{\mu}_{z,t+1}. \quad (37)$$

We suppose that monetary policy, when linearized about steady state, is characterized by the following Taylor rule:

$$\hat{R}_t = r_\pi E_t \hat{\pi}_{t+1} + r_x x_t. \quad (38)$$

The equilibrium of the log-linearly expanded economy is given by Eq. (35) to (38).

2.3 Frisch labor supply elasticity

The magnitude of the parameter, ϕ , in the household utility function plays an important role in the analysis in later sections. This parameter has been the focus of much debate in macroeconomics. Note from Eq. (3) that the elasticity of H_t with respect to the real wage, holding C_t constant, is $1/\phi$. The condition, “holding C_t constant,” could mean that the elasticity refers to the response of H_t to a change in the real wage that is of very short duration, so short that the household’s wealth — and, hence, consumption — is left unaffected. Alternatively, the elasticity could refer to the response of H_t to a change in the real wage that is associated with an offsetting lump-sum transfer payment that keeps wealth unchanged. The debate about ϕ centers on the interpretation of H_t . Under one interpretation, H_t represents the amount of hours worked by a typical person in the labor force. With this interpretation, $1/\phi$ is the Frisch labor supply elasticity.⁶ This is perhaps the most straightforward interpretation of $1/\phi$ given our

⁶ The Frisch labor supply elasticity refers to the substitution effect associated with a change in the wage rate. It is the percent change in a person’s labor supply in response to a change in the real wage, holding the marginal utility of consumption fixed. Throughout this chapter, we assume that utility is additively separable in consumption and leisure, so that constancy of the marginal utility of consumption translates into constancy of consumption.

assumption that the economy is populated by identical households, in which H_t is the labor effort of the typical household. An alternative interpretation of H_t is that it represents the number of people working, and that $1/\phi$ measures the elasticity with which marginal people substitute in and out of employment in response to a change in the wage. Under this interpretation, $1/\phi$ need not correspond to the labor supply elasticity of any particular person. The two different interpretations of H_t give rise to very different views about how data ought to be used to restrict the value of ϕ .

There is an influential labor market literature that estimates the Frisch labor supply elasticity using household level data. The general finding is that, although the Frisch elasticity varies somewhat across different types of people, on the whole the elasticities are very small. Some have interpreted this to mean that only large values of ϕ (say, larger than unity) are consistent with the data. Initially, this interpretation was widely accepted by macroeconomists. However, the interpretation gave rise to a puzzle for equilibrium models of the business cycle. Over the business cycle, employment fluctuates a great deal more than real wages. When viewed through the prism of equilibrium models the aggregate data appeared to suggest that people respond elastically to changes in the wage. But, this seemed inconsistent with the microeconomic evidence that individual labor supply elasticities are in fact small. At the present time, a consensus is emerging that what initially appeared to be a conflict between micro and macro data is really no conflict at all. The idea is that the Frisch elasticity in the micro data and the labor supply elasticity in the macro data represent at best distantly related objects.

It is well known that much of the business cycle variation in employment reflects changes in the quantity of people working, not in the number of hours worked by a typical household. Beginning at least with the work of Rogerson (1988) and Hansen (1985), it has been argued that even if the individual's labor supply elasticity is *zero* over most values of the wage, aggregate employment could nevertheless respond highly elastically to small changes in the real wage. This can occur if there are many people who are near the margin between working in the market and devoting their time to other activities. An example is a spouse who is doing productive work in the home, and yet who might be tempted by a small rise in the market wage to substitute into the market. Another example is a teenager who is close to the margin between pursuing additional education and working, who could be induced to switch to working by a small rise in the wage. Finally, there is the elderly person who might be induced by a small rise in the market wage to delay retirement. These examples suggest that aggregate employment might fluctuate substantially in response to small changes in the real wage, even if the individual household's Frisch elasticity of labor supply is zero over all values of the wage, except the one value that induces a shift in or out of the labor market.⁷

⁷ See Rogerson and Wallenius (2009) for additional discussion and analysis.

The ideas in the previous paragraphs can be illustrated in our model. We adopt the technically convenient assumption that the household has a large number of members, one for each of the points on the line bounded by 0 and 1.⁸ In addition, we assume that a household member only has the option to work full time or not at all. A household member's Frisch labor supply elasticity is zero for almost all values of the wage. Let $l \in [0, 1]$ index a particular member in the family. Suppose this member enjoys the following utility if employed:

$$\log C_t - l^\phi, \phi > 0,$$

and the following utility if not employed:

$$\log C_t.$$

Household members are ordered according to their degree of aversion to work. Those with high values of l have a high aversion (e.g., small children, and elderly or chronically ill people) to work, and those with l near zero have very little aversion. We suppose that household decisions are made on a utilitarian basis, in a way that maximizes the equally weighted integral of utility across all household members. Under these circumstances, efficiency dictates that all members receive the same level of consumption, whether employed or not. In addition, if H_t members are to be employed, then those with $0 \leq l \leq H_t$ should work and those with $l > H_t$ should not. For a household with consumption, C_t , and employment, H_t , utility is, after integrating over all $l \in [0, 1]$:

$$\log C_t - \frac{H_t^{1+\phi}}{1+\phi}, \quad (39)$$

which coincides with the period utility function in Eq. (1). Under this interpretation of the utility function, Eq. (3) remains the relevant first-order condition for labor. In this case, given the wage, W_t/P_t , the household sends out a number of members, H_t , to work until the utility cost of work for the marginal worker, H_t^ϕ , is equated to the corresponding utility benefit to the household, $(W_t/P_t)/C_t$.

Note that under this interpretation of the utility function, H_t denotes a quantity of workers and ϕ dictates the elasticity with which different members of the households enter or leave employment in response to shocks. The case in which ϕ is large corresponds to the case where household members differ relatively sharply in terms of their aversion to work. In this case there are not many members with disutility of work close to that of the marginal worker. As a result, a given change in the wage induces only a small change in employment. If ϕ is very small, then there is a large number of

⁸ Our approach is most similar to the approach of Gali (2010a), although it also resembles the approach taken in the recent work of Mulligan (2001) and Krusell, Mukoyama, Rogerson, and Sahin (2008).

household members close to indifferent between working and not working, and so a small change in the real wage elicits a large labor supply response.

Given that most of the business cycle variation in the labor input is in the form of numbers of people employed, we think the most sensible interpretation of H_t is that it measures numbers of people working. Accordingly, $1/\phi$ is not to be interpreted as a Frisch elasticity, which we instead assume to be zero.

3. SIMPLE MODEL: SOME IMPLICATIONS FOR MONETARY POLICY

Monetary DSGE models have been used to gain insight into a variety of issues that are important for monetary policy. We discuss some of these issues using variants of the simple model developed in the previous section. A key feature of that model is that it is flexible, and can be adjusted to suit different questions and points of view. The classic New Keynesian model, the one with no working capital channel and no materials inputs (i.e., $\gamma = 1$, $\psi = 0$) can be used to articulate the rationale for the Taylor principle. But variants of the New Keynesian framework can also be used to articulate challenges to that principle. Sections 3.1 and 3.2 below describe two such challenges. The fact that the New Keynesian framework can accommodate a variety of perspectives on important policy questions is an important strength. This is because the framework helps to clarify debates and to motivate econometric analyses so that data can be used to resolve those debates.⁹

Sections 3.3 and 3.4 below address the problem of estimating the output gap. The output gap is an important variable for policy analysis because it is a measure of the efficiency with which economic resources are allocated. In addition, New Keynesian models imply that the output gap is an important determinant of inflation, a variable of particular concern to monetary policymakers. We define the output gap as the percent deviation between actual output and potential output, where potential output is output in the Ramsey-efficient equilibrium.¹⁰

We use the classic New Keynesian model to study three ways of estimating the output gap. The first uses the structure of the simple New Keynesian model to estimate the output gap as a latent variable. The second approach modifies the New Keynesian model to include unemployment along the lines indicated by CTW. This modification of the model allows us to investigate the information content of the unemployment rate for the output gap. In addition, by showing one way that unemployment can be integrated into the model, the discussion represents another illustration of the versatility

⁹ For example, the Chowdhury, Hoffmann, and Schabert (2006) and Ravenna and Walsh (2006) papers cited in the previous section, show how the assumptions of the New Keynesian model can be used to develop an empirical characterization of the importance of the working capital channel.

¹⁰ In our model, the Ramsey-equilibrium turns out to be what is often called the “first-best equilibrium,” the one that is not distorted by monopoly power or flexible prices.

of the New Keynesian framework.¹¹ The third approach which is studied in section 3.4 explores the HP filter as a device for estimating the output gap. In the course of the analysis, we illustrate the Bayesian limited information moment matching procedure discussed in the introduction.

3.1 Taylor principle

A key objective of monetary policy is the maintenance of low and stable inflation. The classic New Keynesian model defined by $\gamma = 1$ and $\psi = 0$ can be used to articulate the risk that inflation expectations might become self-fulfilling unless the monetary authorities adopt the appropriate monetary policy. The classic model can also be used to explain the widespread consensus that “appropriate monetary” policy means a monetary policy that embeds the Taylor Principle: a 1% rise in inflation should be met by a greater than 1% rise in the nominal interest rate. This subsection explains how the classic New Keynesian model rationalizes the wisdom of implementing the Taylor principle. However, when we incorporate the assumption of a working capital channel — particularly when the share of materials in gross output is as high as it is in the data — the Taylor principle becomes a *source* of instability. This is perhaps not surprising. When the working capital channel is strong, if the monetary authority raises the interest rate in response to rising inflation expectations, the resulting rise in costs produces the higher inflation that people expect.¹²

It is convenient to summarize the linearized equations of our model here:

$$\hat{R}_t^* = E_t \frac{1}{\gamma} \hat{\mu}_{z,t+1} \quad (40)$$

$$\hat{\pi}_t = \kappa_p [\gamma(1 + \phi)x_t + \alpha_\psi \hat{R}_t] + \beta E_t \hat{\pi}_{t+1} \quad (41)$$

¹¹ For an alternative recent approach to the introduction of unemployment into a DSGE model, see Gali (2010a). Gali demonstrated that with a modest reinterpretation of variables, the standard DSGE model with sticky wages summarized in the next section contains a theory of unemployment. In the model of the labor market used there (it was proposed by Erceg et al. 2000) wages are set by a monopoly union. As a result, the wage rate is higher than the marginal cost of working. Under these circumstances, one can define the unemployed as the difference between the number of people actually working and the number of people that would be working if the cost of work for the marginal person were equated to the wage rate. Gali (2010b) showed how unemployment data can be used to help estimate the output gap, as we do here. The CTW and Gali models of unemployment are quite different. For example, in the text we analyze a version of the CTW model in which labor markets are perfectly competitive, so Gali’s “monopoly power” concept of unemployment is zero in this model. In addition, the efficient level of unemployment in the sense that we use the term here, is zero in Gali’s definition, but positive in our definition. This is because in our model, unemployment is an inevitable by-product of an activity that must be undertaken to find a job. For an extensive discussion of the differences between our model and Gali’s, see Section F in the technical appendix to CTW, which can be found at <http://faculty.wcas.northwestern.edu/~lchrist/research/Riksbank/technicalappendix.pdf>.

¹² Bruckner and Schabert (2003) made an argument similar to ours, although they do not consider the impact of materials inputs, which we find to be important.

$$x_t = E_t[x_{t+1} - (\hat{R}_t - \hat{\pi}_{t+1} - \hat{R}_t^*)] \quad (42)$$

$$\hat{R}_t = r_\pi E_t \hat{\pi}_{t+1} + r_x x_t, \quad (43)$$

where

$$\alpha_\psi = \frac{\psi}{(1 - \psi)\beta + \psi}.$$

The specification of the model is complete when we take a stand on the law of motion for the exogenous shock. We do this in the following subsections as needed.

We begin by reviewing the case for the Taylor principle using the classic New Keynesian model, with $\gamma = 1$, $\psi = 0$. We get to the heart of the argument using the deterministic version of the model, in which $\hat{R}_t^* \equiv 0$. In addition, it is convenient to suppose that monetary policy is characterized by $r_x = 0$. Throughout, we adopt the presumption that the only valid equilibria are paths for $\hat{\pi}_t$, $\hat{R}_t$ and x_t that converge to steady state; that is, 0.¹³ Under these circumstances, Eqs. (41) and (42) can be solved forward as follows:

$$\hat{\pi}_t = \kappa_p \gamma (1 + \phi) x_t + \beta \kappa_p \gamma (1 + \phi) x_{t+1} + \beta^2 \kappa_p \gamma (1 + \phi) x_{t+2} + \dots \quad (44)$$

and

$$x_t = -(\hat{R}_t - \hat{\pi}_{t+1}) - (\hat{R}_{t+1} - \hat{\pi}_{t+2}) - (\hat{R}_{t+2} - \hat{\pi}_{t+3}) - \dots \quad (45)$$

In Eq. (45) we have used the fact that in our setting a path converges to zero if, and only if, it converges fast enough so that a sum like the one in Eq. (45) is well defined.¹⁴ Equation (44) shows that inflation is a function of the present and future output gap. Equation (45) shows that the current output gap is a function of the long term real interest rate (i.e., the sum on the right of Eq. 45) in the model.

Under the Taylor principle, the classic New Keynesian model implies that a rise in inflation expectations launches a sequence of events that ultimately leads to a

¹³ Although our presumption is standard, justifying it is harder than one might have thought. For example, Benhabib, Schmitt-Grohe, and Uribe (2002) presented examples in which some explosive paths for the linearized equilibrium conditions are symptomatic of perfectly sensible equilibria for the actual economy underlying the linear approximations. In these cases, focusing on the nonexplosive paths of the linearized economy may be valid after all if we imagine that monetary policy is a Taylor rule with a particular escape clause. The escape clause specifies that in the event the economy threatens to follow an explosive path, the monetary authority commits to switch to a monetary policy of targeting the money growth rate. There are examples of monetary models in which the escape clause monetary policy justifies the type of equilibrium selection we adopt in the text (see Benhabib et al. 2002 and Christiano & Rostagno, 2001 for further discussion). For a more recent debate about the validity of the equilibrium selection adopted in the text, see McCallum (2009) and Cochrane (2009) and the references they cite.

¹⁴ The reason for this can be seen below, where we show that the solution to this equation is a linear combination of terms like $a\lambda^t$. Such an expression converges to zero if, and only if, it is also summable.

moderation in actual inflation. Seeing this moderation in actual inflation, people's higher inflation expectations would quickly dissipate before they could be a source of economic instability. The way this works is that the rise in the real rate of interest slows spending, causing the output gap to shrink (see Eq. 45). The fall in actual inflation occurs as the reduction in output reduces pressure on resources and drives down the marginal cost of production (see Eq. 41). Strictly speaking, we have just described a rationale for the Taylor principle that is based on learning (for a formal discussion, see McCallum, 2009). Under rational expectations, the posited rise in inflation expectations would not occur in the first place if policy obeys the Taylor principle.

A similar argument shows that if the monetary authority does not obey the Taylor principle, that is, $r_\pi < 1$, then a rise in inflation expectations can be self-fulfilling. This is not surprising, since in this case the rise in expected inflation is associated with a fall in the real interest rate. According to Eq. (45) this produces a rise in the output gap. By raising marginal cost, the Phillips curve, (Eq. 44), implies that actual inflation rises. Seeing higher actual inflation, people's higher inflation expectations are confirmed. In this way, with $r_\pi < 1$ a rise in inflation expectations becomes self-fulfilling by triggering a boom in output and actual inflation. It is easy to see that with $r_\pi < 1$ many equilibria are possible. A drop in inflation expectations can cause a fall in output and inflation. Inflation expectations could be random, causing random fluctuations between booms and recessions.¹⁵

In this way, the classic New Keynesian model has been used to articulate the idea that the Taylor principle promotes stability, while absence of the Taylor principle makes the economy vulnerable to fluctuations in self-fulfilling expectations.

The preceding results are particularly easy to establish formally under the assumption of rational expectations. We continue to maintain the simplifying assumption, $r_x = 0$. We reduce the model to a single second order difference equation in inflation. Substitute out for $\hat{R}_t$ in Eqs. (41) and (42) using Eq. (43). Then, solve Eq. (41) for x_t and use this to substitute out for x_t in Eq. (42). These operations result in the following second-order difference equation in $\hat{\pi}_t$:

$$\hat{\pi}_t + [\kappa_p \gamma (1 + \phi)(r_\pi - 1) - (\kappa_p \alpha_\psi r_\pi + \beta) - 1] \hat{\pi}_{t+1} + (\kappa_p \alpha_\psi r_\pi + \beta) \hat{\pi}_{t+2} = 0.$$

¹⁵ Clarida, Gali, and Gertler (1999; CGG) argued that the high inflation of the 1970s in many countries can be explained as reflecting the failure to respect the Taylor principle in the early 1970s. Christiano and Gust (2000) criticized this argument on the grounds that one did not observe a boom in employment in the 1970s. Christiano and Gust argued that even if one thought of the 1970s as also a time of bad technology shocks (fuel costs and commodity prices soared then), the CGG analysis predicts that employment should have boomed. Christiano and Gust presented an alternative model, a "limited participation" model, which has the same implications for the Taylor principle that the CGG model has. However, the Christiano and Gust model has a very different implication for what happens to real allocations in a self-fulfilling inflation episode. Because of the presence of an important working capital channel, the self-fulfilling inflation episode is associated with a recession in output and employment. Thus, Christiano and Gust concluded that the 1970s might well reflect the failure to implement the Taylor principle, but only if the analysis is done in a model different from the CGG model.

The general set of solutions to this difference equation can be written as follows:

$$\hat{\pi}_t = a_0 \lambda_1^t + a_1 \lambda_2^t,$$

for arbitrary a_0, a_1 . Here, $\lambda_i, i = 1, 2$, are the roots of the following second-order polynomial:

$$1 + [\kappa_p \gamma (1 + \phi)(r_\pi - 1) - (\kappa_p \alpha_\psi r_\pi + \beta + 1)]\lambda + (\kappa_p \alpha_\psi r_\pi + \beta)\lambda^2 = 0.$$

Thus, there is a two-dimensional space of solutions to the equilibrium conditions (i.e., one for each possible value of a_0 and a_1). We continue to apply our presumption that among these solutions, only the ones in which the variables converge to zero (i.e., to steady state) correspond to equilibria. Thus, uniqueness of equilibrium requires that both λ_1 and λ_2 be larger than unity in absolute value. In this case, the unique equilibrium is the solution associated with $a_0 = a_1 = 0$. If one or both of $\lambda_i, i = 1, 2$ are less than unity in absolute value, then there are many solutions to the equilibrium conditions that are equilibria. We can think of these equilibria as corresponding to different, self-fulfilling, expectations.

The following result can be established for the classic New Keynesian model, with $\gamma = 1$ and $\psi = 0$. The model economy has a unique equilibrium if, and only if $r_\pi > 1$ (see, e.g., Bullard & Mitra, 2002). This is consistent with the intuition about the Taylor principle discussed above.

We now reexamine the case for the Taylor principle when there is a working capital channel. The reason the Taylor principle works in the classic New Keynesian model is that a rise in the interest rate leads to a fall in inflation by curtailing aggregate spending. But, with a working capital channel, $\psi > 0$, an increase in the interest rate has a second effect. By raising marginal cost (see Eq. 41), a rise in the interest rate places upward pressure on inflation. If the working capital channel is strong enough, then monetary policy with $r_\pi > 1$ may “add fuel to the fire” when inflation expectations rise. The sharp rise in the nominal rate of interest in response to a rise in inflation expectations may actually cause the inflation that people expected. In this way the Taylor principle could actually be destabilizing. Of course, for this to be true requires that the working capital channel be strong enough. For a small enough working capital channel (i.e., small ψ) implementing the Taylor principle would still have the effect of inoculating the economy from destabilizing fluctuations in inflation expectations.

Whether the presence of the working capital channel overturns the wisdom of implementing the Taylor principle is a numerical question. We must assign values to the model parameters and investigate whether one or both of λ_1 and λ_2 are less than unity in absolute value. If this is the case, then implementing the Taylor principle does not stabilize inflation expectations. Throughout, we set:

$$\beta = 0.99, \xi_p = 0.75, r_\pi = 1.5.$$

The discount rate is 4%, at an annual rate and the value of ξ_p implies an average time between price reoptimization of one year. In addition, monetary policy is characterized by a strong commitment to the Taylor principle. We consider two values for the interest rate response to the output gap, $r_x = 0$ and $r_x = 0.1$. For robustness, we also consider a version of Eq. (43) in which the monetary authority reacts to current inflation.

We do not have a strong prior about the parameter, ϕ , that controls the disutility of labor (see Section 2.3), so we consider two values, $\phi = 1$ and $\phi = 0.1$. To have a sense of the appropriate value of γ , we follow Basu (1995). He argued, using manufacturing data, that the share of materials in gross output is roughly 1/2. Recall that the steady state of our model coincides with the solution to Eq. (28), so that

$$\frac{i}{c+i} = 1 - \gamma.$$

Thus, Basu's empirical finding suggests a value for γ close to 1/2.¹⁶ The instrumental variables results in Ravenna and Walsh (2006) suggest that a value of the working capital share, ψ , in a neighborhood of unity is consistent with the data.

Figure 1 displays our results. The upper row of figures provides results for the case in Eq. (43), in which the policy authority reacts to the one-quarter-ahead expectation of inflation, $E_t \hat{\pi}_{t+1}$. The lower row of figures corresponds to the case where the policy-maker responds instead to current inflation, $\hat{\pi}_t$. The horizontal and vertical axes indicate a range of values for γ and ψ , respectively. The gray areas correspond to the parameter values where one or both of λ_i , $i = 1, 2$ are less than unity in absolute value. Technically, the steady-state equilibrium of the economy is said to be "indeterminate" for parameterizations in the gray area. Intuitively, the gray area corresponds to parameterizations of our economy in which the Taylor principle does not stabilize inflation expectations. The white areas in the figures correspond to parameterizations where implementing the Taylor principle successfully stabilizes the economy.

Consider the 1,1 and 1,2 graphs in Figure 1 first. Note that in each case, $\psi = 0$ and $\gamma = 1$ are points in the white area, consistent with the earlier discussion. However, a very small increase in the value of ψ puts the model into the gray area. Moreover, this is true regardless of the value of γ . For these parameterizations the aggressive response of the interest rate to higher inflation expectations only produces the higher inflation that people anticipate. We can see in the 1,3 and 1,4 graphs of the first row, that $r_x > 0$ greatly reduces the extent of the gray area. Still, for $\gamma = 0.5$ and ψ near unity the model is in the gray area and implementing the Taylor principle would be counterproductive.

¹⁶ Actually, this is a conservative estimate of γ . Had we not selected v to extinguish monopoly power in the steady state, our estimate of γ would have been lower. See Basu (1995) for more discussion of this point.

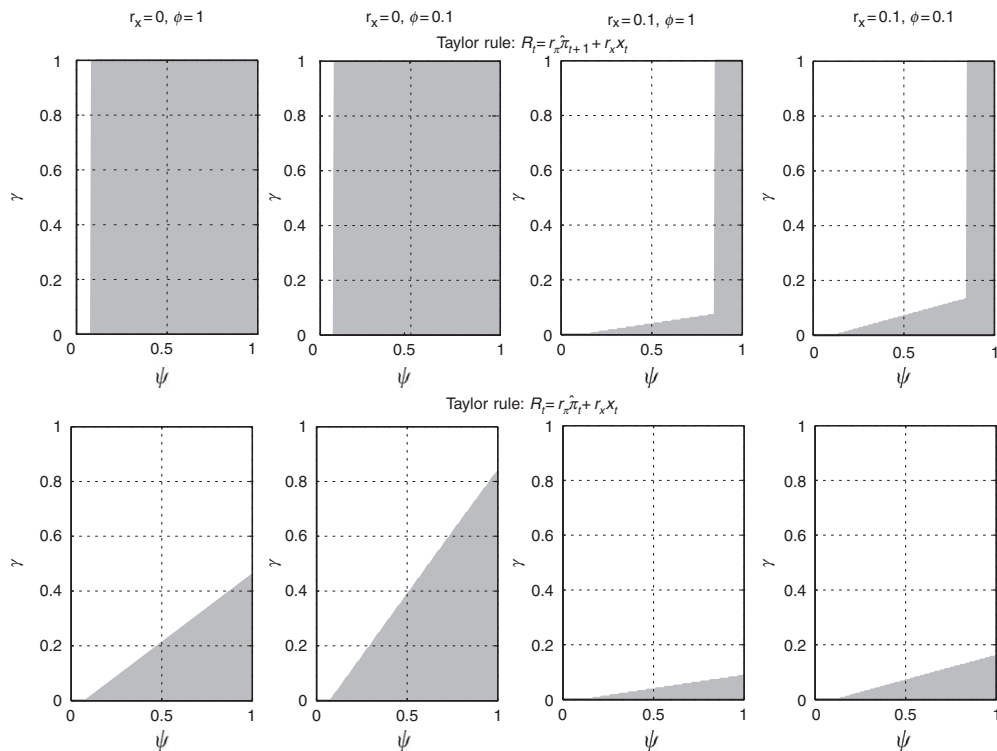


Figure 1 Indeterminacy regions for model with working capital channel and materials inputs. Note: Gray area is region of indeterminacy and white area is region of determinacy.

Now consider the bottom row of graphs. Note that if $\gamma = 1$ then the model is always in the determinacy region. That is, for the economy to be vulnerable to self-fulfilling expectations, it must not only be that there is a substantial working capital channel, but it must also be that materials are a substantial fraction of gross output. The 2,2 graph shows that with $\gamma = 0.5$, $\phi = 0.1$ and ψ above roughly 0.6, the model is in the gray area. When ϕ is substantially higher, the first graph from the left indicates that the gray area is smaller. Note that with $r_x > 0$, the gray area has almost shrunk to zero, according to the two last graphs.

We conclude from this analysis that in the presence of a working capital channel, sharply raising the interest rate in response to higher inflation could actually be counter-productive. This is more likely to be the case when the share of materials inputs in gross output is high. When this is so, one cannot rely exclusively on the Taylor principle to ensure stable inflation and output performance. In the example, responding strongly to the output gap (or, actual rather than expected inflation) could restore stability. However,

in practice the output gap is hard to measure.¹⁷ At best, the policy authority can respond to variables that are correlated with the output gap. Studying the implications for determinacy of responding to such variables would be an interesting project, but would take us beyond the scope of this chapter. Still, the discussion illustrates how DSGE models can be useful for thinking about important monetary policy questions.

3.2 Monetary policy and inefficient booms

In recent years, there has been extensive discussion about the interaction of monetary policy and economic volatility, in particular, asset price volatility. Prior to the recent financial turmoil, a consensus had developed that monetary policy should not actively seek to stabilize asset prices. The view was that in any case, a serious commitment to inflation targeting — one that implements the Taylor principle — would stabilize asset markets automatically.¹⁸ The idea is that an asset price boom is basically a demand boom, the presumption being that the boom is driven by optimism about the future, and not primarily by current actual developments. A boom that is driven by demand should, according to the conventional wisdom, raise production costs and, hence, inflation. The monetary authority that reacts vigorously to inflation then automatically raises interest rates and helps to stabilize asset prices.

When this scenario is evaluated in the classic New Keynesian model, we find that the boom is not necessarily associated with a rise in prices. In fact, if the optimism about the future concerns the expectations about cost saving new technologies, forward-looking price-setters may actually reduce their prices. This is the finding of Christiano, Ilut, Motto, and Rostagno (2008, 2010), which we briefly summarize here.

To capture the notion of optimism about the future, suppose that the time series representation of the log-level of technology is as follows:

$$\log z_t = \rho_z \log z_{t-1} + u_t, u_t = \varepsilon_t + \xi_{t-1}, \quad (46)$$

so that the steady state of z_t is unity. In Eq. (46), u_t is an iid shock, uncorrelated with past $\log z_t$. The innovation in technology growth, u_t , is the sum of two orthogonal processes, ε_t and ξ_{t-1} . The time subscript on these two variables represents the date when they are known to private agents. Thus, at time $t - 1$ agents become aware of a component of u_t , namely ξ_{t-1} . At time t they learn the rest, ε_t . For example, the initial “news” about u_t , ξ_{t-1} , could in principle be entirely false, as would be the case when $\varepsilon_t = -\xi_{t-1}$.

Substituting Eq. (46) into Eq. (40):

$$\hat{R}_t^* = E_t[\log z_{t+1} - \log z_t] = (\rho_z - 1) \log z_t + \xi_t, \quad (47)$$

¹⁷ For further discussion of this point, see Sections 3.3 and 3.4.

¹⁸ See Bernanke and Gertler (2000).

where $\gamma = 1$ since we now consider the classic New Keynesian model.¹⁹ Our system of equilibrium conditions is Eq. (47) with Eqs. (41), (42) and (43). We set $\psi = 0$ (i.e., no working capital channel) and $r_x = 0$. We adopt the following parameter values:

$$\beta = 0.99, \phi = 1, r_x = 0, r_\pi = 1.5, \rho_z = 0.9, \xi_p = 0.75.$$

We perform a simulation in which news arrives in period t that technology will jump one percent in period $t + 1$, i.e., $\xi_t = 0.01$. The value of ε_t is set to zero. We find that hours worked in period t increases by 1%. This rise is entirely inefficient because in the first-best equilibrium hours it does not respond at all to a technology shock, whether it occurs in the present or it is expected to occur in the future (see Eq. 28). Interestingly, inflation falls in period t by 10 basis points, at an annual rate.²⁰ Current marginal cost does rise (see Eq. 34), but current inflation nevertheless falls because of the fall in expected future marginal costs.

The efficient monetary policy sets $\hat{R}_t = \hat{R}_t^*$ which, according to Eq. (47), means the interest rate should rise when a positive signal about the economy occurs. A policy that applies the Taylor principle in this example moves policy in exactly the wrong direction in response to ξ_t . By responding to the fall in inflation, policy not only does not raise the interest rate — as it should — but it actually reduces the interest rate in response to the fall in inflation. By reducing the interest rate in the period of a positive signal about the future, policy overstimulates the economy creating excessive volatility.

So, the classic New Keynesian model can be used to challenge the conventional wisdom that an inflation-fighting central bank automatically moderates economic volatility. But, is this just an abstract example without any relevance? In fact, the typical boom-bust episode is characterized by low or falling inflation (see Adalid & Detken, 2007). For example, during the U.S. booms of the 1920s and the 1980s and 1990s, inflation was low. This fact turns the conventional wisdom on its head and leads to a conclusion that matches our numerical example: an inflation-fighting central bank amplifies boom-bust episodes.

A full evaluation of the ideas in this subsection requires a more elaborate model, preferably one with financial variables such as the stock market. In this way, one could assess the impact on a broader set of variables in boom-bust episodes. In addition, one could evaluate what other variables the monetary authority might look at to avoid contributing to the type of volatility described in this example. We presume that it is not helpful to simply say that the monetary authority should set $\hat{R}_t = \hat{R}_t^*$, because in practice this may require more information than is actually available. A more fruitful

¹⁹ To see why we replaced $\hat{\mu}_{z,t+1}$ in Eq. (40) by $\log z_{t+1} - \log z_t$, note first $\hat{\mu}_{z,t} = \frac{\mu_{z,t} - \mu_z}{\mu_z} = \mu_{z,t} - 1$, because in steady state $\mu_z \equiv z_t/z_{t-1} = 1/1 = 1$. Then, $1 + \hat{\mu}_{z,t} = \mu_{z,t}$. Take the log of both sides and note, $\log \mu_{z,t} = \log(1 + \hat{\mu}_{z,t}) \simeq \hat{\mu}_{z,t}$. But, $\log \mu_{z,t} = \log z_t - \log z_{t-1}$.

²⁰ Because inflation is zero in steady state, $\hat{\pi}_t = \pi_t - 1$. This was converted to annualized basis points by multiplying by 40,000.

approach may be to find variables that are correlated with $\hat{R}_t^*$, so that these may be included in the monetary policy rule. For further discussion of these issues, see Christiano et al. (2008).

3.3 Using unemployment to estimate the output gap

Here, we investigate the use of DSGE models to estimate the output gap as a latent variable. We explore the usefulness of including data on the rate of unemployment in this exercise. Section 3.3.1 describes a scalar statistic for characterizing the information content of the unemployment rate for the output gap, and Section 3.3.2 describes the model used in the analysis. As in the previous subsection, we work with a version of the classic New Keynesian model. In particular, we assume intermediate good producers do not use materials inputs or working capital.²¹ We introduce unemployment into the model following the approach in CTW.

Section 3.3.3 describes how we use data to assign values to the model parameters. This section may be of independent interest because it shows how a moment-matching procedure like the one proposed in Christiano and Eichenbaum (1992a) can be recast in Bayesian terms. Section 3.3.4 presents our substantive results. Based on our simple estimated model with unemployment, we find that including unemployment has a substantial impact on our estimate of the output gap for the U.S. economy. We summarize our findings at the end of Section 3.3.4 where we also indicate several caveats to the analysis.

3.3.1 A measure of the information content of unemployment

As a benchmark, we compute the projection of the output gap on present, future, and past observations on output growth:

$$x_t = \sum_{j=-\infty}^{\infty} h_j \Delta y_{t-j} + \varepsilon_t^y, \quad (48)$$

where h_j is a scalar for each j and ε_t^y is uncorrelated with Δy_{t-s} for all s .²² The projection that also involves unemployment can be expressed as follows:

$$x_t = \sum_{j=-\infty}^{\infty} h_j \Delta y_{t-j} + \sum_{j=-\infty}^{\infty} h_j^u u_{t-j} + \varepsilon_t^{y,u}.$$

Here, h_j^u is a scalar for each j and $\varepsilon_t^{y,u}$ is uncorrelated with Δy_{t-s} , u_{t-s} for all s . We define the information content of unemployment for the output gap by the ratio,

²¹ That is, we set $\gamma = 1$ and $\psi = 0$.

²² In practice only a finite amount of data is available. As a result, the projection involves a finite number of lags where the number of lags varies with t . The Kalman smoother solves the projection problem in this case.

$$r^{\text{two-sided}} \equiv \frac{E(\varepsilon_t^{y,u})^2}{E(\varepsilon_t^y)^2}.$$

The lower the ratio, the greater the information in unemployment for the gap. We also compute the analogous variance ratio, $r^{\text{one-sided}}$, corresponding to the one-sided projection involving only current and past observations on the explanatory variables.²³ The one-sided projection is the one that is relevant to assess the information content of unemployment for policymakers working in real time. Our measure of information does not incorporate sampling uncertainty in parameters. The variances used to construct $r^{\text{two-sided}}$ and $r^{\text{one-sided}}$ assume the parameters are known with certainty and that the only uncertainty stems from the fact that the gap cannot be constructed using the data available to the econometrician.

3.3.2 The CTW model of unemployment

We convert the usual three equation log-linear representation of the New Keynesian model into a model of unemployment by adding one equation. This reduced-form log-linear system is derived from explicit microeconomic foundations in CTW. That paper also shows how our model of unemployment can be integrated into a medium-sized DSGE model such as the one in Section 4.

In the CTW model, finding a job requires exerting effort. Because effort only increases the probability of finding a job, not everyone who looks for a job actually finds one. The unemployed are people who look for a job without success. The unemployment rate is the number unemployed, expressed as fraction of the labor force. As in the official definition, the labor force is the number of people employed plus the number unemployed.

Since effort is unobserved and privately costly, perfect insurance against idiosyncratic labor market outcomes is not possible. As a result, the unemployed are worse off than the employed. In this way, the model captures a key reason that policymakers care about unemployment: a rise in unemployment imposes a direct welfare cost on the families involved. In this respect, our model differs from other work that integrates unemployment into monetary DSGE models.²⁴ In those models, individuals have perfect insurance against labor market outcomes.

We now describe the shocks and the linearized equilibrium conditions of the model. In previous sections of this chapter, the efficient level of hours worked is

²³ In the analysis below, we compute the projections in two ways. When we apply the filter to the data to extract a time series of x_t , we use the Kalman smoother. To compute the weights in the infinite projection problem, we use standard spectral methods described in, for example, Sargent (1979, Chapter 11). The spectral weights can also be computed by numerical differentiation of the output of the Kalman smoother with respect to the input data. We verified that the two methods produce the same results as long as the number of observations is large and t lies in the middle of the data set.

²⁴ For a long list of references, see CTW.

constant, and so the output gap can be expressed simply as the deviation of the number of people working from that constant (see Eq. 33). In this section, the efficient number of people working is stochastic. We denote the deviation of this number from steady state by h_t^* . We continue to assume that the steady state of our economy is efficient, so that $\hat{H}_t$ and h_t^* represent percent deviations from the same steady state values. The output gap is now:

$$x_t = \hat{H}_t - h_t^*.$$

The object, h_t^* , is driven by disturbances to the disutility of work, as well as by disturbances to the technology that converts household effort into a probability of finding a job. These various disturbances to the efficient level of employment cannot be disentangled using the data we assume are available to the econometrician. We refer to h_t^* as a labor supply shock. We hope that this label does not generate confusion. In our context this shock summarizes a broader set of disturbances than simply the one that shifts the disutility of labor. We adopt the following time series representation for the labor supply shock:

$$h_t^* = \lambda h_{t-1}^* + \varepsilon_t^{h^*}, \quad (49)$$

where $\varepsilon_t^{h^*}$ is a zero mean, iid process uncorrelated with h_{t-s}^* , $s > 0$ and $E(\varepsilon_t^{h^*})^2 = \sigma_{h^*}^2$. In the version of the CTW model studied here, h_t^* is orthogonal to all the other shocks.

We assume the technology shock is a logarithmic random walk:

$$\Delta \log z_t = \varepsilon_t^z, \quad (50)$$

where Δ denotes the first difference operator. The object, ε_t^z , is a mean-zero, iid disturbance that is not correlated with $\log z_{t-s}$, $s > 0$. We denote its variance by $E(\varepsilon_t^z)^2 = \sigma_z^2$. The empirical rationale for the random walk assumption is discussed in Section 4.1.²⁵

According to CTW, the interest rate in the first-best equilibrium is given by:

$$\hat{R}_t^* = E_t[\Delta \log z_{t+1} + h_{t+1}^* - h_t^*]. \quad (51)$$

Log consumption in the first best equilibrium is (apart from a constant term) the sum of $\log z_t$ and h_t^* . So, according to Eq. (51), $\hat{R}_t^*$ corresponds to the anticipated growth rate of (log) consumption. This reflects the CTW assumption that utility is additively

²⁵ Another way to assess the empirical basis for the random walk assumption exploits the simple model's implication that the technology shock can be measured using labor productivity. One measure of labor productivity is given by the ratio of real US GDP to a measure of total hours. The first-order autocorrelation of the quarterly logarithmic growth rate of this variable for the period, 1951Q1 to 2008Q4 is -0.02 . The same first-order autocorrelation is 0.02 when calculated using output per hour for the nonfarm business sector. These results are consistent with our random walk assumption.

separable and logarithmic in consumption. We also suppose there is a disturbance, μ_t , that enters the Phillips curve as follows:

$$\hat{\pi}_t = \kappa^p x_t + \beta E_t \hat{\pi}_{t+1} + \mu_t, \quad (52)$$

where $\kappa^p > 0$. Here, κ^p denotes the slope of the Phillips curve in terms of the output gap. This is not to be confused with κ_p in Eqs. (41) and (35), which is the slope of the Phillips curve in terms of marginal cost. Our representation of the Phillips curve shock is given by

$$\mu_t = \chi \mu_{t-1} + \varepsilon_t^\mu, \quad (53)$$

where $E(\varepsilon_t^\mu)^2 = \sigma_\mu^2$. The intertemporal equation, Eq. (42), is unchanged from before. Finally, we suppose that there is an iid disturbance, M_t , that enters the monetary policy rule in the following way:

$$\hat{R}_t = \rho_R \hat{R}_{t-1} + (1 - \rho_R)[r_\pi E_t \hat{\pi}_{t+1} + r_x x_t] + M_t, \quad (54)$$

where $E(\varepsilon_t^M)^2 = \sigma_M^2$. The four exogenous shocks in the model are orthogonal to each other at all leads and lags.

Let the unemployment gap, u_t^g , denote the deviation between actual unemployment and efficient unemployment, when both are expressed in percent deviation from their (common) nonstochastic steady state. The CTW model implies:

$$u_t^g = -\kappa^g x_t, \kappa^g > 0, \quad (55)$$

where κ^g is a function of underlying structural parameters. The previous expression resembles “Okun’s law.” If actual unemployment is one percentage point higher than its efficient level, then output is $1/\kappa^g$ percent below its efficient level. Discussions of Okun’s law often suppose that $1/\kappa^g$ lies in a range of 2 to 3 (see, e.g., Abel & Bernanke, 2005). The unemployment rate in the efficient equilibrium, u_t^* , has the following representation:

$$u_t^* = -\omega h_t^*, \omega > 0.$$

In the CTW model, the factors that increase labor supply also increase the intensity of job finding effort, and this is the reason unemployment in the efficient equilibrium falls. The harder people look for a job, the sooner they find what they are looking for. According to the previous two equations, the actual unemployment rate, u_t , satisfies the following equation:

$$u_t = -[\kappa^g x_t + \omega h_t^*]. \quad (56)$$

Absent the presence of the labor supply shock, the efficient level of unemployment would be constant and the actual unemployment rate would represent a direct observation on the output gap.

In sum, the log-linearized equations of the CTW model consist of the usual three equations of the standard New Keynesian model, Eqs. (42), (52), and (54), plus an equation that characterizes unemployment, Eq. (56). In addition, there are the equations that characterize the laws of motion of the exogenous shocks and of the efficient rate of interest.

3.3.3 Limited information Bayesian inference

To investigate the quantitative implications of the model, we must assign values to its parameters. We set values of the economic parameters of the model,

$$\kappa^p, \omega, \kappa^g, r_\pi, r_x, \rho_R, \beta,$$

as indicated in Table 1a. Let θ denote the 6×1 column vector consisting of the parameters governing the stochastic processes:

$$\theta = (\lambda, \chi, \sigma_z, \sigma_{h^*}, \sigma_M, \sigma_\mu)'. \quad (57)$$

We use data on output growth and unemployment to select values for the elements of θ .²⁶ We do this using a version of the limited information Bayesian procedure described in Kim (2002) and in Section 5.2.²⁷ Let γ denote the 11×1 column vector composed of the j^{th} order autocovariance matrix of output growth and unemployment, for $j = 0, 1, 2$. Let $\hat{\gamma}$ denote the corresponding sample estimate based on $T = 232$ quarterly observations, 1951Q1–2008Q4. Hansen's (1982) generalized method of moments analysis (GMM) establishes that for sufficiently large T , $\hat{\gamma}$ is a realization from a Normal distribution with mean equal to the true value of the second moments, γ^0 , and variance, V/T . The results also hold when V is replaced by a consistent sample estimate, $\hat{V}$.²⁸ Our model provides a mapping from θ to γ , which we denote by $\gamma(\theta)$. Hansen's

²⁶ The seasonally adjusted unemployment rate for people 16 years and older was obtained from the Bureau of Labor Statistics and has mnemonic LNS14000000. We use standard real per capita GDP data, as described in Section A of the technical appendix found at http://www.faculty.econ.northwestern.edu/faculty/christiano/research/Handbook/technical_appendix.pdf.

²⁷ The procedure is the Bayesian analog of the moment matching estimation procedure described in Christiano and Eichenbaum (1992a).

²⁸ We compute $\hat{V}$ as follows. Let θ^0 denote the true, but unknown, values of the model parameters. Let $h(\gamma, w_t)$ denote the 11×1 GMM error vector having the property, $Eh(\gamma^0, w_t) = 0$, where $w_t = (\Delta y_t, u_t)'$. Let $g_T(\gamma) \equiv (1/T) \sum h(\gamma, w_t)$ and define $\hat{\gamma}$ by $g_T(\hat{\gamma}) = 0$. Then, $\sqrt{T}(\hat{\gamma} - \gamma^0)$ converges in distribution as $T \rightarrow \infty$ to $N(0, V)$. Here, $V = (D')^{-1} S D^{-1}$, where S denotes the spectral density at frequency zero of $h(\gamma^0, w_t)$, and $D' = \lim_{T \rightarrow \infty} \partial g_T(\gamma) / \partial \gamma'$, where the derivative is evaluated at $\gamma = \hat{\gamma}$ (for a discussion of these convergence results of GMM see, e.g., Hamilton 1994). Our estimator of V is given by $\hat{V} = (\hat{D}')^{-1} \hat{S} \hat{D}^{-1}$. We estimate $\hat{S}$ by $\hat{\Gamma}_0 + (1 - 1/3)(\hat{\Gamma}_1 + \hat{\Gamma}_1') + (1 - 2/3)(\hat{\Gamma}_2 + \hat{\Gamma}_2')$, where $\hat{\Gamma}_j = [\sum_t h(\hat{\gamma}, w_t) h(\hat{\gamma}, w_{t-j})'] / (T - j)$, $j = 0, 1, 2$. Also, $\hat{D}$ is D with unknown true parameters replaced by consistent estimates. An alternative version of our limited information Bayesian strategy, which we did not explore, works with $V(\theta)$, which is the V matrix constructed with the D and S matrices implied by the model when its parameter values are set to θ .

Table 1a Non-estimated Parameters in Simple Model

Parameter	Value	Description
β	0.99	Discount factor
r_π	1.5	Taylor rule: inflation coefficient
r_x	0.2	Taylor rule: output gap coefficient
ρ_R	0.8	Taylor rule: interest rate smoothing coefficient
κ^p	0.11	Slope of Phillips curve
κ^g	0.4	Okun's law coefficient
ω	1.0	Elasticity of efficient unemployment, u^* , w.r.t. efficient hours, h^*

result suggests that, for sufficiently large T , the likelihood of $\hat{\gamma}$ conditional on θ and $\hat{V}$ is given by the following multivariate Normal distribution:²⁹

$$p\left(\hat{\gamma}|\theta; \frac{\hat{V}}{T}\right) = \frac{1}{(2\pi)^{6/2}} \left| \frac{\hat{V}}{T} \right|^{-\frac{1}{2}} \exp \left\{ -\frac{T}{2} (\hat{\gamma} - \gamma(\theta))' \hat{V}^{-1} (\hat{\gamma} - \gamma(\theta)) \right\}. \quad (58)$$

Given a set of priors for θ , $p(\theta)$, the posterior distribution of θ conditional on $\hat{\gamma}$ and $\hat{V}$ is, for sufficiently large T ,

$$p\left(\theta|\hat{\gamma}; \frac{\hat{V}}{T}\right) = \frac{p\left(\hat{\gamma}|\theta; \frac{\hat{V}}{T}\right)p(\theta)}{p\left(\hat{\gamma}; \frac{\hat{V}}{T}\right)}.$$

²⁹ We performed a small Monte Carlo experiment to investigate whether Hansen's asymptotic results are likely to be a good approximation with a sample size, $T = 232$. The results of the experiment make us cautiously optimistic. Our Monte Carlo study used the classic New Keynesian model without unemployment (i.e., Eqs. (42), (50), and (51) with $h_t^* \equiv 0$, and Eqs. (52)–(54). With one exception, we set the relevant economic parameters as in Table 1a. The exception is ρ_R , which we set to zero. In addition, the parameters in θ were set as in the posterior mode for the partial information procedure in Table 1b. With this parameterization, the model implies (after rounding) $\sigma_y = 0.021$, $\rho_1 = \rho_2 = -0.039$. Here, $\sigma_y \equiv E(\Delta y)^2)^{1/2}$, $\rho_i \equiv E(\Delta y_i \Delta y_{t-i}) / \sigma_y^2$, $i = 1, 2$. We then simulated 10,000 data sets, each with $T = 232$ artificial observations on output growth, Δy_t . The mean, across simulated samples, of estimates of σ_y , ρ_1 , ρ_2 , is, respectively, 0.021, -0.039 , and -0.033 . Thus, the results are consistent with the notion that our second moment estimator is essentially unbiased. To investigate the accuracy of Hansen's Normality result, we examined the coverage of 80% confidence intervals computed in the usual way (i.e., the point estimate ± 1.28 times the corresponding sample standard deviation computed in exactly the way specified in the previous footnote). In the case of σ_y , ρ_1 , ρ_2 the 80% confidence interval excluded the true values of the parameters 22.35, 21.87, and 21.39% of the time, respectively. We found these to be reasonably close to the 20% numbers suggested by the asymptotic theory. Related to this, we found little bias in our estimator of the sample standard deviation estimator. In particular, the actual standard deviation of the estimator of σ_y , ρ_1 , and ρ_2 across the 10,000 samples is 0.00098, 0.064, and 0.065. The mean of the corresponding sample estimates is 0.00095, 0.062, and 0.064, respectively. Evidently, the estimator of the sampling standard deviation is roughly unbiased.

The marginal density, $p(\hat{\gamma}; \hat{V}/T)$, as well as the marginal posterior distribution of individual elements of θ can be computed using a standard random walk metropolis algorithm or by using the Laplace approximation.³⁰ In the present application, we use the Laplace approximation. Our moment-matching Bayesian approach has several attractive features. First, it has the advantage of transparency because it focuses on a small number of features of the data. Second, it does not require the assumption that the underlying data are realizations from a Normal distribution, as is the case in conventional Bayesian analyses.³¹ The Normality in Eq. (58) depends on the validity of the central limit theorem, not on Normality of the underlying data. Third, the method has the advantage of computational speed. The matrix inversion and log determinant in Eq. (58) needs to be computed only once. In addition, evaluating a quadratic form like the one in Eq. (58) is computationally very efficient. These computational advantages are likely to be important when searching for the mode of the posterior distribution. Moreover, the advantages may be overwhelmingly important when computing the whole posterior distribution using a standard random walk Metropolis algorithm. In this case, Eq. (58) must be evaluated on the order of hundreds of thousands of times.

Because our econometric method may be of independent interest, we compare the results obtained using it with results based on a conventional full information Bayesian approach. In particular, let Y denote the data on unemployment and output growth used to compute $\hat{\gamma}$ for our limited information Bayesian procedure. In this case, the posterior distribution of θ given Y is

$$p(\theta|Y) = \frac{p(Y|\theta)p(\theta)}{p(Y)},$$

where $p(Y|\theta)$ is the Normal likelihood function and $p(Y)$ is the marginal density of the data. The priors, $p(\theta)$, used in the two econometric procedures are the same and they are listed in Table 1b.

Table 1b reports posterior modes and posterior standard deviations for the parameters, θ . Note how similar the results are between the full and limited information methods. The one difference has to do with λ , the autoregressive parameter for the labor supply shock. The posterior mode for this parameter is somewhat sensitive to which econometric method is used. The standard deviation of the posterior mode of λ is more sensitive to the method used. In all but one case, there appears to be substantial information in the data about the parameters, as measured by the reduction in standard deviation from prior to posterior. The exception is λ . Under

³⁰ For additional discussion of the Laplace approximation, see Section 5.4.

³¹ Failure of Normality in aggregate macroeconomic data is discussed in Christiano (2007).

Table 1b Priors and Posteriors for Parameters of Simple Model

		Prior		Posterior mode	
				[Std. dev. ^a]	
Parameter		Distribution [bounds]	Mean, std. dev. [5% and 95%]	Limited info ^b	Full info ^c
<i>Exogenous processes parameters</i>					
Autocorrelation, labor supply shock	λ	Beta [0, 1]	0.75, 0.15 [0.47, 0.95]	0.71 [0.16]	0.83 [0.08]
Autocorrelation, Phillips curve shock	χ	Beta [0, 1]	0.75, 0.15 [0.47, 0.95]	0.92 [0.01]	0.93 [0.02]
Std. dev., technology shock (%)	σ_z	Inv. gamma [0, ∞]	0.50, 0.40 [0.18, 1.04]	0.62 [0.04]	0.63 [0.04]
Std. dev., labor supply shock (%)	σ_{h^*}	Inv. gamma [0, ∞]	0.50, 0.40 [0.18, 1.04]	0.24 [0.06]	0.19 [0.03]
Std. dev., monetary policy shock (%)	σ_M	Inv. gamma [0, ∞]	0.50, 0.40 [0.18, 1.04]	0.13 [0.01]	0.11 [0.01]
Std. dev., Phillips curve shock (%)	σ_μ	Inv. gamma [0, ∞]	0.50, 0.40 [0.18, 1.04]	0.24 [0.03]	0.25 [0.03]

^aBased on Laplace approximation.^bLimited info refers to our Bayesian moment-matching procedure.^cFull info refers to standard full information Bayesian inference based on the full likelihood of the data.

the limited information procedure, there is little information in the data about this parameter.

We analyze the properties of the model at the mode of the posteriors of θ . Because the Table 1b results are so similar between limited and full information methods, the corresponding model properties are also essentially the same. As a result, we only report properties based on the posterior mode implied by the limited information procedure.

Table 1c reports $\hat{\gamma}$, the empirical second moments underlying the limited information estimator, as well as the corresponding second moments implied by the model. The empirical and model moments are reasonably close. The variance decomposition implied by the model is reported in Table 1d. Most of the variance in output is due to technology shocks and to the disturbance in the Phillips curve. Note that technology shocks have no impact on any of the other variables. This reflects that with our policy rule, the economy's response to a random walk technology shock is efficient and involves no response in the interest rate, inflation, or any labor market variable. The economics of this result is discussed in Section 3.4. In the case of unemployment, the disturbance to the Phillips curve is the principle source of fluctuations. Labor supply shocks turn out to be relatively unimportant as a source of fluctuations. The implications of the latter finding for our results are discussed in the next section.

3.3.4 Estimating the output gap using the CTW model

The implications of our model for the information in the unemployment rate for the output gap is displayed in Table 1e. The row called “posterior mode” reports

$$r^{\text{two-sided}} = 0.11 \text{ and } r^{\text{one-sided}} = 0.09.$$

Table 1c Properties of Simple Model (at Limited Information Posterior Mode) and Data^a

Covariances ($\times 100$)	Model	Data	Covariances ($\times 100$)	Model	Data
Cov. $(\Delta y_t, \Delta y_t)$	0.0099	0.0090	Cov. $(\Delta y_t, \Delta y_{t-2})$	0.0010	0.0017
Cov. (u_t, u_t)	0.0190	0.0220	Cov. $(\Delta y_t, u_{t-2})$	0.0021	0.0033
Cov. $(\Delta y_t, u_t)$	-0.0013	-0.0002	Cov. $(u_t, \Delta y_{t-2})$	-0.0025	-0.0038
			Cov. (u_t, u_{t-2})	0.0174	0.0201
Cov. $(\Delta y_t, \Delta y_{t-1})$	0.0021	0.0030			
Cov. $(\Delta y_t, u_{t-1})$	0.0012	0.0022			
Cov. $(u_t, \Delta y_{t-1})$	-0.0021	-0.0023			
Cov. (u_t, u_{t-1})	0.0184	0.0215			

^aSample: 1951Q1 to 2008Q4. Data series: Δy — real per capita GDP growth, u — unemployment rate.

Table 1d Variance Decomposition of Simple Model (at Limited Information Posterior Mode, in %)

Output growth	Unemployment rate	Nom. interest rate	Inflation rate	Output gap
<i>Technology Shocks</i>				
38.7	0	0	0	0
<i>Monetary Policy Shocks</i>				
17.7	1.8	0.7	0.5	1.9
<i>Labor Supply Shocks</i>				
0.7	3.9	0.1	0	0.3
<i>Phillips Curve Shocks</i>				
42.9	94.3	99.2	99.5	97.8

Table 1e Information About Output Gap in Unemployment Rate, u , Simple Model

Parameter	Two-sided projection			One-sided projection		
	Projection error (%)			Projection error (%)		
	standard deviation $\times 100$			standard deviation $\times 100$		
	u	u	$r^{\text{two-sided}}$	u	u	$r^{\text{one-sided}}$
	Observed	Unobserved		Observed	Unobserved	
Posterior mode	0.74	2.26	0.11	0.79	2.66	0.09
<i>Alternative parameter values</i>						
$\lambda = 0.99999$, $100\sigma_h^* = 0.0015$	0.68	2.24	0.09	0.68	2.64	0.07
$\omega = 0.001$	0.00081	2.26	0.00	0.00084	2.65	0.00
$100\sigma_h^* = 0.001$	0.0036	2.24	0.00	0.0036	2.64	0.00
$100\sigma_h^* = 1$	1.80	2.53	0.51	2.12	2.84	0.56

Note: (i) $r^{\text{two-sided}}$ is the ratio of the two-sided projection error variance when u is observed to what it is when it is not observed. $r^{\text{one-sided}}$ is the analogous object for the case of one-sided projections. For details, see the text. (ii) The posterior mode of the parameters are based on our limited information Bayesian procedure.

Thus, in the case of the two-sided projection, the variance of the projection error in the output gap is reduced by 89% when the unemployment rate is included in the data used to estimate the output gap. The 95% confidence interval for the percent output gap is the point estimate plus and minus 4.4% when the estimate is based only on

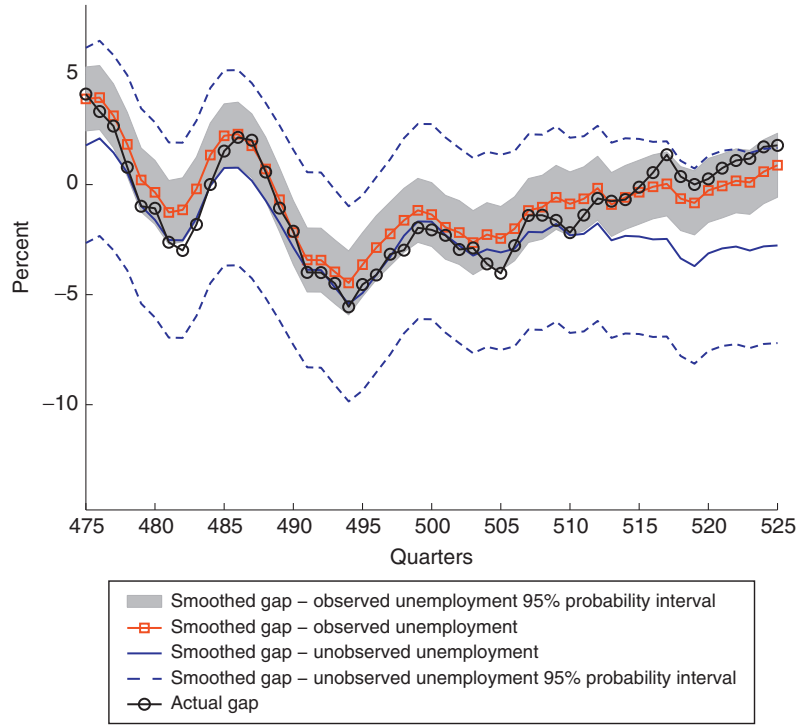


Figure 2 Actual versus smoothed output gap, artificial data.

the output growth data. That interval shrinks by over 60%, to $\pm 1.5\%$ with the introduction of unemployment.³² Figure 2 displays observations 475 to 525 in a simulation of 1000 observations from our model. The figure shows the actual gap as well as estimates-based information sets that include only output growth and output growth plus unemployment. In addition, we display 95% confidence tunnels corresponding to the two information sets.³³ Note how much wider the tunnel is for estimates based on output growth alone.

Our optimal linear estimator of the output gap based on output growth alone (see Eq. 48) is directly comparable to the HP filter as an estimator of the gap.³⁴ The latter is also based on output data alone. The information in Figure 3 allows us to compare these two filters. Panel a shows the filter weights as they apply to the level of output,

³² These observations are based on the following calculations: $1.5 = 0.0074 \times 1.96 \times 100$ and $4.4 = 0.0226 \times 1.96 \times 100$ using the information in Table 1e. Here, 1.96 is the 2.5% critical value for the standard Normal distribution.

³³ The confidence tunnels are constructed by adding and subtracting 1.96 times the standard deviation of the projection error standard deviation implied by the Kalman smoother to the smoothed estimates of the gap. The assumption of Normality implicit in multiplying by 1.96 is justified here because the disturbances in the underlying simulation are drawn from a Normal distribution.

³⁴ We set the smoothing parameter in the HP filter to 1600.

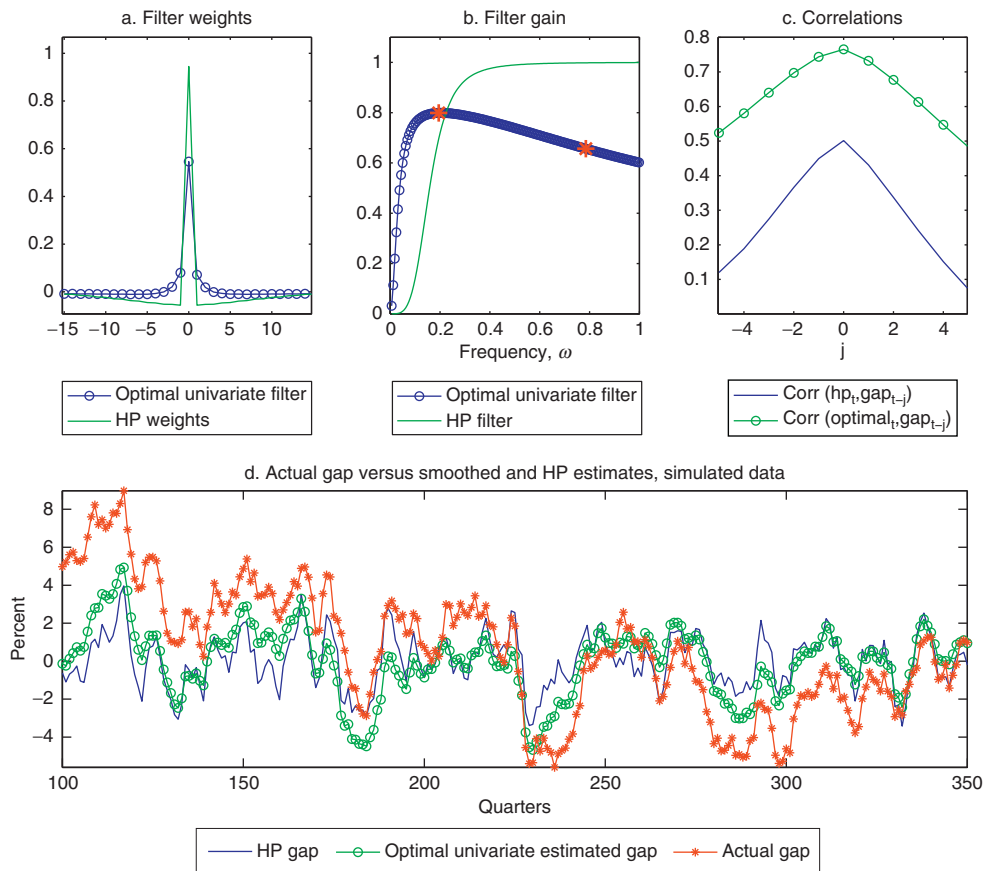


Figure 3 HP Filter and optimal univariate filter for estimating output gap. Note: Stars in panel b indicate business cycle frequencies corresponding to 2 and 8 years.

y_t .³⁵ Note how similar the pattern of weights is, although they are not identical. The filter weights for the HP filter are known to be exactly symmetric. This is not a property of the optimal weights. However, panel a in Figure 3 shows that the optimal filter weights are very nearly symmetric. So, while the phase angle of the HP filter is exactly zero, the phase angle of the optimal filter implied by our model is nearly zero. Panel b in Figure 3 compares the gain of the two filters over a subset of frequencies that includes the business cycle frequencies, whose boundaries are indicated in the figure by stars. Evidently, both are approximately high pass filters. However, the optimal filter

³⁵ We computed the filter weights for the HP filter as well as for Eq. (48) by expressing the filters in the frequency domain and applying the inverse Fourier transform. In the case of Eq. (48), we compute the $\tilde{h}_j$'s in $x_t = \sum_{j=-\infty}^{\infty} h_j \Delta y_{t-j} + e_t^y = \sum_{j=-\infty}^{\infty} \tilde{h}_j y_{t-j} + e_t^y$. We use the result in King and Rebelo (1993) to express the HP filter in the frequency domain.

lets through lower frequency components of the output data and also slightly attenuates the higher frequencies. Panel c displays the cross correlations of the actual output gap with the HP and optimal filters, respectively. This was done in a sample of 1000 artificial observations on output simulated from our model (the optimal filter in a finite sample of data is obtained using the Kalman smoother). Note that both estimates are positively correlated with the actual gap. Of course, the gap is more highly correlated with the optimal estimate of that gap than with the HP filter estimate. Panel d in Figure 3 displays a subsample of our artificial data. We can see directly how similar the two filters are. However, note that there is a substantial low frequency component in the actual gap and this low frequency component is better tracked by the optimal filter. This is consistent with the result in panel b which indicates that the optimal filter allows lower frequency components of output to pass through.

Next, we applied the same statistical procedure to the U.S. data that we used to estimate the output gap in the artificial data. The results are displayed in Figure 4, which displays HP-filtered, log, real, per capita Gross Domestic Product (GDP), as well as the two-sided estimate of the output gap when unemployment is and is not included in the data set used in the projections.³⁶ We have not included confidence tunnels, to

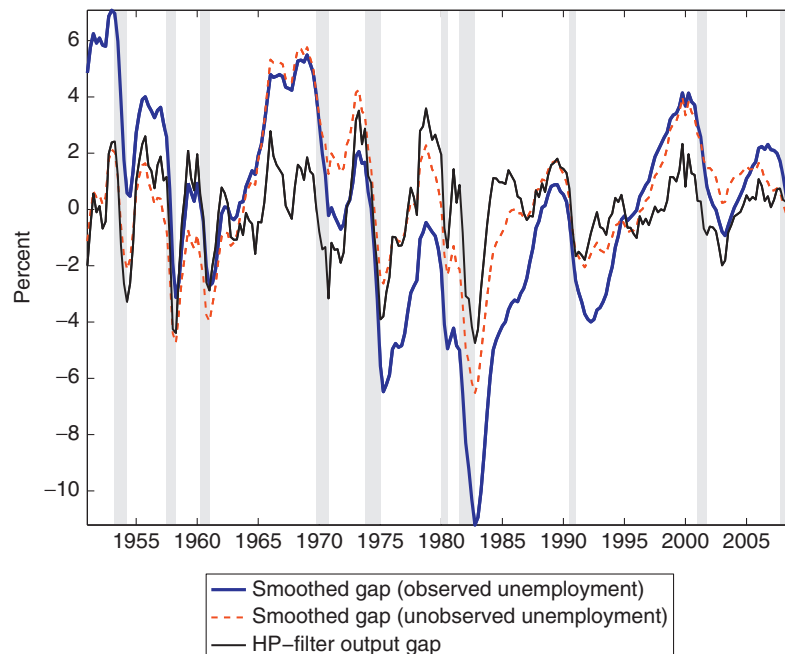


Figure 4 Output gap in U.S. data.

³⁶ Our calculations for Figure 4 are based on the Kalman smoother.

avoid further cluttering the diagram. In addition, the gray areas in the bracket denote the start and end date of recessions, according to the National Bureau of Economic Research (NBER). Several observations are worth making about the results in Figure 4. First, the estimated output gap is always relatively low in a neighborhood of NBER recessions. Second, the gap shows a tendency to begin falling before the onset of the NBER recession. This is to be expected. The NBER typically dates the start of a recession by the first quarter in which the economy undergoes two quarters of negative growth. Given that growth in the U.S. economy is positive on average, the start date of the NBER recession occurs after economic activity has already been winding down for at least a few quarters. This also explains why the HP filter estimate of the gap also typically starts to fall one or two quarters before an NBER recession. Third, consistent with the results in the previous paragraph, the gap estimates based on the HP filter and our estimate based on output data alone produce very similar results. Fourth, the inclusion of unemployment in the data used to estimate the output gap has a quantitatively large impact on the results. The estimated gap is substantially more volatile when unemployment is used and it is also more volatile than the HP filter gap. That the incorporation of unemployment has a big impact is perhaps not surprising, given the posterior mode of our parameters, which implies that labor supply shocks are relatively unimportant. As a result, the efficient unemployment rate, u_t^* , is not very volatile and the actual unemployment rate is a good indicator of the output gap (see Eq. 55).

We gain additional insight into our measures of the gap by examining the implied estimates of potential output. These are presented in Figure 5, which displays actual output as well as our measures of potential output based on using just output and using output and unemployment. Not surprisingly, in view of the results in Figure 4, the estimate of potential that uses unemployment is the smoother one of the two. Our results are similar to the results presented by Justiniano and Primiceri (2008), who also conclude that potential output is smooth.³⁷

Our model is well suited to shed light on the question: Under what circumstances can we expect unemployment to contain useful information about the output gap? The general answer is that if the efficient level of unemployment is constant, then the actual unemployment rate is highly informative, because in this case it represents a direct observation on the output gap. This is documented in three ways in Table 1e. First, we consider the case where the total variance in the labor supply shock, h_t^* , is kept constant, but is reallocated into the very low frequencies. A motivation for this is the

³⁷ Although Sala, Söderström and Trigari (2008) do not specifically display their model's implications for potential output, one can infer from their estimate of the output gap that the measure of potential output implicit in their calculations is also smooth, like the one presented by Justiniano and Primiceri (2008). Except for these two papers, estimates of potential GDP reported in the literature are often more volatile than what we find. See, for example, Walsh's (2005) discussion of Levin, Onatski, Williams, and Williams (2005). See also Kiley (2010) and the sources he cites.

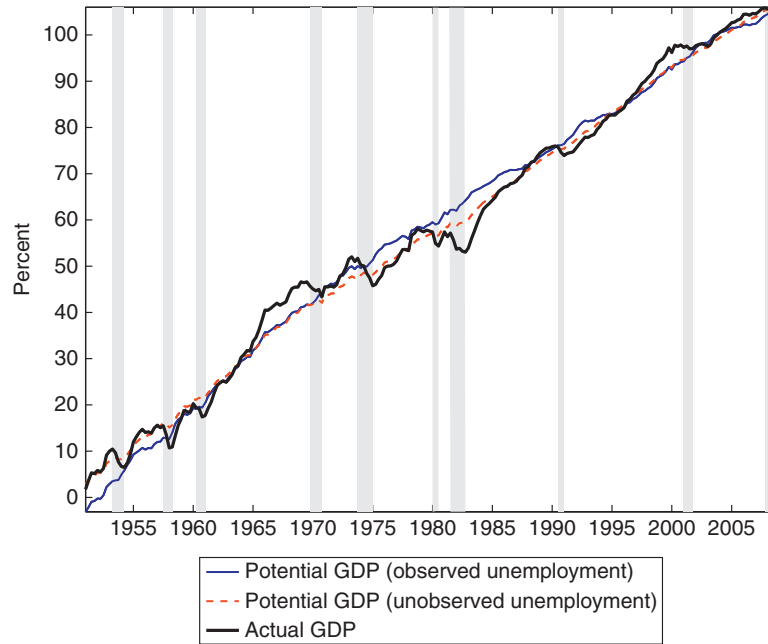


Figure 5 Actual output and two measures of potential output, U.S. data.

finding in Christiano (1988, pp. 266–268) that a low frequency labor supply shock is required to accommodate the behavior of aggregate hours worked. We set $\lambda = 0.99999$ and adjust $\sigma_{h_t}^2$ so that the variance of h_t^* is equal to what is implied by the model at the posterior mode. In this case, the efficient level of employment is a variable that evolves slowly over time.³⁸ As a result, the efficient rate of unemployment itself is slow-moving, so that most of the short-term fluctuations in the actual unemployment rate correspond to movements in the unemployment gap, u_t^g , and, hence in the output gap (recall Eq. 55). Consistent with this intuition, Table 1e indicates that the increase in λ causes $r^{\text{two-sided}}$ and $r^{\text{one-sided}}$ to fall to 0.09 and 0.07, respectively. Similarly, Table 1e also shows that if we reduce the magnitude of ω or of the variance of the labor supply shock itself, then the use of unemployment data essentially removes all uncertainty about the output gap. Finally, the table also shows what happens when we increase the importance of the labor supply shock. In particular, we increased the innovation variance in h_t^* by a factor of 4, from 0.24% to 1.0%. The result of this change on the model is that labor supply shocks now account for 10% of the variance of output growth and 41% of the variance of unemployment. With the efficient level of unemployment more volatile, we can expect that the value of the unemployment

³⁸ This captures the view that the evolution of h_t^* represents demographic and other slowly-moving factors.

rate for estimating the output gap is reduced. Interestingly, according to Table 1e, unemployment is still very informative for the output gap. Despite the relatively high volatility in the labor supply shock, the unemployment rate still reduces the variance of the prediction error for the output gap by about 44–49%.

In sum, the results reported here suggest the possibility that the unemployment rate might be very useful for estimating the output gap. We find that this is likely to be particularly true if the efficient level of unemployment evolves slowly over time. In addition, we found in our estimated model that the HP filter estimate of the gap closely resembles the estimate of the gap that is optimal conditional on our model. All of these observations ought to be viewed as suggestive at best. Because part of our objective here is pedagogic, the observations were made in a very simple setting. It would be interesting to investigate whether they are also true in more complicated environments with more shocks, in which more data are available to the econometrician. Section 3.4 shows that the optimal filter for extracting the output gap is very sensitive to the details of the underlying model. As a consequence, the similarity between the HP filter and the optimal filter found in this section ought to only be treated as suggestive. A final assessment of the relationship between the two filters requires additional experience with a variety of models.

3.4 Using HP-filtered output to estimate the output gap

Section 3.3.4 displayed a model environment with the property that the HP filter is nearly optimal as a device for estimating the output gap. This section shows that the accuracy of the HP filter for extracting the output gap is very sensitive to the details about the underlying model. We demonstrate this point in a simple version of the classic New Keynesian model (i.e., $\gamma = 1$, $\psi = 0$) in which there is only one shock, the technology shock. We show that the HP filter may be positively or negatively correlated with the true output gap, depending on the time series properties of the shock. When the shock triggers strong wealth effects, then output overreacts to the shock, relative to the efficient equilibrium. In this case, the HP-filtered estimate of the gap is positively correlated with the true output gap. If the shock triggers only a weak wealth effect, that correlation is negative.

Our analysis requires a careful review of the economics of the response of employment and output to a technology shock. This is a topic that is of independent interest because it has attracted widespread attention, primarily in response to the provocative paper by Galí (1999).

The linearized equilibrium conditions of the model are given by Eqs. (40)–(43), with $\psi = 0$, $\gamma = 1$. We consider the following two laws of motion for technology:

$$\begin{aligned}\Delta \log z_t &= \rho_z \Delta \log z_{t-1} + \varepsilon_t^z, \text{ “AR (1) in growth rate”} \\ \log z_t &= \rho_z \log z_{t-1} + \varepsilon_t^z, \text{ “AR (1) in levels”}.\end{aligned}$$

These two laws of motion have the same implication for what happens to z_t in the period of a positive realization of ε_t^z . But, they differ sharply in their implications for the eventual impact of a shock on z_t . In the AR(1) in growth rate, a 0.01 shock in ε_t^z drives up z_t by 1%, but creates the expectation that z_t will eventually rise by $1/(1 - \rho_z)\%$. In the AR(1) in levels representation, a jump in z_t is associated with the expectation that z_t will be lower in later periods. We adopt the following parameterization:

$$\beta = 0.99, \rho_z = 0.5, \rho_R = 0, r_x = 0.2, r_\pi = 1.5, \phi = 0.2, \xi_p = 0.75.$$

In the case of the AR(1) in growth rate, a 1% shock up in technology is followed by additional increases, with technology eventually settling at a level that is permanently higher by 2% (see panel d in Figure 6). The response of the efficient level of consumption coincides with the response of the technology shock. Households in this economy experience a big rise in wealth in the moment of the shock. The motive to smooth consumption intertemporally makes them want to set their consumption to its permanently higher level right away. The rise in the rate of interest in the efficient

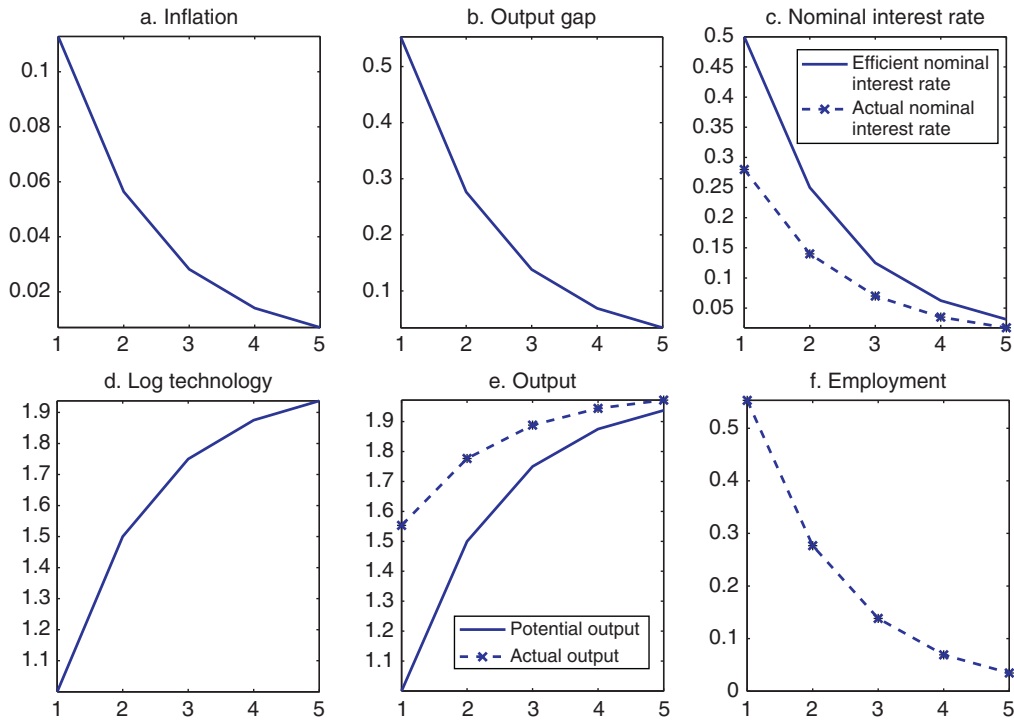


Figure 6 Dynamic response of simple model without capital to a one percent technology shock (AR(1) in growth rate specification).

equilibrium is designed to restrain this potential surge in consumption. This is why it is that in the efficient equilibrium, output (see panel e Figure 6) rises by the same amount as the technology shock, while employment remains unchanged. Now consider the actual equilibrium. According to panel c in the figure, the interest rate rule generates an inefficiently small rise in the rate of interest. As a result, monetary policy fails to fully reign in the surge in consumption demand triggered by the shock. Employment rises and so output rises by more than the technology shock. The increase in employment leads to an increase in costs and, therefore, inflation. The output gap responds positively to the shock and so potential output (i.e., the efficient level of output) is less volatile than the actual level. We can expect that the output gap estimated by the HP filter, which estimates potential output smoothing actual output, will at least be positively correlated with the true output gap.

We simulated a large number of artificial observations using the model and we then HP-filtered the output data.³⁹ Figure 7A displays actual, potential and HP smoothed

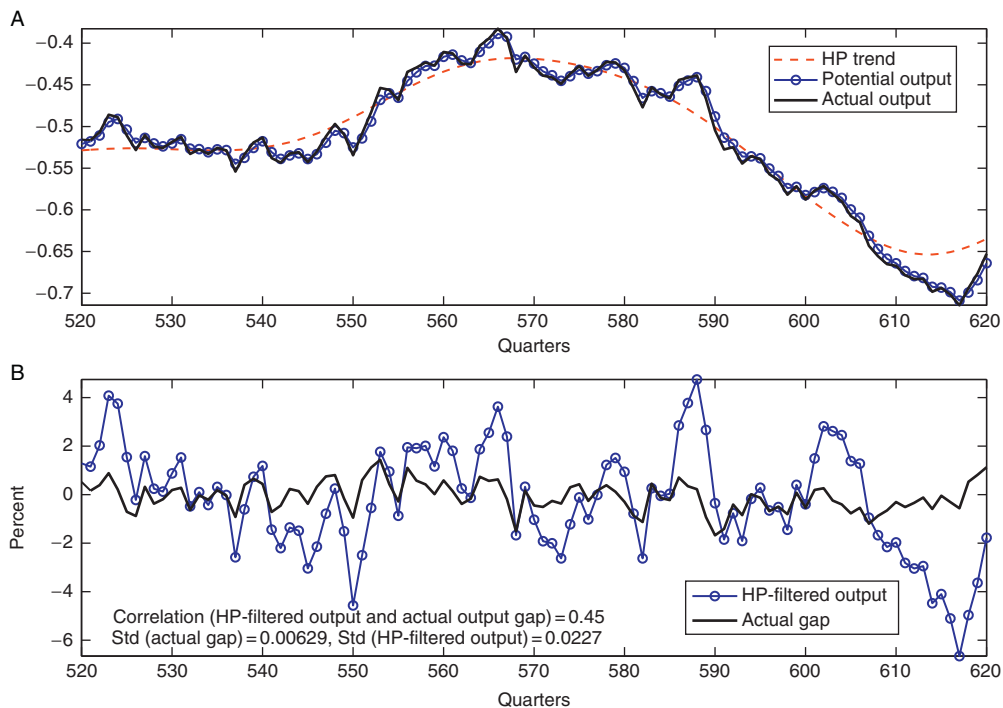


Figure 7 (A) Potential output, actual output and hp trend based on actual output (simulated data) (B) HP filter estimate of output gap versus actual gap (simulated data). AR(1) in growth rate specification.

³⁹ We used the usual smoothing parameter value for quarterly data, 1600.

output. We can see that the HP filter substantially oversmooths the data. However, consistent with the presumption implicit in the HP filter, the actual level of output is (somewhat) more volatile than the corresponding efficient level. Figure 7B displays the actual gap and the HP-estimated gap. Note that they are positively correlated, though the HP filtered gap is too volatile.

Now consider the AR(1) in levels specification of technology. The dynamic response of technology to a 1% disturbance in e_t^z is displayed in panel d of Figure 8. The state of technology is high in the period of the shock, compared to its level anticipated for later periods. As before, the efficient level of consumption mirrors the time path of the technology shock. In the efficient equilibrium, agents expect lower future consumption and so intertemporal smoothing motivates them to cut current consumption relative to its efficient level. The drop in the interest rate in the efficient equilibrium is designed to resist this relative weakness in consumption (see panel c). Put differently, a sharp drop in the interest rate is needed to ensure that demand expands by enough to keep employment unchanged in the face of the technology improvement. In the actual equilibrium, the monetary policy rule cuts the interest rate less aggressively than in the

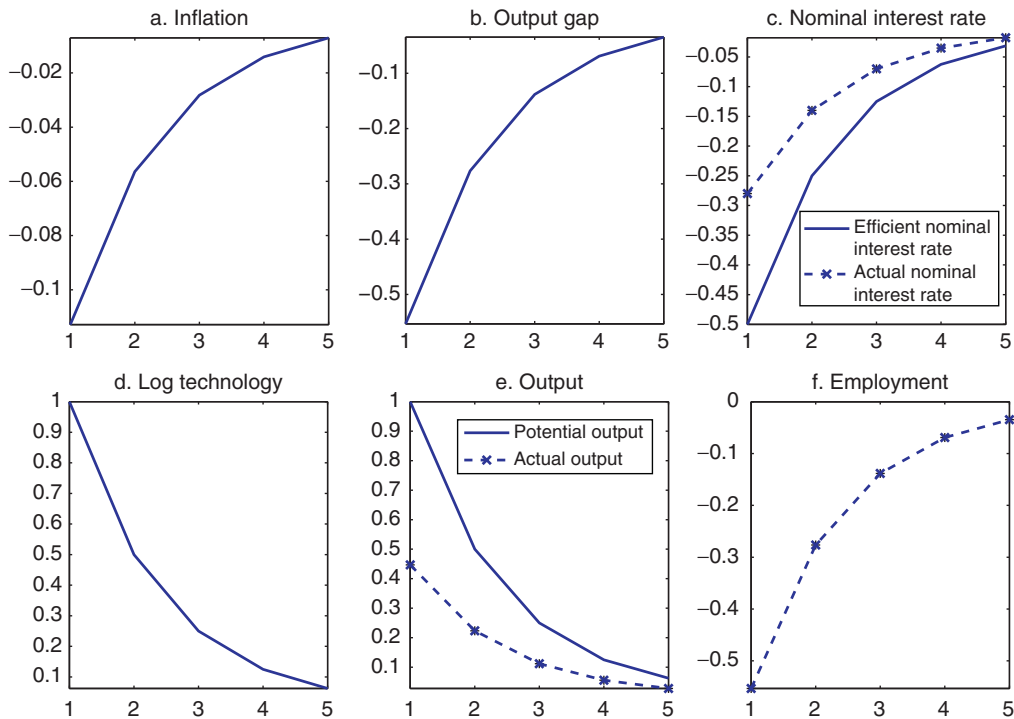


Figure 8 Dynamic response of simple model without capital to a one percent technology shock. AR(1) in levels specification.

efficient equilibrium. The relatively small drop in the interest rate fails to reverse the weakness in demand. As a result, the response of output is relatively weak and employment falls. The fall in employment is associated with a fall in marginal production costs and this explains why inflation falls in response to the technology shock. Figure 9A displays the implications of the AR(1) in levels specification of technology for the HP filter as a way to estimate the output gap. Note how potential output is substantially more volatile than actual output. As an estimator of potential output, the HP filter goes in precisely the wrong direction, by smoothing. Figure 9B compares the HP filter estimate of the output gap with the corresponding actual value. Note how the two are now negatively correlated.

A by-product of the previous discussion is an exploration of the economics of the response of hours worked to a technology shock in the classic New Keynesian model. In that model, hours worked rise in response to a technology shock that triggers a big wealth effect, and falls in response to a technology shock that implies a weak wealth effect. The principle that the hours worked response is greater when a technology shock triggers a large wealth effect survives in more complicated New Keynesian models such as the one discussed in the next section.

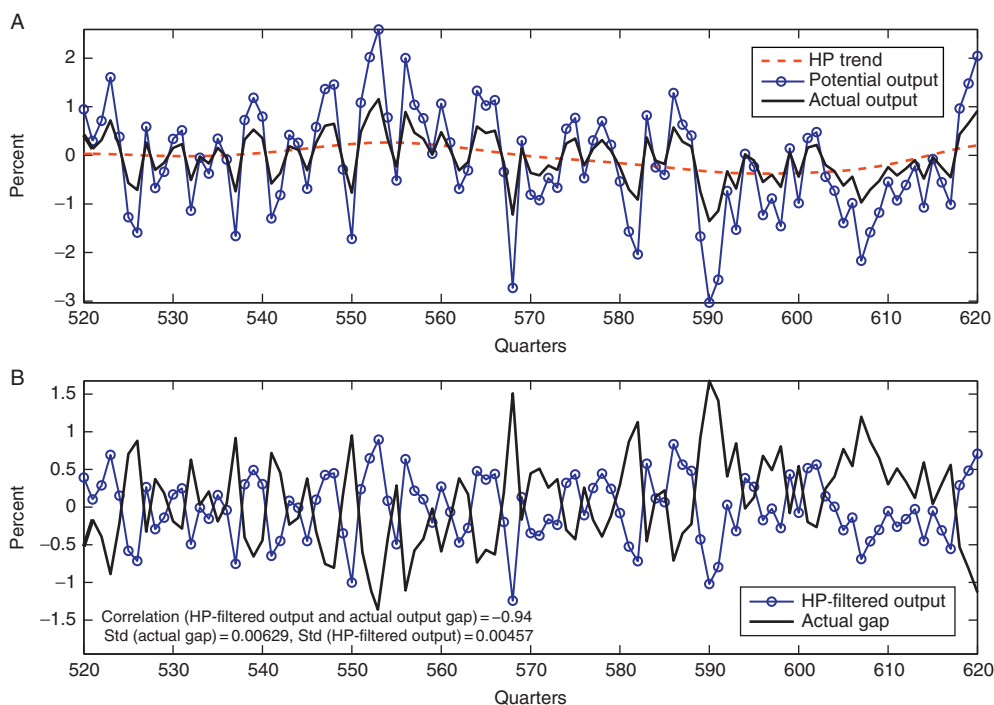


Figure 9 (A) Potential output, actual output and hp trend based on actual output (simulated data) (B) HP filter estimate of output gap versus actual gap (simulated data). AR(1) in levels specification.

4. MEDIUM-SIZED DSGE MODEL

A classic question in economics is: Why do prices take so long to respond to a monetary disturbance and why do real variables react so strongly? Mankiw, writing in the year 2000, maintained that an empirically successful explanation of monetary non-neutrality has confounded economists at least since David Hume wrote “Of Money” in 1752. Moreover, at the time that Mankiw was writing, it looked as though the question would remain unanswered for a considerable time to come. A reason that monetary DSGE models have been so successful in the past decade is that, with a combination of modest price and wage stickiness and various “real frictions,” they roughly reproduce the evidence of monetary non-neutrality that had seemed so hard to match. The purpose of this section and the next two sections is to spell out the basis for this observation in detail. Inevitably, doing this requires a model that is more complicated than the various versions of the simple model studied in the previous sections. In describing the model in this section, we explain the rationale for each departure from the simple model.

The model developed here is a version of the one in CEE. We describe the objectives and constraints of the agents in the model, and leave the derivation of the equilibrium conditions to the technical appendix. This model includes monetary policy shocks, so that it can be used to address the monetary non-neutrality question. In addition, the model includes two technology shocks. As a further check on the model, we follow ACEL in also evaluating the model’s ability to match the estimated dynamic response of economic variables to the two technology shocks.

4.1 Goods production

An aggregate homogeneous good is produced using the technology, (Eq. 5). The first-order condition of the representative, competitive producer of the homogeneous good is given by Eq. (6). Substituting this first-order condition back into Eq. (5) yields the restriction across prices, Eq. (7). Each intermediate good, $i \in (0,1)$, is produced by a monopolist who treats Eq. (6) as its demand curve. The intermediate good producer takes the aggregate quantities, P_t and Y_t , as exogenous.

We use a production function for intermediate good producers that is standard in the literature. It does not use materials inputs, but it does use the services of capital, $K_{i,t}$:

$$Y_{i,t} = (z_t H_{i,t})^{1-\alpha} K_{i,t}^\alpha - z_t^+ \varphi. \quad (59)$$

Here, z_t is a technology shock whose logarithmic first difference has a positive mean and φ denotes a fixed production cost. The economy has two sources of growth: the positive drift in $\log(z_t)$ and a positive drift in $\log(\Psi_t)$, where Ψ_t is the state of an investment specific technology shock discussed later. The object, z_t^+ , in Eq. (59) is defined as follows:

$$z_t^+ = \Psi_t^{\frac{\alpha}{1-\alpha}} z_t.$$

Along a nonstochastic steady-state growth path, Y_t/z_t^+ and $Y_{i,t}/z_t^+$ converge to constants.

The two shocks, z_t and Ψ_t , are specified to be unit root processes in order to be consistent with the assumptions we use in our VAR analysis to identify the dynamic response of the economy to neutral and investment specific technology shocks. We adopt the following time series representations for the shocks:

$$\Delta \log z_t = \mu_z + \varepsilon_t^z, E(\varepsilon_t^z)^2 = \sigma_z^2 \quad (60)$$

$$\Delta \log \Psi_t = (1 - \rho_\Psi)\mu_\Psi + \rho_\Psi \Delta \log \Psi_{t-1} + \varepsilon_t^\Psi, E(\varepsilon_t^\Psi)^2 = \sigma_\Psi^2. \quad (61)$$

Our assumption that the neutral technology shock follows a random walk with drift closely matches the finding in Smets and Wouters (2007) who estimated $\log z_t$ to be highly autocorrelated. The empirical analysis of Prescott (1986) also supported the notion that $\log z_t$ is a random walk with drift. Finally, Fernald (2009) constructed a direct estimate of total factor productivity growth for the business sector. The first-order autocorrelation of quarterly observations covering the period 1947Q2 to 2009Q3 is 0.0034, consistent with the idea of a random walk.

We assume that there is no entry or exit by intermediate good producers. The no entry assumption would be implausible if firms enjoyed large and persistent profits. The fixed cost in Eq. (59) is introduced to minimize the incentive to enter. We set φ so that intermediate good producer profits are zero in steady state. This requires that the fixed cost grow at the same rate as the growth rate of economic output, and this is why φ is multiplied by z_t^+ in Eq. (59). A potential empirical advantage of including fixed costs of production is that, by introducing some increasing returns to scale, the model can in principle account for evidence that labor productivity rises in the wake of a positive monetary policy shock.

In Eq. (59), $H_{i,t}$ denotes homogeneous labor services hired by the i^{th} intermediate good producer. Firms must borrow the wage bill. We follow CEE in supposing that firms borrow the entire wage bill (i.e., $\psi = 1$ and $\nu_t = 0$ in Eq. 9) so that the cost of one unit of labor is given by

$$W_t R_t. \quad (62)$$

Here, W_t denotes the aggregate wage rate and R_t denotes the gross nominal interest rate on working capital loans. The assumption that firms require working capital was introduced by CEE as a way to help dampen the rise in inflation after an expansionary shock to monetary policy. An expansionary shock to monetary policy drives R_t down and, with other things the same, this reduces firm marginal cost. Inflation is dampened because marginal cost is the key input into firms' price-setting decisions. Indirect evidence consistent with the working capital assumption includes the frequently found VAR-based results, suggesting that inflation drops for a little while after a positive monetary policy shock. It is hard to think of an alternative to the working capital

assumption to explain this evidence, apart from the possibility that the estimated response reflects some kind of econometric specification error.⁴⁰

Another motivation for treating interest rates as part of the cost of production has to do with Ball's (1994) "dis-inflationary boom" critique models that do not include interest rates in costs. Ball's critique focuses on the Phillips curve in Eq. (30), which we reproduce here for convenience:

$$\hat{\pi}_t = \beta E_t \hat{\pi}_{t+1} + \kappa_p \hat{s}_t,$$

where $\hat{\pi}_t$ and $\hat{s}_t$ denote inflation and marginal cost, respectively. Also, $\kappa_p > 0$ is a reduced form parameter and β is slightly less than unity. According to the Phillips curve, if the monetary authority announces it will fight inflation by strategies that (plausibly) bring down future inflation more than present inflation, then $\hat{s}_t$ must jump. In simple models $\hat{s}_t$ is directly related to the volume of output (see, e.g., Eq. 34). High output requires more intense utilization of scarce resources, their price goes up, driving up marginal cost, $\hat{s}_t$. Ball (1994) criticized theories that do not include the interest rate in marginal cost on the grounds that we do not observe booms at the start of disinflations. Including the interest rate in marginal cost potentially avoids the Ball critique because the high $\hat{s}_t$ may simply reflect the high interest rate that corresponds to the disinflationary policy, and not higher output.

We adopt the Calvo model of price frictions. With probability ξ_p , the intermediate good firm cannot reoptimize its price, in which case it is assumed to set its price according to the following rule:⁴¹

$$P_{i,t} = \pi P_{i,t-1}. \quad (63)$$

Note that in steady state, firms that do not optimize their prices raise prices at the general rate of inflation. Firms that optimize their prices in a steady-state growth path raise their prices by the same amount. This is why there is no price dispersion in steady state. According to the discussion near Eq. (29), the fact that we analyze the first-order approximation of a DSGE model in a neighborhood of steady state means that we can impose the analog of $p_t^* = 1$.

With probability $1 - \xi_p$ the intermediate good firm can reoptimize its price. Apart from the fixed cost, the i^{th} intermediate good producer's profits are the analog of Eq. (13):

$$E_t \sum_{j=0}^{\infty} \beta^j v_{t+j} [P_{i,t+j} Y_{i,t+j} - s_{t+j} P_{t+j} Y_{i,t+j}],$$

⁴⁰ This possibility was suggested by Sims (1992) and explored further in Christiano et al. (1999). See also Bernanke, Boivin, and Elias (2005).

⁴¹ Equation (63) excludes the possibility that firms index to past inflation. We discuss the reason for this specification in Section 6.2.2.

where s_t denotes the marginal cost of production, denominated in units of the homogeneous good. The object, s_t , is a function only of the costs of capital and labor, and is described in Section C of the technical appendix. Marginal cost is independent of the level of $Y_{i,t}$ because of the linear homogeneity of the first expression on the right of Eq. (59). The first-order necessary conditions associated with this optimization problem are reported in section E of the technical appendix.

Goods market clearing dictates that the homogeneous output good is allocated among alternative uses as follows:

$$Y_t = G_t + C_t + \tilde{I}_t. \quad (64)$$

Here, C_t denotes household consumption, G_t denotes exogenous government consumption, and $\tilde{I}_t$ is a homogenous investment good which is defined as follows:

$$\tilde{I}_t = \frac{1}{\Psi_t} (I_t + a(u_t) \bar{K}_t). \quad (65)$$

The investment goods, I_t , are used by households to add to the physical stock of capital, $\bar{K}_t$.⁴² The remaining investment goods are used to cover maintenance costs, $a(u_t) \bar{K}_t$, arising from capital utilization, u_t . The cost function, $a(\cdot)$, is increasing and convex, and has the property that in steady state, $u_t = 1$ and $a(1) = 0$. The relationship between the utilization of capital, u_t , capital services, K_t , and the physical stock of capital, $\bar{K}_t$, is as follows:

$$K_t = u_t \bar{K}_t.$$

The investment and capital utilization decisions are discussed in Section 4.2. See Section 4.4 for the functional form of the capital utilization cost function. Finally, Ψ_t in Eq. (65) denotes the unit root investment specific technology shock defined in Eq. (61).

4.2 Households

In the model, households supply the factors of production, labor and capital. The model incorporates Calvo-style wage setting frictions along the lines spelled out in Erceg, Henderson, and Levin (2000). Because wages are an important component of costs, wage-setting frictions help slow the response of inflation to a monetary policy shock. As in the case of prices, wage-setting frictions require that there be market power. To ensure that this market power is suffused through the economy and not, for example, concentrated in the hands of a single labor union, we adopt the framework that is now standard in monetary DSGE models. In particular, we adopt a variant

⁴² The notation, I_t , used here should not be confused with materials inputs in Section 2. Our medium-sized DSGE model does not include materials inputs.

of the model in Erceg et al. (2000) by using the analog of the Dixit-Stiglitz type framework used to model price-setting frictions. The assumption that prices are set by producers of specialized goods appears here in the form of the assumption that there are many different specialized labor inputs, $h_{j,t}$, for $j \in (0,1)$. There is a single monopolist that sets the wage for each type, j , of labor service. However, that monopolist's market power is severely limited by the presence of other labor services, $j' \neq j$, that are substitutable for $h_{j,t}$.

The variant of the Erceg et al. (2000) model that we work with follows the discussion in Section 2.3 in supposing that labor is indivisible: people work either full time or not at all.⁴³ That is, $h_{j,t}$ represents a quantity of people and not, say, the number of hours worked by a representative worker.

Section 4.2.1 below discusses the interaction between households and the labor market. The Section 4.2.2 discusses the monopoly wage-setting problem in the model. Section 4.2.3 discusses the representative household's capital accumulation decision, and Section 4.2.4 states the representative household's optimization problem.

4.2.1 Households and the labor market

The “labor” hired by firms in the goods-producing sector is interpreted as a homogeneous factor of production, H_t , supplied by “labor contractors.” Labor contractors produce H_t by combining a range of differentiated labor inputs, $h_{t,j}$, using the following linear homogeneous technology:

$$H_t = \left[\int_0^1 (h_{t,j})^{\frac{1}{\lambda_w}} dj \right]^{\lambda_w}, \lambda_w > 1.$$

Labor contractors are perfectly competitive and take the wage rate, W_t , of H_t as given. They also take the wage rate, $W_{t,j}$, of the j^{th} labor type as given. Contractors choose inputs and outputs to maximize profits,

$$W_t H_t - \int_0^1 W_{t,j} h_{t,j} dj.$$

The first order necessary condition for optimization is given by:

$$h_{t,j} = \left(\frac{W_t}{W_{t,j}} \right)^{\frac{\lambda_w}{\lambda_w - 1}} H_t. \quad (66)$$

Substituting the latter back into the labor aggregator function and rearranging, we obtain:

⁴³ Our approach follows the one in Gali (2010a).

$$W_t = \left[\int_0^1 W_{t,j}^{\frac{1}{1-\lambda_w}} dj \right]^{1-\lambda_w}. \quad (67)$$

Differentiated labor is supplied by a large number of identical households. The representative household has many members corresponding to each type, j , of labor. Each worker of type j has an index, l , distributed uniformly over the unit interval, $[0,1]$, which indicates that worker's aversion to work. A type j worker with index l experiences utility:

$$\log(c_t^e - bC_{t-1}) - l^\phi, \phi > 0,$$

if employed and

$$\log(c_t^{ne} - bC_{t-1}),$$

if not employed. When $b > 0$ the worker's marginal utility of current consumption is an increasing function of the household's consumption in the previous period. Given the additive separability of consumption and employment in utility, the efficient allocation of consumption across workers within the household implies:⁴⁴

$$c_t^e = c_t^{ne} = C_t.$$

The quantity of the j^{th} type of labor supplied by the representative household, $h_{t,j}$, is determined by Eq. (66). We suppose the household sends j -type workers with $0 \leq l \leq h_{t,j}$ to work and keeps those with $l > h_{t,j}$ out of the labor force. The equally weighted integral of utility over all $l \in [0,1]$ workers is

$$\log(C_t - bC_{t-1}) - A \frac{h_{t,j}^{1+\phi}}{1+\phi}.$$

Aggregate household utility also integrates over the unit measure of j -type workers:

$$\log(C_t - bC_{t-1}) - A \int_0^1 \frac{h_{t,j}^{1+\phi}}{1+\phi} dj. \quad (68)$$

Next we explain how $h_{t,j}$ is determined and how the household chooses C_t .

The wage rate of the j^{th} type of labor, $W_{t,j}$, is determined outside the representative household by a monopoly union that represents all j -type workers across all households. The union's problem is discussed in Section 4.2.2.

The presence of $b > 0$ in Eq. (68) is motivated by VAR-based evidence like that displayed in section 6 below, which suggests that an expansionary monetary policy shock triggers (i) a hump-shaped response in consumption and (ii) a persistent

⁴⁴ For an environment in which perfect insurance is not feasible, see CTW.

reduction in the real rate of interest.⁴⁵ With $b = 0$ and a utility function separable in labor and consumption like the one in Eq. (68), (i) and (ii) are difficult to reconcile. An expansionary monetary policy shock that triggers an increase in expected future consumption would be associated with a rise in the real rate of interest, not a fall. Alternatively, a fall in the real interest rate would cause people to rearrange consumption intertemporally, so that consumption is relatively high right after the monetary shock and low later. Intuitively, one can reconcile (i) and (ii) by supposing the marginal utility of consumption is inversely proportional not to the level of consumption, but to its derivative. To see this, it is useful to recall the familiar intertemporal Euler equation implied by household optimization (see, e.g., Eq. 4):

$$\beta E_t \frac{u_{c,t+1}}{u_{c,t}} \frac{R_t}{\pi_{t+1}} = 1.$$

Here, $u_{c,t}$ denotes the marginal utility of consumption at time t . From this expression, we see that a low R_t/π_{t+1} tends to produce a high $u_{c,t+1}/u_{c,t}$; that is, a rising trajectory for the marginal utility of consumption. This illustrates the problematic implication of the model when $u_{c,t}$ is inversely proportional to C_t as in Eq. (68) with $b = 0$. To fix this implication we need a model change with the property that a rising $u_{c,t}$ path implies hump-shaped consumption. A hump-shaped consumption path corresponds to a scenario in which the slope of the consumption path is falling, suggesting that (i) and (ii) can be reconciled if $u_{c,t}$ is proportional to the slope of consumption. The notion that marginal utility is inversely proportional to the slope of consumption corresponds loosely to $b > 0$.⁴⁶ The fact that (i) and (ii) can be reconciled with the assumption of habit persistence is of special interest, because there is evidence from other sources that also favors the assumption of habit persistence; for example, in asset pricing (see, e.g., Boldrin, Christiano, & Fisher, 2001; Constantinides, 1990) and growth (see Carroll, Overland, & Weil, 1997, 2000). In addition, there may be a solid foundation in psychology for this specification of preferences.⁴⁷

⁴⁵ The earliest published statement of the idea that $b > 0$ can help account for (i) and (ii) that we are aware of is Fuhrer (2000).

⁴⁶ In particular, suppose first that lagged consumption in Eq. (68) represents aggregate, economy-wide consumption and $b > 0$. This corresponds to the so-called “external habit” case, where it is the lagged consumption of others that enters utility. In that case, the marginal utility of household C_t is $1/(C_t - bC_{t-1})$, which corresponds to the inverse of the slope of the consumption path, at least if b is large enough. In our model we think of C_{t-1} as corresponding to the household’s own lagged consumption (that is why we use the same notation for current and lagged consumption), the so-called “internal habit” case. In this case, the marginal utility of C_t also involves future terms, in addition to the inverse of the slope of consumption from $t = 1$ to t . The intuition described in the text, which implicitly assumed external habit, also applies roughly to the internal habit case that we consider.

⁴⁷ Anyone who has gone swimming has had the experience of habit persistence. It is usually very hard at first to jump into a swimming pool because it seems so cold. The swimmer who jumps (or is pushed) into the water after much procrastination initially experiences a tremendous shock with the sudden drop in temperature. However, after only a few minutes the new, lower temperature is perfectly comfortable. In this way, the lagged temperature seems to influence one’s experience of current temperature, as in habit persistence.

The logic associated with the previous intertemporal Euler equation suggests that there are other approaches that can at least go part way in reconciling (i) and (ii). For example, Guerron-Quintana (2008) showed that nonseparability between consumption and labor in Eq. (68) can help reconcile (i) and (ii). He pointed out that if the marginal utility of consumption is an increasing function of labor and the model predicts that employment rises in a hump shape after an expansionary monetary shock, then it is possible that consumption rises in a hump shape.

4.2.2 Wages, employment and monopoly unions

We turn now to a discussion of the monopoly union that sets the wage of j -type workers. In each period, the monopoly union must satisfy its demand curve, (Eq. 66), and it faces Calvo frictions in the setting of $W_{t,j}$. With probability $1 - \xi_w$ the union can optimize the wage and with the complementary probability, ξ_w , it cannot. In the latter case, we suppose that the nominal wage rate is set as follows:

$$W_{j,t+1} = \tilde{\pi}_{w,t+1} W_{j,t} \quad (69)$$

$$\tilde{\pi}_{w,t+1} = \pi_t^{\kappa_w} \pi^{(1-\kappa_w)} \mu_{z+}, \quad (70)$$

where $\kappa_w \in (0, 1)$. With this specification, the wage of each type j labor is the same in the steady state. Because the union problem has no state variable, all unions with the opportunity to reoptimize in the current period face the same problem. In particular, such a union chooses the current value of the wage, $\tilde{W}_t$, to maximize:

$$E_t \sum_{i=0}^{\infty} (\beta \xi_w)^i \left[v_{t+i} \tilde{W}_{t+i}^t h_{t+i}^t - A_L \frac{(h_{t+i}^t)^{1+\phi}}{(1+\phi)} \right]. \quad (71)$$

Here, h_{t+i}^t and $\tilde{W}_{t+i}^t$ denote the quantity of workers employed and their wage rate, in period $t+i$, of a union that has an opportunity to reoptimize the wage in period t and does not reoptimize again in periods $t+1, \dots, t+i$. Also, v_{t+i} denotes the marginal value assigned by the representative household to the wage.⁴⁸ The union treats v_t as an exogenous variable. In the expression (71), ξ_w appears in the discounting because the union's period t decision only impacts on future histories in which it cannot reoptimize its wage.

Optimization by all labor unions leads to a simple equilibrium condition, when the variables are linearized about the nonstochastic steady state.⁴⁹ The condition is:

⁴⁸ The object, v_t , is the multiplier on the household budget constraint in the Lagrangian representation of the problem.

⁴⁹ The details of the derivation are explained in Section G of the technical appendix.

$$\Delta_{\kappa_w} \hat{\pi}_{w,t} = \frac{\kappa}{1 + \phi \frac{\lambda_w}{\lambda_w - 1}} \left(\underbrace{-\hat{\psi}_{z^+,t} + \phi \hat{H}_t}_{\text{scaled labor cost of marginal worker}} - \underbrace{\hat{w}_t}_{\text{scaled real wage}} \right) + \beta \Delta_{\kappa_w} \hat{\pi}_{w,t+1}, \quad (72)$$

where

$$\kappa = \frac{(1 - \xi_w)(1 - \beta \xi_w)}{\xi_w}.$$

In Eq. (72), $\hat{\pi}_{w,t}$ is the gross growth rate in the nominal wage rate, expressed in percent deviation from steady state. Also, $\hat{\psi}_{z^+,t}$ represents the percent deviation of the scaled multiplier, $\psi_{z^+,t}$, from its steady-state value. The scaled multiplier is defined as follows:

$$\psi_{z^+,t} \equiv v_t P_t z_t^+,$$

where v_t is the multiplier on the household budget constraint. The first two terms inside the parentheses in Eq. (72) correspond to the marginal cost of labor and the third term, $\hat{w}_t$, corresponds to the real wage. Both the marginal cost of labor and the real wage have been scaled by z_t^+ . Expression (72) has a simple interpretation. The first term in parentheses is related to the cost of working by the marginal worker. When this (scaled) cost exceeds the (scaled) real wage, $\hat{w}_t$, then the monopoly unions currently setting wages place upward pressure on wage inflation. The coefficient multiplying the term in parentheses is also interesting. If the degree of wage and price stickiness are the same; that is, $\xi_w = \xi_p$, then κ takes on the same value as κ_p , the analog of κ in the price Phillips curve, Eq. (35). In this case, the slope of the price Phillips curve in terms of marginal cost is bigger than the slope of the wage Phillips curve, Eq. (72). This reflects that in the slope of the wage Phillips curve, κ is divided by:

$$1 + \phi \frac{\lambda_w}{\lambda_w - 1} > 1.$$

According to this expression, the slope of the wage Phillips curve is smaller if the elasticity of demand for labor, $\lambda_w/(\lambda_w - 1)$ is large and/or if the marginal cost of work, MRS , is sharply increasing in work (i.e., ϕ is large). The intuition for this is as follows. Suppose the j^{th} monopoly union contemplates a particular rise in the nominal wage, for whatever reason. Consider a given slope of the demand for labor. The rise in the wage implies a lower quantity of labor demanded. The steeper the marginal cost curve is, the greater the implied drop in marginal cost. Now consider a given slope of marginal cost. The flatter the slope of demand for the j^{th} type of labor is, the larger the drop in the quantity of labor demanded in response to the given contemplated rise

in the wage. Given the upward-sloping marginal cost curve, this also implies a large fall in marginal cost. Thus, the monopoly union that contemplates a given rise in the wage rate anticipates a larger drop in marginal cost to the extent that the demand curve is elastic and/or the marginal cost curve is steep. But, with other things the same, low marginal cost reduces the incentive for a monopolist to raise its price (i.e., the wage in this case). These considerations are absent in our price Phillips curve, Eq. (35), because marginal cost is constant (i.e., the analog of ϕ is zero).⁵⁰

4.2.3 Capital accumulation

The household owns the economy's physical stock of capital, sets the utilization rate of capital, and rents out the services of capital in a competitive market. The household accumulates capital using the following technology:

$$\bar{K}_{t+1} = (1 - \delta)\bar{K}_t + F(I_t, I_{t-1}) + \Delta_t, \quad (73)$$

where Δ_t denotes physical capital purchased in a competitive market from other households. Since all households are the same in terms of capital accumulation decisions, $\Delta_t = 0$ in equilibrium. We nevertheless include Δ_t so that we can assign a price to installed capital. In Eq. (73), $\delta \in [0, 1]$ and we use the specification suggested in CEE:

$$F(I_t, I_{t-1}) = \left(1 - S\left(\frac{I_t}{I_{t-1}}\right)\right)I_t, \quad (74)$$

where the functional form, S , that we use is described in Section 4.4. In Eq. (74), $S = S' = 0$ and $S'' > 0$ along a nonstochastic steady state growth path. Here, S' and S'' denote the first and second derivatives, respectively, of S .

Let $P_t P_{k',t}$ denote the nominal market price of Δ_t . For each unit of $\bar{K}_{t+1}$ acquired in period t , the household receives X_{t+1}^k in net cash payments in period $t + 1$:

$$X_{t+1}^k = u_{t+1} P_{t+1} r_{t+1}^k - \frac{P_{t+1}}{\Psi_{t+1}} a(u_{t+1}). \quad (75)$$

The first term is the gross nominal period $t + 1$ rental income from a unit of $\bar{K}_{t+1}$. The second term represents the cost of capital utilization, $a(u_{t+1})P_{t+1}/\Psi_{t+1}$. Here, P_{t+1}/Ψ_{t+1} is the nominal price of the investment goods absorbed by capital utilization. That P_{t+1}/Ψ_{t+1} is the equilibrium market price of investment goods follows from the technology specified in Eqs. (64) and (65), and the assumption that investment goods are produced from homogeneous output goods by competitive firms.

The introduction of variable capital utilization is motivated by a desire to explain the slow response of inflation to a monetary policy shock. In any model prices are

⁵⁰ This intuition for why the slope of the wage Phillips curve is flatter with elastic labor demand and/or steep marginal cost is the same as the intuition that firm-specific capital flattens the price Phillips curve (see, e.g., ACEL; Christiano, 2004; de Walque, Smets, & Wouters, 2006; Svein & Weinke, 2005; Woodford, 2004).

heavily influenced by costs. Costs in turn are influenced by the elasticity of the factors of production. If factors can be rapidly expanded with a small rise in cost, then inflation will not rise much after a monetary policy shock. Allowing for variable capital utilization is a way to make the services of capital elastic. If there is very little curvature in the a function, then households are able to expand capital services without much increase in cost.

The form of the investment adjustment costs in Eq. (73) is motivated by a desire to reproduce VAR-based evidence that investment has a hump-shaped response to a monetary policy shock. Alternative specifications include $F \equiv I_t$ and

$$F = I_t - \frac{S''}{2} \left(\frac{I_t}{\bar{K}_t} - \delta \right)^2 \bar{K}_t. \quad (76)$$

Specification (76) has a long history in macroeconomics, and has been in use since at least Lucas and Prescott (1971). To understand why DSGE models generally use the adjustment cost specification in Eq. (74) rather than Eq. (76), it is useful to define the rate of return on investment:

$$R_{t+1}^k = \frac{x_{t+1}^k + \left[1 - \delta + S'' \left(\frac{I_{t+1}}{\bar{K}_{t+1}} - \delta \right) \frac{I_{t+1}}{\bar{K}_{t+1}} - \frac{S''}{2} \left(\frac{I_{t+1}}{\bar{K}_{t+1}} - \delta \right)^2 \right] P_{k',t+1}}{P_{k',t}}. \quad (77)$$

The numerator is the one-period payoff from an extra unit of $\bar{K}_{t+1}$, and the denominator is the corresponding cost, both in consumption units. In Eq. (77), $x_{t+1}^k \equiv X_{t+1}^k / P_{t+1}$ denotes the earnings net of costs. The term in square brackets is the quantity of additional $\bar{K}_{t+2}$ made possible by the additional unit of $\bar{K}_{t+1}$. This is composed of the undepreciated part of $\bar{K}_{t+1}$ left over after production in period $t + 1$, plus the impact of $\bar{K}_{t+1}$ on $\bar{K}_{t+2}$ via the adjustment costs. The object in square brackets is converted to consumption units using $P_{k',t+1}$, which is the market price of $\bar{K}_{t+2}$ denominated in consumption goods. Finally, the denominator is the price of the extra unit of $\bar{K}_{t+1}$.

The price of extra capital in competitive markets corresponds to the marginal cost of production. Thus,

$$\begin{aligned} P_{k',t} &= -\frac{dC_t}{d\bar{K}_{t+1}} = -\frac{dC_t}{dI_t} \times \frac{dI_t}{d\bar{K}_{t+1}} \\ &= \frac{1}{\frac{d\bar{K}_{t+1}}{dI_t}} = \begin{cases} 1 & \text{When } F \text{ is } I \\ \frac{1}{1 - S'' \times \left(\frac{I_t}{\bar{K}_t} - \delta \right)} & \text{When } F \text{ is as in (76),} \end{cases} \end{aligned} \quad (78)$$

where we ignore Ψ_t for now ($\Psi_t \equiv 1$). The derivatives in the first line correspond to marginal rates of technical transformation. The marginal rate of technical transformation between consumption and investment is implicit in Eqs. (64) and (65). The marginal rate of technical transformation between I_t and $\bar{K}_{t+1}$ is given by the capital accumulation equation. The relation in the second line of Eq. (78) is referred to as “Tobin’s q ” relation, where Tobin’s q here corresponds to $P'_{k',t}$. This is the market value of capital (i.e., the marginal cost of capital under our assumption that markets are competitive) divided by the price of investment goods. Here, q can differ from unity due to the investment adjustment costs.

We are now in a position to convey the intuition about why DSGE models have generally abandoned the specification in Eq. (76) in favor of Eq. (73). The key reason has to do with VAR-based evidence (see Section 6 below) that suggests the real interest rate falls persistently after a positive monetary policy shock, while investment responds in a hump-shaped pattern. Any model that is capable of producing this type of response will have the property that the real return on capital, Eq. (77) — for arbitrage reasons — also falls after an expansionary monetary policy shock. Suppose, to begin, that $S'' = 0$, so that there are no adjustment costs at all and $P'_{k',t} = 1$. In this case, the only component in R_t^k that can fall is λ_{t+1}^k , which is dominated by the marginal product of capital. That is, approximately, the rate of return on capital is

$$(1 - \alpha)K_{t+1}^{\alpha-1}H_{t+1}^{1-\alpha} + 1 - \delta.$$

In steady state this object equals $1/\beta$ (ignoring growth), which is roughly 1.03 in annual terms. At the same time, the object, $1 - \delta$, is roughly 0.9 in annual terms, so that the endogenous part of the rate of return of capital is a very small part of that rate of return. As a result, any given drop in the return on capital requires a very large percentage drop in the endogenous part, $K_{t+1}^{\alpha-1}H_{t+1}^{1-\alpha}$. An expansion in investment can bring about this drop, but it has to be a very substantial surge. To see why investment must expand so much, note first that the endogenous part of the rate of return is not only small, but the capital stock receives a weight substantially less than unity in that expression. Second, a model that successfully reproduces the VAR-based evidence that employment rises after a positive monetary policy shock implies that hours worked rises. This pushes the endogenous component of the rate of return up, increasing the burden on the capital stock to bring the rate of return on investment down. For these reasons, models without adjustment costs generally imply a counterfactually strong surge in investment in the wake of a positive shock to monetary policy.

With $S'' > 0$ the endogenous component of the rate of return on capital is much larger. However, in practice models that adopt the adjustment cost specification, Eq. (76), generally imply that the biggest investment response occurs in the period of the shock, and not later. To gain intuition into why this is so, suppose the contrary:

that investment does exhibit a hump-shape response in investment. Equation (78) implies a similar hump-shape pattern in the price of capital, $P_{k',t}$.⁵¹ This is because $P_{k',t}$ is primarily determined by the contemporaneous flow of investment. So, under our supposition about the investment response, a positive monetary policy shock generates a rise in $P_{k',t+1}/P_{k',t}$ over at least several periods in the future. According to Eq. (77), the anticipated future capital gains create an incentive to invest right away. Thus, households would be induced to substitute away from a hump-shaped response, toward one in which the immediate response is much stronger. In practice, this means that in equilibrium, the biggest response of investment occurs in the period of the shock, with later responses converging to zero.

The adjustment costs in Eq. (74) do have the implication that investment responds in a hump-shaped manner. According to Eq. (74), a quick rise in investment over its previous level is costly.

There are other reasons to take the specification in Eq. (74) seriously. Lucca (2006) and Matsuyama (1984) have described interesting theoretical foundations that produce Eq. (74) as a reduced form. For example, in Matsuyama shifting production between consumption and capital goods involves a learning-by-doing process, which makes quick movements in either direction expensive. Also, Matsuyama explains why the abundance of empirical evidence that appears to reject Eq. (76) may be consistent with Eq. (74). Consistent with Eq. (74), Topel and Rosen (1988) argue that data on housing construction cannot be understood without using a cost function that involves the change in the flow of housing construction.

4.2.4 Household optimization problem

The household's period t budget constraint is as follows:

$$P_t \left(C_t + \frac{1}{\Psi_t} I_t \right) + B_{t+1} + P_t P_{k',t} \Delta_t \leq \int_0^1 W_{t,j} h_{t,j} dj + X_t^k \bar{K}_t + R_{t-1} B_t, \quad (79)$$

where $W_{t,j}$ represents the wage earned by the household, B_{t+1} denotes the quantity of risk-free bonds purchased by the household, and R_t denotes the gross nominal interest rate on bonds purchased in period $t - 1$, which pay off in period t . The household's problem is to select sequences, $\{C_t, I_t, \Delta_t, B_{t+1}, \bar{K}_{t+1}\}$, to maximize Eq. (68) subject to the wage process selected by the monopoly unions, Eqs. (73), (75), and (79).

⁵¹ Note from Eq. (78) that the price of capital increases as investment rises above its level in steady state, which is the level required to just meet the depreciation in the capital stock. Our assertion that the price of capital follows the same hump-shaped pattern as investment after a positive monetary policy shock reflects our implicit assumption that the shock occurs when the economy is in a steady state. This will be true on average, but not at each date.

4.3 Fiscal and monetary authorities and equilibrium

We suppose that monetary policy follows a Taylor rule of the following form:

$$\log\left(\frac{R_t}{R}\right) = \rho_R \log\left(\frac{R_{t-1}}{R}\right) + (1 - \rho_R) \left[r_\pi \log\left(\frac{\pi_{t+1}}{\pi}\right) + r_y \log\left(\frac{gdp_t}{gdp}\right) \right] + \varepsilon_{R,t}, \quad (80)$$

where $\varepsilon_{R,t}$ denotes an iid shock to monetary policy. As in CEE and ACEL, we assume that the period t realization of $\varepsilon_{R,t}$ is not included in the period t information set of the agents in our model. This ensures that our model satisfies the restrictions used in the VAR analysis to identify a monetary policy shock. In Eq. (80), gdp_t denotes scaled real GDP defined as follows:

$$gdp_t = \frac{G_t + C_t + I_t/\Psi_t}{z_t^+}. \quad (81)$$

We adopt the model of government consumption suggested in Christiano and Eichenbaum (1992a):

$$G_t = g z_t^+.$$

In principle, g could be a random variable, although our focus in this paper is just on monetary policy and technology shocks. So, we set g to a constant. Lump-sum transfers are assumed to balance the government budget.

An equilibrium is a stochastic process for the prices and quantities with the property that the household and firm problems are satisfied, and goods and labor markets clear.

4.4 Adjustment cost functions

We adopt the following functional forms. The capacity utilization cost function is

$$a(u) = 0.5b\sigma_a u^2 + b(1 - \sigma_a)u + b((\sigma_a/2) - 1), \quad (82)$$

where b is selected so that $a(1) = a'(1) = 0$ in steady state and σ_a is a parameter that controls the curvature of the cost function. The closer σ_a is to zero, the less curvature there is and the easier it is to change utilization. The investment adjustment cost function takes the following form:

$$S(x_t) = \frac{1}{2} \left\{ \exp \left[\sqrt{S''} (x_t - \mu_{z^+} \mu_\Psi) \right] + \exp \left[-\sqrt{S''} (x_t - \mu_{z^+} \mu_\Psi) \right] - 2 \right\}, \quad (83)$$

$$= 0, x = \mu_{z^+} \mu_\Psi.$$

where $x_t = I_t/I_{t-1}$ and $\mu_{z^+} \mu_\Psi$ is the growth rate of investment in steady state. With this adjustment cost function, $S(\mu_{z^+} \mu_\Psi) = S'(\mu_{z^+} \mu_\Psi) = 0$. Also, $S'' > 0$ is a parameter having the property that it is the second derivative of $S(x_t)$ evaluated at $x_t = \mu_{z^+} \mu_\Psi$. Because of the nature of the above adjustment cost functions, the curvature parameters have no impact on the model's steady state.

5. ESTIMATION STRATEGY

Our estimation strategy is a Bayesian version of the two-step impulse response matching approach applied by Rotemberg and Woodford (1997) and CEE. We begin with a discussion of the two steps. After that, we discuss the computation of a particular weighting matrix used in the analysis.

5.1 VAR step

We estimate the dynamic responses of a set of aggregate variables to three shocks, using standard VAR methods. The three shocks are the monetary policy shock; the innovation to the permanent technology shock, z_t ; and the innovation to the investment specific technology shock, Ψ_t . The estimated contemporaneous and 14 lagged responses in each of $N = 9$ macroeconomic variables to the three shocks are stacked in a vector, $\hat{\psi}$. These macroeconomic variables are a subset of the variables that appear in the VAR. The additional variables in our VAR pertain to the labor market. We use this augmented VAR to facilitate comparison between the analysis in this chapter and in other research where we integrate labor market frictions into a monetary DSGE model.⁵² We denote the vector of variables in the VAR by Y_t , where⁵³

$$\underbrace{Y_t}_{14 \times 1} = \begin{pmatrix} \Delta \ln(\text{relative price of investment}_t) \\ \Delta \ln(\text{real } GDP_t/\text{hours}_t) \\ \Delta \ln(GDP \text{ deflator}_t) \\ \text{unemployment rate}_t \\ \text{capacity utilization}_t \\ \ln(\text{hours}_t) \\ \ln(\text{real } GDP_t/\text{hours}_t) - \ln(W_t/P_t) \\ \ln(\text{nominal } C_t/\text{nominal } GDP_t) \\ \ln(\text{nominal } I_t/\text{nominal } GDP_t) \\ \text{vacancies}_t \\ \text{job separation rate}_t \\ \text{job finding rate}_t \\ \log(\text{hours}_t/\text{labor force}_t) \\ \text{Federal Funds Rate}_t \end{pmatrix}. \quad (84)$$

An extensive general review of identification in VARs appears in Christiano et al. (1999). The specific technical details of how we compute impulse response functions

⁵² See Christiano, Trabandt and Walentin (2010b). The variables, GDP, hours, C, I and the labor force, are expressed in per capita terms.

⁵³ See Section A of the technical appendix for details about the data.

imposing the shock identification are reported in ACEL.⁵⁴ We estimate a two-lag VAR using quarterly data that are seasonally adjusted and cover the period 1951Q1 to 2008Q4. Our identification assumptions are as follows. The only variable that the monetary policy shock affects contemporaneously is the federal funds rate. We make two assumptions to identify the dynamic response to the technology shocks: (i) the only shocks that affect labor productivity in the long run are the two technology shocks, and (ii) the only shock that affects the price of investment relative to consumption is the innovation to the investment specific shock. All of these identification assumptions are satisfied in our model.

Our data set extends over a long range, while we estimate a single set of impulse response functions and model parameters. In effect, we suppose that there has been no parameter break over this long period. Whether or not there has been a break is a question that has been debated. For example, it has been argued that the parameters of the monetary policy rule have not been constant over this period. We do not review this debate here. Implicitly, our analysis sides with the conclusions of those that argue that the evidence of parameter breaks is not strong. For example, Sims and Zha (2006) argued that the evidence is consistent with the idea that monetary policy rule parameters have been unchanged over the sample. Christiano et al. (1999) argued that the evidence is consistent with the proposition that the dynamic effects of a monetary policy shock have not changed during this sample. Standard lag-length selection criteria led us to work with a VAR with 2 lags.⁵⁵

The number of elements in $\hat{\psi}$ corresponds to the number of impulses estimated. Since we consider the contemporaneous and 14 lag responses in the impulses, there are in principle 3 (i.e., the number of shocks) times 9 (number of variables) times 15 (number of responses) = 405 elements in $\hat{\psi}$. However, we do not include in $\hat{\psi}$ the 8 contemporaneous responses to the monetary policy shock that are required to be zero by our monetary policy identifying assumption. Taking this into account, the vector $\hat{\psi}$ has 397 elements.

According to standard classical asymptotic sampling theory, when the number of observations, T , is large, we have

$$\sqrt{T}(\hat{\psi} - \psi(\theta_0))^a \overset{\sim}{\sim} N(0, W(\theta_0, \zeta_0)),$$

⁵⁴ The identification assumption for the monetary policy shock by itself imposes no restriction on the VAR parameters. Similarly, Fisher (2006) showed that the identification assumptions for the technology shocks, when applied without simultaneously applying the monetary shock identification, also imposes no restriction on the VAR parameters. However, ACEL showed that when all the identification assumptions are imposed at the same time, then there are restrictions on the VAR parameters. We found that the test of the overidentifying restrictions on the VAR fails to reject the null hypothesis that the restrictions are satisfied at the 5% critical level.

⁵⁵ We considered VAR specifications with lag length 1, 2, ..., 12. The Schwartz and Hannan-Quinn criteria indicate that a single lag in the VAR is sufficient. The Akaike criterion indicates 12 lags, but we discounted that result.

where θ_0 represents the true values of the parameters that we estimate. The vector, ζ_0 , denotes the true values of the parameters of the shocks that are in the model, but that we do not formally include in the analysis. We find it convenient to express the asymptotic distribution of $\hat{\psi}$ in the following form:

$$\hat{\psi} \stackrel{d}{\sim} N(\psi(\theta_0), V(\theta_0, \zeta_0, T)), \quad (85)$$

where

$$V(\theta_0, \zeta_0, T) \equiv \frac{W(\theta_0, \zeta_0)}{T}.$$

5.2 Impulse response matching step

In the second step of our analysis, we treat $\hat{\psi}$ as “data” and we choose a value of θ to make $\psi(\theta)$ as close as possible to $\hat{\psi}$. As discussed in Section 3.3.3 and following Kim (2002), we refer to our strategy as a limited information Bayesian approach. This interpretation uses Eq. (85) to define an approximate likelihood of the data, $\hat{\psi}$, as a function of θ :

$$\begin{aligned} f(\hat{\psi}|\theta) &= \left(\frac{1}{2\pi}\right)^{\frac{N}{2}} |V(\theta_0, \zeta_0, T)|^{-\frac{1}{2}} \\ &\times \exp \left[-\frac{1}{2} (\hat{\psi} - \psi(\theta))' V(\theta_0, \zeta_0, T)^{-1} (\hat{\psi} - \psi(\theta)) \right]. \end{aligned} \quad (86)$$

In Eq. (86), N denotes the number of elements in $\hat{\psi}$. As we explain next, we treat the value of $V(\theta_0, \zeta_0, T)$ as a known object. Under these circumstances, the value of θ that maximizes the above function represents an approximate maximum likelihood estimator of θ . It is approximate for two reasons: (i) the central limit theorem underlying Eq. (85) only holds exactly as $T \rightarrow \infty$ and (ii) the value of $V(\theta_0, \zeta_0, T)$ that we use is guaranteed to be correct only for $T \rightarrow \infty$.

Treating the function, f , as the likelihood of $\hat{\psi}$, it follows that the Bayesian posterior of θ conditional on $\hat{\psi}$ and $V(\theta_0, \zeta_0, T)$ is

$$f(\theta|\hat{\psi}) = \frac{f(\hat{\psi}|\theta)p(\theta)}{f(\hat{\psi})}, \quad (87)$$

where $p(\theta)$ denotes the priors on θ and $f(\hat{\psi})$ denotes the marginal density of $\hat{\psi}$:

$$f(\hat{\psi}) = \int f(\hat{\psi}|\theta)p(\theta)d\theta.$$

As usual, the mode of the posterior distribution of θ can be computed by simply maximizing the value of the numerator in Eq. (87), since the denominator is not a function of θ . The marginal density of $\hat{\psi}$ is required when we want an overall measure of the fit of our model and when we want to report the shape of the posterior marginal distribution of individual elements in θ . To compute the marginal likelihood, we can use a standard random walk metropolis algorithm or a Laplace approximation. We explain the latter in Section 5.4. The results that we report are based on a standard random walk Metropolis algorithm resulting in a single Monte Carlo Markov Chain of length 600,000. The first 100,000 draws were dropped and the average acceptance rate in the chain is 27%. We confirmed that the chain is long enough so that all the statistics reported in the paper have converged. Section 6.3 compares results based on the Metropolis algorithm with the results based on the Laplace approximation.

5.3 Computation of V

A crucial ingredient in our empirical methodology is the matrix, $V(\theta_0, \zeta_0, T)$. The logic of our approach requires that we have at least an approximately consistent estimator of $V(\theta_0, \zeta_0, T)$. A variety of approaches are possible here. We use a bootstrap approach. Using our estimated VAR and its fitted disturbances, we generate a set of M bootstrap realizations for the impulse responses. We denote these by ψ_i , $i = 1, \dots, M$, where ψ_i denotes the i^{th} realization of the 397×1 vector of impulse responses.⁵⁶ Consider

$$\bar{V} = \frac{1}{M} \sum_{i=1}^M (\psi_i - \bar{\psi})(\psi_i - \bar{\psi})', \quad (88)$$

where $\bar{\psi}$ is the mean of ψ_i , $i = 1, \dots, M$. We set $M = 10,000$. The object, $\bar{V}$, is a 397 by 397 matrix, and we assume that the small sample (in the sense of T) properties of this way (or any other way) of estimating $V(\theta_0, \zeta_0, T)$ are poor. To improve small sample efficiency, we proceed in a way that is analogous to the strategy taken in the estimation of frequency-zero spectral densities (see Newey & West, 1987). In particular, rather than working with the raw variance-covariance matrix, $\bar{V}$, we instead work with $\hat{\bar{V}}$:

$$\hat{\bar{V}} = f(\bar{V}, T).$$

⁵⁶ To compute a given bootstrap realization, ψ_i , we first simulate an artificial data set, $Y_1, \dots, Y_T$. We do this by simulating the response of our estimated VAR to an iid sequence of 14×1 shock vectors that are drawn randomly with replacement from the set of fitted shocks. We then fit a 2-lag VAR to the artificial data set using the same procedure used on the actual data. The resulting estimated VAR is then used to compute the impulse responses, which we stack into the 397×1 vector, ψ_i .

The transformation, f , has the property that it converges to the identity transform, as $T \rightarrow \infty$. In particular, $\hat{\bar{V}}$ dampens some elements in $\bar{V}$, and the dampening factor is removed as the sample grows large. The matrix, $\hat{\bar{V}}$, has on its diagonal the diagonal elements of $\bar{V}$. The entries in $\hat{\bar{V}}$ that correspond to the correlation between the l^{th} lagged response and the j^{th} lagged response in a given variable to a given shock equals the corresponding entry in $\bar{V}$, multiplied by

$$\left[1 - \frac{|l-j|}{n}\right]^{\theta_{1,T}}, l, j = 0, 1, \dots, n.$$

Here, n denotes the number of estimated impulse response lags. Now consider the components of $\bar{V}$ that correspond to the correlations between components of different impulse response functions, either because a different variable is involved or because a different shock is involved, or both. We dampen these entries in a way that is increasing in τ , the separation in time of the two impulses. In particular, we adopt the following dampening factors for these entries:

$$\beta_T \left[1 - \frac{|\tau|}{n}\right]^{\theta_{2,T}}, \tau = 0, 1, \dots, n.$$

We suppose that

$$\beta_T \rightarrow 1, \theta_{i,T} \rightarrow 0, \text{ as } T \rightarrow \infty, i = 1, 2,$$

where the rate of convergence is whatever is required to ensure consistency of $\hat{\bar{V}}$. These conditions leave completely open what values of β_T , $\theta_{1,T}$, $\theta_{2,T}$ we use in our sample. At one extreme, we have

$$\beta_T = 0, \theta_{1,T} = \infty,$$

and $\theta_{2,T}$ unrestricted. This corresponds to the approach in CEE and ACEL, in which $\hat{\bar{V}}$ is simply a diagonal matrix composed of the diagonal components of $\bar{V}$. At the other extreme, we could set β_T , $\theta_{1,T}$, $\theta_{2,T}$ at their $T \rightarrow \infty$ values, in which $\hat{\bar{V}} = \bar{V}$. Here, we work with the approach taken in CEE and ACEL. This has the important advantage of transparency. It corresponds to selecting θ so that the model implied impulse responses lie inside a confidence tunnel around the estimated impulses. When nondiagonal terms in $\bar{V}$ are also used, then the estimator aims not just to put the model impulses inside a confidence tunnel about the point estimates, but it is also concerned about the pattern of discrepancies across different impulse responses. Precisely how the off-diagonal components of $\bar{V}$ give rise to concerns about cross-impulse response patterns of discrepancies is virtually impossible to understand intuitively. This is both because $\bar{V}$ is an enormous matrix and because it is not $\bar{V}$ that enters our criterion but its inverse.

5.4 Laplace approximation of the posterior distribution

The Metropolis algorithm for computing the posterior distribution can be time intensive, and it may be useful — at least in the intermediate stages of a research project — to use the Laplace approximation instead. In Section 6.3, we show that the two approaches generate similar results in our application, although one cannot rely on this being true in general.

To derive the Laplace approximation to $f(\theta|\hat{\psi})$, define

$$g(\theta) \equiv \log f(\hat{\psi}|\theta) + \log p(\theta).$$

Let θ^* denote the mode of the posterior distribution and define the following Hessian matrix:

$$g_{\theta\theta} = -\frac{\partial^2 g(\theta)}{\partial\theta\partial\theta'} \Big|_{\theta=\theta^*}.$$

Note that the matrix, $g_{\theta\theta}$, is an automatic by-product of standard gradient methods for computing the mode, θ^* . The second-order Taylor series expansion of g about $\theta = \theta^*$ is

$$g(\theta) = g(\theta^*) - \frac{1}{2}(\theta - \theta^*)' g_{\theta\theta}(\theta - \theta^*),$$

where the slope term is zero if θ^* is an interior optimum, which we assume. Then,

$$f(\hat{\psi}|\theta)p(\theta) \approx f(\hat{\psi}|\theta^*)p(\theta^*) \exp \left[-\frac{1}{2}(\theta - \theta^*)' g_{\theta\theta}(\theta - \theta^*) \right].$$

Note that

$$\frac{1}{(2\pi)^{\frac{m}{2}}} |g_{\theta\theta}|^{\frac{1}{2}} \exp \left[-\frac{1}{2}(\theta - \theta^*)' g_{\theta\theta}(\theta - \theta^*) \right],$$

where m denotes the number of elements in θ . The last expression is the m -variable Normal distribution for the m random variables, θ , with mean θ^* and variance-covariance matrix, $g_{\theta\theta}^{-1}$. By the standard property of a density function,

$$\int \frac{1}{(2\pi)^{\frac{m}{2}}} |g_{\theta\theta}|^{\frac{1}{2}} \exp \left[-\frac{1}{2}(\theta - \theta^*)' g_{\theta\theta}(\theta - \theta^*) \right] d\theta = 1. \quad (89)$$

Bringing together the previous results, we obtain:

$$\begin{aligned} f(\hat{\psi}) &= \int f(\hat{\psi}|\theta)p(\theta)d\theta \\ &\approx \int f(\hat{\psi}|\theta^*)p(\theta^*) \exp \left[-\frac{1}{2}(\theta - \theta^*)' g_{\theta\theta}(\theta - \theta^*) \right] d\theta \\ &= (2\pi)^{\frac{m}{2}} |g_{\theta\theta}|^{-\frac{1}{2}} f(\hat{\psi}|\theta^*)p(\theta^*), \end{aligned}$$

by Eq. (89). We now have the marginal distribution for $\hat{\psi}$. We can use this to compare the fit of different models for $\hat{\psi}$. In addition, we have an approximation to the marginal posterior distribution for an arbitrary element of θ , say θ_i :

$$\theta_i \sim N\left(\theta_i^*, [g_{\theta\theta}^{-1}]_{ii}\right),$$

where $[g_{\theta\theta}^{-1}]_{ii}$ denotes the i^{th} diagonal element of the matrix, $g_{\theta\theta}^{-1}$.

6. MEDIUM-SIZED DSGE MODEL: RESULTS

We first describe our VAR results. We then turn to the estimation of the DSGE model. Finally, we study the ability of the DSGE model to replicate the VAR-based estimates of the dynamic response of the economy to three shocks.

6.1 VAR results

We briefly describe the impulse response functions implied by the VAR. The solid line in Figures 10–12 indicate the point estimates of the impulse response functions, while the gray area displays the corresponding 95% probability bands.⁵⁷ Inflation and the interest rate are in annualized percent terms, while the other variables are measured in percent. The solid lines with squares and the dashed lines will be discussed when we review the DSGE model estimation results.

6.1.1 Monetary policy shocks

We make five observations about the estimated dynamic responses to an about 50 basis point shock to monetary policy, displayed in Figure 10. Consider first the response of inflation. Two important things to note here are the price puzzle and the delayed and gradual response of inflation.⁵⁸ In the very short run the point estimates indicate that inflation moves in a seemingly perverse direction in response to the expansionary monetary policy shock. This transitory drop in inflation in the immediate aftermath of a monetary policy shock has been widely commented upon, and has been dubbed the “price puzzle.” Christiano et al. (1999) reviewed the argument that the puzzle may be the outcome of the sort of econometric specification error suggested by Sims (1992), and found evidence consistent with that view. Here, we follow ACEL and CEE in taking the position that there is no econometric specification error. Although

⁵⁷ The probability interval is defined by the point estimate of the impulse response, ± 1.96 times the square root of the relevant term on the diagonal of $\hat{V}$ reported in Eq. (88).

⁵⁸ Here, we have borrowed Mankiw’s (2000) language, “delayed and gradual,” to characterize the nature of the response of inflation to a monetary policy shock. Although Mankiw wrote 10 years ago and he cites a wide range of evidence, Mankiw’s conclusion about how inflation responds to a monetary policy shock resembles our VAR evidence very closely. Mankiw argued that the response of inflation to a monetary policy shock is gradual in the sense that it does not peak for 9 quarters.

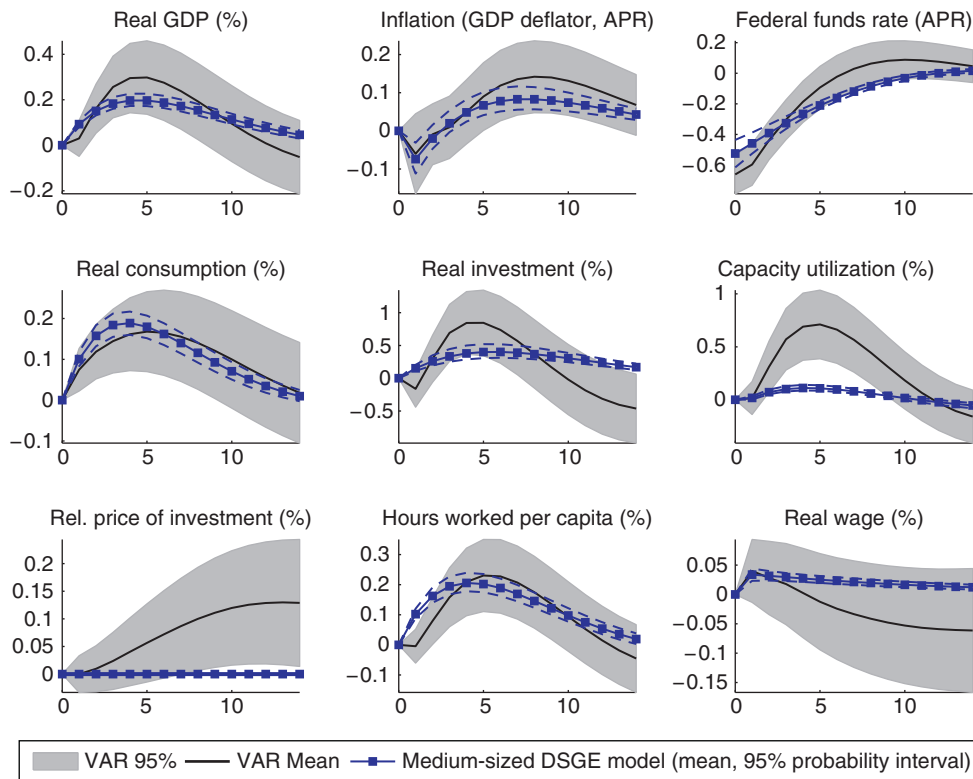


Figure 10 Dynamic responses of variables to a monetary policy shock.

the price puzzle is not statistically significant in our VAR estimation, it nevertheless deserves comment because it has potentially great economic significance. For example, the presence of a price puzzle in the data complicates the political problem associated with using high interest rates as a strategy to fight inflation. High interest rates and the consequent slowdown in economic growth are politically painful. If the public sees the high interest rate strategy producing higher inflation in the short run, support for the strategy may evaporate unless the price puzzle has been explained.⁵⁹

Regarding the delayed and gradual response of inflation to a monetary policy shock, note how inflation reaches a peak after two years. Of course, the wide confidence

⁵⁹ There is an important historical example of this political problem. In the early 1970s, at the start of the Great Inflation in the United States, Arthur Burns was chairman of the U.S. Federal Reserve and Wright Patman was chairman of the U.S. House Committee on Banking and Currency. Patman had the opinion that, by raising costs of production, high interest rates increase inflation. Patman's belief had enormous significance because he was influential in writing the wage and price control legislation at the time. He threatened Burns that if Burns tried to raise interest rates to fight inflation, Patman would see to it that interest rates were brought under the control of the wage-price control board (see "The Lasting, Multiple Hassles of Topic A," *Time Magazine*, Monday, April 9, 1973).

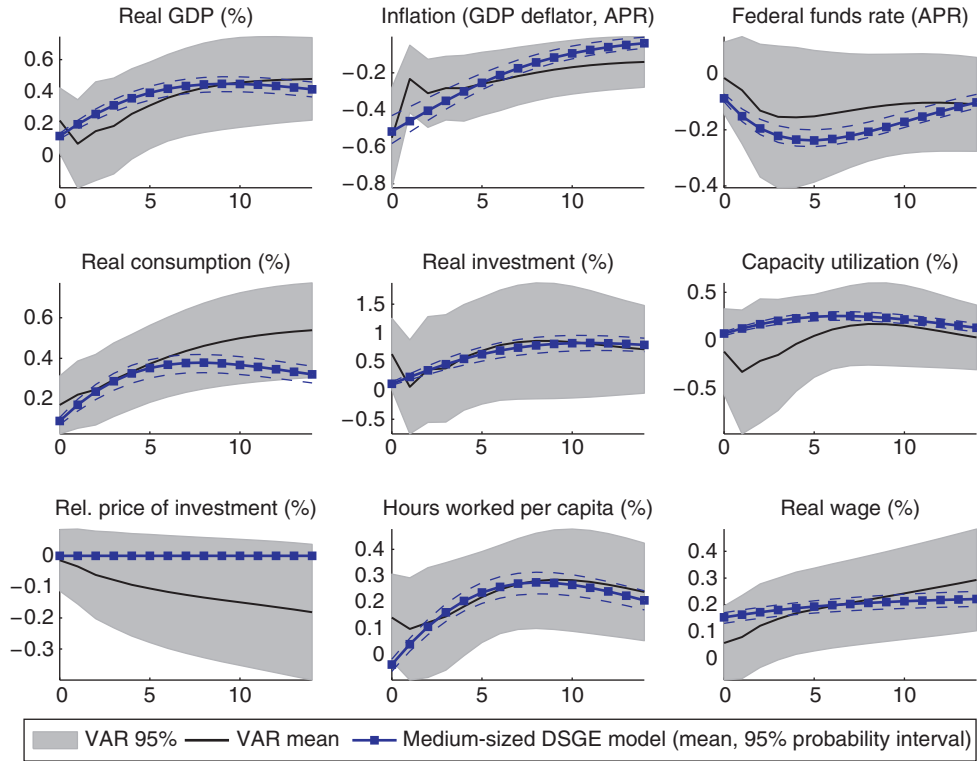


Figure 11 Dynamic responses of variables to a neutral technology shock.

intervals indicate that the exact timing of the peak is not precisely determined. However, the evidence does suggest a sluggish response of inflation. This is consistent with the views of others, arrived at by other methods, about the slow response of inflation to a monetary policy shock. As noted in the introduction to Section 4, it has been argued that the slow inflation response is a major puzzle for macroeconomics. For example, Mankiw (2000) argued that with price frictions of the type used here, the only way to explain the delayed and gradual response of inflation to a monetary policy shock is to introduce a degree of stickiness in prices that exceeds by far what can be justified based on the micro evidence. For this reason, when we study the ability of our models to match the estimated impulse response functions, we must be wary of the possibility that this is done only by making prices and wages counterfactually sticky. In addition, we must be wary of the possibility that the econometrics leans too hard on other features (such as variable capital utilization) to explain the gradual and delayed response of inflation to a monetary policy shock.

The second observation about the results in Figure 10 is that output, consumption, investment, and hours worked all display a slow, hump-shaped response to a monetary policy shock, peaking a little over one year after the shock. As emphasized in Section 4,

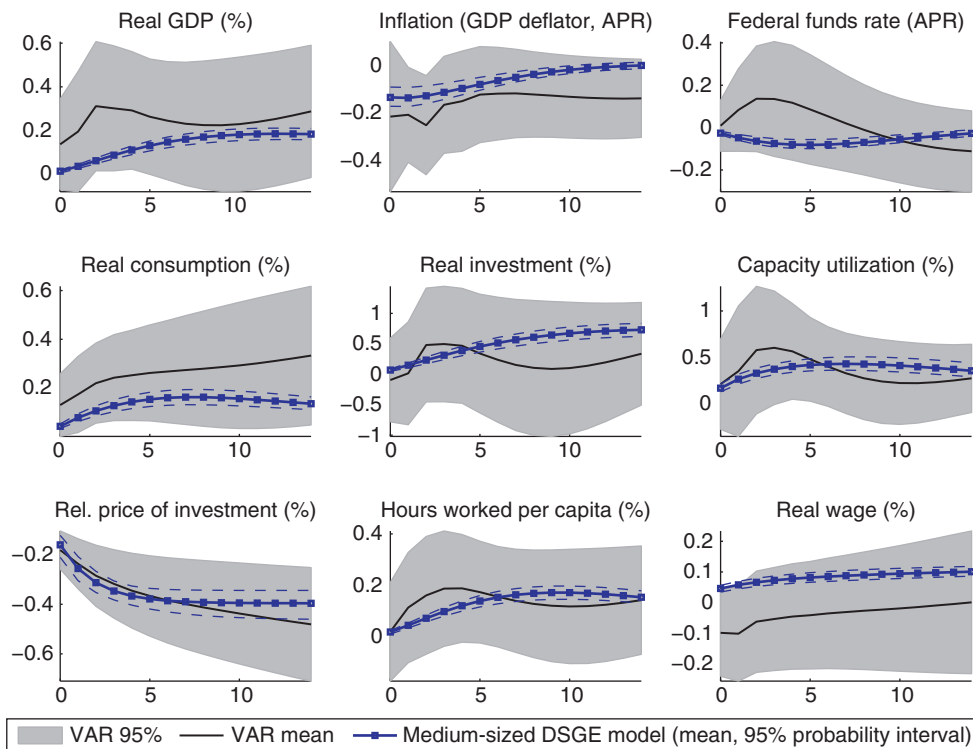


Figure 12 Dynamic responses of variables to an investment specific technology shock.

these hump-shaped observations are the reason that researchers introduce habit persistence and costs of adjustment in the flow of investment into the baseline model. Our third observation about the results in Figure 10 is that the effect of the monetary shock on the interest rate is roughly gone after two years, while the economy continues to respond well after that. This suggests that to understand the dynamic effects of a monetary policy shock, one must have a model that displays considerable sources of internal propagation.

A fourth observation concerns the response of capacity utilization. Recall from the discussion of Section 4 that the magnitude of the empirical response of this variable represents an important discipline on the analysis. In effect, those data constrain how heavily we can lean on variable capital utilization to explain the slow response of inflation to a monetary policy shock. The evidence in Figure 10 suggests that capacity utilization responds very sharply to a positive monetary policy shock. For example, utilization rises three times as much as employment, in percent terms. In interpreting this finding, we must bear in mind that the capital utilization numbers we have are for the manufacturing sector. To the extent that these data are influenced by the durable part of manufacturing, they may overstate the volatility of capacity utilization generally in the economy.

Our fifth observation about Figure 10 concerns the price of investment. In our model, this price is, by construction, unaffected by shocks other than those to the technology for converting homogeneous output into investment goods. Figure 10 indicates that the price of investment rises in response to an expansionary monetary policy shock, contrary to our model. This suggests that it would be worth exploring modifications to the technology for producing investment goods so that the trade-off between consumption and investment is nonlinear.⁶⁰ Under these conditions, the rise in the investment to consumption ratio that appears to occur in response to an expansionary monetary policy shock would be associated with an increase in the price of investment.

6.1.2 Technology shocks

Figures 11 and 12 display the responses to neutral and investment specific technology shocks, respectively. Overall, the confidence intervals are wide. The width of these confidence intervals should be no surprise in view of the nature of the question being addressed. The VAR is informed that there are two shocks in the data which have a long run effect on labor productivity, and it is being asked to determine the dynamic effects of these shocks on the data. To understand the challenge that such a question poses, imagine gazing at a data plot and thinking how the technology shocks might be detected visually. It is no wonder that in many cases, the VAR response is, “I don’t know how this variable responds.” This is what the wide confidence intervals tell us. For example, little can be said about the response of capacity utilization to a neutral technology shock.

Although confidence intervals are often wide, there are some responses that are significant. For example, there is a significant rise in consumption, output, and hours worked in response to a neutral shock. A particularly striking result in Figure 11 is the immediate drop in inflation in the wake of a positive shock to neutral technology. This drop has led some researchers to conjecture that the rapid response of inflation to a technology shock spells trouble for sticky price/sticky wage models. We investigate this conjecture in the next section.

6.2 Model results

6.2.1 Parameters

Parameters whose values are set a priori are listed in Table 2. We found that when we estimated the parameters κ_w and λ_w , the estimator drove them to their boundaries. This is why we simply set λ_w to a value near unity and we set $\kappa_w = 1$. The steady-state value of inflation (a parameter in the monetary policy rule and the price and wage updating equations), the steady-state government consumption to output ratio, and the steady-state growth rate of the investment specific technology were chosen to coincide with

⁶⁰ For example, instead of specifying a resource constraint in which $C_t + I_t$ appears, we could adopt one in which C_t and I_t appear in a CES function, i.e., $\left[a_1 C_t^{1/\rho} + a_2 I_t^{1/\rho} \right]^\rho$. The standard linear specification is a special case of this one, with $a_1 = a_2 = \rho = 1$.

Table 2 Non-Estimated Parameters in Medium-Sized DSGE Model

Parameter	Value	Description
α	0.25	Capital share
δ	0.025	Depreciation rate
β	0.999	Discount factor
π	1.0083	Gross inflation rate
η_g	0.2	Government consumption to GDP ratio
P_k'	1	Relative price of capital
κ_w	1	Wage indexation to π_{t-1}
λ_w	1.01	Wage markup
ξ_w	0.75	Wage stickiness
μ_z	1.0041	Gross neutral technology growth
μ_ψ	1.0018	Gross investment technology growth

their corresponding sample means in our data set.⁶¹ The growth rate of neutral technology was chosen so that, conditional on the growth rate of investment specific technology, the steady-state growth rate of output in the model coincides with the corresponding sample average in the data. We set $\xi_w = 0.75$, so that the model implies wages are reoptimized once a year on average. We did not estimate this parameter because we found that it is difficult to separately identify the value of ξ_w and the curvature parameter of household labor disutility, ϕ .

The parameters for which we report priors and posteriors are listed in Table 3. Note first that the degree of price stickiness, ξ_p , is modest. The time between price reoptimizations implied by the posterior mean of this parameter is a little less than 3 quarters. The amount of information in the likelihood, Eq. (86), about the value of ξ_p is substantial. The posterior standard deviation is roughly one-third the size of the prior standard deviation and the posterior 95% probability interval is a quarter of the width of the corresponding prior probability interval. Generally, the amount of information in the likelihood about all the parameters is large in this sense. An exception to this pattern is the coefficient on inflation in the Taylor rule, r_π . There appears to be relatively little information about this parameter in the likelihood. Note that ϕ is estimated to be quite small, implying a consumption-compensated labor supply elasticity for the household of around 8. Such a high elasticity would be regarded as empirically implausible if

⁶¹ In our model, the relative price of investment goods represents a direct observation of the technology shock for producing investment goods.

Table 3 Prior and Posteriors of Parameters for Medium-Sized DSGE Model

Parameter		Prior		Posterior ^a
		Distribution [bounds]	Mean, std. dev. [5% and 95%]	Mean, std. dev. [5% and 95%]
Price-setting parameters				
Price stickiness	ξ_p	Beta [0, 0.8]	0.50, 0.15 [0.23, 0.72]	0.62, 0.04 [0.56, 0.68]
Price markup	λ_f	Gamma [1.01, ∞]	1.20, 0.15 [1.04, 1.50]	1.20, 0.08 [1.06, 1.32]
Monetary authority parameters				
Taylor rule: Interest smoothing	ρ_R	Beta [0, 1]	0.80, 0.10 [0.62, 0.94]	0.87, 0.02 [0.85, 0.90]
Taylor rule: Inflation coefficient	r_π	Gamma [1.01, 4]	1.60, 0.15 [1.38, 1.87]	1.43, 0.11 [1.25, 1.59]
Taylor rule: GDP coefficient	r_y	Gamma [0, 2]	0.20, 0.15 [0.03, 0.49]	0.07, 0.03 [0.02, 0.11]
Household parameters				
Consumption habit	b	Beta [0, 1]	0.75, 0.15 [0.47, 0.95]	0.77, 0.02 [0.74, 0.80]
Inverse labor supply elasticity	ϕ	Gamma [0, ∞]	0.30, 0.20 [0.06, 0.69]	0.12, 0.03 [0.08, 0.16]
Capacity adjustment costs curv.	σ_a	Gamma [0, ∞]	1.00, 0.75 [0.15, 2.46]	0.30, 0.08 [0.16, 0.44]
Investment adjustment costs curv.	S''	Gamma [0, ∞]	12.00, 8.00 [2.45, 27.43]	14.30, 2.92 [9.65, 18.8]
Shocks				
Autocorr. investment technology	ρ_ψ	Uniform [0, 1]	0.50, 0.29 [0.05, 0.95]	0.60, 0.08 [0.48, 0.72]
Std. dev. neutral tech. shock (%)	σ_z	Inv. Gamma [0, ∞]	0.20, 0.10 [0.10, 0.37]	0.22, 0.02 [0.19, 0.25]
Std. dev. invest. tech. shock (%)	σ_ψ	Inv. Gamma [0, ∞]	0.20, 0.10 [0.10, 0.37]	0.16, 0.02 [0.12, 0.20]
Std. dev. monetary shock (APR)	σ_R	Inv. Gamma [0, ∞]	0.40, 0.20 [0.21, 0.74]	0.51, 0.05 [0.44, 0.58]

^aBased on standard random walk metropolis algorithm. 600 000 draws, 100 000 for burn-in, acceptance rate 27%.

Table 4 Medium-Sized DSGE Model Steady State at Posterior Mean for Parameters

Variable	Standard model	Description
k/γ	7.73	Capital to GDP ratio (quarterly)
c/γ	0.56	Consumption to GDP ratio
i/γ	0.24	Investment to GDP ratio
H	0.63	Steady-state labor input
R	1.014	Gross nominal interest rate (quarterly)
R^{real}	1.006	Gross real interest rate (quarterly)
r^k	0.033	Capital rental rate (quarterly)
A_L	2.22	Slope, labor disutility

it were interpreted as the elasticity of supply of hours by a representative agent. However, as discussed in Section 2.3, this is not our interpretation. Table 4 reports steady-state properties of the model, evaluated at the posterior mean of the parameters.

6.2.2 Impulse responses

We now comment on the DSGE model impulse responses displayed in Figures 10–12. The line with solid squares in the figures display the impulse responses of our model, at the posterior mean of the parameters. The dashed lines display the 95% probability interval for the impulse responses implied by the posterior distribution of the parameters. These intervals are in all cases reasonably tight, reflecting the tight posterior distribution on the parameters as well as the natural restrictions of the model.

Our estimation strategy in effect selects a model parameterization that places the model-implied impulse response functions as close as possible to the center of the gray area, while not suffering too much of a penalty from the priors. The estimation criterion is less concerned about reproducing VAR-based impulse response functions where the gray areas are the widest.

Consider Figure 10, which displays the response of standard macroeconomic variables to a monetary policy shock. Note how well the model captures the delayed and gradual response of inflation. In the model it takes two years for inflation to reach its peak response after the monetary policy shock. Importantly, the model even captures the price puzzle phenomenon, according to which inflation moves in the “wrong” direction initially. This apparently perverse initial response of inflation is interpreted by the model as reflecting the reduction in labor costs associated with the cut in the nominal rate of interest. The notable result here is that the slow response of inflation to a monetary policy shock is explained with a modest degree of wage and price-setting frictions. In addition, the gradual and delayed response of inflation is not due to an excessive or counterfactual

increase in capital utilization. Indeed, the model substantially understates the rise in capital utilization. While on its own this is a failure of the model, the weak utilization response does draw attention to the apparent ease with which the model is able to capture the inertial response of inflation to a monetary shock.

The model also captures the response of output and consumption to a monetary policy shock reasonably well. However, the model apparently does not have the flexibility to capture the relatively sharp fall and rise in the investment response, although the model responses lie inside the gray area. The relatively large estimate of the curvature in the investment adjustment cost function, S'' , suggests that to allow a greater response of investment to a monetary policy shock would cause the model's prediction of investment to lie outside the gray area in the first couple of quarters. These findings for monetary policy shocks are broadly similar to those reported in CEE and ACEL.

Figure 11 displays the response of standard macroeconomic variables to a neutral technology shock. Note that the model is reasonably successful at reproducing the empirically estimated responses. The dynamic response of inflation is particularly notable, in light of the estimation results reported in ACEL. Those results suggest that the sharp and precisely estimated drop in inflation in response to a neutral technology shock is difficult to reproduce in a model like ours. In describing this problem for their model, ACEL expressed a concern that the failure reflects a deeper problem with sticky price models.⁶² They suggested that perhaps the emphasis on price- and wage-setting frictions, largely motivated by the inertial response of inflation to a monetary shock, is shown to be misguided by the evidence that inflation responds rapidly to technology shocks.⁶³ Our results suggest a far more mundane possibility.

There are two key differences between our model and the one in ACEL that allow it to reproduce the response of inflation to a technology shock more or less exactly without hampering its ability to account for the slow response of inflation to a monetary policy shock. First, in our model there is no indexation of prices to lagged inflation (see Eq. 63). ACEL follows CEE in supposing that when firms cannot optimize their price, they index it fully to lagged aggregate inflation. The position of our model on price indexation is a key reason why we can account for the rapid fall in inflation after a neutral technical shock while ACEL cannot. We suspect that our way of treating indexation is a step in the right direction from the point of view of the microeconomic data. Micro observations suggest that individual prices do not change for extended periods of time. A second distinction between our model and the one in ACEL is that we

⁶² See Paciello (2009) for another discussion of this point.

⁶³ The concern is reinforced by the fact that an alternative approach, one based on information imperfections and minimal price/wage-setting frictions, seems like a natural one for explaining the puzzle of the slow response of inflation to monetary policy shocks and the quick response to technology shocks (see Maćkowiak and Wiederholt, 2009; Mendes, 2009; and Paciello, 2009). Dupor, Han, and Tsai (2009) suggested more modest changes in the model structure to accommodate the inflation puzzle.

specify the neutral technology shock to be a random walk (see Eq. 60), while in ACEL the growth rate of the estimated technology shock is highly autocorrelated. In ACEL, a technology shock triggers a strong wealth effect, which stimulates a surge in demand that places upward pressure on marginal cost and thus inflation.

Figure 12 displays dynamic responses of macroeconomic variables to an investment specific shock. The DSGE model fits the dynamics implied by the VAR well, although the confidence intervals are large.

6.3 Assessing VAR robustness and accuracy of the Laplace approximation

It is well known that when the start date or number of lags for a VAR are changed, the estimated impulse response functions change. In practice, one hopes that the width of probability intervals reported in the analysis is a reasonable rule-of-thumb guide to the degree of nonrobustness. In Figures 13–15 we display all the estimated impulse response functions from our VAR when we apply a range of different start dates and lag lengths. The VAR point estimates used in our estimation exercise are displayed

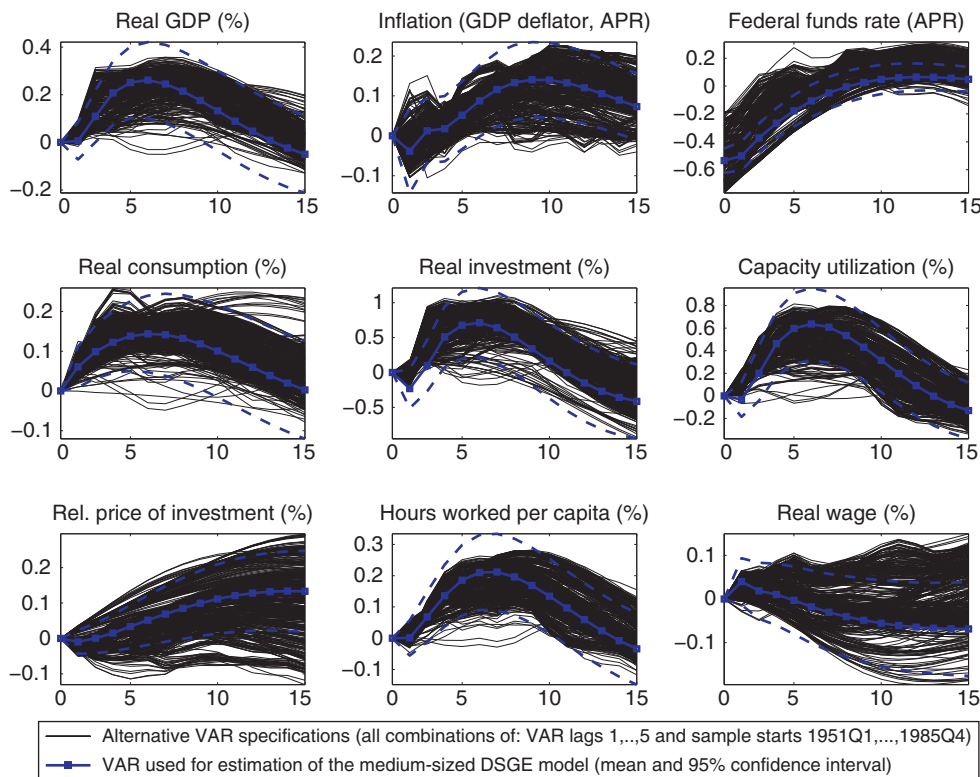


Figure 13 VAR specification sensitivity: Response to a monetary policy shock.

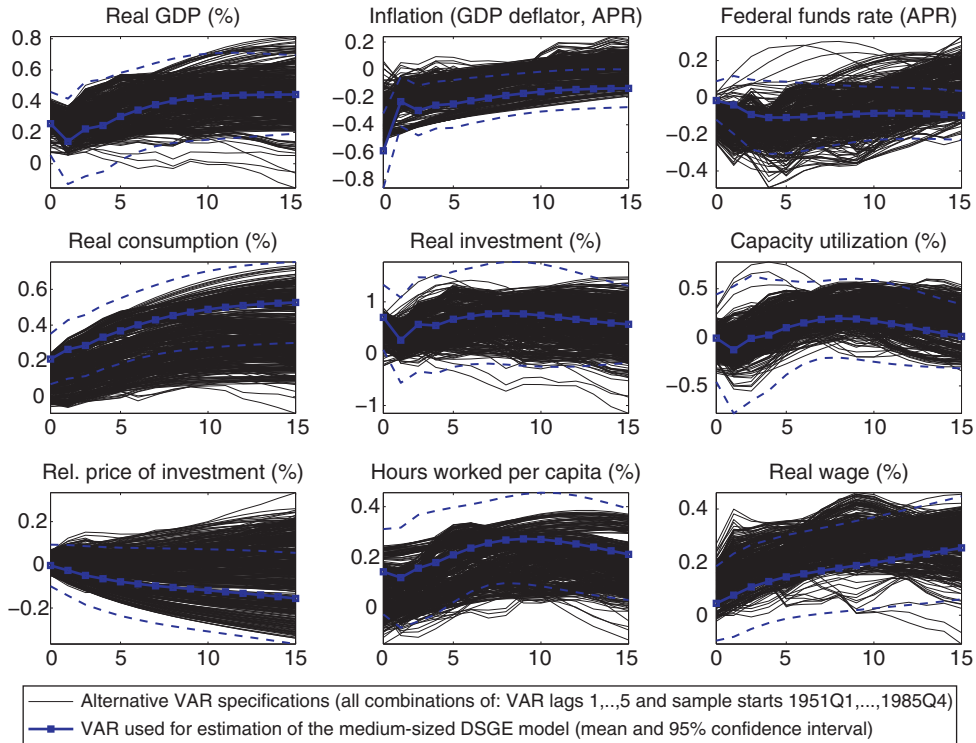


Figure 14 VAR specification sensitivity: Neutral technology shock.

in Figures 13–15 in the form of the solid line with solid squares. The 95% probability intervals associated with the impulse response functions used in our estimation exercise are indicated by the dashed lines. According to the figures, the degree of variation across different samples and lag lengths corresponds roughly to the width of probability intervals. Although results do change across the perturbed VARs, the magnitude of the changes are roughly what is predicted by the rule-of-thumb. In this sense, the degree of nonrobustness in the VAR is not great.

Finally, Figure 16 displays the priors and posteriors of the model parameters. The posteriors are computed by two methods: the random walk Metropolis method, and the Laplace approximation described in Section 5.4. It is interesting that the Laplace approximation and the results of the random walk Metropolis algorithm are very similar. These results suggest that one can save substantial amounts of time by computing the Laplace approximation during the early and intermediate phases of a research project. At the end of the project, when it is time to produce the final draft of the manuscript, one can then perform the time-intensive random walk Metropolis calculations.

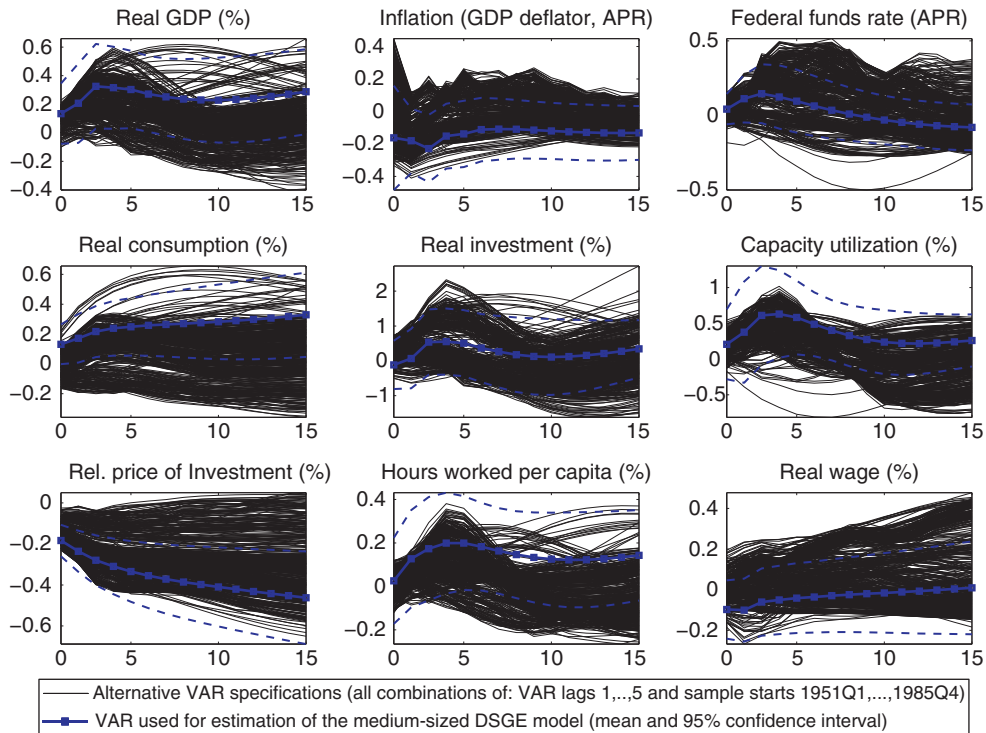


Figure 15 VAR specification sensitivity: Investment specific technology shock.

7. CONCLUSION

The literature on DSGE models for monetary policy is too large to review in all its detail in this chapter. Necessarily, we have been forced to focus on only a part. Relatively little space has been devoted to the limitations of monetary DSGE models. A key challenge is posed by the famous statistical rejections of the intertemporal Euler equation that lies at the heart of DSGE models (see, e.g., Hansen & Singleton, 1983). These rejections of the “IS equation” in the New Keynesian model pose a challenge for that model’s account of the way shocks propagate through the economy. At the same time, the Bayesian impulse response matching technique that we apply suggests that the New Keynesian model is able to capture the basic features of the transmission of three important shocks.⁶⁴ An outstanding question is how to resolve these apparently conflicting pieces of information.

Also, we have been able to do little in the way of reviewing the new frontiers for monetary DSGE models. The recent financial turmoil has accelerated work to

⁶⁴ In our empirical analysis we have not reported our VAR’s implications for the importance of the three shocks that we analyzed. However, ACEL documents that these shocks together account for well over 50% of the variation of macroeconomic time series like output, investment and employment.

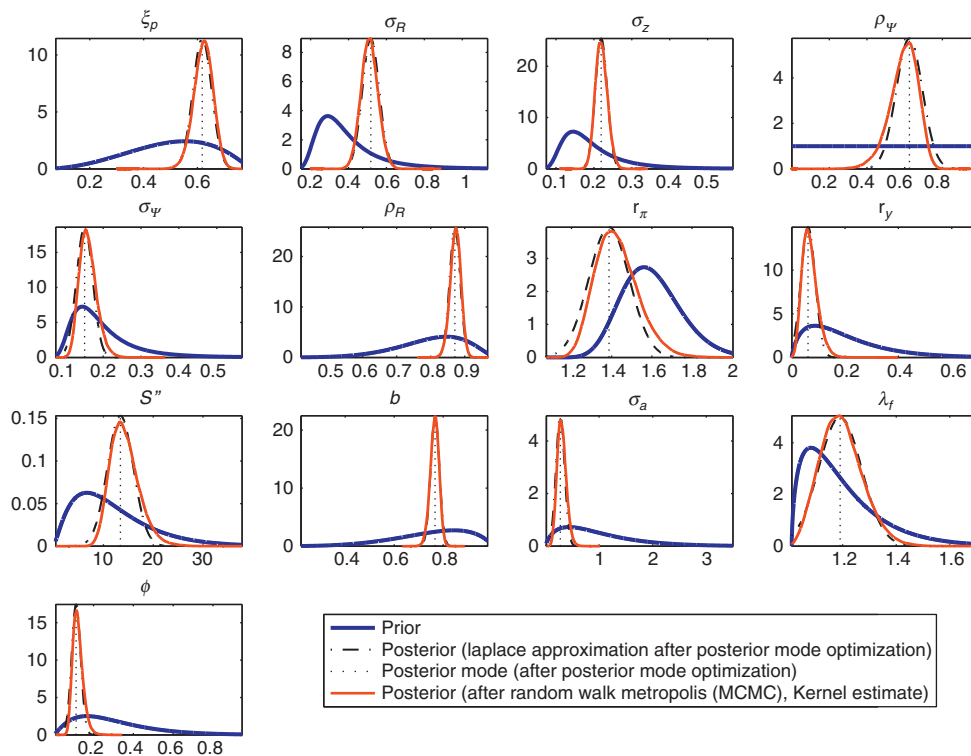


Figure 16 Priors and posteriors of estimated parameters of the medium-sized DSGE model.

introduce a richer financial sector into the New Keynesian model. With this addition, the model is able to address important policy questions that cannot be addressed by the models described here. How should monetary policy respond to an increase in interest rate spreads? How should we think about the recent “unconventional monetary policy” actions, in which the monetary authority purchases privately issued liabilities such as mortgages and commercial paper? The models described in this chapter are silent on these questions. However, an exploding literature too large to review here has begun to introduce the modifications necessary to address them.⁶⁵ The labor market is another frontier of new model development. We have presented a rough sketch of the approach in CTW, but the literature merging the best of labor market research with monetary DSGE models is too large to survey here.⁶⁶ Still, these new developments

⁶⁵ For a small sampling, see, for example, Bernanke, Gertler, and Gilchrist (1999); Christiano, Motto, and Rostagno (2003, 2009); Cúrdia and Woodford (2009); and Gertler and Kiyotaki (2010).

⁶⁶ A small open economy model with financial and labor market frictions, estimated by full information Bayesian methods, appears in Christiano, Trabandt, and Walentin (2010c). Important other papers on the integration of unemployment and other labor market frictions into monetary DSGE models include Gali (2010a); Gertler, Sala, and Trigari (2008); and Thomas (2009).

ensure that monetary DSGE models will remain an active and exciting area of research for the foreseeable future.

REFERENCES

- Abel, A.B., Bernanke, B., 2005. *Macroeconomics*, fifth ed. Pearson Addison Wesley, Upper Saddle River, NJ.
- Adalid, R., Detken, C., 2007. European Central Bank, Working Paper No. 732, Liquidity shocks and asset price boom/bust cycles. (February).
- Altig, D., Christiano, L.J., Eichenbaum, M., Lindé, J., 2005. Firm-specific capital, nominal rigidities and the business cycle. NBER, Working Paper 11034.
- Ball, L., 1994. Credible disinflation with staggered price setting. *Am. Econ. Rev.* 84 (March), 282–289.
- Barth III, M.J., Ramey, V.A., 2002. The cost channel of monetary transmission. In: Bernanke, B., Rogoff, K.S. (Eds.), *NBER chapters, NBER macroeconomics annual 2001*. 16, MIT Press, Cambridge, MA, pp. 199–256.
- Basu, S., 1995. Intermediate goods and business cycles: Implications for productivity and welfare. *Am. Econ. Rev.* 85 (3), 512–531.
- Benhabib, J., Schmitt–Grohé, S., Uribe, M., 2002. Chaotic interest-rate rules. *Am. Econ. Rev.* 92 (2), 72–78.
- Bernanke, B., Boivin, J., Elias, P.S., 2005. Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *Q. J. Econ.* 120 (1), 387–422.
- Bernanke, B., Gertler, M., 2000. Monetary policy and asset price volatility. NBER, Working Paper No. 7559.
- Bernanke, B., Gertler, M., Gilchrist, S., 1999. The financial accelerator in a quantitative business cycle framework. In: Taylor, J.B., Woodford, M. (Eds.), *Handbook of macroeconomics*. Elsevier Science, North-Holland, Amsterdam, pp. 1341–1393.
- Boldrin, M., Christiano, L.J., Fisher, J.D.M., 2001. Habit persistence, asset returns, and the business cycle. *Am. Econ. Rev.* 91 (1), 149–166.
- Bruckner, M., Schabert, A., 2003. Supply-side effects of monetary policy and equilibrium multiplicity. *Econ. Lett.* 79 (2), 205–211.
- Bullard, J., Mitra, K., 2002. Learning about Monetary policy rules. *J. Monetary Econ.* 49, 1105–1129.
- Calvo, G.A., 1983. Staggered prices in a utility-maximizing framework. *J. Monetary Econ.* 12 (3), 383–398.
- Carroll, C.D., Overland, J., Weil, D.N., 1997. Comparison utility in a growth model. *J. Econ. Growth* 2 (4), 339–367.
- Carroll, C.D., Overland, J., Weil, D.N., 2000. Saving and growth with habit formation. *Am. Econ. Rev.* 90 (3), 341–355.
- Chernozhukov, V., Hong, H., 2003. An MCMC approach to classical estimation. *J. Econom.* 115 (2), 293–346.
- Chowdhury, I., Hoffmann, M., Schabert, A., 2006. Inflation dynamics and the cost channel of monetary transmission. *Eur. Econ. Rev.* 50 (4), 995–1016.
- Christiano, L.J., 1988. Why does inventory investment fluctuate so much? *J. Monetary Econ.* 21, 247–280.
- Christiano, L.J., 1991. Modeling the liquidity effect of a money shocks. *Federal Reserve Bank of Minneapolis Quarterly Review* 15 (1), 1–34.
- Christiano, L.J., 2004. Firm-specific capital and aggregate inflation dynamics in Woodford's model. Manuscript.
- Christiano, L.J., 2007. Discussion of Del Negro, Schorfheide, Smets and Wouters. *J. Bus. Econ. Stat.* 25 (2), 143–151.
- Christiano, L.J., Eichenbaum, M., 1992a. Current real business cycle theories and aggregate labor market fluctuations. *Am. Econ. Rev.* 82 (3), 430–450.
- Christiano, L.J., Eichenbaum, M., 1992b. Liquidity effects and the monetary transmission mechanism. *American Economic Review, Papers and Proceedings* 82, 346–353.

- Christiano, L.J., Eichenbaum, M., Evans, C.L., 1996, The effects of monetary policy shocks: evidence from the flow of funds, *The Review of Economics and Statistics*, Vol. 78, No. 1 (February), pp. 16–34.
- Christiano, L.J., Eichenbaum, M., Evans, C.L., 1999. Monetary policy shocks: What have we learned and to what end? In: Taylor, J.B., Woodford, M. (Eds.), *Handbook of macroeconomics*. Elsevier Science, North-Holland, Amsterdam, pp. 65–148.
- Christiano, L.J., Eichenbaum, M., Evans, C.L., 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. *J. Polit. Econ.* 113 (1), 1–45.
- Christiano, L.J., Gust, C., 2000. The expectations trap hypothesis. Federal Reserve Bank of Chicago Economic Perspectives 24, 21–39.
- Christiano, L.J., Ilut, C., Motto, R., Rostagno, M., 2008. Monetary policy and stock market boom bust cycles. European Central Bank working paper number 955.
- Christiano, L.J., Ilut, C., Motto, R., Rostagno, M., 2010, Monetary Policy and Stock Market Booms, paper presented to conference sponsored by Federal Reserve Bank of Kansas City, ‘Macroeconomic Challenges: the Decade Ahead,’ at Jackson Hole, Wyoming, August 26–28.
- Christiano, L.J., Motto, R., Rostagno, M., 2003. The great depression and the Friedman-Schwartz hypothesis. *J. Money Credit Bank.* 35 (6) December, Part 2, 1119–1197.
- Christiano, L.J., Motto, R., Rostagno, M., 2009. Financial factors in business cycles. Manuscript.
- Christiano, L.J., Rostagno, M., 2001. Money growth monitoring and the Taylor rule. NBER, Working Paper, No. 8539.
- Christiano, L.J., Trabandt, M., Walentin, K., 2010a. Involuntary unemployment and the business cycle. NBER, Working Paper No. 15801.
- Christiano, L.J., Trabandt, M., Walentin, K., 2010b. A monetary business cycle model with labor market frictions. Northwestern University Manuscript.
- Christiano, L.J., Trabandt, M., Walentin, K., 2010c. Introducing financial frictions and unemployment into a small open economy model. Sveriges Riksbank, Working Paper No. 214.
- Clarida, R., Gali, J., Gertler, M., 1999. The science of monetary policy: A New Keynesian perspective. *J. Econ. Lit.* 37, 1661–1707.
- Cochrane, J.H., 2009. Can learnability save New-Keynesian models? *J. Monetary Econ.* 56, 1109–1113.
- Constantinides, G.M., 1990. Habit formation: A resolution of the equity premium puzzle. *J. Polit. Econ.* 98 (3), 519–543.
- Cúrdia, V., Woodford, M., 2009. Credit spreads and monetary policy. NBER, Working Paper No. 15289.
- de Walque, G., Smets, F., Wouters, R., 2006. Firm-specific production factors in a DSGE model with Taylor price setting. *International Journal of Central Banking* 2 (3), 107–154.
- Dupor, B., Han, J., Tsai, Y.C., 2009. What do technology shocks tell us about the New Keynesian paradigm? *J. Monetary Econ.* 56 (4), 560–569.
- Erceg, C.J., Henderson, D.W., Levin, A.T., 2000. Optimal monetary policy with staggered wage and price contracts. *J. Monetary Econ.* 46, 281–313.
- Fernald, J., 2009. A quarterly, utilization-adjusted series on total factor productivity. Federal Reserve Bank of San Francisco Manuscript.
- Fisher, J., 2006. The dynamic effects of neutral and investment-specific technology shocks. *J. Polit. Econ.* 114 (3), 413–451.
- Fuerst, T.S., 1992. Liquidity, loanable funds, and real activity. *J. Monetary Econ.* 29 (1), 3–24.
- Fuhrer, J.C., 2000. Habit formation in consumption and its implications for monetary policy models. *Am. Econ. Rev.* 90 (3), 367–390.
- Gali, J., 1999. Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? *Am. Econ. Rev.* 89 (1), 249–271.
- Gali, J., 2010a. The return of the wage Phillips curve. CREI Manuscript.
- Gali, J., 2010b. Unemployment fluctuations and stabilization policies: A New Keynesian perspective. Department of Economics, University of Copenhagen Zeuthen lectures delivered March 19–20 to Zeuthen Workshop on Macroeconomics.
- Gertler, M., Kiyotaki, N., 2010. Financial intermediation and credit policy in business cycle analysis. Manuscript.

- Gertler, M., Sala, L., Trigari, A., 2008. An estimated monetary DSGE model with unemployment and staggered nominal wage bargaining. *J. Money Credit Bank.* 40 (8), 1713–1764.
- Guerron-Quintana, P., 2008. Refinements on macroeconomic modeling: The role of non-separability and heterogeneous labor supply. *J. Econ. Dyn. Control* 32, 3613–3630.
- Hamilton, J., 1994. Time series analysis. Princeton University Press, Princeton, NJ.
- Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029–1054.
- Hansen, L.P., Singleton, K.J., 1983. Stochastic consumption, risk aversion, and the temporal behavior of asset returns. *J. Polit. Econ.* 91 (2), 249–265.
- Hansen, G.D., 1985. Indivisible labor and the business cycle. *J. Monetary Econ.* 16 (3), 309–327.
- Hodrick, R.J., Prescott, E.C., 1997. Postwar U.S. business cycles: An empirical investigation. *J. Money Credit Bank.* 29 (1), 1–16 February.
- Justiniano, A., Primiceri, G., 2008. Potential and natural output. Northwestern University Manuscript.
- Kiley, M.T., 2010. Output gaps. Federal Reserve Board of Governors Finance and Economics Discussion Series, Working Paper 2010–27.
- Kim, J.Y., 2002. Limited information likelihood and Bayesian analysis. *J. Econom.* 107, 175–193.
- King, R.G., Rebelo, S., 1993. Low frequency filtering and real business cycles. *J. Econ. Dyn. Control* 17 (1–2), 207–231.
- Krusell, P., Mukoyama, T., Rogerson, R., Sahin, A., 2008. Aggregate implications of indivisible labor, incomplete markets and labor market frictions. NBER, Working Paper No. 13871.
- Kwan, Y.K., 1999. Asymptotic Bayesian analysis based on a limited information estimator. *J. Econom.* 88, 99–121.
- Levin, A.T., Onatski, A., Williams, J.C., Williams, N., 2005. Monetary policy under uncertainty in micro-founded macroeconomic models. *NBER Macroeconomics Annual* 20, 229–287.
- Lucas Jr., R.E., Prescott, E.C., 1971. Investment under uncertainty. *Econometrica* 39 (5), 659–681.
- Lucca, D.O., 2006. Essays in investment and macroeconomics. Northwestern University, Department of Economics, Ph.D. dissertation.
- Maćkowiak, B., Wiederholt, M., 2009. Optimal sticky prices under rational inattention. *Am. Econ. Rev.* 99 (3), 769–803.
- Mankiw, N.G., 2000. The inexorable and mysterious tradeoff between inflation and unemployment. NBER, Working Paper 7884.
- Matsuyama, K., 1984. A learning effect model of investment: An alternative interpretation of Tobin's q. Northwestern University, Manuscript.
- McCallum, B.T., 2009. Inflation determination with Taylor rules: Is New-Keynesian analysis critically flawed? *J. Monetary Econ.* 56, 1101–1108.
- Mendes, R., 2009. Information, expectations, and the business cycle. Research Department, Bank of Canada Manuscript.
- Mulligan, C.B., 2001. Aggregate implications of indivisible labor. *Advances in Macroeconomics* 1 (1) Article 4, Berkeley Electronic Press, pp. 1–33.
- Newey, W.K., West, K.D., 1987. A simple, positive semi-definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55 (3), 703–708.
- Paciello, L., 2009. Does inflation adjust faster to aggregate technology shocks than to monetary policy shocks?. Einaudi Institute for Economics and Finance, Working Paper.
- Prescott, E.C., 1986. Theory ahead of business-cycle measurement. *Carnegie-Rochester Conference Series on Public Policy* 25 (1), 11–44.
- Prescott, E.C., 1998. Needed: A theory of total factor productivity. *Int. Econ. Rev.* 39 (3), 525–551.
- Ravenna, F., Walsh, C.E., 2006. Optimal monetary policy with the cost channel. *J. Monetary Econ.* 53, 199–216.
- Rogerson, R., 1988. Indivisible labor, lotteries and equilibrium. *J. Monetary Econ.* 21 (1), 3–16.
- Rogerson, R., Wallenius, J., 2009. Micro and macro elasticities in a life cycle model with taxes. *J. Econ. Theory* 144, 2277–2292.

- Rotemberg, J., Woodford, M., 1997. An optimization-based econometric framework for the evaluation of monetary policy. In: Bernanke, B., Rotemberg, J. (Eds.), *NBER macroeconomics annual*. MIT Press, Cambridge, MA.
- Sala, L., Söderström, U., Trigari, A., 2008. Monetary policy under uncertainty in an estimated model with labor market frictions. *J. Monetary Econ.* 55 (5), 983–1006.
- Sargent, T., 1979. *Macroeconomic theory*. Academic Press, New York.
- Sims, C.A., 1992. Interpreting the macroeconomic time series facts: The effects of monetary policy. *Eur. Econ. Rev.* 36 (5), 975–1000.
- Sims, C.A., Zha, T., 2006. Were there regime switches in U.S. monetary policy? *Am. Econ. Rev.* 96 (1), 54–81.
- Smets, F., Wouters, R., 2007. Shocks and frictions in US business cycles. *Am. Econ. Rev.* 97 (3), 586–606.
- Sveen, T., Weinke, L., 2005. New perspectives on capital, sticky prices, and the Taylor principle. *J. Econ. Theory* 123 (1), 21–39.
- Thomas, C., 2009. Search frictions, real rigidities and inflation dynamics. Bank of Spain Manuscript.
- Topel, R., Rosen, S., 1988. Housing investment in the United States. *J. Polit. Econ.* 96 (4), 718–740.
- Walsh, C., 2005. Discussion of Levin, Onatski, Williams and Williams. 20, MIT Press, Cambridge, MA *NBER macroeconomics annual*.
- Woodford, M., 2003. *Interest and prices: Foundations of a theory of monetary policy*. Princeton University Press, Princeton, NJ.
- Woodford, M., 2004. Inflation and output dynamics with firm-specific investment. Manuscript.
- Yun, T., 1996. Nominal price rigidity, money supply endogeneity, and business cycles. *J. Monetary Econ.* 37 (2), 345–370.

This page intentionally left blank

How Has the Monetary Transmission Mechanism Evolved Over Time?★

Jean Boivin*, Michael T. Kiley** and Frederic S. Mishkin†

*Bank of Canada, HEC Montréal and National Bureau of Economic Research

**Board of Governors of the Federal Reserve System

†Graduate School of Business, Columbia University and National Bureau of Economic Research

Contents

1. Introduction	370
2. The Channels of Monetary Transmission	374
2.1 Neoclassical channels	374
2.2 Investment-based channels	376
2.2.1 <i>Direct interest-rate channels</i>	376
2.2.2 <i>Tobin's q</i>	377
2.2.3 <i>Previous empirical literature on investment-based channels</i>	377
2.3 Consumption-based channels	379
2.3.1 <i>Wealth effects</i>	379
2.3.2 <i>Intertemporal substitution effects</i>	379
2.3.3 <i>Previous empirical literature on consumption-based channels</i>	379
2.4 International-trade based channels	380
2.4.1 <i>Exchange rate channel</i>	380
2.5 Non-neoclassical channels: The credit view	380
2.5.1 <i>Effects on credit supply from government interventions in credit markets</i>	381
2.5.2 <i>Bank-based channels</i>	382
2.5.3 <i>Balance sheet channel</i>	383
3. Why the Monetary Transmission Mechanism may have Changed	385
3.1 Institutional changes in credit markets	385
3.2 Changes in the Way Expectations Are Formed	388
4. Has the Effect of Monetary Policy on the Economy Changed? Aggregate Evidence	388
4.1 Modeling the monetary transmission mechanism	389
4.2 Existing evidence	391
4.3 New Evidence	392
4.3.1 <i>New FAVAR-based evidence</i>	393
4.3.2 <i>Comparison with the VAR approach</i>	393
4.3.3 <i>Multidimensional effects of monetary policy</i>	396

★ We thank Dalibor Stevanovic and Dane Vrabac for excellent research assistance. Our analysis benefited substantially from the comments of Ray Fair, Ben Friedman, and participants in the *Key Developments in Monetary Economics* conference held at the Federal Reserve Board in October 2009. Jean Boivin acknowledges financial support from the National Science Foundation (SES-0518770) and the Social Sciences and Humanities Research Council of Canada. The views expressed are those of the authors and do not reflect the views of any institutions with which they are affiliated.

5. What Caused the Monetary Transmission Mechanism to Evolve?	396
5.1 FAVAR-based evidence	397
5.1.1 <i>The expectation channel</i>	397
5.1.2 <i>The balance sheet channel</i>	398
5.2 Evidence from a completely specified structural model	399
5.2.1 <i>The model</i>	400
5.2.2 <i>Changes in monetary transmission</i>	405
5.2.3 <i>Signs of changing credit conditions</i>	411
5.2.4 <i>Monetary policy and the transmission of other shocks</i>	412
6. Implications for the Future Conduct of Monetary Policy	415
Appendix	416
Estimation of the DSGE Model	416
References	418

Abstract

We discuss the evolution in macroeconomic thought on the monetary policy transmission mechanism and present related empirical evidence. The core channels of policy transmission — the neoclassical links between short-term policy interest rates, other asset prices such as long-term interest rates, equity prices, and the exchange rate, and the consequent effects on household and business demand — have remained steady from early policy-oriented models (like the Penn-MIT-SSRC MPS model) to modern dynamic, stochastic general equilibrium (DSGE) models. In contrast, non-neoclassical channels, such as credit-based channels, have remained outside the core models. In conjunction with this evolution in theory and modeling, there have been notable changes in policy behavior (with policy more focused on price stability) and in the reduced form correlations of policy interest rates with activity in the United States. Regulatory effects on credit provision have also changed significantly. As a result, we review the empirical evidence on the changes in the effect of monetary policy actions on real activity and inflation and present new evidence, using both a relatively unrestricted factor-augmented vector autoregression (FAVAR) and a DSGE model. Both approaches yield similar results: Monetary policy innovations have a more muted effect on real activity and inflation in recent decades as compared to the effects before 1980. Our analysis suggests that these shifts are accounted for by changes in policy behavior and the effect of these changes on expectations, leaving little role for changes in underlying private-sector behavior (outside shifts related to monetary policy changes).
JEL classification: E5, E4, E2, E3

Keywords

Monetary Policy
 Monetary Transmission
 Stability

1. INTRODUCTION

The monetary transmission mechanism is one of the most studied areas of monetary economics for two reasons. First, understanding how monetary policy affects the economy is essential to evaluating what the stance of monetary policy is at a particular point in time. Even

if a central bank's policy instrument, for example, the federal funds rate in the United States, is low, monetary policy may well be restrictive because of effects that monetary policy has had on other asset prices and quantities. Second, in order to decide on how to set policy instruments, monetary policymakers must have an accurate assessment of the timing and effect of their policies on the economy. To make this assessment, they need to understand the mechanisms through which monetary policy impacts real economic activity and inflation.

Over the last thirty years there have been dramatic changes in the way financial markets operate. In addition, the conduct of monetary policy has also changed in dramatic ways, with an increased focus on achieving price stability. And research in monetary economics has stimulated new thinking on how monetary policy can affect the economy, leading to further evolution in our understanding of the monetary transmission mechanism. All of these developments suggest that there is a strong possibility that there have been changes in the monetary transmission mechanism.

A first look at the data shows notable differences in the reduced-form correlations between aggregate economic activity or various components of private expenditure and the short-term nominal policy interest rate in the United States in the most recent decades from the correlations that prevailed in the decades prior to the Volcker disinflation and numerous regulatory changes that occurred in the late 1970s and early 1980s. Figure 1 plots the correlation between the growth rates of output (real GDP), four components of private expenditure (nondurables and services consumption, durables consumption, residential investment, and nonresidential investment), and the nominal federal funds rate (both lagged and led four quarters) for the periods from 1962Q1–1979Q3 (the dark grey bars) and from 1984Q1–2008Q4 (the light gray bars). The correlations shifted notably across these periods: In the earlier sample, growth in aggregate activity and expenditure was negatively correlated with the nominal federal funds rate, especially with lags of the nominal federal funds rate; in the latter sample, growth in aggregate activity and expenditure was positively correlated with the nominal funds rate, especially with leads of the nominal federal funds rate.

These changes may suggest changes in the effects of interest rate movements on demand; indeed, an uneducated look at the positive correlation between output growth and the nominal interest rate in recent decades might lead an observer to suggest, naively, that efforts by the monetary authority to bring about stronger economic growth should raise short-term interest rates, not lower them. Alternatively, these changes may reflect changes in the behavior of policymakers — for example, a more systematic approach that focuses on stability in inflation and economic activity, which implies a positive correlation between the policy interest rate and economic growth due to a policymaker's tendency to lean against strengthening in demand.

We start our analysis by reviewing the various channels of monetary policy transmission and how our understanding of them has changed. We then discuss how developments in the financial markets and the conduct of monetary policy may have caused these transmission mechanisms to change. This discussion is followed by our summary of empirical work on the evolution of the monetary transmission mechanism and our independent analysis,

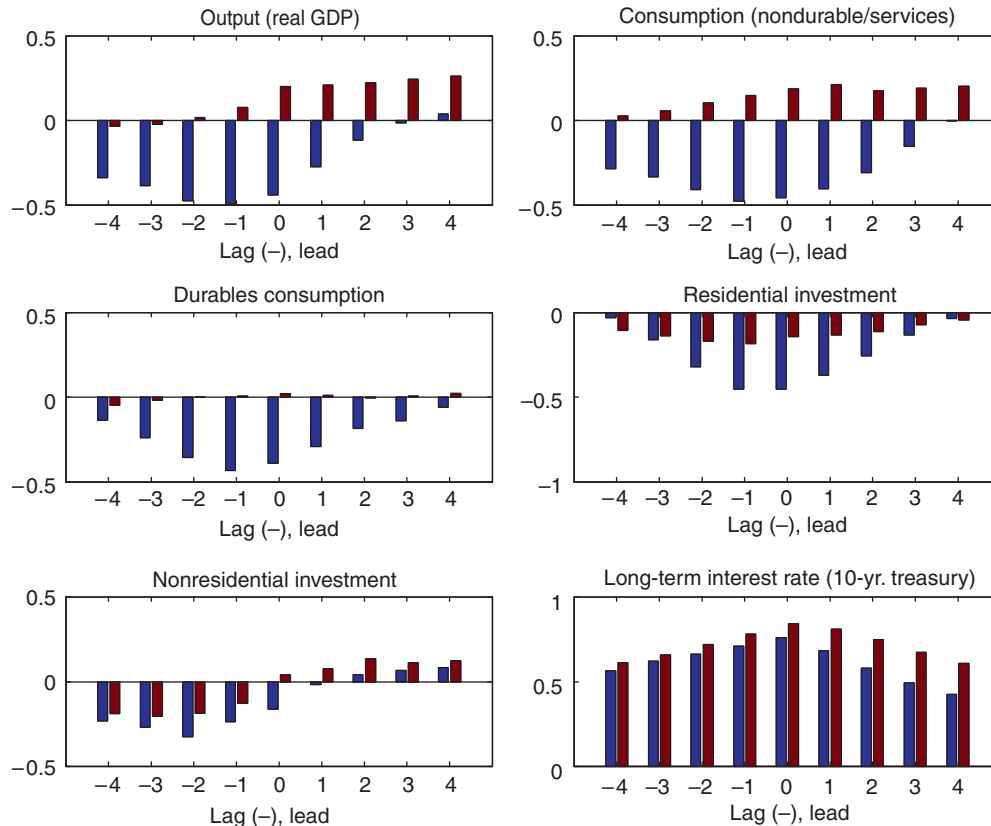


Figure 1 Correlation between measures of activity and demand (log-differences), the long-term interest rate, and the nominal federal funds rate. The dark grey bars denote the correlation between the nominal funds rate (lagged/led) and the data series indicated for the 1962Q1–1979Q3 sample period; the light grey bars denote the same correlations for the 1984Q1–2008Q4 time period.

where we focus on the potential pitfalls associated with alternative identification strategies, the changes in statistical relationships that appear most robust, and structural interpretations of changes in the links between short-term interest rate movements and real activity.

Our analysis is structured around two approaches. The first is based on vector autoregressions (VARs). In this part of our analysis, we build on, for example, the survey by Christiano, Eichenbaum, and Evans (1999) by expanding their analysis to include the more recent FAVAR approach (e.g., Bernanke, Boivin, & Elias, 2005), which allows consideration of a larger set of information. This shift leads to an analysis of a larger range of economic variables; one particular area on which we focus is inflation expectations, as our overall analysis will lead us to conclude that shifts in the management of expectations may be among the most important changes in the relationship between monetary policy and aggregate economic activity. We also emphasize the changes in the effects of monetary policy shocks,

or lack thereof, much more than the earlier literature.¹ Subsequent to this analysis, we present a structural analysis using a DSGE model. This form of analysis allows us to consider changes in monetary policy effects and the impact of changes in monetary policy behavior to ensure that shifts in reduced-form correlations are not simply related to changes in policy behavior as noted by Lucas (1976). We will also consider a number of plausible structural changes through this lens. Our analysis in this vein builds on, for example, Smets and Wouters (2007) and the increasing use of such structural models at central banks (e.g., Christoffel, Coenen, & Warne, 2008; Edge, Kiley, & Laforge, 2007, 2008, 2010).

These two approaches, VAR and DSGE, span the range from relatively unstructured to highly structured. An intermediate approach, adopted in, for example, Akhtar and Harris (1987), Friedman (1989), Mauskopf (1990), and Fair (2004) specifies equations for various categories of expenditure using information from economic theory on the plausible set of determinants and “Cowles Commission” econometrics. Our results will in many ways be similar to those from this literature, which largely concluded that the evidence for changes in the monetary policy transmission mechanism was limited. Nonetheless, we see aspects of our analysis as representing a substantial step forward both in exploiting a large set of information and imposing only limited identifying assumptions (as in the FAVAR approach) and in moving to the other extreme to try to address the Lucas critique and consider the “management of expectations,” which Woodford (2003) emphasized is a primary transmission channel.

Several results stand out from our review and analysis. First, changes at the macroeconomic level are difficult to detect: Relatively unrestricted approaches using macroeconomic data, such as analyses using VARs, suffer from the curse of dimensionality and have reached different conclusions regarding the importance of time variation in the links between monetary policy and macroeconomic activity; more restricted structural approaches are more controversial. Nonetheless, the data do suggest certain changes that are important for monetary transmission. Overall, the responses of measures of real activity and prices have become smaller and more persistent since 1984. Also, changes in government regulation and financial innovations related to housing finance in the United States seem to have altered the response of residential investment to changes in monetary policy in recent decades from that in earlier periods (and studies examining a range of countries have noted the importance of such changes around the world). Perhaps more clearly in the data for the United States, changes in the behavior of monetary policy have anchored inflation expectations and altered the transmission of other shocks to activity and inflation significantly. Finally, the overall importance of non-neoclassical, or credit-type, channels of monetary policy remains difficult to empirically assess with macroeconomic data and models, perhaps

¹ With that said, Christiano, Eichenbaum, and Evans (1999) did examine changes in the effect of monetary policy shocks, and found only limited evidence for such changes conditional on the size of the policy shock; we will reach similar conclusions along some dimensions. These authors did find much smaller shocks in recent samples, as we will.

because the theoretical guidance for this type of macroeconomic empirical research has been limited. This area is likely to be a very active, and hopefully fertile, area of research in coming years. We use our analysis to discuss directions for such research about the conduct of monetary policy in the aftermath of the current financial crisis.

2. THE CHANNELS OF MONETARY TRANSMISSION

Monetary transmission can be categorized into two basic types: neoclassical channels in which financial markets are perfect and non-neoclassical channels that involve financial market imperfections, which are usually referred to as the credit view.

In our upcoming discussion, we will take as given that the monetary authority's policy instrument, at least in normal times, involves direct control over a short-run interest rate (e.g., the federal funds rate in the United States). We also assume that nominal wage and price rigidities imply that variations in the nominal policy interest rate affect the real interest rate directly. Our discussion of the effects from policy settings to real activity hence focuses on how variation in the short-term nominal policy rate feeds through to the real interest rate and other asset prices, thereby influencing spending. Table 1 provides a summary of the channels we discuss.²

An important feature of many of the transmission mechanisms we discuss is that it is the real (rather than the nominal) interest rate that affects other asset prices and spending in (many) transmission channels. In addition, the entire expected path of interest rates, not solely the current value, influences asset prices and spending. Both of these factors give rise to an important role for expectations in the effects of monetary policy actions, as policy strategies can influence both the expected course of nominal interest rates and the outlook for inflation and hence real interest rates.³ Indeed, Woodford (2003) suggested that the "management" of expectations is the primary responsibility of a monetary authority. We discuss the important role of expectations at several points in our analysis; we also highlight channels in which nominal, rather than real, interest rates play a special role.

2.1 Neoclassical channels

The traditional channels of monetary policy transmission are built upon the core models of investment, consumption, and international trade behavior developed

² Mishkin (1995) covered similar ground, Taylor (1995) emphasized neoclassical channels, and Bernanke and Gertler (1995) emphasized credit channels.

³ That the real interest rate rather than the nominal rate affects spending provides an important mechanism for how monetary policy can stimulate the economy, even if nominal interest rates hit the zero lower bound, which happened during persistent deflationary episodes and as has occurred recently around the world. With nominal interest rates at a floor of zero, a commitment to future expansionary monetary policy can lower long-term interest rates and raise expected inflation, lowering real interest rates and stimulating spending (e.g., Eggertsson & Woodford, 2003). For example, the Federal Reserve's FOMC statements have indicated since 2009 that the federal funds rate would be kept at very low values for an extended period of time.

Table 1 Monetary Policy Transmission Channels

Channel	Description	Incorporation in policy models
Neoclassical channels		
Interest rate/ cost-of- capital/ Tobin's q	Changes in short-term policy rates affect the user cost of capital for consumer and business investment.	Standard in large-scale models (like the MPS or FRB/US models, Fair, 2004) and DSGE models.
Wealth effects	Changes in short-term interest rates affect discounted present values and/or Tobin's q for various types of assets, and these changes in the market value of assets induce changes in consumption.	Standard in the large-scale models (MPS or FRB/US, Fair, 2004). Standard in DSGE models, but not separated from intertemporal substitution effects.
Intertemporal substitution	Changes in short-term interest rates affect the slope of the consumption profile.	Absent from traditional large-scale. Standard in DSGE models, but not separated from wealth effects.
Exchange rate effects	Changes in short-run policy interest rates induce changes in the exchange rate through uncovered-interest parity and/or portfolio balance effects.	Standard in large-scale models. Incorporated in international DSGE models (e.g., Erceg, Guerrieri, & Gust, 2006).
Non-neoclassical channels		
Regulation-induced credit effects	Restrictions on financial institutions (e.g., deposit rate ceilings, credit restrictions) affect spending.	Incorporated empirically for relevant periods in some large-scale models (e.g., MPS model).
Bank-based channels	Banks play a special role addressing problems of asymmetric information. Thus, decreases in bank's lending capacity impact spending.	Not explicitly incorporated in most large-scale models or DSGE models.
Balance sheet channel	Changes in net worth associated with the asset price effects of monetary actions influence external finance premia facing firms and households.	Not explicitly incorporated in most large-scale models. Increasingly incorporated in DSGE models, often along the lines suggested in Bernanke, Gertler, and Gilchrist (1999).

during the mid-twentieth century: the neoclassical models of investment of Jorgenson (1963) and Tobin (1969); the life cycle/permanent income models of consumption of Brumberg and Modigliani (1954), Ando and Modigliani (1963), and Friedman (1957); and the international IS/LM-type models of Mundell (1963) and Fleming (1962). We categorize these primary channels using this framework, and hence distinguish by channels that directly affect investment, consumption, and international trade. For investment, the key channels are the direct interest rate channel operating through the user cost of capital and the closely related Tobin's q channel; for consumption, the channels operate through wealth effects and intertemporal substitution effects. For trade, the direct channel operates through the exchange rate. We look at each of these in turn.

2.2 Investment-based channels

2.2.1 Direct interest-rate channels

The most traditional channel of monetary transmission that has been embedded in macroeconomic models involves the impact of interest rates on the cost of capital and hence on business and household investment spending (e.g., residential and consumer durables investment). Standard neoclassical models of investment demonstrate that the user cost of capital is a key determinant of the demand for capital, whether it be investment goods, residential housing, or consumer durables.⁴ The user cost of capital (u_c) can be written as:

$$u_c = p_c[(1 - \tau)i - \pi_c^e + \delta]$$

where p_c is the relative price of new capital, i is the nominal interest rate, π_c^e is the expected rate of price appreciation of the capital asset, and δ is the depreciation rate. The user cost formula also allows for the deductibility of the interest rate (which is particularly important in the United States where mortgage interest is deductible) by adjusting the nominal interest rate by the marginal tax rate τ . Regrouping terms, the user cost of capital can be rewritten in terms of after-tax real interest rate, $(\tau - 1)i - \pi^e$, and the expected real rate of appreciation of the capital asset, $\pi_c^e - \pi^e$, where π^e is the expected inflation rate.

$$u_c = p_c[\{(1 - \tau)i - \pi^e\} - \{\pi_c^e - \pi^e\}]$$

Several factors are important in determining the effects of monetary policy operating through these direct, user-cost channels. The first regards the horizon over which interest rates influence spending. Because capital assets are long-lived and the adjustment of these stocks involves costs (of planning, procurement, installation, etc.), businesses and households take the long view when factoring variation in interest rates into their investment decisions. As a result, the real interest rate and the expected real appreciation of the

⁴ The classic reference is Jorgenson (1963).

capital asset that influence spending will typically be related to the expected life of the asset, which is often very long. In traditional econometric models, this link typically is formalized through direct inclusion of a long-term interest rate in the user-cost formula, rather than a short-term interest rate. In the recent generation of microfounded models, often called DSGE models, this link typically arises through a dynamic intertemporal optimality condition for investment that makes spending depend on the expected sequence of short-term interest rates going forward (as we will present next).

With the monetary policy instrument being a short-term interest rate, this discussion makes clear that the monetary transmission mechanism involves the link between short- and long-term interest rates through some version of the expectations hypothesis of the term structure. When monetary policy raises short-term interest rates, long-term interest rates also tend to rise because they are linked to future short-term rates; consequently the user cost of capital rises and the demand for the capital asset falls. The decline in the demand for the capital asset leads to lower spending on investment in these assets causing aggregate spending and demand to decline.

2.2.2 Tobin's q

The investment decisions of firms and households can also be considered in the framework of James Tobin (1969). For business investment, Tobin (1969) defined q as the market value of firms divided by the replacement cost of capital. When q is high, the market price of firms is high relative to the replacement cost of capital, and new plant and equipment capital is cheap relative to the market value of firms. Companies can then issue stock and get a high price for it relative to the cost of the facilities and equipment they are buying. As a result, investment spending will rise, because firms can buy a lot of new investment goods with only a small issue of stock. In principle, similar reasoning could be applied to household investment decisions.

Tobin's q theory can be linked to the user cost of capital approach, as shown by, for example, Hayashi (1982). Indeed, the q -formulation dominates formal micro-based modeling efforts and the DSGE literature mentioned previously, in large part because the formal links between q -theory and the user-cost approach in the dynamic adjustment cost approach of Hayashi (1982), allow for convenient analytical expressions in such models. In addition, the q -approach does add a degree of richness, as it emphasizes that there is a direct link between stock prices and investment spending. In practice, Tobin's q leads to another channel of monetary transmission: When monetary policy is eased and interest rates lowered, the demand for stocks increases and stock prices rise, leading to increased investment spending and aggregate demand.

2.2.3 Previous empirical literature on investment-based channels

The user-cost channel described earlier is a standard feature of large scale macroeconomic models used for forecasting and policy analysis in the United States such as the

MPS model developed in the 1970s (Brayton & Mauskopf, 1985) and the more recent FRB/US model used at the Federal Reserve (e.g., Reifschneider, Tetlow, & Williams, 1999). It is also a standard feature in large scale macroeconometric models developed at central banks for other countries. Examples include the ECB's Area Wide Model (Fagan, Henry, & Mestre, 2005; Bank of England's Quarterly Model; Harrison et al., 2005). The q -representation of this channel is the baseline model of investment decisions in DSGE models used at central banks (e.g., the EDO model of the Federal Reserve Board; Edge et al., 2007, 2008, 2010; Kiley, 2010), the New Area Wide Model of the ECB (Christoffel et al., 2008), and ToTEM at the Bank of Canada (Murchison & Rennison, 2006).

This channel of monetary policy transmission is an important one in these models — investment spending is the bulk of the near-term response to changes in the short-term policy rate. This finding has long been true in models employed at central banks (e.g., see the comparison of central bank models reported in Smets, 1995). Nonetheless, the long-run sensitivity of investment to changes in the user cost of capital is controversial, and the short-run elasticities can be estimated to be quite small in data for the United States and other countries — findings which have led some (e.g., Bernanke & Gertler, 1995) to question the primacy of this channel. For example, for residential housing using U.S. data, the long-run elasticities range from -0.2 to -1.0 (e.g., Case, 1986; Hanushek & Quigley, 1980; Henderson & Ioannnides, 1986; McCarthy & Peach, 2002; and Reifschneider et al., 1999); short-run elasticities, which may be more important for monetary policy questions, are modest (especially abstracting from regulation-induced credit market effects in the United States prior to the early 1980s). For business investment, the estimated range of elasticities is also considerable: Chirinko (1993) summarized evidence for the United States and said that “the response of investment to price variables tends to be small and unimportant relative to quantity variables”; Fagan et al. (2005) reported for the Euro Area elasticity after one year of less than 0.1%. Estimates for consumer durables are scant, but also tend to be small in the short-run; for example, the short-run semielasticity of consumer durables investment reported in Taylor (1993) lies close to zero.

The second term of the user cost of capital, the expected real rate of appreciation of the capital asset, $\pi_c^e - \pi^e$, provides an additional way for monetary policy to affect investment spending, whether it is by businesses or households. Changes in these expectations can have an important effect on the user cost of capital and thus on spending, and this has been particularly emphasized for the housing market by Case and Shiller (2003). When monetary policy tightens and interest rates rise, housing prices soften because the demand for housing declines through the user-cost transmission mechanism described earlier. Expectations of future tightening of monetary policy could therefore lower the expected real rate of appreciation of housing prices, raising the *current* user cost of capital, which would then lead to a decline in the demand for housing and residential construction.

2.3 Consumption-based channels

2.3.1 *Wealth effects*

Standard applications of the life cycle hypothesis of saving and consumption, first developed by Brumberg and Modigliani (1954) and later augmented by Ando and Modigliani (1963), indicate that consumption spending is determined by the lifetime resources of consumers, which includes wealth, whether from stock, real estate or other assets. Expansionary monetary policy in the form of lower short-term interest rates will stimulate the demand for assets such as common stocks and housing, driving up their prices; alternatively (and equivalently), lower interest rates lower the discount rate applied to the income and service flows associated with stocks, homes, and other assets, driving up their price. The resulting increase in total wealth will then stimulate household consumption and aggregate demand. Standard life cycle wealth effects operating through asset prices are thus an important element in the monetary transmission mechanism.

2.3.2 *Intertemporal substitution effects*

A second consumption-based channel reflects intertemporal substitution effects. Indeed, this channel is central to the models in the DSGE tradition mentioned earlier. In this channel, changes in short-term interest rates alter the slope of the consumption profile, so that lower interest rates induce higher consumption today. In DSGE models, this channel naturally arises through the models' use of the standard consumption Euler equation linking the marginal rate of substitution between current and future consumption with the real interest rate.

2.3.3 *Previous empirical literature on consumption-based channels*

The wealth effect has had a prominent role in macroeconometric models, such as the ones used at the Federal Reserve for policy analysis. This view is embedded in the macroeconometric models used at the Federal Reserve Board and elsewhere, in which the long-run marginal propensity to consume out of wealth in the United States is currently estimated to be between 3 and 4 cents per dollar, for both housing wealth and stock market wealth. Fair (2004) reported a wealth effect of similar size for the United States.^{5,6} Catte, Girouard, Price, and Andre (2004), in a study of OECD countries, found that the long-run marginal propensity to consume out of financial wealth ranges from 0.01 in Italy

⁵ The life cycle view that wealth effects are the same for all types of wealth is controversial, with some research indicating that housing wealth has a greater effect on consumption than non-housing wealth, with other research finding the opposite. For a survey of this literature, see Mishkin (2007).

⁶ An overview of the monetary transmission mechanism in the FRB/US model is in Reifschneider, Tetlow, and Williams (1999). The wealth effects estimated by the staff of the Federal Reserve Board have varied importantly over time. As discussed in Brayton and Mauskopf (1985), in the MPS model (the predecessor to FRB/US), the propensity to spend real estate wealth ranged from an estimate in the 1970s of 2.9 cents per dollar to an estimate in the 1980s of 8.4 cents. The source of that variation appears to have been a lack of variation in the ratio of real estate wealth to aggregate income. In contrast, historical fluctuations in stock market wealth have been sufficient to allow a more precise estimation of the propensity to spend that wealth. The Board staff's estimates of this propensity have stayed within a narrow range of 3 to 4 cents per dollar for the past 40 years.

to 0.07% in Japan; their estimate of the OECD average is about 0.035, and their estimate for the United States is 0.03. With that said, the short-run wealth effects are even smaller, and monetary policy can only influence wealth in the short-run; as a result, the wealth effect has played an important role in modeling efforts, but has played a secondary role to direct interest rate channels of investment in most modeling efforts (e.g., the summary of central bank models in Smets, 1995).⁷

The intertemporal-substitution channel is also typically modest in the short run – as the sensitivity of the slope of the consumption profile to short-term interest rates is typically estimated to be small, mainly through the inclusion of habit persistence (e.g., Christoffel, et al., 2008; Edge et al., 2007; Smets & Wouters, 2007). This finding is directly related to a large empirical literature: Hall (1988) and subsequent research tended to uniformly suggest modest intertemporal substitution. For this reason, it is perhaps not surprising that econometric models discussed in the previous paragraph have typically not emphasized this channel; for example, this channel of monetary transmission has not been a factor in the Federal Reserve's MPS or FRB/US models and was not included in the ECB's Area Wide Model (e.g., Fagan et al., 2005).

2.4 International-trade based channels

2.4.1 Exchange rate channel

When the central bank lowers interest rates, the return on domestic assets falls relative to foreign assets. As a result, the value of domestic assets relative to other currency assets falls, and the domestic currency depreciates. The lower value of the domestic currency makes domestic goods cheaper than foreign goods, leading to expenditure switching and a rise in net exports. The rise in net exports then adds directly to aggregate demand. Therefore, the exchange rate channel plays an important role in how monetary policy affects the economy. In this regard, two factors are important. First, the sensitivity of the exchange rate to interest rate movements is important. For example, it was not uncommon on earlier, econometric models for the estimated sensitivities to be small, implying a small channel; whereas models that impose uncovered interest parity tend to find a larger role for this channel. Second, smaller, more open economies tend to see larger effects through this channel.⁸

2.5 Non-neoclassical channels: The credit view

We call channels that arise because of market imperfections (other than those associated with nominal wage and price rigidities) *non-neoclassical transmission mechanisms*. Such channels can arise either from government interference in markets or through imperfections in private markets, such as asymmetric information or market segmentation that leads to barriers to efficient financial markets functioning. In general, these

⁷ Lettau and Ludvigson (2004) emphasized the difference between short- and long-run movements in wealth and movements in consumption, albeit not in the context of the monetary transmission mechanism per se.

⁸ For examples, see Bryant, Hooper, and Mann (1993); Smets (1995); and Taylor (1993).

non-neoclassical transmission mechanisms involve market imperfections in credit markets and have been named the “credit view.” There are three basic non-neoclassical channels that we discuss here: effects on credit supply from government interventions in credit markets, the bank-based channels (through lending and bank capital), and the balance sheet channel (affecting both firms and households).

2.5.1 Effects on credit supply from government interventions in credit markets

Governments often interfere with the free functioning of credit markets in order to achieve certain policy objectives such as redistribution or encouraging particular types of investment. In the United States, government intervention has been particularly important in housing finance in order to encourage home ownership.

Up until the 1980s, the U.S. government had set up a system in which thrift institutions, particularly savings and loan associations, were the primary issuers of residential mortgages. As a result of regulatory constraints, these institutions primarily made long-term, fixed-rate mortgage loans in their local areas using funds provided by local time deposits (see, e.g., McCarthy & Peach, 2002). Government regulations also were geared to helping these thrifts attract deposit funding, enabling them to make more mortgage loans, by establishing ceilings on the interest rates on deposits under Regulation Q, and allowing thrifts to pay 25 basis points (0.25 percentage points) more on their deposits than commercial banks.

The regulatory requirements that thrifts issue long-term mortgages and Regulation Q ceilings led to an important channel of monetary transmission involving credit supply. When the Federal Reserve tightened policy and raised interest rates, there were two effects that led to a decline in the supply of credit to the mortgage market. First higher short-term rates would increase the cost of funds for the thrifts, while income from fixed-rate mortgages is slow to change, leading to a contraction in net interest income. The resulting weakening of thrifts balance sheets would then result in a decreased willingness to issue mortgages, thus causing a contraction in credit supply.

Even more important, higher short-term rates would often lead to rates that were higher than the deposit rate ceilings, causing depositors to withdraw their funds from thrifts and commercial banks in order to put them into higher yielding securities. This loss of deposits from the banking system, a process called “disintermediation,” restricted the amount of funds that banks and thrifts could lend and caused a sharp contraction in mortgage credit and hence in residential construction activity.

The credit rationing channel described here indeed was important in macroeconomic models pre-1980 (e.g., the description of the MPS model in Brayton & Mauskopf, 1985), although their effects partly operated through the timing, rather than overall magnitude, of the impact of monetary policy actions on spending. Starting in the early 1980s, the Regulation Q deposit rate ceilings were gradually eliminated with complete abandonment by 1986, so disintermediation from this government intervention in credit markets is no longer an important channel of monetary transmission.

2.5.2 Bank-based channels

There are two distinct bank-based transmission channels. In both, banks play a special role in the transmission process because bank loans are imperfect substitutes for other funding sources.

The first is the traditional *bank lending channel*. According to this view, banks play a special role in the financial system because they are especially well suited to solve asymmetric information problems in credit markets. Because of a bank's special role, certain borrowers will not have access to credit markets unless they borrow from banks. As long as there is no perfect substitutability of retail bank deposits with other sources of funds, the bank lending channel operates as follows. Expansionary monetary policy, which increases bank reserves and bank deposits, increases the quantity of bank loans available. Because many borrowers are dependent on bank loans to finance their activities, this increase in loans will cause investment and consumer spending to rise.⁹

An important implication of the bank lending channel is that monetary policy will have a greater effect on expenditure by smaller firms, which are more dependent on bank loans, than it will on large firms, which can get funds directly through stock and bond markets (and not only through banks.) Though the bank lending channel has been supported in empirical work (e.g., Gertler & Gilchrist, 1993, 1994; Kashyap & Stein, 1995; Peek & Rosengren, 1995a,b, 1997), other research has raised doubts about the bank lending channel (see Ramey, 1993; Romer & Romer, 1989). Lown and Morgan (2002) reported results that suggest that bank lending may have an important role in macroeconomic fluctuations, but also found that the bank lending channel for monetary policy changes may be quite small. Iacoviello and Minetti (2008) presented results that suggest the presence of a bank-lending channel for households in countries where mortgage finance is more bank dependent. Overall, the literature on the bank lending channel has focused on evidence showing its potential importance, but little work has been developed to provide an overall assessment of the *macroeconomic* importance of this channel, rather than its importance for certain classes of firms or banks, or for certain episodes.

A separate bank channel is called the *bank capital channel*. In this channel, the state of banks' and other financial intermediaries' balance sheets has an important impact on lending. A fall in asset prices can lead to losses in banks' loan portfolios; alternatively, a decline in credit quality, because borrowers are less able or willing to pay back their loans, may also reduce the value of bank assets. The resulting losses in bank assets can result in a diminution of bank capital, as has occurred during the recent financial crisis. The shortage of bank capital can then lead to a cutback in the supply of bank credit, as external financing for banks can be costly, particularly during a period of declining asset

⁹ For surveys of the literature on the bank lending channel, see Bernanke and Gertler (1995) and Peek and Rosengren (1995b).

prices, implying that the most cost-effective way for banks to increase their capital to asset ratio is to shrink their asset base by cutting back on lending. This “deleveraging” process means that bank-dependent borrowers are now no longer able to get credit so they will cut back their spending and aggregate demand will fall.¹⁰

Expansionary monetary policy can lead to improved bank balance sheets in two ways. First, lower short-term interest rates tend to increase net interest margins and lead to higher bank profits that result in an improvement in bank balance sheets over time. Second, expansionary monetary policy can raise asset prices and lead to immediate increases in bank capital. In the bank capital channel, expansionary monetary policy boosts bank capital, lending, and hence aggregate demand by enabling bank-dependent borrowers to spend more.

The bank lending and bank capital channels have typically not been built into either large-scale macroeconomic models or DSGE models used in policy analysis. Despite this, awareness of the bank lending and bank capital channels has played an important role in the conduct of monetary policy in recent years. This was true in the early 1990s, when Alan Greenspan talked about “headwinds” in the economy as a result of the deterioration in bank balance sheets (see, e.g., the discussion of credit channels and the MPS model in Mauskopf, 1990, and the description of the early 1990s in Reifschneider, Stockton, & Wilcox, 1997). While research documenting the policy process recently remains a future topic, the importance of these channels in the recent crisis has been emphasized by policy-makers and popular accounts (e.g., Mishkin, 2008; Wessel, 2009). Moreover, research is now focused explicitly on incorporating such channels in mainstream models used in policy analysis (e.g., Angeloni & Faia, 2010; Gerali, Neri, Sessa, & Signoretti, 2009; Gertler & Kiyotaki, 2010; Meh & Moran, 2008).

2.5.3 Balance sheet channel

Like the bank lending channel, the balance sheet channel arises from the presence of asymmetric information problems in credit markets. When an agent’s net worth falls, adverse selection and moral hazard problems increase in credit markets. Lower net worth means that the agent has less collateral, increasing adverse selection and increasing the incentive to boost risk-taking, thus exacerbating the moral hazard problem. As a result, lenders will be more reluctant to make loans (either by demanding higher risk premia or curtailing the quantity lent), leading to a decline in spending and aggregate demand. A particularly convenient, and widely adopted, model of this type is the financial accelerator framework of Bernanke and Gertler (1989) and Bernanke, Gertler, and Gilchrist (1999), in which lower net worth increases the problems associated with asymmetric information in debt financing, increasing the external finance premium.

¹⁰ See Van den Heuvel (2002) and Peek and Rosengren (2010).

Monetary policy affects firms' balance sheets in several ways. As we have seen, contractionary monetary policy leads to a decline in asset prices, particularly equity prices, which lowers the net worth of firms. Contractionary monetary policy therefore causes adverse selection and moral hazard problems to worsen, which leads to a decline in lending, spending, and aggregate demand. Another way that monetary policy can affect firms' balance sheets is through cash flow, the difference between cash receipts and cash expenditure. Contractionary monetary policy, which raises interest rates, causes firms' interest payments to rise, causing cash flow to fall. With less cash flow, the firm has fewer internal funds and must raise funds externally. Because external funding is subject to asymmetric information problems and hence an external finance premium, additional reliance on external funds boosts the cost of capital, curtailing lending, investment, and economic activity. An interesting feature of the cash flow channel is that *nominal* interest rates affect firms' cash flow, in contrast to the role of the real interest rate emphasized in neoclassical channels. Furthermore, the short-term interest rate plays a special role in this transmission mechanism, because interest payments on short-term (rather than long-term) debt typically have the greatest impact on firms' cash flow. Different variants of the balance sheet channel have been recently considered in investigating optimal monetary policy in the presence of credit frictions. Examples include Curdia and Woodford (2009) and Carlstrom, Fuerst, and Paustian (2009).

These types of balance sheet channels also affect households. For example, an increase in house prices leads to more potential collateral for the homeowner, which may improve both the amount and terms of credit available to these households. In other words, higher house prices can reduce the external finance premium or relax constraints on the quantity of credit available to a household. In principle, other assets affecting household net worth could similarly affect the cost of external funds or quantity of credit available to households. A number of empirical studies have suggested that changes in home values have had important effects on households' access to credit and spending (Benito, Thompson, Waldron, & Wood, 2006; Hatzius, 2005). Some modeling efforts, and associated empirical work, have also found support for financial accelerator mechanisms related to housing and household expenditures (e.g., Iacoviello, 2005; Iacoviello & Neri, 2010). The importance of rising house prices in relaxing credit constraints and stimulating consumer spending is clearly dependent on how costly it is to withdraw housing equity and thus on the efficiency of mortgage markets that enable homeowners to overcome credit constraints. In countries with better developed mortgage markets, consumer spending may be more sensitive to increases in house prices.¹¹ Indeed, Calza,

¹¹ Major differences exist across mortgage markets in advanced industrial countries (Calza, Monacelli, & Stracca, 2007). Mortgage markets in the United States are considered to be among the most developed; in some other countries, mortgage lending is hobbled by relatively weak bankruptcy laws and difficulties in seizing collateral. In Italy, for example, where procedures to repossess collateral are lengthy and expensive, the average loan-to-value ratio on mortgages is relatively low (50% vs. 70% for the United States), and the ratio of mortgage debt to GDP is likewise low (15% vs. 70% for the United States).

Monacelli, and Stracca (2007) found that the correlation of consumption growth with changes in house prices is higher in economies with more developed mortgage finance systems; Iacoviello and Minetti (2008) presented evidence that the balance sheet channel affecting households is stronger in countries with less developed mortgage finance systems.¹²

3. WHY THE MONETARY TRANSMISSION MECHANISM MAY HAVE CHANGED

The survey of the different channels of monetary transmission provides two primary reasons the monetary transmission mechanism may have changed over time: structural changes in the economy, particularly credit markets, and the interaction between changes in monetary policy actions and the way expectations are formed.¹³

3.1 Institutional changes in credit markets

Changes in the institutional structure of credit markets have the potential to alter the monetary transmission mechanism, particularly by affecting market imperfections that are the source of the non-neoclassical channels.

One major change in credit markets over the years was the removal in the 1980s of many of the restrictive regulations that limited thrifts to making long-term fixed-rate mortgages and limited the interest rate that financial institutions could pay on their deposits with

¹² Another way of looking at how the balance sheet channel may affect consumer spending is through liquidity effects on consumer durable and housing expenditure, which have been found to be an important factor affecting aggregate demand during the Great Depression (Mishkin, 1978). In the liquidity effects view, balance sheet effects work through their impact on consumers' desire to spend rather than on lenders' desire to lend. Because of asymmetric information about their quality, consumer durables and housing are very illiquid assets (Mishkin, 1976). If, as a result of a bad income shock, consumers need to sell their consumer durables or housing to raise money, they would expect a big loss because they would not get the full value of these assets in a distress sale. (This is the manifestation of the "lemons" problem described by Akerlof, 1970.) In contrast, if consumers expect a higher likelihood of finding themselves in financial distress, they would rather hold fewer illiquid consumer durable or housing assets and more liquid financial assets. When consumers have a large amount of financial assets relative to their debts, their estimate of the probability of financial distress is low, and they will be more willing to purchase consumer durables or housing. Expansionary monetary policy boosts the value of financial assets; consumer durable expenditure and housing purchases then rise because consumers have a more secure financial position and a lower estimate of suffering financial distress. Liquidity effects therefore lead to another household balance sheet channel for monetary transmission.

¹³ One potentially important factor affecting monetary transmission arises from the increased pace of globalization and consequent increased openness of the United States and other economies. With traded goods becoming a more important sector of the economy, exchange rate movements have the potential to have a larger effect on aggregate spending. Hence the exchange rate channel of monetary transmission may have become more important over time. Such changes are likely even more important for small economies. However, these are likely less important for the United States, despite the increased importance of trade, because the *net* effect on aggregate demand from international trade following a monetary policy innovation tends to be close to zero, as the exchange-rate induced movements in exports tend to be offset, on net, by shifts in imports related to the accompanying changes in domestic demand. For this reason, we do not focus on this channel much in our subsequent empirical work.

deposit rate ceilings. The result of this financial liberalization is that the disintermediation process in which higher interest rates led to a reduced supply of mortgages from thrift institutions is no longer operational. Large swings in credit supply in the mortgage market resulting from an increase in interest rates that limited the ability of mortgage-issuing institutions to acquire funds are thus no longer a feature of the economy. Since this channel of monetary policy transmission was important prior to 1980, its absence currently would weaken the impact of monetary policy on residential construction.

In addition, the growth of securitization in the mortgage market also weakened credit supply effects in the mortgage market and tightened the link between market interest rates and interest rates on residential mortgages. As shown in Table 2, banks provided most home mortgage credit in the 1966–1970 period; however, the GSEs (and securitization) grew increasingly important over the next two decades, funding a nearly equal share of home mortgage credit by 1986–1990; and GSEs came to dominate such credit provision over the 2004–2008 period, with other sources accounting for a share similar to that of banks.

In the pre-1980s period, residential construction fell very quickly in response to tighter monetary policy, while mortgage rates responded gradually. In contrast, after the 1980s, mortgage rates responded more quickly and persistently to changes in monetary policy. As a result, monetary policy now primarily affects housing through pricing channels rather than through credit supply restrictions, as was the case before 1980 (Mauskopf, 1990; McCarthy & Peach, 2002). This means that the response of residential construction is more delayed than in earlier periods and is smaller initially. However, these changes are significant shifts only in the short-run timing of responses.

The second major change in the credit markets is that improvements in information technology have improved the efficiency of credit markets, allowing a wider set of institutions to become engaged in extending credit. Particularly noteworthy in this regard is the growth of securitization, the transformation of otherwise illiquid financial assets (such as residential mortgages, auto loan, small business loans, and credit card receivables), which have typically been the bread and butter of banking institutions, into marketable securities. Securitization has led to the enormous expansion of the

Table 2 Sources of Funding for Home Mortgages

	GSEs	Banks	Other
1966–1970	0.04	0.71	0.25
1986–1990	0.39	0.45	0.16
2004–2008	0.43	0.31	0.26

Source: Federal Reserve Board Flow of Funds Accounts. Banks refer to banks, savings and loans, and credit unions; the GSEs refers to GSEs and agency and GSE-backed mortgage pools. The data come from the Flow of Funds accounts produced by the Federal Reserve.

so-called “shadow banking system,” in which bank lending has been replaced by lending via the securities market. The growth of the shadow banking system has had two enormous impacts. First, it has enabled borrowers to bypass banks to get credit. The result has been a shrinking share of credit that is provided by banks. Second, the shadow banking system has led, at least until the recent financial crisis, to wider access to credit by a larger percentage of the population, which is sometimes referred to as the “democratization of credit.”

The first impact, which indicates that the banking system is playing a smaller role in credit markets, suggests that the bank lending and bank capital channels may be less important than they were previously. However, the relative strength of these channels, at least in typical times, has always been a subject of controversy (as discussed earlier), and hence there is little evidence documenting variation over time in the importance of this channel; for example, Miron, Romer, and Weil (1994) examined a long span of U.S. history and find little evidence that the changing nature of financial markets has affected the importance of the bank lending channel, in part because they find very limited evidence for such a channel. However, we have recently seen a substantial shrinkage in the shadow banking system as a result of the recent financial crisis; it is certainly possible, and perhaps probable, that the bank lending and bank capital channels may become more important than they have been in recent years.

The second impact, the democratization of credit, has led to much easier access to credit. For example, in the United States, down-payment requirements have been falling, along with refinancing costs, and the use of credit scoring has widened access to housing and other loans (e.g., Edelberg, 2006). These developments may have increased the role of balance sheet channels for households, perhaps increasing the responsiveness of consumer spending to changes in house prices (e.g., Aoki, Proudman, & Vlieghe, 2002). But a greater balance sheet channel could be offset by other impacts of increased access to credit. For example, better household access to credit could lower the sensitivity of consumer spending to transitory income shocks – as suggested by Dynan, Elmendorf, and Sichel (2006), who found evidence that the sensitivity of consumption to transitory income shocks has fallen in the United States since the mid-1980s. Support for that view also comes from microeconomic evidence that households use mortgage refinancing to buffer their spending from income shocks (Hurst & Stafford, 2004), and that the propensity to refinance mortgages has increased as a result of structural changes in the mortgage market, such as the development of credit scoring (Bennett, Peach, & Peristiani, 2001). Such a decreased sensitivity to transitory income shocks could reduce the responsiveness of spending to monetary policy shifts indirectly by altering the impact on consumption of income.

3.2 Changes in the Way Expectations Are Formed

While we have only touched on expectations in our survey so far, one of the most important shifts in the practice of monetary policy, and hence potentially in its transmission to activity and inflation, is the manner in which the “management of expectations” has become an important tool of monetary authorities throughout the world. Shifts in the behavior of the monetary authority can importantly affect the transmission mechanism. These effects have two forms, both of which are likely to be quantitatively important. First, expenditures depend directly on the expected path of policy rates through the influence of this path on asset prices; for example, if a rise in the policy rate is expected to be more persistent, the expectations hypothesis of the term structure indicates that the impact on long-term interest rates will be larger than if it is expected to be temporary. Second, the nature of the policy rule can have important feedback effects through its influence on expected spending and inflation; for example, policy behavior that responds strongly to deviations of output from potential and deviations of inflation from desired levels will lead to greater stability in expectations for income and inflation, and hence greater stability in actual spending and inflation. Indeed, some research has emphasized the potential importance of changes in policy behavior of this type in shifts in the aggregate impact of monetary policy actions (e.g., Boivin & Giannoni, 2006). We will examine the potential evidence for such changes in some detail in Section 5.

While the potential importance of the expectations channel is especially apparent in simple New Keynesian models (e.g., Woodford, 2003) and their DSGE descendants, the potential for large quantitative effects is not confined to this class of models. For example, the approach in Taylor (1993), which emphasized expectations channels but is less strictly tied to specific microeconomic optimization problems, also allows for potentially powerful effects. And the reduced-form expectations approach followed in the most common version of the FRB/US model (based on small VAR systems for expectations formation) also allows for potentially large effects; indeed, Reifschneider et al. (1999) reported that more than one-half of the effect of monetary policy on activity over the first year of a change in the federal funds rate reflects the expectations channel, rather than direct interest rate, wealth, or exchange rate channels.

4. HAS THE EFFECT OF MONETARY POLICY ON THE ECONOMY CHANGED? AGGREGATE EVIDENCE

As discussed in the previous section, there are potentially many developments that could have implied a change in the way monetary policy transmits to the economy. But before investigating the causes, the first set of questions to investigate is whether

the effect monetary policy on the economy — in particular real activity, prices, and their key components — has changed in a meaningful way over time and how. In this section we review the existing results on this question and provide some new evidence.

4.1 Modeling the monetary transmission mechanism

One crude and simple way to measure the effect of monetary policy on a variable of interest is to regress this variable on the monetary policy instrument as well as, perhaps, additional control variables. The estimated coefficient on the policy instrument is interpreted as the sensitivity of that variable to monetary policy and changes in this sensitivity as suggestive of a change in the transmission of monetary policy. However, since, in the context of such regressions, exogenous sources of policy changes are not clearly isolated and causality is not well established, these are not the only potential interpretations. For instance, the estimated coefficients might instead be capturing the response of monetary policy to these variables, rather than the opposite, as intended. Indeed, we highlighted the pitfalls of reasoning from such reduced-form correlations in our presentation of the shifts in raw correlations in Section 1.¹⁴

To be able to go beyond this reduced form evidence and establish a causal link, the main general strategy used in the literature consists of using what is believed to be an exogenous source of variation in the monetary policy instrument and tracing out its effect on key variables capturing the aggregate behavior of the economy. This is typically achieved in the context of a system of equation where just enough restrictions are imposed to identify the exogenous source of variations in monetary policy, but that is otherwise left free of *a priori* assumptions on the structure of the economy. This has the virtue of providing robust estimates of the effect of monetary policy, in the sense that they are consistent with a large class of linear structural models. However, that also means that while these models are useful to document the effect of monetary policy on the economy, their use in determining the cause of the change, a question we take up in the next section, is more limited.

¹⁴ A variant of this reduced-form approach has examined the evolution of the transmission of monetary on disaggregated categories of expenditures; these studies typically involve either regressions of the expenditure category of interest on the short-term policy rate or on other interest rates, with auxiliary reduced-form equations specified to link these other interest rates to the short-term policy rate. They generally find a reduced interest sensitivity of residential investment (e.g., Dynan, Elmendorf, & Sichel, 2006; Friedman, 1989; Mauskopf, 1990). These studies uniformly attribute this shift to financial deregulation and financial innovations. Results for other expenditure categories are ambiguous. For consumption, Friedman (1989) reported lower interest sensitivity, while Akhtar and Harris (1987) and Mauskopf (1990) reported no change. For nonresidential investment, Mauskopf (1990) reported lower interest sensitivity, Akhtar and Harris (1987) reported no change, and Friedman (1989) reported some increase in sensitivity. Overall, we are hesitant to place too much weight on these studies given their reduced-form approach. Nonetheless, we read these studies, and others, as somewhat ambiguous, with perhaps moderate evidence of reduced interest sensitivity in the aggregate, most likely reflecting the regulatory and other changes in mortgage finance that have eliminated the credit restrictions associated with disintermediation following interest rate increases in the period prior to the 1980s.

The class of empirical models considered in the literature, and that we discuss later, can all be seen as special cases of a general FAVAR model.¹⁵ It is thus useful to start by first introducing this general class of empirical models. In its general form, FAVAR has the following state-space representation:

$$X_t = \Lambda F + e_t \quad (1)$$

$$F_t = A(L)F_{t-1} + u_t \quad (2)$$

where X_t denotes a potentially long vector of observed macroeconomic indicators of interest, F_t is a vector of potentially unobserved variables governing the comovements of the observable macroeconomic variables, e_t is a variable specific observational error, and finally u_t are innovations that are linear combinations of the structural macroeconomic shocks, one of which is the monetary policy shock. Equation (1) states that observable macroeconomic indicators are *potentially* imperfect measures of the latent macroeconomic forces. Equation (2) states that the evolution of the comovements among the macroeconomic indicators is governed by a set of common factors,, that follow a VAR.

This empirical setup is appealing because it is consistent with a large class of linear rational expectation structural models and can accommodate various assumptions about the information set available to the agents, the monetary authority, or the econometrician.¹⁶ By far, the most common approach in the literature we survey later is to assume that the set of relevant fundamental macroeconomic concepts, such as real activity, inflation, and interest rate, is perfectly observed by the econometrician. In that case, Eq. (1) boils down to a set of identities, is observed, and the system (1)–(2) collapse to Eq. (2), which becomes a standard VAR in terms of observable macroeconomic indicators. All the VARs that have been used to investigate the effect of monetary policy can thus be seen as a special case of system (1)–(2). They differ by the macroeconomic indicators they choose to include in.

Uncovering the monetary transmission mechanism within this empirical framework requires imposing restrictions to identify a structural shock corresponding to an exogenous change in monetary policy from. In the special case of VARs, these restrictions often amount in practice to restrictions on the contemporaneous responses among the variables. Given an identification scheme, the effect of the monetary policy on the variables in a standard VAR consists of computing from the estimated model the dynamic effect of an identified monetary policy shock. This can be computed for any variables included in a VAR.

The standard VAR approach assumes that the dynamics of the macroeconomy can effectively be summarized by a handful of observable macroeconomic indicators. One reason why this might be unrealistic is that the true concepts of interest, such as real

¹⁵ See Bernanke, Boivin, and Elias (2005).

¹⁶ See Boivin and Giannoni (2008) for an illustration.

activity and inflation, might not be perfectly measured by any observable macroeconomic indicators. In that case, wrongly asserting that a specific measure corresponds to a particular theoretical concept might lead to biased estimates of the effect of monetary policy. Proper estimation would require recognizing the presence of such observational errors and once this is recognized, a potentially large set of macroeconomic indicators could conceivably carry useful information about the true state of the economy. This is why Bernanke et al. (2005) proposed the more general FAVAR framework characterized by Eqs. (1) and (2).

While retaining the flavor of the VAR, the general FAVAR framework allows relaxing the assumptions that the relevant theoretical concepts of interest are known and perfectly observed by the econometricians. Instead, it treats observable variables as noisy indicators of the true but unobservable state of the economy. The same (typically recursive) identification scheme used in the standard VAR framework can be implemented in FAVAR. Moreover, by expanding the size of, the universe of potentially useful information can be exploited, and the dynamic effect of monetary policy on any of these indicators can be documented.

4.2 Existing evidence

A change in the transmission of monetary policy means that some of the parameters of system (1)–(2) have changed over time, which, from a reduced-form perspective, could manifest itself by a change in the correlation of the policy instrument and the variable of interest. To evaluate the existence and the importance of changes in this transmission mechanism, existing studies have used one of the following three strategies: (1) estimate an empirical model over different subsamples; (2) estimate an empirical model treating (some subsets of) the parameters as time-varying latent processes (typically assumed to evolve according to random walk); or (3) estimate a regime switching version of an empirical model where (some subset of) the parameters can stochastically switch between different, regime-dependent, values.

Boivin and Giannoni (2002, 2006), estimated a VAR over two samples corresponding to the pre- and post-Volcker periods (pre- and post-1979:4) and identify the monetary policy shock using a recursive identification scheme. They find that exogenous changes in monetary policy have had a smaller effect in the post-Volcker period; for instance, they report that the through response of output in the post-1979:4 period is about a quarter of that in the previous period. Primiceri (2005), Galí and Gambetti (2009), and Canova and Gambetti (2009) used time-varying VARs with random walk coefficients to allow for a much richer evolution of the transmission of monetary policy. Galí and Gambetti (2009) also found that the effect of demand-type shocks on real activity and inflation has fallen over time, although they do not separate out the effect of policy shock per se. On the other hand, Primiceri (2005), based on a

recursively identified VAR, reported little change in the transmission of monetary policy over the last fifty years. A similar conclusion is reached by Canova and Gambetti (2009) using a similar strategy except that monetary policy shocks are identified through sign restrictions. However, they also find that over the last decade, real activity has become more responsive to monetary policy shocks *on impact*.

A careful look at the relationship between the strategy adopted and the results obtained provides some clues that are useful to sort out this conflicting evidence. For instance, the results of Canova and Gambetti (2009) that real activity has become more responsive to monetary policy shock post-1990 is in sharp contrast to the results of Boivin and Giannoni (2002, 2006), who found the opposite for the post-1980 period. Part of the explanation might be due to the fact that the evolution of the monetary policy transmission is more complex than what can be captured by the split sample estimation strategy, with a single break date, as assumed in Boivin and Giannoni (2002, 2006). Clearly, the way the monetary policy shock is identified also plays a role. For instance, Canova and Gambetti (2009) found that the main change in the effect of monetary policy is on impact, yet, this impact effect is constrained to be zero under the identification scheme of Boivin and Giannoni (2002, 2006). It is important to note, however, that Canova and Gambetti (2009) left the impact response of real activity at the cost of only obtaining partial identification. That is, since the sign restrictions they used only produced set identification, the impulse response functions they reported are in general not to a pure policy shock, but to some combination of structural shocks that include the policy shock. Finally, these studies use empirical models based on a handful of macroeconomic variables. The omitted information can, in principle, explain why the conclusions are not robust to the way time-variation is modeled or the policy shock is identified.

However, we would highlight a couple of studies that helps shed light on these issues and will be related to our findings — t Galí, Lopez-Salido, and Valles (2003) and Galí and Gambetti (2009). The latter study uses a structural VAR approach and finds smaller effects of “demand” shocks on activity in the post-1980 period, similar to the findings of Boivin and Giannoni (2002, 2006) for policy shocks. Both studies also found evidence of a larger effect on output (and hours, where larger implies a smaller fall in hours) to innovations in productivity. We will find very similar results in our structural DSGE approach, and our findings are driven entirely by changes in monetary policy behavior and the effect of such changes on the transmission of shocks to activity (and inflation).

4.3 New Evidence

Given that the literature is ambiguous about the importance of the changes in the transmission of monetary policy to the aggregate economy, it is useful to revisit this question empirically in the context of the FAVAR framework.

4.3.1 New FAVAR-based evidence

As mentioned earlier, one potential worry with the VAR-based evidence is that it could be contaminated by the omission of some important variables from the analysis. The solution that would consist of simply adding more variables to the VAR becomes quickly impractical, especially when we are looking for the presence of changes in recent history that require the use of short time series. One way to potentially address these issues is to identify the effect of monetary policy within the more general FAVAR framework where the information from literally hundreds of macroeconomic indicators can be exploited. To our knowledge, the evolution of the monetary transmission mechanism has not been systematically investigated through the lens of such framework.

The FAVAR we consider is based on a data set comprised of a total of 181 macroeconomic indicators, of 182 variables of both at the quarterly (58) — a subset of which are the quarterly variables used in the previous VAR analysis(124) — and at the monthly (58) frequencies (123). These include mainly real activity, price, and interest rate measures, but also exchange rates, stock prices, and money and credit aggregates. The analysis is carried out at the monthly frequency. In the benchmark specification we use, there are 5 factors and 3 lags. The identification of monetary policy is the FAVAR equivalent of the recursive identification used in previous VAR analyses; that is, monetary policy is assumed to respond contemporaneously to real GDP, the PCE deflator, and the unemployment rate, but none of these variables can respond to monetary policy within the period.¹⁷ The contemporaneous response of all other variables is left unrestricted. We estimate the model over two subsamples: 1962:1–1979:9 and 1984:1–2008:12.

The impulse responses of real GDP, the PCE deflator, and the federal funds rate to a monetary policy shock are reported in Figure 2. Confidence intervals on the post-1984 estimate of the impulse response functions are also reported. The results suggest that the magnitude of responses of real GDP was greater in the pre-1979Q3 than in the post-1984Q1 period, but the response in the later period seems more delayed and persistent. For the response of the PCE deflator, it appears to have been considerably reduced in the post-1984Q1 periods, compared to the earlier period.

4.3.2 Comparison with the VAR approach

Given that the existing literature has investigated these types of questions in the context of VARs, it is interesting at this stage to compare the FAVAR results for aggregate price and real activity measures with those that would be obtained from a VAR. Based on preliminary exploration, we have found the VAR results to be very sensitive to the

¹⁷ See Stock and Watson (2005) for a complete discussion of alternative identification strategies, and their implementation, in a FAVAR context.

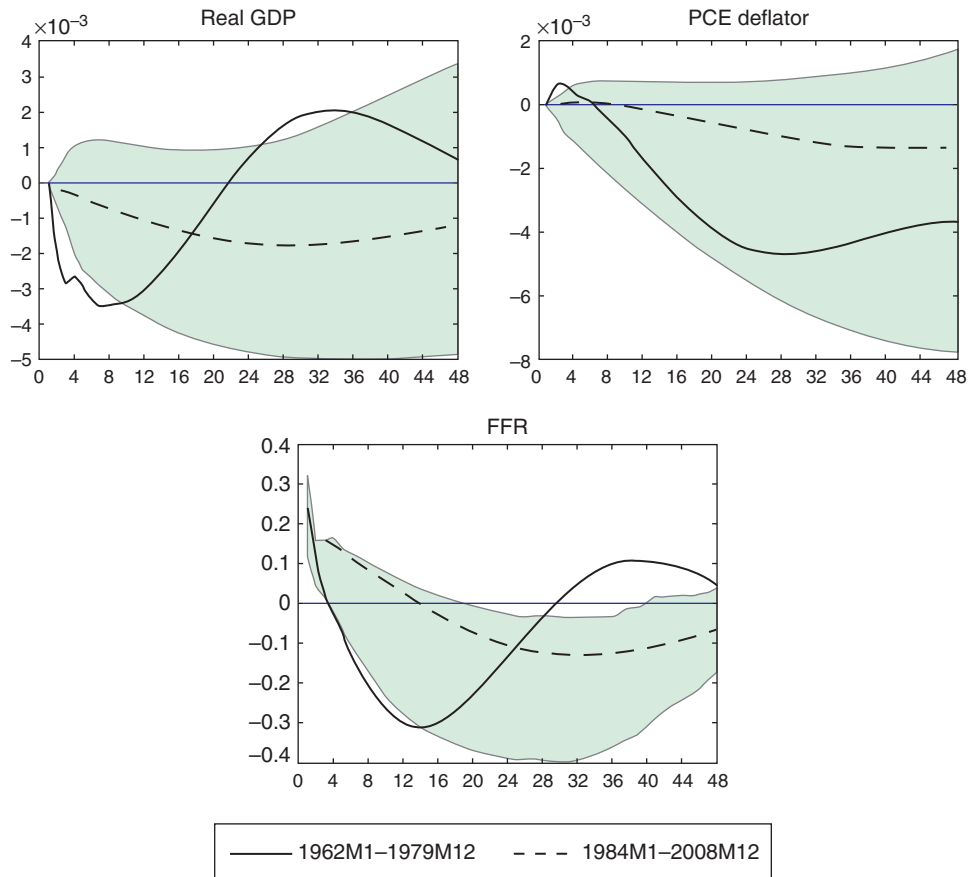


Figure 2 FAVAR-evidence of the aggregate effect of monetary policy. Impulse response functions to a 25 bp surprise increase in the Fed funds rate, estimated from the FAVAR model described in the text. Shaded areas represent the 95% confidence interval on the post-1984 estimates.

specification and the price puzzle — the fact that prices increase following a tightening of monetary policy — to be pervasive across the different periods. The inclusion of an index of commodity prices to the VAR did not resolve the price puzzle. However, the inclusion of a measure of expected inflation in the VAR leads to estimates that eliminate the price puzzle in the post-1984:1 period. The fact that adding more information to the VAR — through the inclusion of an expected inflation measure — helps eliminate the price puzzle in the later sample and that the FAVAR does not display the puzzle, leads us to believe that it is indeed an anomaly of the simpler VAR specification as opposed to a genuine feature of the economy.

Our benchmark VAR specification is thus based on a subset of the macroeconomic indicators used in the previous FAVAR analysis that comprises quarterly data on real

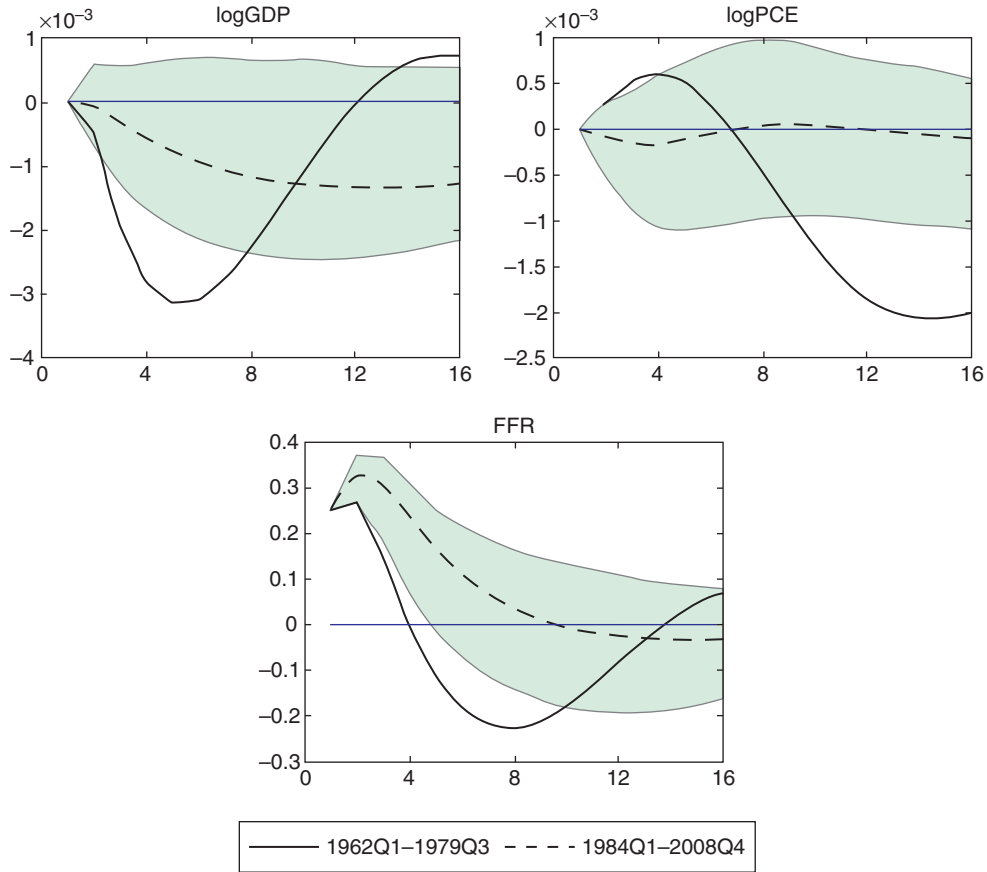


Figure 3 VAR-based evidence of the aggregate effect of monetary policy. Impulse response functions to a 25 bp surprise increase in the Fed funds rate, estimated from the benchmark VAR. Shaded areas represent the 95% confidence interval on the post-1984 estimates.

GDP, the personal consumption expenditure deflator, a commodity price index, a measure of expected inflation, and the federal funds rate.¹⁸ We identify the monetary policy shock recursively with the federal funds rate ordered last, the VAR equivalent of the identification strategy used in the FAVAR. The impulse responses of real GDP, the PCE deflator, and the federal funds rate to a monetary policy shock are reported in Figure 3.

The results for aggregate real activity are very similar to those based on the FAVAR. This is particularly remarkable since the FAVAR is estimated at the monthly

¹⁸ We use a three-year ahead expected inflation measure extracted from the term structure of interest rate (see Section 5 for details on this measure). The variables in the VAR are all in log-level, except the federal funds rate and expected inflation which are in levels.

frequency and uses an entirely different set of information. The response of aggregate prices is, however, markedly different. While the VAR suggests that the pre-1979Q3 response of aggregate prices is also of a greater magnitude than for the post-1984Q1 period, unlike the FAVAR results, the response in the earlier period displays an important price puzzle. This would be consistent with the fact that the VAR is omitting important information that the FAVAR succeeds at extracting from a large set of indicators.

4.3.3 Multidimensional effects of monetary policy

The FAVAR framework also allows us to document the changing effect of monetary on a wide range of macroeconomic indicators. This provides an interesting way to check if the conclusions we have reached are also valid for disaggregated components or alternative measures of real activity and prices.

These results are reported in Figure 4. These results highlight the fact that the general pattern uncovered so far seems to be shared by a large set of relevant indicators. Alternative measures of aggregate prices, such as the Consumer Price Index (CPI) and core CPI, show a reduction in the effect of monetary policy shock post-1984Q1. Measures related to real activity also display a behavior broadly consistent with real GDP in response to monetary policy. The response of industrial production, capacity utilization, employment, housing starts, and orders for durable goods are somewhat smaller initially in the first year to two years, while their shape suggests a more protracted response after 1980 (albeit with wide confidence intervals). Consumer credit displays a similar pattern.

Overall, our reading of the new evidence is as follows. There is some evidence of an evolution in the response of prices and expenditure categories to monetary policy, both in terms of magnitude and timing. The effect of monetary policy on aggregate real activity seems to have become smaller in the post-1984 period compared to the earlier period, and perhaps more persistent as well (although the latter is difficult to assess, given the confidence intervals at longer horizons).

5. WHAT CAUSED THE MONETARY TRANSMISSION MECHANISM TO EVOLVE?

The previous results are suggestive of an evolving effect of monetary policy on real activity, or on some of its components, and inflation. But to understand the policy implications of these changes, we need to know the reasons for these changes and which particular channels are involved.

To try to isolate the source of the evolution of monetary policy, and which of its channels might have been involved, we consider two broad approaches. First, the changing response of some particular variables to monetary policy might be informative

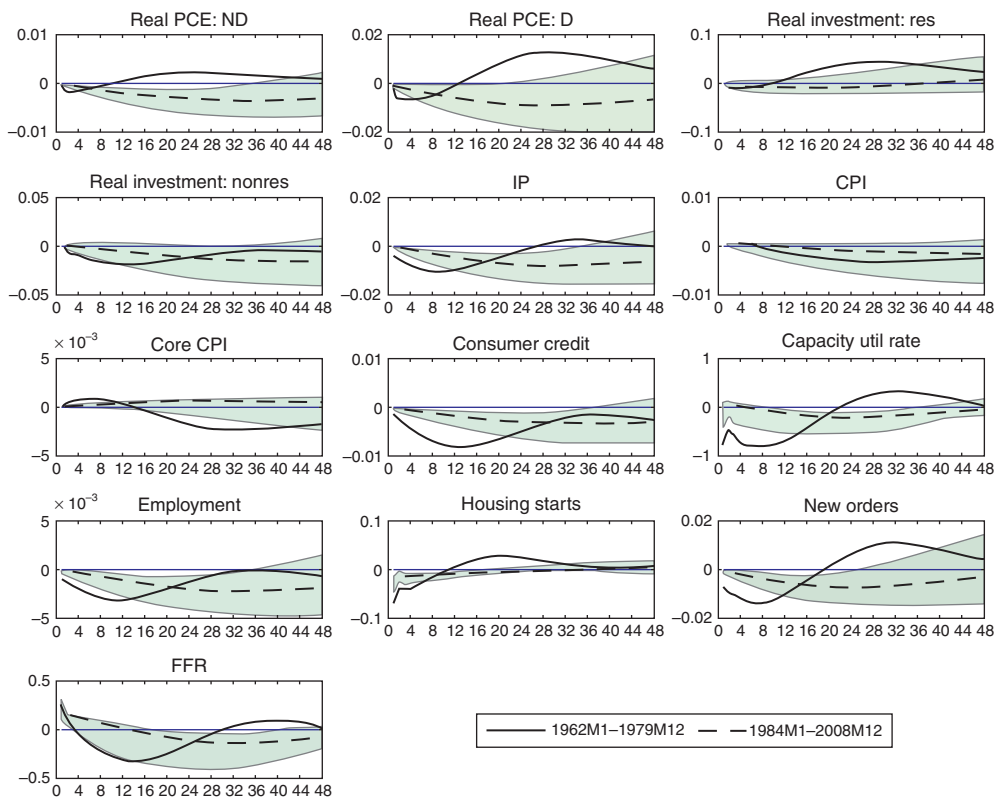


Figure 4 Multidimensional effects of monetary policy. Impulse response functions to a 25 bp surprise increase in the federal funds rate, estimated from the FAVAR model described in the text. Shaded areas represent the 95% confidence interval on the post-1984 estimates.

about the changing nature of some specific channels. In particular, a differentiated response of inflation expectations to the same policy shock over time could be informative about the varying strength of the expectation channel. Or a changing response of the external finance premium to a policy shock, and how this evolution correlates with the growth of the shadow banking sector, might be suggestive of a change in the strength in the lending channel. The second strategy is to consider lessons regarding the evolution of the monetary transmission mechanism that can be gleaned from a structural model.

5.1 FAVAR-based evidence

5.1.1 The expectation channel

As discussed in Section 2, one channel through which monetary policy exerts its influence on the economy is by the effect it might have on private sector expectations.

One way to investigate the potentially changing role of the expectation channel is to document the responses of measures of expected inflation to monetary policy shocks over different periods. We consider a total of four alternative measures of expected inflation. The first four are constructed from term structure of nominal interest rate. They are constructed, as in Canova and Gambetti (2010), as the predicted value from a regression of realized PCE inflation at a given horizon on a constant and the corresponding forward nominal interest rate. We do this for horizons 1, 3, 5, and 10. The logic behind such measures is simple — changes in far forward rates must primarily reflect changes in inflation, as real interest rates presumably converge to some “normal” value at far horizons (as argued in, e.g., Gürkaynak, Sack, & Swanson, 2005). The next measure is survey-based and includes the one-year ahead expectation of CPI inflation from the Michigan Survey, which is available monthly since 1978. The other two are the one-year ahead expectation of CPI inflation and of the GDP implicit deflator inflation from the Survey of Professional Forecasters, available quarterly, respectively, since 1981 and 1970.

Figure 5 reports the FAVAR-based estimates of the responses of these alternative measures of expected inflation to monetary policy shock. The results suggest a considerable reduction in the effect of monetary policy shock on expected inflation based on the term structure or from the Michigan survey. Because of their availability, an estimate of the responses of expected inflation from the Survey of Professional forecasters is not available in the earlier period. However, consistent with the other measures, they respond very little to monetary policy shock in the post-1979Q4 period. In sum, we conclude that the evidence suggests a better anchoring of inflation expectations in the period following the Volcker disinflation.

5.1.2 The balance sheet channel

The balance channel of monetary policy suggests that monetary policy can exert an influence on the economy by affecting the balance sheets of firms and consumers and in turn, their access to credit. The FAVAR results above already noted a reduction in the response of credit following a policy innovation, although this finding does not distinguish between the responses of the supply of credit and the demand for credit.

To investigate empirically how the strength of other balance sheet channels might have evolved over time, we document the responses of alternatives measures of external finance premium to a monetary policy shock across different periods. We consider the spread between the yields on AAA or BAA corporate bonds and corresponding U.S. Treasury bonds, for maturities of 1, 3, 5, and 10 years, as well as the external finance premium measure of Gilchrist, Ortiz, and Zakrajsek (2009).¹⁹ All of these measures are included in the data set, X_t , used to estimate the FAVAR.

¹⁹ Gilchrist, Ortiz, and Zakrajsek (2009) constructed their measure from a portfolio of bonds prices on outstanding senior unsecured debt issued by a large panel of nonfinancial firms.

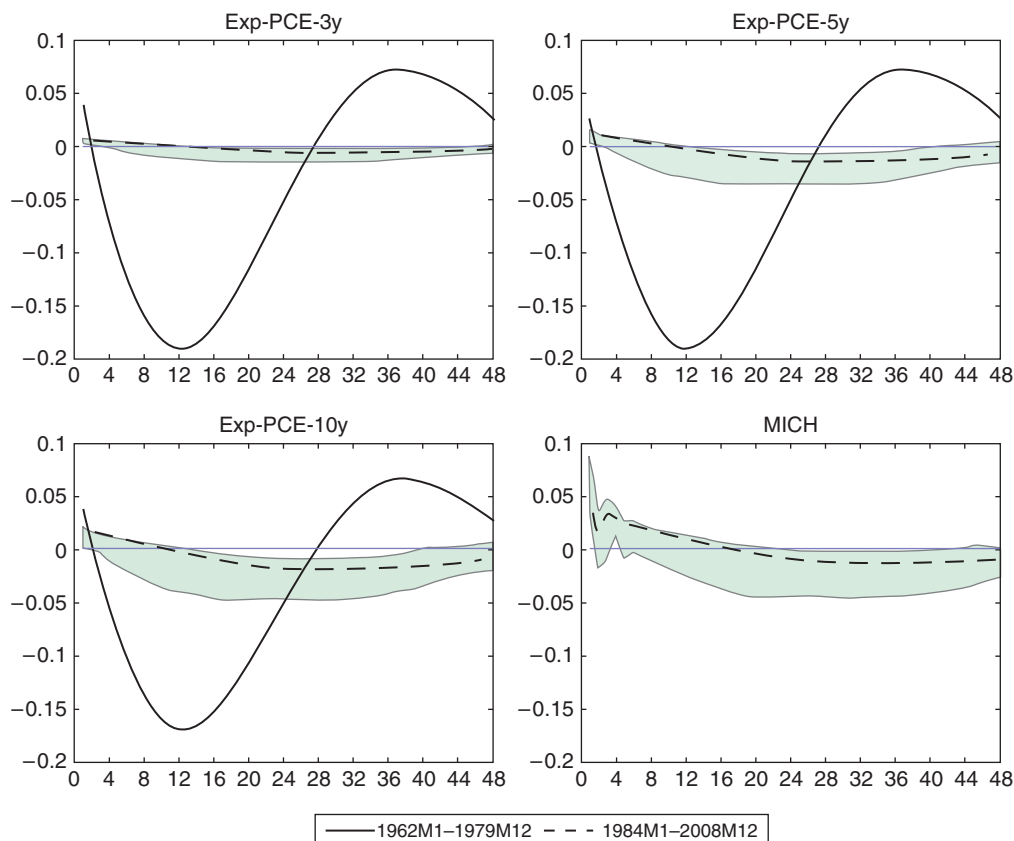


Figure 5 FAVAR-based evidence of the effect of monetary policy on expected inflation. Impulse response functions to a 25 bp surprise increase in the federal funds rate, estimated from the FAVAR model described in the text. Note that the Survey of Professional Forecasters measure of expected inflation is not available at the monthly frequency. Shaded areas represent the 95% confidence interval on the post-1984 estimates.

Results are reported in Figure 6. The overall conclusion that seems to emerge is that the magnitude of response of corporate spreads is somewhat smaller in the first year in the recent period while perhaps also being more persistent during this period. Unfortunately, because of data availability, we do not have an estimate of the response for Gilchrist et al. (2009) measure of external finance premium for the pre-1979:9 period. But its response in the post-1984:1 appears consistent with the one obtained for the other measures.

5.2 Evidence from a completely specified structural model

We now turn to a discussion about how the transmission process may have evolved in a structural model. We employ a relatively standard New Keynesian DSGE

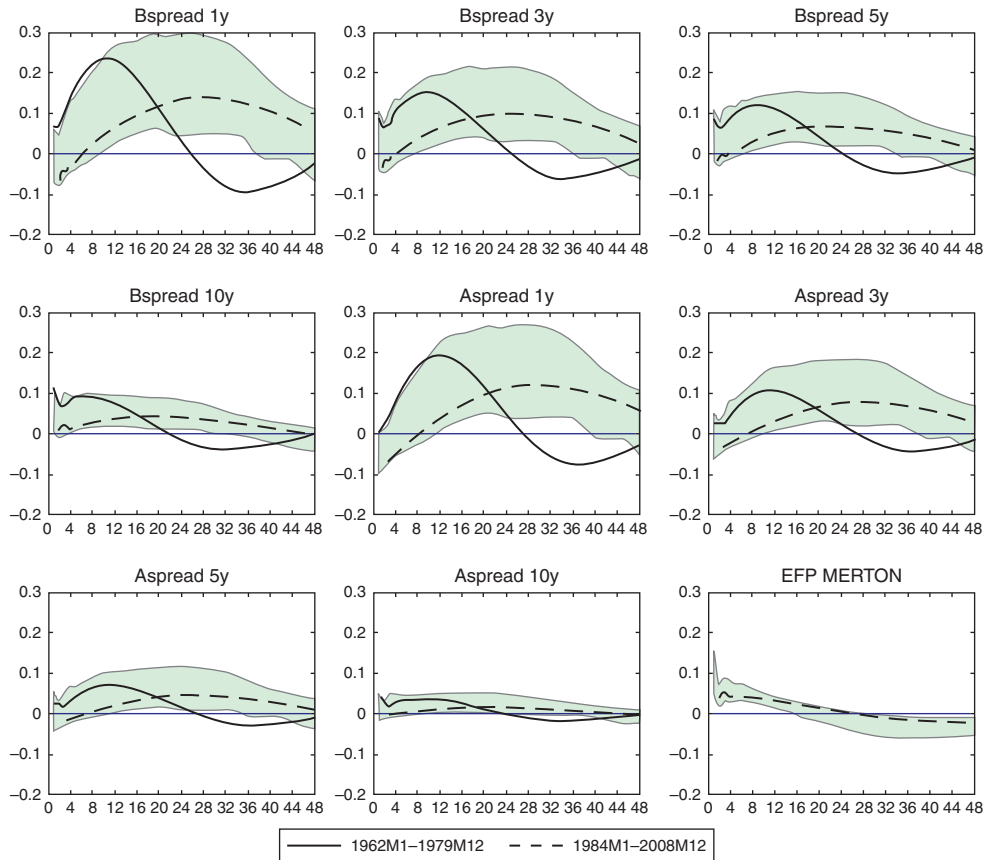


Figure 6 FAVAR-based evidence of the effect of monetary policy on external finance premium. Impulse response functions to a 25 bp surprise increase in the federal funds rate, estimated from the FAVAR model described in the text. Note that the Survey of Professional Forecasters measure of expected inflation is not available at the monthly frequency. Shaded areas represent the 95% confidence interval on the post-1984 estimates.

model. This framework has three key features that allow us to build on our FAVAR-based analysis: it allows a discussion of structural features, including monetary policy behavior; it emphasizes the potential role for expectations management in influencing monetary transmission, as highlighted in the New Keynesian literature; and it is a framework used widely in research and policy environments, as discussed earlier.

5.2.1 The model

The starting point for our specification is the model of Smets and Wouters (2007). We extend the model along two dimensions. First, we disaggregate investment spending

into consumer durable expenditures, residential investment, and business investment, as in the Federal Reserve's EDO DSGE model (Edge et al., 2007, 2008, 2010); such a disaggregation allows our analysis to connect with the large literature we summarized earlier that examines the impact of monetary policy on these spending categories. In addition, we add a financial accelerator, inspired by Bernanke et al. (1999) and following Gilchrist et al. (2009) closely; this addition allows some consideration of a credit (non-neoclassical) channel.

As the basic framework follows these earlier contributions closely, we present the model briefly and in its log-linear form. Table 3 presents the list of model variables.

The IS block of the model consists of the optimality conditions governing consumption ($c(t)$) and investment (in durables, $d(t)$, residential investment, $h(t)$, and nonresidential investment, $i(t)$) decisions.

$$c(t) - \frac{\zeta}{1+\zeta}c(t-1) - \frac{\zeta}{1+\zeta}Ec(t+1) = \frac{1-\zeta}{1+\zeta}[r(t) - E\pi(t+1) + b(t)] \quad (3)$$

$$\Xi(t) = \frac{1}{1+B}\Xi(t-1) + \frac{B}{1+B}E\Xi(t+1) + \frac{1}{1+B}\frac{1}{\Phi_{\Xi}}q_{\Xi}(t) + \varepsilon_{\Xi}(t), \Xi = d, h, i \quad (4)$$

$$Er\mathbf{k}_{\Xi}(t+1) = (1 - \varepsilon_{\Xi})\text{Emp } \mathbf{k}_{\Xi}(t+1) + \varepsilon_{\Xi}E\mathbf{q}_{\Xi}(t+1) - \mathbf{q}_{\Xi}(t), \Xi = d, h, i \quad (5)$$

$$mpk_{\Xi}(t) = -k_{\Xi}(t-1) + \frac{1}{1+\zeta}(c(t) - \zeta c(t-1)), \Xi = d, h \quad (6)$$

$$mpk_i(t) = -k_i(t-1) - z(t) + l(t) + w(t) \quad (7)$$

These equations are standard. Consumption depends upon future and past consumption (where habit persistence yields the inclusion of the latter), the policy interest rate ($r(t)$) minus expected inflation ($\pi(t+1)$) and a risk premium shock ($b(t)$); investment in each category depends upon $q(t)$ for the relevant type of capital (and an i.i.d. shock to the q /investment relation ($\varepsilon_{\Xi}(t)$ for durables, residential investment, and nonresidential investment); and q is a function of the risk-premium adjusted short rate ($rk(t)$) and the marginal product ($mpk(t)$) of the associated type of capital — where these marginal products are determined by the economy's production function for business capital (which includes variable utilization, $z(t)$) and the households preferences for consumer durable and residential capital. Of particular interest are the capital adjustment cost parameters (Φ), as these determine the short-run elasticities of investment expenditures to q and hence govern (in part) the responsiveness of such expenditures to monetary policy.

The financial block consists of the equations determining the endogenous risk premia and the evolution of the net worth of the agents who finance investment projects. Following the Bernanke et al. (1999) financial accelerator framework, these premia

Table 3 Variables in DSGE Model
Endogenous variables in DSGE model

Variable name	Symbol
Expenditure components and GDP	
Consumption (ex. durables)	$c(t)$
Consumption, durables	$d(t)$
Residential investment	$h(t)$
Business investment	$i(t)$
GDP (output)	$y(t)$
Productive inputs and household stocks	
Business, residential, and durables capital	$k_i(t)$
Hours worked	$l(t)$
Capital utilization	$z(t)$
Financial market variables	
Tobin's q — business, residential, and durables capital	$q_i(t)$
Marginal product — business, residential, and durables capital	$mpk_i(t)$
Return to business, residential, and durables capital	$rk_i(t)$
Policy interest rate	$r(t)$
Long-term interest rate	$r_l(t)$
Long-term expected return to business, residential, and durables capital	$rk_{l,i}(t)$
Net worth to finance business, residential, and durables capital	$n_i(t)$
Inflation and wages	
Inflation	$\pi(t)$
Real wage	$w(t)$
Expenditure components and GDP	
Residual demand	$g(t)$
Productive potential	
Productivity	$a(t)$
Financial market	
Economy-wide risk premium	$b(t)$

Table 3 Variables in DSGE Model—cont'd
Endogenous variables in DSGE model

Variable name	Symbol
Spread (exog.) for business, residential, and durables	spread(t)
Term premium	tp(t)
Shocks in DSGE model	
Expenditure components and GDP	
Residual demand	$e_g(t)$
Productive potential and markups	
Productivity	$e_a(t)$
Price markup	$e_\pi(t)$
Financial market	
Monetary policy	$e_r(t)$
Economy-wide risk premium	$e_b(t)$
Spread (exog.) for business, residential, and durables	$e_{spread}(t)$
Term premium	$e_{tp}(t)$
Investment shocks — business, residential, and consumer durables	$e_i(t)$

depend upon the net worth of the agents financing such projects ($n(t)$) and the amount of capital expenditures financed (i.e., on leverage). We arbitrarily assume that each type of project is financed by a different class of entrepreneurs, implying that the risk premia are specific to each investment type; a more natural framework may have been to have a set of financing constraints jointly influencing all household expenditures, as in models like Iacoviello (2005).

$$Erk_{\Xi}(t+1) - [r(t) - E\pi(t+1) + b(t)] = v_{\Xi}[q_{\Xi}(t) + k_{\Xi}(t) - n_{\Xi}(t)] + spread_{\Xi}(t), \quad (8)$$

$$\Xi = d, h, i$$

$$n_{\Xi}(t) = \frac{K_{\Xi}}{N_{\Xi}} rk_{\Xi}(t) - \left(1 - \frac{K_{\Xi}}{N_{\Xi}}\right) E_{t-1} rk_{\Xi}(t) + \theta n_{\Xi}(t), \quad \Xi = d, h, i \quad (9)$$

The spread terms represent exogenous movements in the risk premia associated with financing investment. In these equations, the parameters v , which govern the sensitivity of the external finance premia to variations in the leverage associated with each type

of investment ($q + k - n$), provide the only non-neoclassical channel in this model. As a result, pure credit-type channels are not present, and we will highlight the implications of this absence for our empirical findings and subsequent research.

The supply block consists of the resource constraints — the GDP identity, the production function depending on business capital, hours, and utilization; an optimality condition for capital utilization; and the capital accumulation equations — and Phillips curves determining price and wage inflation.

$$y(t) = c_y c(t) + d_y d(t) + h_y h(t) + i_y i(t) + g_y g(t) \quad (10)$$

$$y(t) = a[k(t-1) + z(t)] + (1-a)l(t) + a(t) \quad (11)$$

$$z(t) = \frac{1-\psi}{\psi} mpk_i(t) \quad (12)$$

$$k_{\Xi}(t) = (1 - \delta_{\Xi})k_{\Xi}(t-1) + \delta_{\Xi}\Xi(t), \Xi = d, h, i \quad (13)$$

$$\pi(t) = \frac{1}{1+B}\pi(t-1) + \frac{B}{1+B}E\pi(t+1) - \frac{1}{1+B}K[y(t) - l(t) - w(t)] \quad (14)$$

$$w(t) - \pi(t) = \frac{1}{1+B} \left[w(t-1) + \pi(t-1) + \frac{B}{1+B}E \left[w(t+1) + \pi(t+1) \right] \right. \\ \left. + \frac{1}{1+B} \kappa \left[\frac{1}{1+\zeta} (c(t) - \zeta c(t-1)) + l(t) - w(t) \right] \right] \quad (15)$$

The nominal interest rate is set by the monetary authority according to a simple policy rule involving price inflation and a traditional output gap, defined as the deviation of output from the level consistent with labor input and utilization at their long-run levels.

$$r(t) = \rho_r r(t-1) + (1 - \rho_r)(r_y[y(t) - a(t) - ak(t-1)] + r_{\pi}\pi(t)) + \varepsilon_r(t) \quad (16)$$

Finally, we include an equation for a long-term interest rate ($r_l(t)$), based on the expectations hypothesis and an exogenous term premia ($tp(t)$). We also consider an expectations-hypothesis-based equation for a long-term bond associated with entrepreneurs financing of investment, as the data we will use on interest-rate spreads are based on long-term debt:

$$r_l(t) = B_r E r_l(t+1) + (1 - B_r)r(t) + tp(t) \quad (17)$$

$$Er_{k_{\Xi,1}}(t+1) = B_r Er_{k_{\Xi}}(t+1) + (1 - B_r)Er_{k_{\Xi,1}}(t+2) \quad (18)$$

We estimate the model for two periods, 1962Q1–1979Q3 and 1984Q1–2008Q4; these samples are two of the periods we have emphasized in our VAR analysis. We use twelve data series: the growth rates (in real terms) of GDP, nondurables and

services consumption, durable consumption, residential investment, and nonresidential investment; detrended hours per capita; GDP price inflation; the nominal federal funds rate; the nominal yield on a 10-year Treasury; and external finance premia measured as the difference between a composite yield on corporate BBB bonds and the 10-year Treasury, the difference between a mortgage rate and the 5-year Treasury, and the difference between the interest rate on automobile loans and the 5-year Treasury.

Table 4 presents calibrated parameters. We choose conventional values: expenditure shares for consumption of nondurables and durables of about 2/3, residential and business investment shares of about 4 and 12%, and a residual demand (e.g., government) expenditure share of 18%; a quarterly discount rate of 1%; a depreciation rate for consumer durables double that of business investment and quadruple that of residential investment; a leverage rate in the financial accelerator of 2; a Phillips curve slope just below 0.1; a capital share in production of 35%; and other parameters (governing utilization and capital returns) similar to values from, for example, Gilchrist et al. (2009). Table 5 presents prior distributions over the estimated parameters most critical for the effects of monetary policy — the adjustment cost parameters determining the sensitivity of investment spending to fundamentals (i.e., q), the parameters determining the sensitivity of risk premia to leverage, and the parameters in the monetary policy rule — along with estimates for the posterior mode and their standard deviations for each sample. (A more complete description of our estimation approach and results are presented in the appendix in this chapter.)

Three points are evident from these estimation results. First, the parameters of the monetary policy rule are substantially more reactive to inflation (r_π) and output (r_y) in the 1984Q1–2008Q4 sample. Second, the standard deviations of the exogenous shock processes are in several cases, including the monetary policy rule (σ_r), lower in the 1984Q1–2008Q4 sample. Both of these results suggest better policy behavior and echo findings in other studies, most notably Clarida, Gali, and Gertler (2000). Third, the parameters governing the shape of the investment demand schedules (Φ_Ξ) and the response of risk premia to economic conditions (v) are only modestly different across samples. Combining these results, the most significant changes appear to be in monetary policy behavior, not private-sector parameters. We now turn to the implications of these results for the evolution of monetary transmission.

5.2.2 Changes in monetary transmission

Our FAVAR analysis yielded three conclusions: The effect of monetary policy on output appears somewhat smaller at a one-to-two year horizon in the most recent sample, but the response of output at more distant horizons is not lower and may be more persistent, although standard errors are large at such horizons; inflation responds less to monetary policy in the recent sample; and credit and risk spreads may respond to policy

Table 4 Calibrated Parameters for DSGE Model

c_y	0.59
d_y	0.075
h_y	0.04
i_y	0.115
g_y	0.18
K/N (for d,h,i)	2
B	0.99
α	0.35
δ_i	0.025
δ_h	0.05
δ_d	0.0125
κ	0.08
$RK_x, x=d,h,i$	$1/B - (1 - \delta_x)$
$\varepsilon_x, x=d,h,i$	$RK_x/(RK_x+1 - \delta_x)$
θ	0.95
ψ	0.5

Table 5 Estimated Parameters for DSGE Models

	Prior Distribution	Posterior 1966q1–1979q3				Posterior 1984q1–2008q4	
		Mean	S.D.	Mode	S.D.	Mode	S.D.
ζ	Beta	0.75	0.20	0.79	0.10	0.35	0.12
Φ_i	Normal	2.00	0.50	2.14	0.52	0.57	0.40
Φ_d	Normal	2.00	0.50	3.08	0.39	2.71	0.41
Φ_h	Normal	2.00	0.50	2.95	0.53	0.83	0.25
v_i	Beta	0.05	0.01	0.05	0.10	0.05	0.01
v_d	Beta	0.05	0.01	0.06	0.03	0.05	0.01
v_h	Beta	0.05	0.01	0.04	0.02	0.04	0.01
ρ_r	Beta	0.75	0.10	0.72	0.00	0.72	0.05
r_π	Normal	1.50	0.25	1.00	0.00	1.71	0.23
r_y	Normal	0.50	0.25	0.01	0.00	0.20	0.03
σ_r	Invg	0.10	2.00	0.20	0.01	0.08	0.02

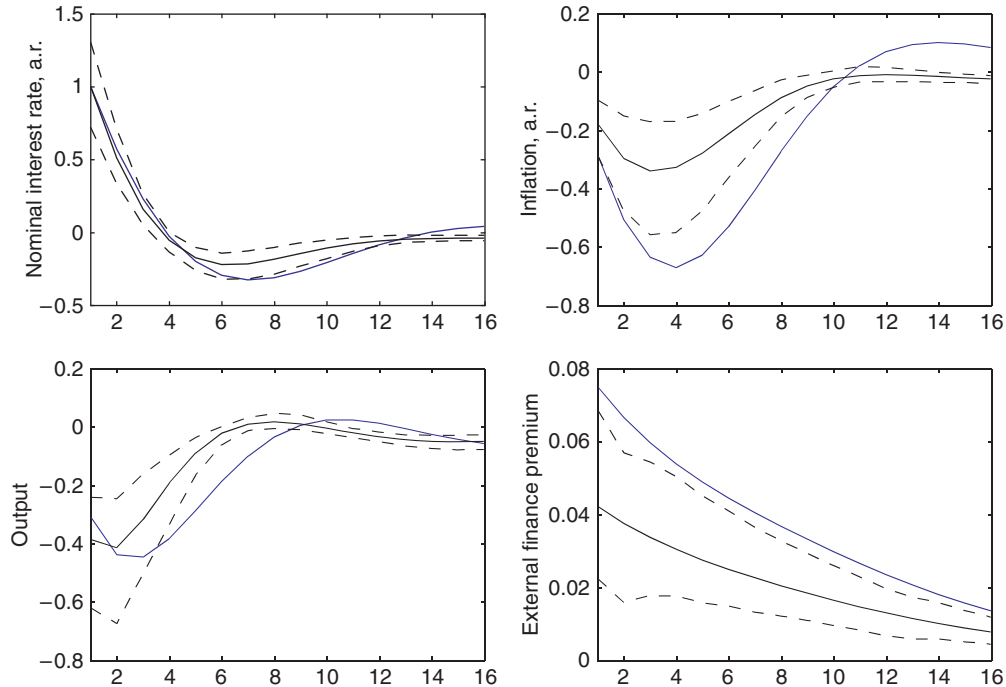


Figure 7 DSGE model-based evidence of the effect of monetary policy in two sample periods. Impulse response functions to a 100 bp (a.r.) surprise increase in the fed funds rate in the DSGE model described in the text. The grey, solid line is the response at the 1962:1–1979:3 sample period parameter estimates; the black line is the response for the 1984:1–2008:4 sample period; and the black, dashed lines are the 90% credible set around these estimates. The units on the x-axis represent quarters.

actions less in recent samples in the short run, although responses are estimated in the FAVAR to follow a more drawn out trajectory in the recent sample.

Figure 7 presents the DSGE-based impulse responses following a 100 basis point increase (annual rate) in the federal funds rate (e.g., 25 basis points at a quarterly rate) for the two sample periods, along with the 90% credible set (the dashed lines) around the 1984:1–2008:4 sample period response, for the federal funds rate, inflation, output, and the credit spread associated with business investment (where the comparable data is the 10-year BBB corporate bond spread from Figure 6). The results conform reasonably closely with the FAVAR results: inflation responds much less to the policy innovation in the recent sample (the black line) relative to the response in the 1962:1–1979:3 sample (the grey line); output responds less to a shock to the federal funds rate in the 1984:1–2008:4 sample (the black line) than in the earlier sample (the grey line), especially at horizons from one to two years; and the risk spread response is also more modest in the recent sample. (However, we should emphasize

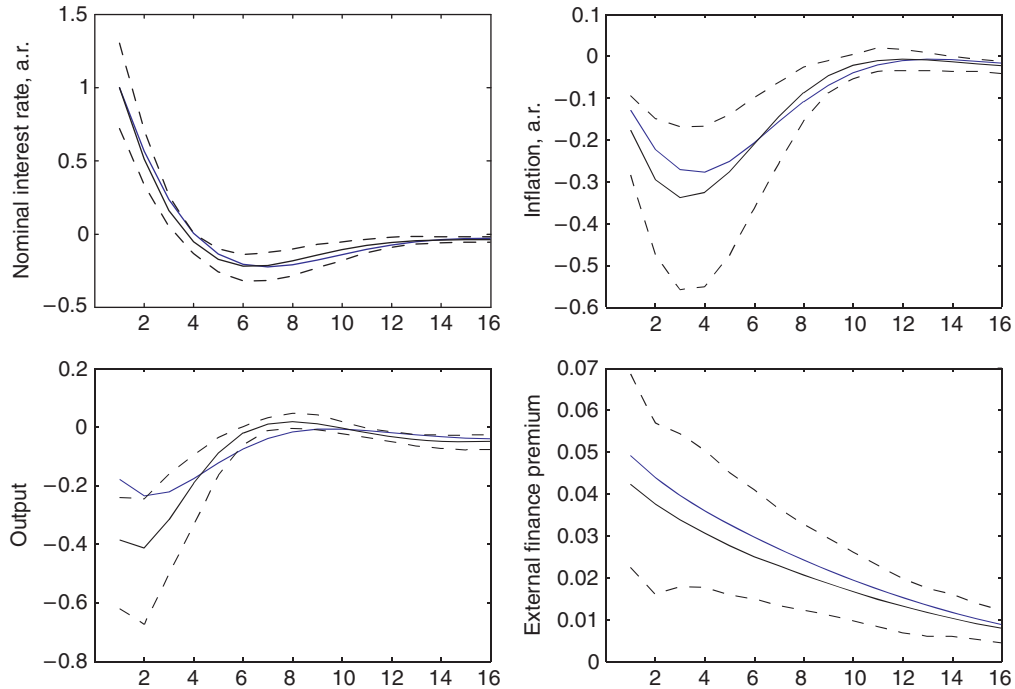


Figure 8 DSGE model-based evidence of the effect of monetary policy in two sample periods, change in policy parameters. Impulse response functions to a 100 bp (a.r.) surprise increase in the federal funds rate in the DSGE model described in the text. The grey, solid line is the response at the 1962Q1–1979Q3 sample period parameter estimates, with the monetary policy parameters set to their 1984Q1–2008Q4 sample period estimates; the black line is the response for the 1984Q1–2008Q4 sample period; and the black, dashed lines are the 90% credible set around these estimates. The units on the x-axis represent quarters.

that there are important differences from the FAVAR responses — most notably that inflation and output jump following a policy innovation, whereas the identifying assumption underlying the FAVAR responses excludes this possibility).

Because all changes in the model arise from a change in some structural parameter, the estimates for each sample can be used to identify the source of these shifts in policy transmission. The first candidate is the changes in the monetary policy parameters. Figure 8 presents the response for the 1984Q1–2008Q4 sample and the response that would arise using the 1984Q1–2008Q1 policy parameters with all of the other structural parameters at the values estimated for the earlier sample. As can be clearly seen, the shift in policy parameters brings the responses closely in line, indicating that the change in monetary policy behavior can account for the changes in the responses of inflation, output, and risk spreads.

Indeed, the changes in other parameters imply little change in the responses of inflation, activity, and risk spreads to a policy innovation. For example, Figure 9 presents the 1984Q1–2008Q4 responses and credible sets and the response that would arise using the 1984Q1–2008Q1 parameters governing the sensitivity of risk spreads (ν) with all of the other structural parameters at the values estimated for the earlier sample; Figure 10 presents the 1984Q1–2008Q4 responses and credible sets and the response that would arise using the 1984Q1–2008Q1 parameters governing the slope of the investment demand schedules (Φ_{Ξ}) with all of the other structural parameters at the values estimated for the earlier sample. In each case, the differences in impulse responses across the two samples remain intact, implying that little of the change in the responsiveness of activity, inflation, or risk spreads stems from private-sector behavior.

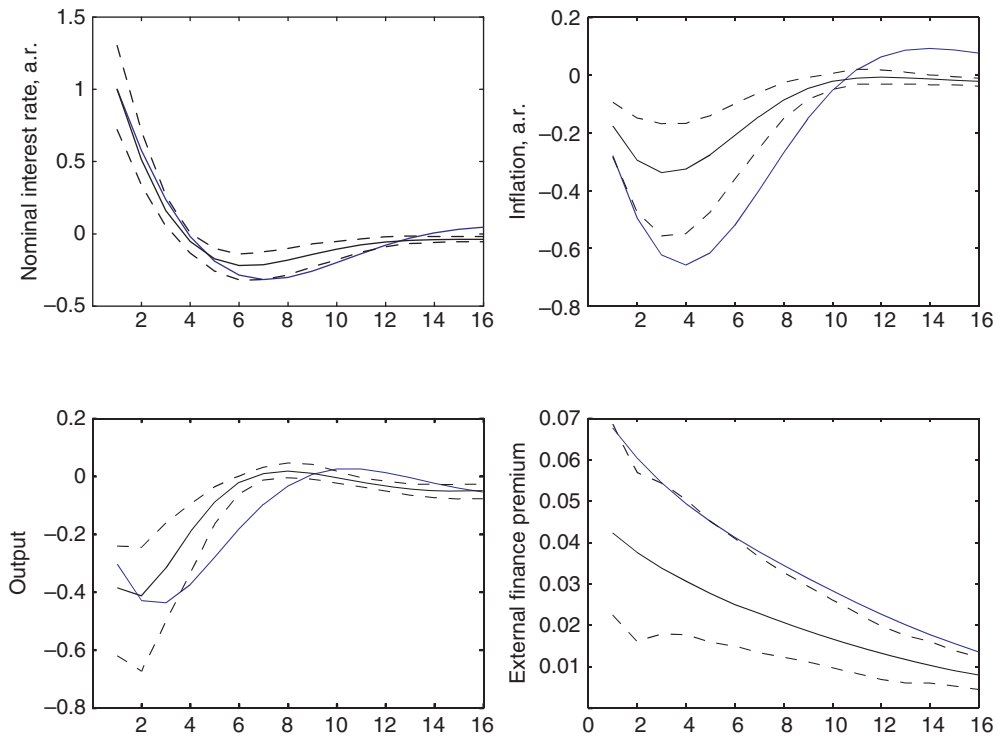


Figure 9 DSGE model-based evidence of the effect of monetary policy in two sample periods, change in risk-premia (financial accelerator) parameters. Impulse response functions to a 100 bp (a.r.) surprise increase in the federal funds rate in the DSGE model described in the text. The grey, solid line is the response at the 1962Q1–1979Q3 sample period parameter estimates, with the risk-premia/financial accelerator parameters set to their 1984Q1–2008Q4 sample period estimates; the black line is the response for the 1984Q1–2008Q4 sample period; and the black, dashed lines are the 90% credible set around these estimates. The units on the x-axis represent quarters.

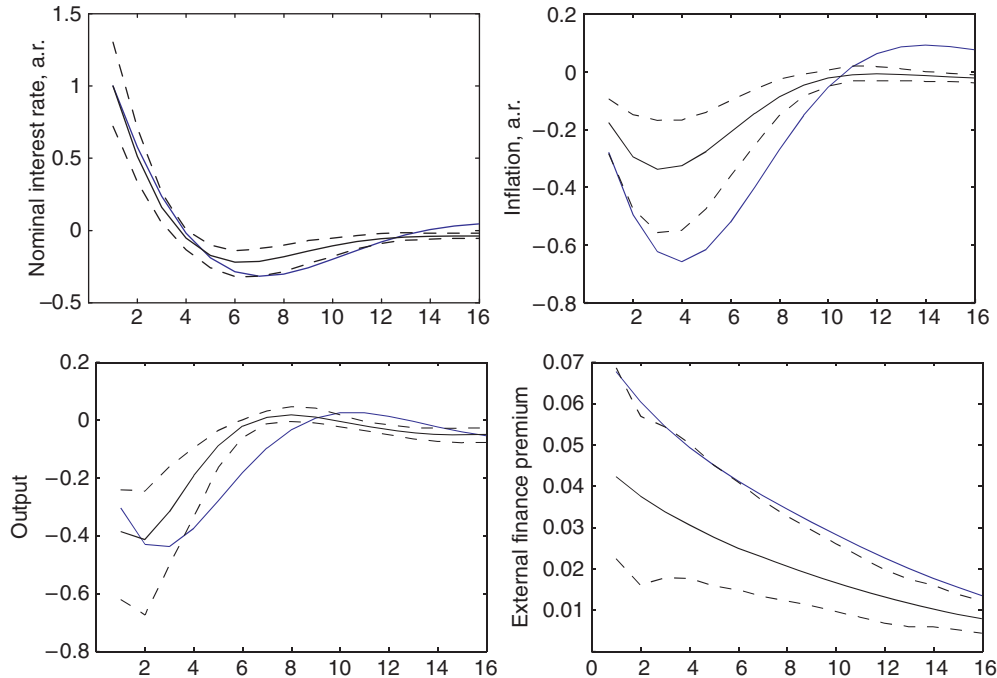


Figure 10 DSGE model-based evidence of the effect of monetary policy in two sample periods, change in investment demand schedule (adjustment cost) parameters. Impulse response functions to a 100 bp (a.r.) surprise increase in the federal funds rate in the DSGE model described in the text. The grey, solid line is the response at the 1962 1–1979Q3 sample period parameter estimates, with the adjustment cost parameters set to their 1984Q1–2008Q4 sample period estimates; the black line is the response for the 1984Q1–2008Q4 sample period; and the black, dashed lines are the 90% credible set around these estimates. The units on the x-axis represent quarters.

Finally, we should emphasize that the shifts in monetary policy behavior we detect are similar to the findings in many other studies (e.g., Boivin & Giannoni, 2006). With that said, this finding has attracted some controversy; for example, Sims and Zha (2006) followed a less structured approach than ours and showed that, under some assumptions, the finding of a shift in monetary policy behavior, other than the variance of the shock, is not always clear. Moreover, Smets and Wouters (2007), employing a similar methodology to our DSGE approach, found no changes in monetary policy behavior. In part this difference likely reflects their specification of the monetary policy rule (which responds to the deviation of output from flexible price output, rather than from a production-function gap as we choose) and other aspects of their specification. Indeed, any analysis with a structural model will be impacted by all the model's assumptions. With that said, other DSGE-based analyses reach similar conclusions regarding monetary policy behavior (e.g., Arestis, Chortareas, & Tsoukalas, 2010). More generally, our findings are in line with

those from the policy-rule literature, most notably Taylor (1999) and Clarida et al. (2000), giving us confidence in our qualitative conclusions.

5.2.3 *Signs of changing credit conditions*

A straight read of these results might suggest that there have not been significant changes over time in the importance of non-neoclassical channels, as the parameters of the financial accelerator in our DSGE model have not changed in a way notable enough to have effects on the economy's response to monetary policy shifts. However, we interpret our estimation results as suggesting that the most widely adopted version of this channel in quantitative macroeconomic models, the Bernanke et al. (1999) financial accelerator mechanism, does not provide much information on such changes. One possible reason for this is that the effects of financial accelerator-type mechanisms are nonlinear, and not picked up in the linear framework we consider.²⁰ Another possibility is that the financial accelerator framework largely works through an amplification mechanism associated with external finance premia, and ignores other aspects of credit provision.

Indeed, one of the results from our DSGE analysis points in this direction. Specifically, an economically significant change related to residential investment is evident in the change in the standard deviation of the shock to its Eq. (4) (which, as shown in Appendix, drops from just above 2 to a minuscule value near 0), implying that more of the fluctuations in residential investment in the recent sample represent movements along this curve, rather than deviations from the model's implied relationship. Of course, the model does not include the quantity rationing induced by regulation or other non-price channels of monetary transmission, and hence this type of model is perhaps not especially informative about shifts in such channels across the sample periods we consider. Nonetheless, the smaller variance of shocks to this relationship in the 1984Q1–2008Q4 sample period is consistent with the findings in McCarthy and Peach (2002), who reported a closer association of housing market variables with interest rates and other neoclassical fundamentals for that period.

Indeed, we can even be a little bit more suggestive; for example, periods of credit rationing associated with disintermediation from falling deposits at savings and loans associations during the 1962Q1–1979Q3 period were identified in Brayton and Mauskopf (1985) to have occurred in 1966Q3–Q4, 1969Q3–1970Q3, and 1974Q1–1975Q1. For these periods, the mean “off-equation” movements in residential investment (e.g., shock) equaled -3% (with a t-statistic of -4.3); in comparison, the mean of such shocks in periods without credit constraints was 0.7% (with a t-statistic of 2.3).²¹ In other words, these shocks appear tightly linked to credit conditions, with credit rationing especially important for residential investment. Moreover,

²⁰ Gilchrist, Ortiz, and Zakrajsek (2009) found very small accelerator effects following a monetary policy innovation for their estimated DSGE model; much of the importance of financial shocks in their results stems from the exogenous shocks to the financial sector, rather than through the endogenous propagation mechanism.

²¹ This shock process is estimated, implying a mean of about zero (0.1; and a small t-statistic for the entire sample of 0.4).

the decline in the importance of this shock in the 1984Q1–2008Q4 sample points to a lessening in the importance of credit per se.

5.2.4 Monetary policy and the transmission of other shocks

Our analysis highlights the central role of monetary policy behavior in the evolution of the transmission of monetary shocks.

The endogenous aspect of monetary policy is crucial for the behavior of activity and inflation following other shocks, which implies that the shift in behavior on the part of monetary policymakers could have even more significant effects on the nature of economic fluctuations through its impact on the effect of fundamentals other than policy shocks. We consider the response of the policy interest rate, output, and inflation to a one standard deviation shock to productivity and to the economy-wide risk premia; these shocks are the most important shocks for output fluctuations in our model and in other similar DSGE models (e.g., Smets & Wouters, 2007, and especially to Federal Reserve and ECB policy models described in Kiley, 2010, and Christoffel et al., 2008). Moreover, these two shocks illustrate the nature of policy trade-offs facing policymakers: An improvement in productivity increases output and lowers inflation, and a policymaker will hope to accommodate the improvement in output and stabilize inflation in response to such a shock; in contrast, an increase in the risk premia will depress output and inflation, and a policymaker will aim to offset these effects and stabilize both output and inflation. In our comparisons, we compare the impulse responses using the 1984Q1–2008Q4 parameter estimates and altering the policy parameters to those for the earlier period in our alternative case.

Figure 11 presents the impulse responses following a productivity shock. The differences implied by the change in policy rule are dramatic — output responds more, and inflation less, to the innovation in productivity. Figure 12 presents the results for a risk premia shock. In this case, the more active policy response to inflation in the recent sample period serves to stabilize both activity and inflation. These figures show that the most important impact of monetary policy is through its affect on how other shocks are transmitted to inflation and activity. And the nature of our findings — that monetary policy has shifted since the early 1980s to a stance that accommodates productivity innovations and resists demand-side fluctuations — is consistent with other results in the literature, most clearly those of Galí and Gambetti (2009), who investigated similar issues using a less structural (VAR-based) approach.

In both cases, the changes in the responses to shocks move in the direction of more desirable economic outcomes. And in both cases, the large effects of the change in the policy rule are driven by *expectations*. This can be seen by looking at the response of the nominal interest rate, which is actually *smaller* under the more active policy in both cases, because the commitment to a highly reactive policy stabilizes expectations and hence actual outcomes without generating additional *realized* volatility in the nominal interest rate.

These results drive home two points that are central to the evolution of economist's understanding of the monetary transmission mechanism.

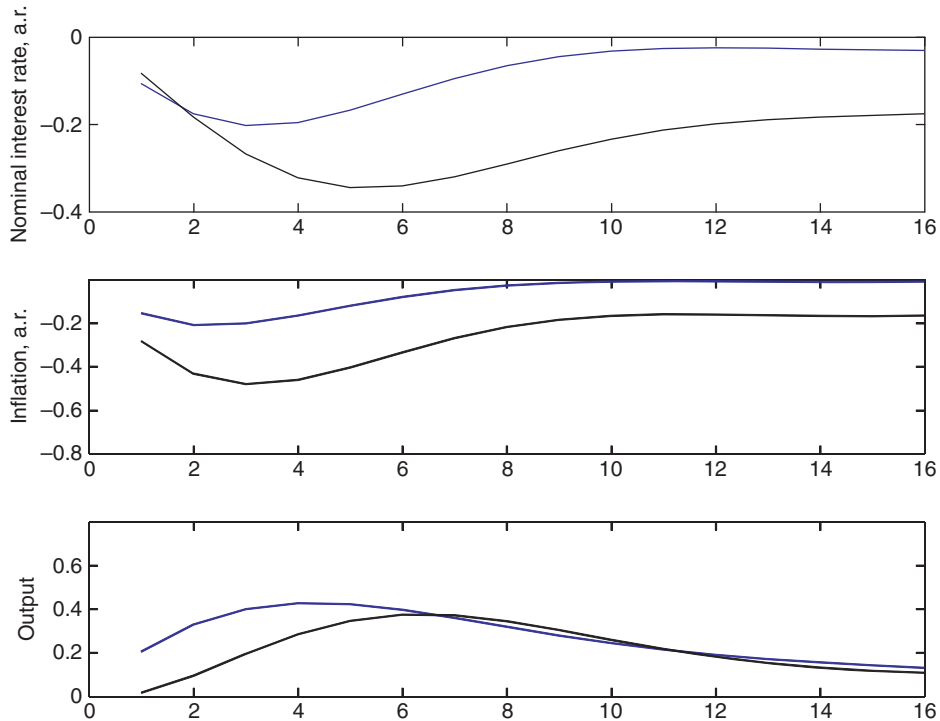


Figure 11 DSGE model-based evidence of the effect of monetary policy: Change in policy-rule parameters and productivity shock. Impulse response functions to a one-standard deviation surprise increase in productivity. The grey, solid line is the response at the 1984Q1–2008Q4 sample period parameter estimates and the black, solid line is the response if the parameters of the monetary policy rule are set to the values estimated for the 1962Q1–1979Q3 sample period. The units on the x-axis represent quarters.

First, the systematic component of policy and its effect on the macroeconomic response to a wide range of shocks is the principle mechanism through which monetary policy affects inflation and activity. This transmission channel is the primary focus of modern studies of the effects of monetary policy, following the large literature on the effects of policy rules on economic performance (e.g., the literature summarized in Chapter 15 on policy rules by Taylor & Williams, 2010) and the emphasis on managing expectations through systematic behavior presented in Woodford (2003). This evolution is significant, as the primary focus of policy discussions of the transmission mechanism in the past has centered on model analyses or simulations that focus on exogenous paths for policy with at most glancing attention to the systematic nature of policy, expectations formation, and the transmission of policy actions. (Examples that did not emphasize expectations include the equation-by-equation approach of Akhtar & Harris, 1987; Friedman, 1989; analyses with the Federal Reserve's MPS model; Mauskopf, 1990; or the central bank comparisons, representing a large number of policy models employed in the early-to-mid 1990s presented in Smets, 1995).

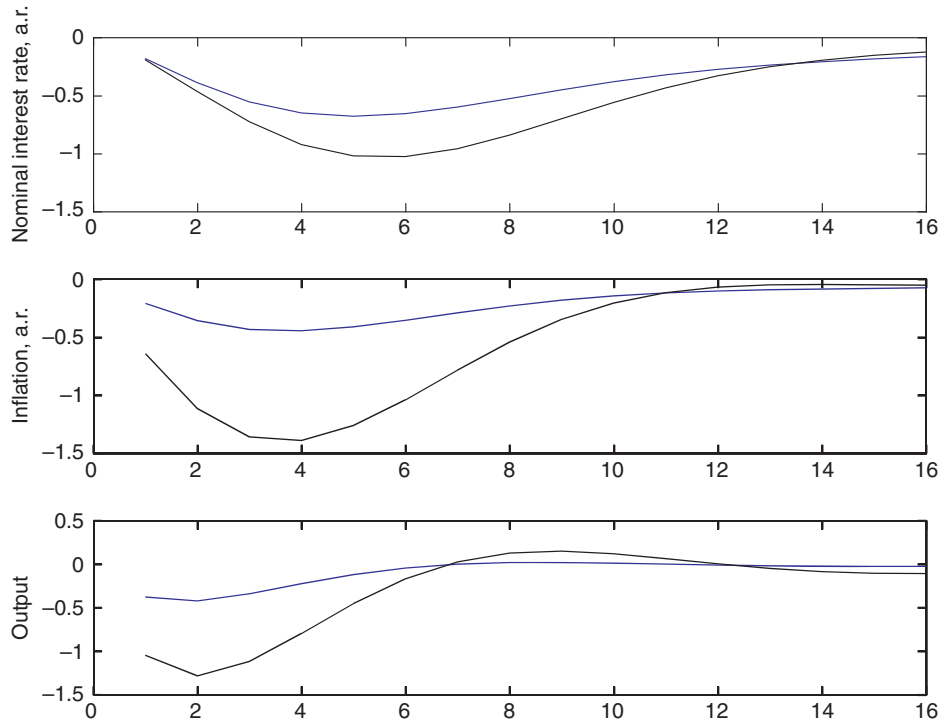


Figure 12 DSGE model-based evidence of the effect of monetary policy: Change in policy-rule parameters and risk-premium shock. Impulse response functions to a one standard deviation surprise increase in the economy-wide risk premium. The grey, solid line is the response at the 1984Q1–2008Q4 sample period parameter estimates and the black, solid line is the response if the parameters of the monetary policy rule are set to the values estimated for the 1962Q1–1979Q3 sample period. The units on the x-axis represent quarters.

Second, a greater emphasis on inflation stabilization is likely to lead to greater stability in inflation but not necessarily in output, as a focus on price stability will accommodate increases in output reflecting productivity advances and resist such movements due to fluctuations in risk premia or some other demand factors. This latter point suggests that studies that look to identify the importance of changes in the transmission mechanism related to policy behavior through the lens of overall output stability may fail to find strong evidence. This may help explain, in part, the diversity of findings in this area (e.g., the different conclusions of Boivin & Giannoni, 2006; Canova & Gambetti, 2009). The subtle difference between overall output stability and stability in output around an efficient level; that is, the notion that policymakers should design policy to accommodate productivity movements while resisting inefficient movements due to risk premia, has also represented an important evolution in understanding regarding how the monetary transmission mechanism should be used to promote price stability.

6. IMPLICATIONS FOR THE FUTURE CONDUCT OF MONETARY POLICY

Looking back over our summary of related literature, four findings are apparent.

First, the neoclassical channels — direct interest rate effects on investment spending, wealth, and intertemporal substitution effects on consumption, and the trade effects through the exchange rate — have remained the core channels in macroeconomic modeling. The literature on time variation in the strength of these channels has not suggested large changes over time.

Second, the macroeconomic literature on non-neoclassical channels in general equilibrium models is sparse. Most analyses of the potential importance of, for example, bank-based channels have focused on heterogeneous effects on different classes of borrowers or lenders that could signal a potential role for such channels, without moving on to the macroeconomic consequences. Macroeconomic models that incorporate such channels, most notably a balance sheet channel like that of Bernanke et al. (1999), find only modest effects on monetary transmission from these factors, as we found in our DSGE model exercises. Indeed, the variation in external finance premia in such empirical models seems more important as a source of shocks driving fluctuations than as an endogenous transmission mechanism.

Third, there have been large changes in the regulatory structure in the United States and other countries, and these changes have had important implications for the transmission of monetary policy actions to residential investment. In particular, residential investment is now more tied to interest rates rather than credit availability. Some aspects of these shifts are apparent in our macro-based approaches. For example, credit seems to respond more slowly and by a smaller amount to policy shifts in our FAVAR analysis (Figure 5) in the period after 1982; similarly, shocks to the residential investment equation in our DSGE model are tightly linked to periods of credit rationing in the pre-1980 period, but of minimal importance in the later sample. With that said, these results only hint at the role of credit *per se* in different periods, and fail to speak to the more global issue of the role of financial frictions in economic fluctuations.

Finally, monetary policy has become substantially more focused on inflation stabilization, and this shift has importantly affected the volatility of inflation and the response of output to non-monetary disturbances. Indeed, the systematic component of monetary policy is the most important monetary factor in economic fluctuations, and the evolution of economist's understanding of how to use the monetary transmission mechanism through a systematic focus on price stability is one of the central shifts in policy behavior and macroeconomic modeling over the last quarter century. These results emerge quite clearly from our structural model analysis, and echo findings from some studies that impose less structure.

This summary leaves two extremely important outstanding questions for research. One is the role of non-neoclassical channels in our understanding of economic fluctuations and monetary policy. The literature in this area remains thin, and this thinness reflects difficulty in specifying the relevant mechanisms and finding the supporting

empirical evidence. While we are able to hint at the importance of such channels at times in the past, this area is currently very active and will undoubtedly yield future insights (e.g., Angeloni & Faia, 2009; Gerali et al., 2009; Gertler & Kiyotaki, 2010; Meh & Moran, 2008). Indeed, the global financial crisis that began in mid-2007 illustrated that the intersection of banking, finance, and macroeconomics is as important as ever.

The course of policy following the crisis has also shown the importance of understanding aspects of the monetary transmission mechanism further. In particular, policy rates have been brought to near zero in the United States, Europe, and Japan. As a result, the importance of managing expectations in such an environment has been brought to the fore. In addition, central banks around the world have engaged in “quantitative easing” or “large-scale asset purchases” in an effort to impart additional impetus to activity. But both the empirical and theoretical channels (e.g., Bernanke, Reinhart, and Sack, 2004; Clouse et al., 2000) associated with such actions remain far less developed than desirable. Also, some of the policy recommendations in reaction to the crisis, such as the implementation of macroprudential regulations, are likely to have an important impact on the evolution of the monetary transmission mechanism going forward.

APPENDIX

Estimation of the DSGE Model

The DSGE model presented in the main text is estimated using Bayesian methods using the observable variables mentioned in the text: the real growth rates of GDP; nondurables and services consumption (excluding housing services); residential investment; nonresidential fixed investment; the percent change in the GDP deflator; hours worked in the non-farm business sector (divided by the civilian noninstitutional population, and detrended with the HP filter); the nominal federal funds rate; the nominal yield on the 10-year Treasury; and risk spreads on corporate bonds, the car loan rate, and the fixed mortgage rate. The estimated parameters are found by maximizing the log posterior function, which combines the prior information on the parameters with the likelihood of the data. We assume a small amount of measurement error on all the observable data, except the nominal funds rate and nominal Treasury yield. (The degree of assumed measurement error is reported in the Appendix.)

Each of the exogenous processes follow autoregressive, AR(1), processes. The estimated AR(1) coefficients and standard deviations for all of the shocks to these processes are also presented in Table A1 for the two subsamples considered in the text.²²

²² The 1984Q1–2008Q4 sample is conditioned on observations for 1983Q1–1983Q4 (to initialize the Kalman filter); these observations are not used in the computation of the likelihood. The first year of observations condition the filter for the 1962Q1–1979Q3 sample; data availability do not allow conditioning on earlier data for that case.

Appendix Additional Estimated Parameters for DSGE Model

	Prior Distribution	Mean	S.D.	Posterior 1962Q1–1979Q3		Posterior 1984Q1–2008Q4	
				Mode	S.D.	Mode	S.D.
ρ_a	Beta	0.50	0.20	0.89	0.13	0.91	0.02
ρ_{tp}	Beta	0.50	0.20	0.46	0.00	0.84	0.03
ρ_b	Beta	0.50	0.20	0.65	0.41	0.88	0.03
ρ_g	Beta	0.50	0.20	0.45	0.37	0.50	0.25
ρ_{fi}	Beta	0.50	0.10	0.65	1.52	0.53	0.09
ρ_{fd}	Beta	0.50	0.10	0.56	0.10	0.57	0.09
ρ_{fh}	Beta	0.50	0.10	0.49	0.61	0.75	0.09
σ_a	Invg	0.10	2.00	0.73	0.19	0.36	0.08
σ_π	Invg	0.10	2.00	0.23	0.04	0.18	0.03
σ_b	Invg	0.10	2.00	0.87	1.77	0.16	0.04
σ_g	Invg	0.10	2.00	0.78	0.88	0.05	0.02
σ_{fi}	Invg	0.10	2.00	0.53	1.44	0.05	0.02
σ_{fd}	Invg	0.10	2.00	0.06	0.06	0.57	0.19
σ_{fh}	Invg	0.10	2.00	0.37	2.52	0.19	0.07
σ_i	Invg	0.10	2.00	3.70	0.39	2.27	0.27
σ_d	Invg	0.10	2.00	1.47	0.32	1.49	0.19
σ_h	Invg	0.10	2.00	2.28	0.19	0.05	0.02
σ_{tp}	Invg	0.10	2.00	0.05	0.00	0.02	0.00
Measurement errors on observables for DSGE model							
$\sigma_{\Delta y}$	0.36			σ_l		0.70	
$\sigma_{\Delta c}$	0.18			σ_{fd}		0.13	
$\sigma_{\Delta i}$	1.00			σ_{fh}		0.13	
$\sigma_{\Delta d}$	1.00			σ_{fi}		0.13	
$\sigma_{\Delta h}$	1.00			$\sigma_{\Delta p}$		0.18	

REFERENCES

- Akerlof, G., 1970. The market for “lemons”: quality, uncertainty and the market mechanism. *Q. J. Econ.* 84, 488–500.
- Akhtar, M.A., Harris, E., 1987. Monetary policy influence on the economy An empirical analysis. *Federal Reserve Bank of New York, Quarterly Review* (Winter), 19–34.
- Ando, A., Modigliani, F., 1963. The “life cycle” hypothesis of saving: aggregate implications and tests. *Am. Econ. Rev.* 53 (1), 55–84, Part 1 (Mar., 1963).
- Angeloni, I., Faia, E., 2009. Tale of two policies: Prudential regulation and monetary policy with fragile banks. *Mimeo*, October 29.
- Aoki, K., Proudman, J., Vlieghe, G., 2002. Houses as collateral: Has the link between house prices and consumption in the U.K. changed? *Federal Reserve Bank of New York Economic Policy Review* 8 (1), 163–177.
- Arestis, P., Chortareas, G., Tsoukalas, J.D., 2010. Money and information in a new neoclassical synthesis framework. *Econ. J.* 120, 101–128 (February).
- Benito, A., Thompson, J.N.R., Waldron, M., Wood, R., 2006. House prices and consumer spending. *Bank of England Quarterly Bulletin*, Summer (2006), 142–152.
- Bennett, P., Peach, R., Peristiani, S., 2001. Structural change in the mortgage market and the propensity to refinance. *J. Money Credit Bank.* 33 (4), 954–976 (November).
- Bernanke, B.S., Boivin, J., Elias, P., 2005. Measuring the effects of monetary policy: A factor-augmented vector autoregressive (FAVAR) approach. *Q. J. Econ.* 120 (1), 387–422.
- Bernanke, B.S., Gertler, M., 1989. Agency costs, net worth, and business fluctuations. *Am. Econ. Rev.* 79 (1), 14–31.
- Bernanke, B.S., Gertler, M., 1995. Inside the black box: The credit channel of monetary policy transmission. *J. Econ. Perspect.* 9 (4), 27–48.
- Bernanke, B.S., Gertler, M., Gilchrist, S., 1999. The financial accelerator in a quantitative business cycle framework. In: Taylor, J.B., Woodford, M. (Eds.), *Handbook of macroeconomics*. 1, Elsevier, Amsterdam, pp. 1341–1393 (Part 3).
- Bernanke, B.S., Reinhart, V.R., Sack, B.P., 2004. Monetary policy alternatives at the zero bound: An empirical assessment. *Brookings Pap. Econ. Act.* 35, 1–100 (2004–2).
- Boivin, J., Giannoni, M., 2002. Assessing changes in the monetary transmission mechanism: A VAR approach. *Federal Reserve Bank of New York Economic Policy Review* 8 (1), 97–111.
- Boivin, J., Giannoni, M.P., 2006. Has monetary policy become more effective? *Rev. Econ. Stat.* 88 (3), 445–462.
- Boivin, J., Giannoni, M., 2008. *Global Forces and Monetary Policy Effectiveness*. NBER Working Papers 13736, National Bureau of Economic Research, Inc.
- Brayton, F., Mauskopf, M., 1985. The Federal Reserve Board-MPS Quarterly econometric model of the U.S. economy. *Econ. Model.* 2, 170–292 (July).
- Brumberg, R.E., Modigliani, F., 1954. Utility analysis and the consumption function: An interpretation of cross-section data. In: Kurihara, K. (Ed.), *Post-Keynesian economics*. Rutgers University Press, New Brunswick, NJ.
- Bryant, R.C., Hooper, P., Mann, C.L. (Eds.), 1993. *Evaluating policy regimes: New research in empirical macroeconomics*. The Brookings Institution, Washington D.C.
- Calza, A., Monacelli, T., Stracca, L., 2007. Mortgage markets, collateral constraints, and monetary policy: Do institutional factors matter? *CEPR Discussion Papers* 6231.
- Canova, F., Gambetti, L., 2009. Structural changes in the US economy: Is there a role for monetary policy? *J. Econ. Dyn. Control* 33, 477–490.
- Canova, F., Gambetti, L. (2010). Do expectations matter? The Great Moderation revisited. *Am. Econ. J. Macroeconomics* 183–205.
- Carlstrom, C., Fuerst, T., Paustian, M., 2009. Optimal monetary policy in a model with agency costs. Paper presented at the Financial Markets and Monetary Policy Conference, sponsored by the Federal Reserve Board, and the Journal of Money, Credit and Banking, June 45.

- Case, K.E., 1986. The market for single-family homes in the Boston area. *New England Economic Review* 38–48 (May).
- Case, K.E., Shiller, R.J., 2003. Is there a bubble in the housing market? *Brookings Pap. Econ. Act.* 2003 (2), 299–342.
- Catte, P., Girouard, N., Price, R., Andre, C., 2004. Housing markets, wealth and the business cycle. OECD Economics Department, Working Papers.
- Chirinko, R.S., 1993. Business fixed investment spending: Modeling strategies, empirical results, and policy implications. *J. Econ. Lit.* 31 (4), 1875–1911.
- Christiano, L., Eichenbaum, M., Evans, C., 1999. Monetary policy shocks: What have we learned and to what end? In: Woodford, M., Taylor, J. (Eds.), *Handbook of macroeconomics*. Elsevier/North-Holland, Amsterdam.
- Christoffel, K., Coenen, G., Warne, A., 2008. The new area-wide model of the euro area A micro-founded open-economy model for forecasting and policy analysis. European Central Bank, Working Paper Series 944.
- Clarida, R., Galí, J., Gertler, M., 2000. Monetary policy rules and macroeconomic stability: Evidence and some theory. *Q. J. Econ.* 115, 147–180.
- Clouse, J., Henderson, D., Orphanides, A., Small, D., Tinsley, P., 2000. Monetary policy when the nominal short-term interest rate is zero. Federal Reserve Board Finance and Economics Discussion Series No. 2000-51.
- Curdia, V., Woodford, M., 2009. Credit spreads and monetary policy. Paper presented at the Financial Markets and Monetary Policy Conference, sponsored by the Federal Reserve Board, and the Journal of Money, Credit and Banking, June 45.
- Dynan, K., Elmendorf, D.W., Sichel, D.E., 2006. Can financial innovation help to explain the reduced volatility of economic activity? *J. Monetary Econ.* 53, 123–150 (January).
- Edelberg, W., 2006. Risk-based pricing of interest rates for consumer loans. *J. Monetary Econ.* 53, 2283–2298 (November).
- Edge, R.M., Kiley, M.T., Laforge, J.P., 2007. Documentation of the Research and Statistics Division's estimated DSGE model of the U.S. economy: 2006 version. Board of Governors of the Federal Reserve System (U.S.) Finance and Economics Discussion Series 2007-53.
- Edge, R., Kiley, M.T., Laforge, J.P., 2008. Natural rate measures in an estimated DSGE model of the U.S. economy. *J. Econ. Dyn. Control* 32, 2512–2535.
- Edge, R.M., Kiley, M.T., Laforge, J.P., 2010. A comparison of forecast performance between Federal Reserve staff forecasts, simple reduced-form models, and a DSGE model. *Journal of Applied Econometrics*. John Wiley & Sons, Ltd., 25 (4), 720–754.
- Eggertsson, G.B., Woodford, M., 2003. The Zero Bound on Interest Rates and Optimal Monetary Policy. *Brookings Papers on Economic Activity*, Economic Studies Program, The Brookings Institution 34 (1), 139–235.
- Erceg, C.J., Guerrieri, L., Gust, C., 2006. SIGMA: A new open economy model for policy analysis. *International Journal of Central Banking* 2 (1), March.
- Fagan, G., Henry, J., Mestre, R., 2005. An area-wide model for the euro area. *Econ. Model.* 22 (1), 39–59.
- Fair, R.C., 2004. Estimating how the macroeconomy works. Harvard University Press, Cambridge, MA.
- Fleming, J.M., 1962. Domestic financial policies under fixed and under floating exchange rates (*Politiques financières intérieures avec un système de taux de change fixe et avec un système de taux de change fluctuant*) (*Política financiera interna bajo sistemas de tipos de cambio fijos o de tipos de cambio fluctuantes*). Staff Papers — International Monetary Fund 9 (3), 369–380 (Nov.).
- Friedman, B.M., 1989. Changing effects of monetary policy on real economic activity. *Monetary Policy Issues in 1990's*. Federal Reserve Bank of Kansas City Symposium Proceedings.
- Friedman, M., 1957. A theory of the consumption function. Princeton University Press, Princeton, NJ.
- Galí, J., Gambetti, L., 2009. On the sources of the Great Moderation. Economics Working Papers 1041. Department of Economics and Business, Universitat Pompeu Fabra, Revised Jun 2007.
- Galí, J., Lopez-Salido, D., Valles, J., 2003. Technology shocks and monetary policy: Assessing the Fed's performance. *J. Monetary Econ.* 50, 723–743 (May).

- Gerali, A., Neri, S., Sessa, L., Signoretti, F., 2009. Credit and banking in a DSGE model of the Euro Area (493 KB). Paper presented at the Financial Markets and Monetary Policy Conference, sponsored by the Federal Reserve Board, and the Journal of Money, Credit and Banking, June 45.
- Gertler, M., Gilchrist, S., 1993. The role of credit market imperfections in the monetary transmission mechanism: Arguments and evidence. *Scand. J. Econ.* 95 (1), 43–64.
- Gertler, M., Gilchrist, S., 1994. Monetary policy, business cycles, and the behavior of small manufacturing firms. *Q. J. Econ.* 109 (2), 309–340.
- Gertler, M., Kiyotaki, N., 2010. Financial Intermediation and Credit Policy in Business Cycle Analysis. In: Friedman, B.M., Woodford, M. (Eds.), *Handbook of monetary economics*. 3A, Elsevier/North-Holland, Amsterdam (Chapter 11).
- Gilchrist, S., Oriz, A., Zakrajsek, E., 2009. Credit risk and the macroeconomy: Evidence from an estimated DSGE model. Paper presented at the Financial Markets and Monetary Policy Conference, sponsored by the Federal Reserve Board, and the Journal of Money, Credit and Banking, June 45.
- Gürkaynak, R.S., Sack, B., Swanson, E., 2005. The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models. *Am. Econ. Rev.* 95 (1), 425–436.
- Hall, R.E., 1988. Intertemporal substitution in consumption. *J. Polit. Econ.* 96 (2), 339–357 (Apr.).
- Hanushek, E.A., Quigley, J.M., 1980. What is the price elasticity of housing demand? *Rev. Econ. Stat.* 62 (3), 449–454 (Aug.).
- Harrison, R., Nikolov, K., Quinn, M., Ramsay, G., Thomas, R., Scott, A., 2005. The Bank of England Quarterly Model. Bank of England.
- Hatzius, J., 2005. Housing holds the key to Fed policy. Goldman Sachs, New York Goldman Sachs Global Economics Paper No. 137.
- Hayashi, F., 1982. Tobin's marginal q and average q : A neoclassical interpretation. *Econometrica* 50 (1), 213–224 (Jan.).
- Henderson, J.V., Ioannides, Y.M., 1986. Tenure choice and the demand for housing. *Economica*, New Series 53 (210), 231–246 (May).
- Hurst, E., Stafford, F., 2004. Home is where the equity is: Mortgage refinancing and household consumption. *J. Money Credit Bank.* 36 (6), 985–1014 (Dec.).
- Iacoviello, M., 2005. House prices, borrowing constraints, and monetary policy in the business cycle. *Am. Econ. Rev.* 95 (3), 739–764.
- Iacoviello, M., Minetti, R., 2008. The credit channel of monetary policy: Evidence from the housing market. *J. Macroecon.* 30 (1), 69–96.
- Iacoviello, M., Neri, S., 2010. Housing market spillovers: Evidence from an estimated DSGE model. *Am. Econ. J. Macroeconomics*. 2 (2), 125–164.
- Jorgenson, D., 1963. *Am. Econ. Rev.* 53 (2), 247–259. Papers and Proceedings of the Seventy-Fifth Annual Meeting of the American Economic Association, (May).
- Kashyap, A.K., Stein, J.C., 1995. The impact of monetary policy on bank balance sheets. *Carnegie-Rochester Conference Series on Public Policy* 42 (1), 151–195.
- Kiley, M.T., 2010. Output gaps. Federal Reserve Board Finance and Economics Discussion Series.
- Lettau, M., Ludvigson, S.C., 2004. Understanding trend and cycle in asset values: Reevaluating the wealth effect on consumption. *Am. Econ. Rev.* 94 (1), 276–299.
- Lown, C.S., Morgan, D.P., 2002. Credit effects in the monetary mechanism. *Econ. Policy Rev.* 217–235 (May).
- Lucas, R., 1976. Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy* 1, 19–46.
- Maukopf, E., 1990. The transmission channels of monetary policy: How have they changed? *Federal Reserve Bulletin* 76 (12), 985.
- McCarthy, J., Peach, R.W., 2002. Monetary policy transmission to residential investment. *Federal Reserve Bank of New York Economic Policy Review* 8 (1), 139–158.
- Meh, C., Moran, K., 2008. The role of bank capital in the propagation of shocks. Bank of Canada, Working Papers 08-36.

- Miron, J.A., Romer, C.D., Weil, D.N., 1994. Historical perspectives on the monetary transmission mechanism. In: Mankiw, N.G. (Ed.), *Monetary policy*. National Bureau of Economic Research, Inc, pp. 263–306.
- Mishkin, F.S., 1976. Illiquidity, consumer durable expenditure, and monetary policy. *Am. Econ. Rev.* 66 (4), 642–654 (September).
- Mishkin, F.S., 1978. The household balance-sheet and the Great Depression. *J. Econ. Hist.* 38, 918–937 (December).
- Mishkin, F.S., 1995. Symposium on the monetary transmission mechanism. *J. Econ. Perspect.* 9 (4), 3–10.
- Mishkin, F.S., 2007. Housing and the monetary transmission mechanism. In: *Housing, Housing Finance, and Monetary Policy, 2007 Jackson Hole Symposium*. Federal Reserve Bank of Kansas City, Kansas City, MO, pp. 359–413.
- Mishkin, F.S., 2008. Global financial turmoil and the world economy. Speech given at the Caesarea Forum of the Israel Democracy Institute, Eliat, Israel, July 2, 2008, <http://www.federalreserve.gov/newsevents/speech/mishkin20080702a.htm>.
- Mundell, R.A., 1963. Capital mobility and stabilization policy under fixed and flexible exchange rates. *Can. J. Econ.* 29, 475–485.
- Murchison, S., Rennison, A., 2006. ToTEM: The Bank of Canada's new quarterly projection model. Bank of Canada, Technical Reports 97.
- Peek, J., Rosengren, E., 1995a. The capital crunch: Neither a borrower nor a lender be source. *J. Money Credit Bank.* 27 (3), 625–638 (Aug.).
- Peek, J., Rosengren, E.S., 1995b. Is bank lending important for the transmission of monetary policy? An overview. *New England Economic Review* 3–11 (Nov.).
- Peek, J., Rosengren, E.S., 1997. The international transmission of financial shocks: The case of Japan. *Am. Econ. Rev.* 87 (4), 495–505 (Sept.).
- Peek, J., Rosengren, E.S., 2010. The role of banks in the transmission of monetary policy. In: Berger, A., Molyneux, P., Wilson, J. (Eds.), *The Oxford handbook of banking*. Oxford University Press, Oxford, UK.
- Primiceri, G., 2005. Why inflation rose and fell: Policymakers' beliefs and US postwar stabilization policy. National Bureau of Economic Research, Inc., NBER Working Papers 11147.
- Ramey, V.A., 1993. How important is the credit channel in the transmission of monetary policy?. National Bureau of Economic Research, Inc., NBER Working Papers 4285.
- Reifschneider, D., Stockton, D.J., Wilcox, D.W., 1997. Econometric models and the monetary policy process. *Carnegie-Rochester Conference Series on Public Policy* 47 (Dec.), 1–37.
- Reifschneider, D., Tetlow, R., Williams, J., 1999. Aggregate disturbances, monetary policy, and the macroeconomy: The FRB/US perspective. *Federal Reserve Bulletin* 1–19 (Jan.).
- Romer, C.D., Romer, D.H., 1989. Does monetary policy matter? A new test in the spirit of Friedman and Schwartz. *NBER Macroeconomics Annual* 1989 4, 121–184. National Bureau of Economic Research, Inc.
- Sims, C.A., Zha, T., 2006. Were there regime switches in U.S. monetary policy? *Am. Econ. Rev.* 96 (1), 54–81.
- Smets, F., 1995. Central bank macroeconomic models and the monetary policy transmission mechanism. BIS (1995). In: *Financial structure and the monetary policy transmission mechanism*. Central Bank, p. 394 (March).
- Smets, F., Wouters, R., 2007. Shocks and frictions in US business cycles: A Bayesian DSGE approach. *Am. Econ. Rev.* 97 (3), 586–606.
- Stock, J.H., Watson, M.W., 2005. Implications of dynamic factor models for VAR analysis. National Bureau of Economic Research, Inc., NBER Working Papers 11467.
- Taylor, J.B., 1993. *Macroeconomic policy in a world economy: From econometric design to practical operation*. W. W. Norton, New York.
- Taylor, J.B., 1995. The monetary transmission mechanism: An empirical framework. *J. Econ. Perspec.* 9 (4), 11–26.
- Taylor, J.B., 1999. A historical analysis of monetary policy rules. In: Taylor, J.B. (Ed.), *Monetary policy rules*. National Bureau of Economic Research, Inc, pp. 319–348.

- Taylor, J.B., Williams, J.C., 2010. Simple and robust rules for monetary policy. In: Friedman, B.M., Woodford, M. (Eds.), *Handbook of monetary economics*. 3B, Elsevier/North-Holland, Amsterdam (Chapter 15).
- Tobin, J., 1969. A general equilibrium approach to monetary theory. *J. Money Credit Bank*. 1 (1), 15–29 (Feb.).
- Van Den Heuvel, S., 2002. Does bank capital matter for monetary transmission? *Econ. Policy Rev.* May. [Updated and reprinted in *The Evolving Financial System and Public Policy*, Bank of Canada, December 2004.].
- Wessel, D., 2009. In Fed we trust: Ben Bernanke's war on the great panic. *Crown Business*, New York.
- Woodford, M., 2003. *Interest and prices*. University Press, Princeton, NJ.

Inflation Persistence[☆]

Jeffrey C. Fuhrer

Federal Reserve Bank of Boston

Contents

1. Introduction	424
1.1 Early inflation models and the empirical necessity of lagged inflation	425
1.2 Rational expectations and inflation persistence: An introduction to some of the issues	426
2. Defining and Measuring Reduced-Form Inflation Persistence	431
2.1 Defining reduced-form persistence	431
2.2 Measuring reduced-form inflation persistence	433
2.2.1 <i>The data</i>	434
2.2.2 <i>Unit root tests</i>	435
2.2.3 <i>First-order autocorrelations</i>	438
2.2.4 <i>Autocorrelation functions</i>	439
2.2.5 <i>Dominant root of the univariate time series process</i>	440
2.3 Evidence of changing reduced-form persistence in the United States	442
2.3.1 <i>Unknown breakpoint tests for univariate ARs</i>	443
2.3.2 <i>Pre-war inflation persistence under the gold standard</i>	444
2.3.3 <i>A parameterized characterization of reduced-form persistence</i>	445
2.3.4 <i>Multivariate evidence of changes in reduced-form inflation persistence</i>	446
2.4 International evidence of changing reduced-form persistence	447
2.5 Conclusions from the reduced-form evidence	448
3. Structural Sources Of Persistence	449
3.1 Inherited and intrinsic persistence	449
3.2 An alternative decomposition of persistence: Disinflations and supply shocks	450
3.3 Persistence in the Calvo/Rotemberg model	451
3.4 The analytics of inflation persistence: "Inherited" and "intrinsic" persistence	452
3.4.1 <i>The baseline case</i>	452
3.4.2 <i>More complex cases</i>	452
3.4.3 <i>Shocks to the Euler equation</i>	453
3.4.4 <i>The pivotal role of the coefficient on x_t</i>	454
3.4.5 <i>Hybrid models of inflation and "intrinsic" persistence: Including lagged inflation</i>	456
3.4.6 <i>The persistence of the driving process</i>	459
3.5 Persistence in models of "trend inflation"	461
3.5.1 <i>Cogley and Sbordone's measure of trend inflation</i>	462

[☆] The author thanks the editors, Fabià Gumbau-Brisa, Denny Lie, Giovanni Olivei, Scott Schuh, Oz Shy, Raf Wouters, participants at the Federal Reserve Bank of Boston's lunchtime workshop, and participants at the European Central Bank's October 2009 Handbook conference for helpful comments, and Timothy Cogley for providing data.

3.6	Using a DSGE model to interpret structural sources of persistence	463
3.7	Persistence in state-dependent models of inflation	469
3.8	Persistence in sticky-information models	470
3.9	Persistence in learning models	471
4.	Inference About Persistence In Small Samples: “Anchored Expectations” and Their Implications for Inflation Persistence	473
5.	Microeconomic Evidence On Persistence	478
5.1	Persistence in micro data: U.S. evidence	478
5.2	Persistence in micro data: Euro Area evidence	480
5.3	More on aggregation and persistence	480
6.	Conclusions	482
	References	483

Abstract

This chapter examines the concept of inflation persistence in macroeconomic theory. It begins by defining persistence — emphasizing the difference between reduced-form and structural persistence. It then examines a number of empirical measures of reduced-form persistence, considering the possibility that persistence may have changed over time. The chapter then examines the theoretical sources of persistence, distinguishing “intrinsic” from “inherited” persistence, and deriving a number of analytical results on persistence, emphasizing the influence of the monetary policy regime. It summarizes the implications for persistence from the literature on imperfect information models, learning models, and so-called “trend inflation models,” providing some new results throughout. Finally, it summarizes the results on persistence from the many studies of disaggregated price data.

JEL classification: E31, E52

Keywords

Autocorrelation
Inflation Persistence
Phillips Curve

1. INTRODUCTION

What is “persistence”? Why is it important to macroeconomists and policymakers? In broad terms, persistence is the economic analog of inertia in physics. Inertia may be defined as the resistance of a body to changing its velocity (direction and rate of speed) unless acted upon by an external force. This law is often paraphrased as “a body at rest will remain at rest unless acted upon by an external force,” which is one example of the principle. Newton’s second law of motion, $F = ma$ or $a = \frac{F}{m}$, captures this idea algebraically: The magnitude of the force required to produce a given change of velocity (acceleration) is proportional to the body’s mass. The more mass the body has, the more force is required to accelerate it. In this sense, the mass of a body calibrates its inertia or persistence.

While no analogy is perfect, this one works reasonably well at an intuitive level. An economic variable is said to be persistent if, other things being equal, it shows a tendency to stay

near where it has been recently, absent other economic forces that move it elsewhere. In the case of inflation, the rate of change of the price level tends to remain constant (inflation tends to be persistent) in the absence of an economic “force” to move it from its current level. This chapter provides more precise definitions of inflation persistence in the following sections, but this physical analogy may serve to provide motivating intuition.

1.1 Early inflation models and the empirical necessity of lagged inflation

For many decades, economists assumed that inflation was an inertial or persistent economic variable. The concept of the sacrifice ratio — the number of point-years of elevated unemployment required to reduce inflation by a percentage point — implies that inflation does not move freely but requires significant economic effort in the form of elevated unemployment or lost output to reduce its level.¹ The early incarnations of the accelerationist Phillips curve modeled the apparent inertia in inflation by including lags of inflation. A canonical example of these specifications is Gordon’s “triangle model” of inflation, here replicated in simplified form (Gordon, 1982):²

$$\pi_t = \sum_{i=1}^k a_i \pi_{t-i} - b(U_t - \bar{U}) + c x_t + \varepsilon_t. \quad (1)$$

Inflation, π_t depends on its own lags (normally constrained to sum to one to reflect the Friedman-Phelps accelerationist principle), a measure of real activity (here the deviation of unemployment, U_t , from the nonaccelerating inflation rate of unemployment; NAIRU), and supply-shifters such as key relative price shifts, summarized in x_t . In such a model, inflation moves gradually, partially anchored by its recent history, in response to real activity and supply shocks. These variables may be persistent, in which case inflation will “inherit” some of their persistence. A key question is whether and why inflation has its own or “intrinsic” persistence, beyond that inherited from U_t and x_t (or perhaps ε_t , if it is also serially correlated). If inflation exhibits intrinsic persistence, then a model of inflation may require the equivalent of the lags in Eq. (1).

In this early literature, the theoretical justification for including lags of inflation was as a proxy for expected inflation and a proxy for contracting and other price-setting frictions. As an empirical matter, the lags helped the model fit the data. To see this last point simply, consider the R^2 for estimates of Gordon-style Phillips curves with and without the lags of inflation in Table 1. The specification is $\pi_t = \sum_{i=1}^4 \alpha_i \pi_{t-i} + \sum_{j=1}^2 \beta_j U_{t-j} + \sum_{k=1}^2 \gamma_k \Delta r_{t-k}^o + C$, where π_t is the quarterly percentage change in the core CPI, U is the civilian unemployment rate, and r^o is the relative price of oil. The sum of the α_i ’s is constrained to one in some of the estimates.³

¹ See Gordon, King, and Modigliani (1982) for the first study that uses the term “sacrifice ratio.” This study followed on the work of Arthur Okun (1977).

² See Friedman (1968) and Phelps (1968) for the earliest explications of the accelerationist Phillips curve.

³ Note that this constraint is not statistically significant in these regressions — as indicated in the Table 1, the R^2 ’s are identical in the constrained and unconstrained cases, and the p -value for the test of these restrictions exceeds 0.8 for both samples.

Table 1 R^2 for Gordon-style Phillips curves
Model

	R^2
Core CPI, 1966:Q1–1984:Q4	
With lags, $\sum \alpha_i = 1$	0.74
With lags, $\sum \alpha_i \neq 1$	0.74
Without lags	0.24
Core CPI 1985:Q1–2008:Q4	
With lags, $\sum \alpha_i = 1$	0.79
With lags, $\sum \alpha_i \neq 1$	0.79
Without lags	0.09
Core PCE, 1966:Q1–1984:Q4	
With lags, $\sum \alpha_i = 1$	0.76
With lags, $\sum \alpha_i \neq 1$	0.77
Without lags	0.39
Core PCE 1985:Q1–2008:Q4	
With lags, $\sum \alpha_i = 1$	0.72
With lags, $\sum \alpha_i \neq 1$	0.72
Without lags	0.16

Points from this table include: (1) the lags of inflation are empirically critical, whatever they may represent structurally; and (2) as a consequence, it is critical to understand what these lags represent structurally.⁴

1.2 Rational expectations and inflation persistence: An introduction to some of the issues

The introduction of Muth's (1961) theory of rational expectations into the macroeconomics literature and the consequent move toward explicit modeling of expectations posed considerable challenges in modeling prices and inflation. In the earliest rational expectations models of Lucas (1972) and Sargent and Wallace (1975), the price level was a purely forward-looking or expectations-based variable like an asset price, which

⁴ The results in Table 1 are completely invariant to the choice of the lag length for lagged inflation. For lag lengths up to 24, we obtain nearly identical coefficient sums, R^2 's, and nonbinding unit sum constraints. However, if we begin the later sample in 1997, results change dramatically. The constraint on the sum of the lag coefficients is now binding, and the R^2 's drop to 0.25 or below, a result highlighted in Williams (2006). This apparent shift in reduced-form persistence is discussed in more detail later.

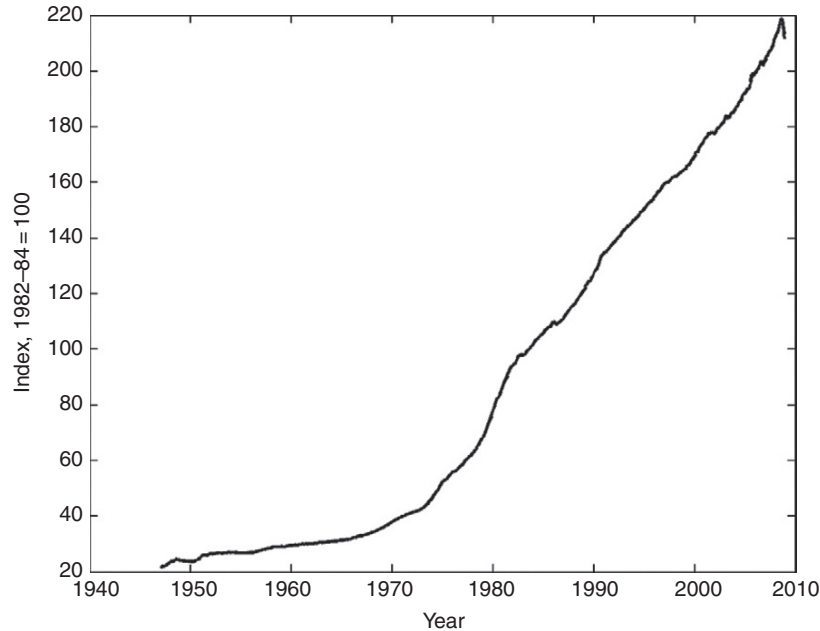


Figure 1 The consumer price index.

in these models implied that prices were flexible, and could “jump” in response to shocks. It was difficult at first to reconcile the very smooth, continuous behavior of measured aggregate price indexes such as the consumer price index with the flexible-price implications of these early rational expectations models. Figure 1 displays the consumer price index for the post-war period.

A number of economists recognized the tension between the obvious persistence in the price level data and the lack of persistence implied by these early rational expectations models. Fischer (1977), Gray (1977), Taylor (1980), Calvo (1983), and Rotemberg (1982, 1983) developed a sequence of models that rely on nominal price contracting in attempts to impart a data-consistent degree of inertia to the price level in a rational expectations setting. The overlapping contracts of Taylor and Calvo/Rotemberg were successful in doing this, allowing contracts negotiated in period t to be affected by contracts set in neighboring periods, which would remain in effect during the term of the current contract.⁵ The subsequent trajectory of macroeconomic research drew heavily on these seminal contributors, who had neatly reconciled rational expectations with inertial (or persistent) macroeconomic time series.

⁵ For example, a four-period contract negotiated in period t would be influenced by the contracts negotiated in the previous three periods, as well as by the contracts expected to be negotiated in the following three periods.

However, in the early 1990s, a number of authors discovered that these rational expectations formulations held less satisfying implications for the change in the price level, that is, the rate of inflation. Ball (1994) demonstrated that such models could imply a counterfactual “disinflationary boom” – the central bank could engineer a disinflation that caused output to rise rather than contract. Buiter and Jewitt (1981) and Fuhrer and Moore (1992, 1995) showed that Taylor-type contracting models implied a degree of inflation persistence that was far lower than was apparent in inflation data of the post-war period to that point. To build intuition, compare Eq. (1) with a Calvo- or Taylor-type inflation equation⁶

$$\pi_t = E_t \pi_{t+1} + \gamma \tilde{y}_t + \varepsilon_t. \quad (2)$$

Implicitly, this makes inflation, π_t , a function of all expected future output gaps, $\tilde{y}_t$, (assuming that $E_t \varepsilon_{t+i} = 0 \forall i > 0$, or equivalently that ε_t has no autocorrelation).⁷ Inflation will indeed inherit the persistence in output, but nothing beyond that. The lags in the Gordon triangle model are gone. Inflation is completely forward looking: Following a shock to output, inflation can jump immediately in response. If output exhibits no persistence (i.e., there are no “real rigidities”), then neither will inflation. In contrast, Eq. (1) implies that inflation depends on all *past* output gaps and shocks. As in the forward-looking model, inflation inherits the persistence in output, but the lagged inflation terms mean that inflation cannot jump in response to a shock to output — inflation exhibits additional persistence (or inertia) in the sense that in response to shocks, inflation has a tendency to remain near its most recent values.

To explore the dynamic implications of this forward-looking specification, we embed Eq. (2) in a skeletal macro model. To complete the model, we include a rudimentary IS curve and a simplified policy rule

$$\begin{aligned} \tilde{y}_t &= \rho \tilde{y}_{t-1} - a(r_t - E_t \pi_{t+1}) \\ r_t &= b(\pi_t - \bar{\pi}), \end{aligned} \quad (3)$$

where r_t is the short-term policy rate controlled by the central bank, the inflation target is denoted $\bar{\pi}$, and the equilibrium real rate is suppressed for convenience.⁸ We will consider cases with zero and nonzero output persistence, that is, $\rho = 0$ or $\rho \neq 0$. The policy response to inflation gaps is calibrated by parameter b , which is set to 1.5 throughout.

Consider first the case in which output is not persistent. Inflation is at its target and the output gap is zero. For comparison with the large literature that attempts to measure the economy’s response to an identified monetary policy shock, consider a one unit positive temporary shock to the policy rate r_t . With no inertia in output or inflation, both output and inflation are

⁶ See Roberts (1997) for a derivation that shows the approximate equivalence of these formulations.

⁷ This is made explicit below in Eq. (18), but one can obtain the result by inspection through repetitive substitution of the definition of future inflation into Eq. (2).

⁸ Note that the derivation of the Calvo Phillips curve of Eq. (2) assumes a zero steady-state inflation rate, whereas Eq. (3) allow for the possibility of a nonzero steady-state inflation rate. Cogley and Sbordone (2008) derived the appropriate Phillips curve for the case in which the central bank pursues a nonzero inflation rate. For the purposes of this exercise, the inaccuracy in the approximation is likely small.

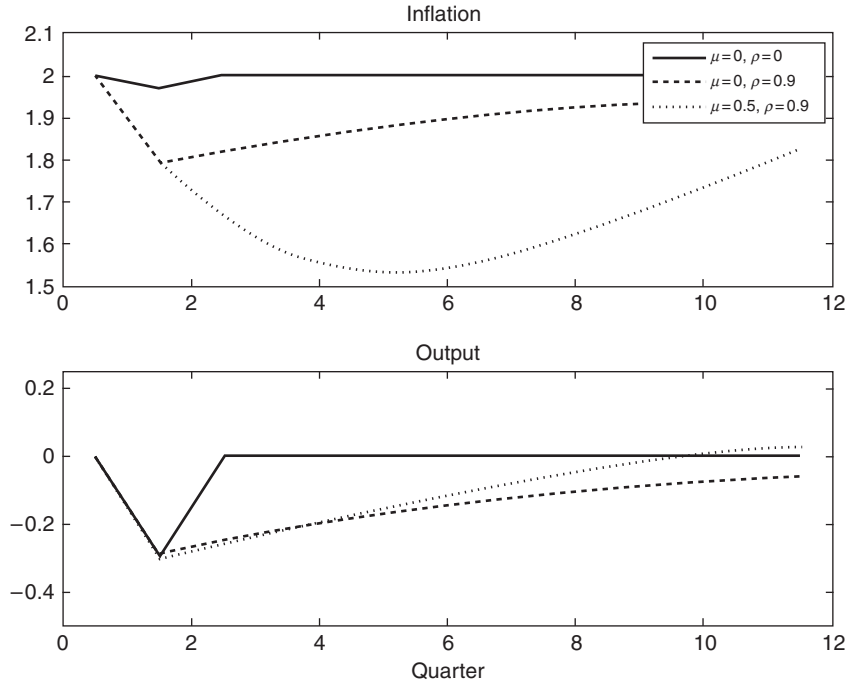


Figure 2 Response of inflation to a monetary policy shock in the simple Calvo model.

perturbed below their steady state for one period. In period 2, both return to their steady state. The solid lines in Figure 2 display the results of this rather uninteresting exercise.

When output is persistent (here $\rho = 0.9$), the dynamics are a bit more complex. Starting from the same steady state, consider a one unit positive shock to the policy rate. The results of this simulation are depicted in the dashed lines in the same figure. Output is depressed below its steady state in the first period, and because of its own persistence, remains below its steady state for some time. For Eq. (2) to hold, it must be that expected inflation lies above current inflation whenever the output gap is negative. Because output is persistent, the expected change in inflation must be positive for as long as output remains negative. But ultimately, inflation will have to return to its original steady state. Thus, inflation must immediately jump down below its steady-state level, and then rise to it from below. In this simulation, inflation exhibits dynamics that are reminiscent of the exchange rate in the famous Dornbusch (1976) overshooting model.⁹

For the sake of comparison, the dotted line in Figure 2 shows the outcome when the inflation equation includes a lag of inflation. Here, the equation is a “hybrid” equation that mixes both forward- and backward-looking elements, of the form

$$\pi_t = \mu\pi_{t-1} + (1 - \mu)E_t\pi_{t+1} + \gamma\tilde{y}_t + \varepsilon_t \quad (4)$$

⁹ This overshooting property is examined in Fuhrer and Moore (1995), and in Estrella and Fuhrer (2002).

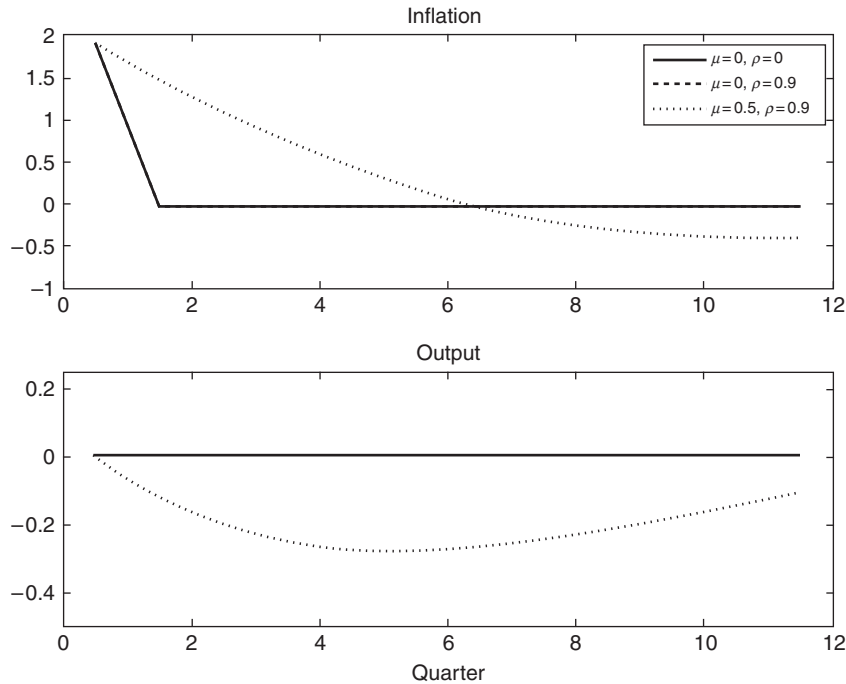


Figure 3 Credible disinflations in the simple Calvo model.

with $\mu = 0.5$ in Figure 2. Now, in response to the monetary policy shock, output behaves approximately as before, but inflation declines gradually over several quarters, exhibiting the more typical “hump-shaped” response found in the literature on monetary vector autoregressions (VARs).¹⁰ It is no longer the case that expected inflation must exceed current inflation for as long as the output gap is negative.

While much of economists’ intuition about inflation persistence is obtained from responses to identified monetary policy shocks, as in the simulations above, considerable interest also centers on the behavior of inflation in response to a central-bank-engineered disinflation. The work of Ball (1994), Fuhrer and Moore (1992) and others emphasizes this aspect of inflation dynamics. The differences across specifications in the behavior of inflation in a fully credible disinflation are as striking as those in response to monetary policy shocks as previously illustrated. In the solid and heavy dashed lines of Figure 3, inflation is forward looking ($\mu = 0$). The central bank announces a permanent and fully credible reduction in its target inflation rate from 2% to 0 at time 1. As Figure 3 indicates, regardless of how persistent the output, the inflation rate jumps to its new equilibrium in the period after the announcement, with no disruption to output, so that the solid and

¹⁰ See Christiano, Eichenbaum, and Evans (2005) for representative results. Note that the output response differs modestly from the dashed line because the path of inflation, and thus the path of the short-term real rate, differs.

heavy dashed lines coincide in the figure.¹¹ In marked contrast, when lagged inflation is added to the inflation equation, inflation declines gradually to its new long-run equilibrium, with a concomitant decline in output during the transition.

Many would view the dynamics of the purely forward-looking specification as strikingly counterfactual.¹² Counterfactual or not, one needs to understand the dynamics of inflation to pursue appropriate monetary policy. A knowledge of the reduced-form behavior of inflation is not sufficient. The central bank needs to understand the sources of inflation dynamics — in these examples, whether it arises from the persistence of output, which may in turn arise from the behavior of monetary policymakers, or from persistence intrinsic to the price-setting process. A third key source of persistence not examined in these exercises, but explored in detail next, is the behavior of the central bank. Either through the vigor (or lack thereof) of its systematic response to deviations of inflation from its current target, or in the low-frequency movement in its inflation target, it can exert significant influence on the persistence of inflation.¹³ These simulations make clear why the issue of persistence is of more than passing interest to macroeconomists and policymakers.

To be sure, understanding why and when inflation may be persistent can be more complicated than these simple examples suggest. For instance, the examples leave out the possibility that inflation may respond differently to a disinflation engineered by a monetary authority with imperfect credibility; they also abstract from imperfect knowledge of the economy that might lead to learning on the part of private agents. All of these situations could alter the dynamic implications of the simple examples. We return to these considerations in the following section.

2. DEFINING AND MEASURING REDUCED-FORM INFLATION PERSISTENCE

The previous discussion suggests that it will be useful to distinguish reduced-form from structural inflation persistence. Reduced-form persistence will refer to an empirical property of an observed inflation measure, without economic interpretation. Structural persistence will refer to persistence that arises from identified economic sources. A key objective of recent inflation research has been to map observed or reduced-form persistence into the underlying economic structures that produce it. To a significant extent, I will argue that this challenge remains.

2.1 Defining reduced-form persistence

There is no definitive measure of reduced-form persistence. As the following sections discuss, researchers have employed a variety of measures to capture the idea that

¹¹ If the inflation target were pre-announced, the inflation rate would jump to its new target in the first period, again with no disruption to output.

¹² Estrella and Fuhrer (2002) examined counterfactual implications of this class of models.

¹³ The movement over time of the central bank's inflation goal is the focus of a considerable amount of research, see for example Ireland (2007) and Cogley and Sbordone (2008).

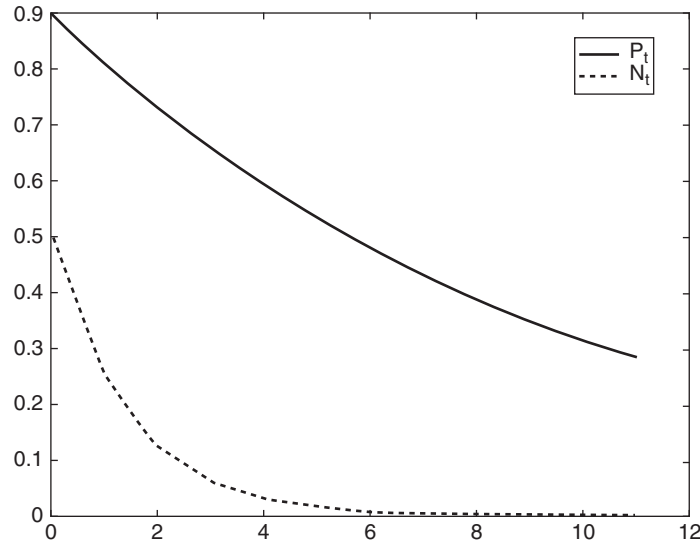


Figure 4 Hypothetical autocorrelation functions.

inflation responds gradually to shocks, or remains close to its recent history.¹⁴ Most of the measures of inflation persistence derive from the autocorrelation function for inflation, so it will be useful to define it here. The i^{th} autocorrelation, ρ_i , of a stationary variable, x_t —the correlation of the variable with its own i^{th} lag, x_{t-i} —may be expressed as¹⁵

$$\rho_i = \frac{E(x_t x_{t-i})}{V(x_t)}, \quad (5)$$

where $V(x)$ is the variance of x . The correlations are of course bounded between -1 and 1 . The variable's autocorrelation function is correspondingly defined as the vector of correlations of current period x with each of its own lags x_{t-i} from $i = 1$ to k :

$$A = [\rho_1, \dots, \rho_k]. \quad (6)$$

A time series will be said to be relatively persistent if its correlations with its own past decay slowly. Thus, a graphical depiction of this measure of a series' reduced-form persistence is provided by a plot of the variable's autocorrelation function. In Figure 4, the correlation of the persistent series, P_t , with its own lags declines gradually from 0.9 to 0.3 over 12 periods. In contrast, the less persistent time series, N_t , shows a more rapid decline in its autocorrelation function from 0.5 to 0 in about eight periods. To a large

¹⁴ This definition implicitly assumes that inflation is *positively* correlated with its own lags, an assumption that holds up well over most of post-war history. More generally, a time series may be deemed persistent if the absolute value of its autocorrelations is high, so that a strongly *negatively* autocorrelated series would also be characterized as persistent.

¹⁵ For convenience, this definition assumes that x_t is a mean zero series.

extent, the alternative measures of persistence surveyed in Section 2.2 provide alternative ways of quantifying the rate at which inflation's autocorrelations decay.

The analytical representation of the autocorrelation function will also be useful in this discussion. For example, if the variable x_t is defined as a first-order autoregressive process

$$x_t = ax_{t-1} + \varepsilon_t \quad (7)$$

$-1 < a < 1$, then its autocorrelation function is simply

$$A = [a, a^2, \dots, a^k]. \quad (8)$$

From Eq. (8) it is clear that the autocorrelations of x_t die out geometrically at the rate determined by the autoregressive parameter a . The analytical expressions for the autocorrelations of inflation in more complex structural models with rational expectations are derived next.

Some researchers define persistence as the extent to which shocks that occurred in the past continue to have an effect on current inflation (see, e.g., Cogley, Primiceri, & Sargent, 2010). This concept is related to the autocorrelation function. The more correlated inflation is with its distant past, the more shocks that perturb inflation in the distant past will be reflected in current inflation. More formally, adopting a simple first-order autoregressive model for inflation from Eq. (7), one can iterate the equation backward to obtain the moving-average representation of inflation

$$\begin{aligned} \pi_t &= \varepsilon_t + a(\pi_{t-2} + \varepsilon_{t-1}) \\ &= \varepsilon_t + a\varepsilon_{t-1} + a^2\varepsilon_{t-2} + a^3\varepsilon_{t-3} + \dots, \end{aligned} \quad (9)$$

which shows that the larger a is, and thus the more slowly inflation's autocorrelations decay, the larger the influence of past shocks on current inflation — equivalently, shocks have a more persistent effect on inflation.¹⁶

2.2 Measuring reduced-form inflation persistence

Because there is little agreement in the extant literature on how best to measure persistence, we examine a battery of measures that attempt to capture the persistence in inflation:

- Conventional unit root tests
- Autocorrelation function of the inflation series (as defined in Eq. (6))
- First autocorrelation of the inflation series
- Dominant root of the univariate autoregressive inflation process (defined later)
- Sum of the coefficients from a univariate AR for inflation

¹⁶ This algebra provides a more rigorous justification for the argument that old-style Phillips curves like those in Eq. (1) add persistence to inflation beyond that inherited from output — the lags of inflation imply persistent effects of past shocks on current inflation.

- Unobserved component decompositions of inflation that estimate the relative contributions of “permanent” and “transitory” components of inflation (e.g., the IMA (1,1) and related models proposed by Stock & Watson, 2007)

Because the autocorrelation function summarizes much of the information in a time series, it may be the best overall measure of persistence. But researchers often desire a single number that captures the overall persistence implied by the full autocorrelation function; this desire motivates many of the scalar measures itemized in the previous list. One can find overlap across the results from these tests, but, to be sure, the results are neither uniform nor unambiguous. Throughout, we examine these measures of inflation persistence across a number of subsamples, providing suggestive evidence as to whether persistence has changed over time.¹⁷

2.2.1 *The data*

For the purposes of this chapter, we will focus on three key inflation measures. Each will be defined as 400 times the log change in the corresponding price index. The three indexes are the GDP deflator, the consumer price index (CPI), and the personal consumption expenditures chain-type price index (PCE). In some cases, we will examine the so-called “core” versions of the CPI and PCE, that is, the price indexes that exclude food and energy prices. We denote these series throughout as “CPI-X” and “PCE-X.” These series abstract from the high-frequency noise that can be introduced by volatile food and energy price series.¹⁸

Figure 5 displays the three overall series in the top panel, with corresponding core series in the bottom panel. Table 2 provides summary statistics for the three inflation series over several sample periods. Evident in the figure and echoed in Table 2 are the drop in the level of all inflation measures since the early 1980s (the top panel of the table); the decline in the variance of all measures, consistent with the so-called “Great Moderation” (the second panel of the table); the high correlation among the three series (the third panel of the table); and the lesser volatility of the core series. Less evident in Figure 5 but clear in the bottom panel of Table 2 is the decline in the correlations across the series. In particular, while the CPI and PCE remain highly correlated, the correlations between the core and overall measures have declined, as have the correlations between the GDP deflator and the consumer price measures. Despite the relatively high correlations among these series, the measures of persistence presented below will show some noticeable differences across the series.

¹⁷ We do not execute a full battery of unknown continuous breakpoint tests for each of the measures, as the goal of this chapter is not to provide definitive evidence on the precise timing of changes in persistence. However, Section 2.3.1 provides the results of an unknown breakpoint test in the univariate AR representation of inflation.

¹⁸ As indicated in Figure 5, the core price series are available only starting in the late 1950s.

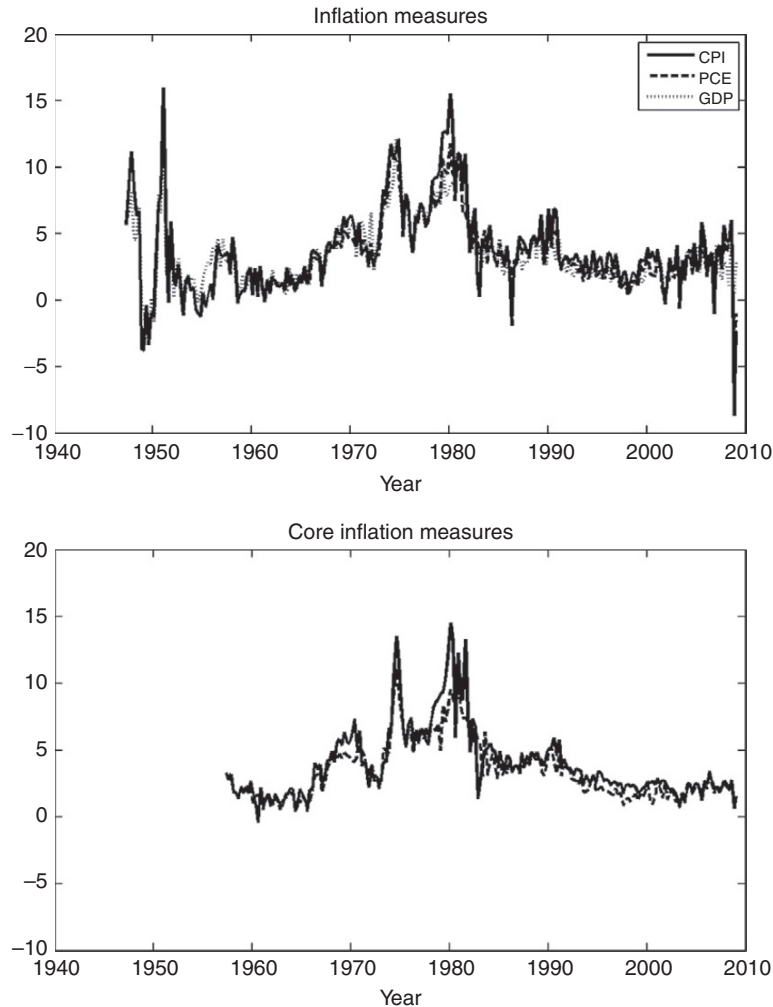


Figure 5 Inflation data.

2.2.2 Unit root tests

The first test of persistence is a unit root test. If inflation contains a unit root, its persistence is unquestionably large (infinite) and its variance can be unbounded.¹⁹ Many papers test for a unit root in inflation (see Ball & Cecchetti, 1990; Barsky, 1987); prior to the 1990s, the results tend to suggest a unit root in inflation. In more recent years, a researcher would be less likely to reject stationarity. Most monetary models would suggest that the more

¹⁹ A series with a unit root has infinite “memory,” in the sense that a shock in period t has influence on all periods $t + k$, $k > 0$. More formally, if $x_t = x_{t-1} + \varepsilon_t$, then $x_t = \sum_{i=0}^{\infty} \varepsilon_{t-i}$. Thus, any shock to a series with a unit root persists forever.

Table 2 Summary statistics, inflation measures

Table 2 Summary statistics, inflation measures

Series	Means				
	59–08	59–84	85–08	95–08	
CPI	4.0	5.0	3.0	2.5	
PCE	3.6	4.5	2.6	2.1	
GDP	3.6	4.6	2.5	2.2	
CPI-X	4.0	4.9	3.0	2.3	
PCE-X	3.5	4.4	2.5	1.8	
Series	Variances				
	59–08	59–84	85–08	95–08	
CPI	9.4	12.7	3.7	4.3	
PCE	6.4	8.6	2.2	2.2	
GDP	5.7	7.8	0.95	0.83	
CPI-X	7.2	10.9	1.2	0.31	
PCE-X	4.7	6.4	1.2	0.25	
Correlation matrix, 1959–2008					
	CPI	PCE	GDP	CPI-X	PCE-X
CPI	1.00				
PCE	0.96	1.00			
GDP	0.86	0.92	1.00		
CPI-X	0.86	0.85	0.86	1.00	
PCE-X	0.81	0.90	0.92	0.91	1.00
Correlation matrix, 1985–2008					
	CPI	PCE	GDP	CPI-X	PCE-X
CPI	1.00				
PCE	0.94	1.00			
GDP	0.57	0.70	1.00		
CPI-X	0.48	0.56	0.53	1.00	
PCE-X	0.40	0.62	0.65	0.87	1.00

Table 3 Unit root tests for inflation measures (*p*-values, null = series has unit root) 1966–2008

Series	ADF	Phillips-Perron
CPI	0.14	0.0069
PCE	0.37	0.045
GDP	0.33	0.025
1985–2008		
Series	ADF	Phillips-Perron
CPI	0.00	0.00
PCE	0.07	0.00
GDP	0.31	0.00

vigorous attention to inflation on the part of central banks around the world in recent decades is responsible for this change. In addition, the relatively limited variation in recent years in central banks' inflation targets may have reduced what appeared to be a unit root component of inflation (in this regard, see Cogley & Sbordone, 2008; Stock & Watson, 2007).

Table 3 provides univariate tests for the null in the inflation series that contains a unit root, for a long sample (1966–2008), and for a “Great Moderation” sample (1985–2008). The results in the table are somewhat ambiguous. For the most recent decades, one typically develops strong rejections of the null of a unit root, although this varies somewhat depending on the inflation series and on the test employed. For the longer sample, the Phillips-Perron test rejects the null, although not always very strongly. The ADF test fails to reject for all three inflation series.

What should one make of these tests? Certainly for the past 25 years, the U.S. central bank has behaved as if it had a specific, relatively stable, and low inflation goal, although it has so far chosen not to announce that goal. If it does indeed have such a goal, then theory suggests that the U.S. inflation rate will not have a unit root — or that if it does have a unit root, the variance of that component of inflation will be quite small.²⁰ This reasoning, combined with the more frequent rejections of the unit root null in recent decades, suggests that one may safely assume that inflation does not contain an important unit root, at least not under current monetary policy.

²⁰ This statement is somewhat too strong, as a very gradual drift downward in the target inflation rate could manifest itself as a unit root component of inflation, albeit with a relatively small variance. See Stock and Watson (2007) for an inflation model that formalizes this reasoning. See also Cogley and Sbordone (2008) for a model that examines the importance of a “trend inflation” component of inflation that may have accounted for the near-unit root in inflation in the 1970s. They associate this component with time variation in the Federal Reserve's implicit inflation goal and find that it accounts for much of the low-frequency variation in inflation. Their results are discussed in detail in Sections 3.5 and 3.6.

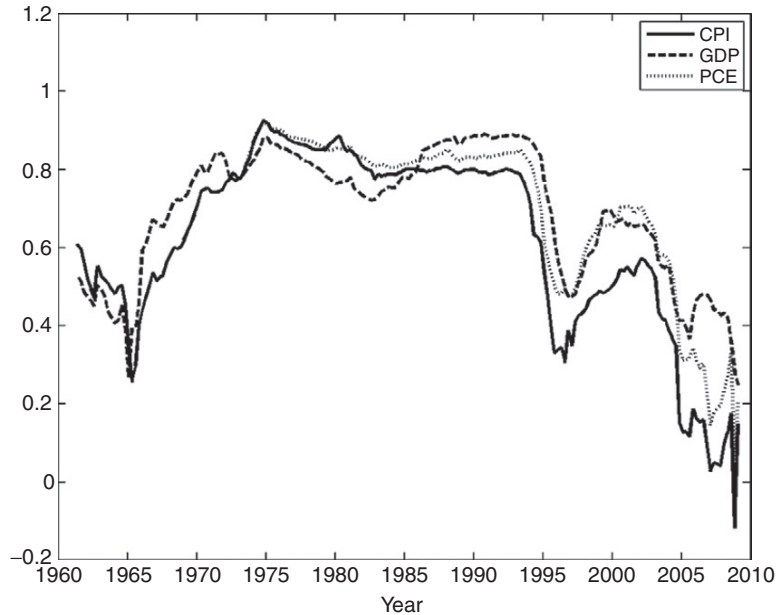


Figure 6 Rolling-sample estimates of the first-order autocorrelation coefficient.

2.2.3 First-order autocorrelations

Section 2 suggests the first-order autocorrelation coefficient for the series as a simple measure of persistence, a measure used, for example, in Pivetta and Reis (2007). Figure 6 extends and expands their Figure 2a, presenting rolling-sample estimates of the first-order autocorrelation coefficient for three aggregate inflation measures.²¹ Figure 6 shows that all three inflation measures display similar time variation in their autocorrelation. All rise in the 1970s to 0.8 or higher, and remain there until the mid-1990s, at which point the correlation drops to 0.5 or 0.6. A third decline is evident in the mid-2000s, with first-order autocorrelation dropping to very low levels indeed, between 0 and 0.4. This relatively recent decline in the first-order autocorrelation makes it difficult to know whether the decline in persistence reflects an enduring feature of inflation or a small-sample phenomenon that may be the result of a temporary period characterized by small shocks to inflation. A Monte Carlo exercise addresses this question in Section 4

Overall, these simple measures support the conclusion that inflation is not currently well characterized as a process with a unit root, although in earlier decades it exhibited behavior that may have been consistent with the presence of a unit root. The

²¹ The rolling-sample estimates presented here employ a 14-year window, the same as in Pivetta and Reis.

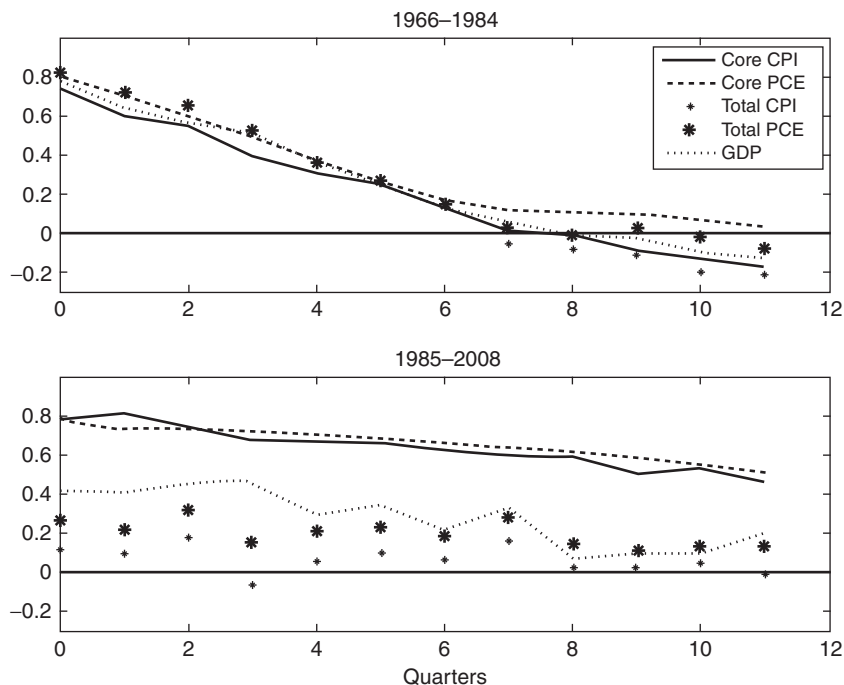


Figure 7 Autocorrelations of inflation data, various measures and samples.

autocorrelations also make it clear that there has been considerable time variation in inflation's reduced-form persistence.

2.2.4 Autocorrelation functions

Extending the previous section's results, Figure 7 displays the sample autocorrelation function for five of the key measures of inflation (CPI-X, PCE-X, CPI, PCE, and the GDP deflator) for two sample periods, 1966Q1–1984Q4 and 1985Q1–2008Q4.²² As Figure 7 suggests, over the first half of the past 43 years, inflation exhibited considerable persistence in this reduced-form sense, according to all measures. Since the mid-1980s, roughly corresponding to the onset of the “Great Moderation,” the persistence of inflation has declined for some, but not all, measures.²³ The two measures of “core” inflation show little reduction in persistence in the most recent sample, a feature that is echoed in the results for the dominant root in the next

²² I use both core and total measures here because, as discussed in the next subsection, the influence of key relative prices on the reduced-form persistence of inflation in the latter half of the sample can be significant. The beginning of the early sample is chosen to coincide with the first use of the federal funds rate as the Federal Reserve's policy instrument. Extending the sample back to 1959:Q2, the earliest date for which the PCE price series is available, produces results that are qualitatively similar.

²³ See Benati (2008) for an empirical investigation into changing inflation persistence.

section. The ambiguity of these results and the source of any reduction that may have occurred are discussed in some detail in Section 2.4.

2.2.5 Dominant root of the univariate time series process

An alternative measure of inflation persistence is the dominant root implied by the univariate autoregressive process for inflation. In particular, if the autoregressive representation of inflation is of lag length k

$$\pi_t = c_1\pi_{t-1} + \dots + c_k\pi_{t-k} + \varepsilon_t, \quad (10)$$

then the companion matrix for the state-space representation of π_t is

$$C \equiv \begin{bmatrix} c_1 & \dots & c_k \\ I_{k-1 \times k-1} & 0_{k-1 \times 1} \end{bmatrix}, \quad (11)$$

and the root of C with the largest magnitude is the (dominant) root of interest.²⁴

Table 4 summarizes the results for our three measures of inflation for a variety of samples.²⁵ It suggests a high degree of persistence over the past 25 years, with a modest decline for some measures as the earlier decades are dropped from the sample. In all cases, the precision of the estimate in the later samples declines, with the estimated standard error increasing by a factor of three to four in most cases. The results are somewhat dependent on the inflation measure; studies that focus on the GDP deflator (such as Pivetta & Reis, 2007) may well uncover less evidence of a decline in reduced-form persistence. Both the CPI and the PCE measures show a more pronounced decline in persistence, particularly in the post-1995 subsample. The standard errors in many cases make it difficult to reject the hypothesis of stable persistence across these samples.²⁶ Another widely used measure of reduced-form persistence is the sum of the autoregressive (AR) coefficients, which approximates the long-run impulse response to a unit shock $c(1) \equiv \sum_{i=1}^k c_i$. This measure is provided in the right-hand-most column of Table 4. As Table 4 illustrates, the mapping

²⁴ It is well known that least-squares estimates of these AR parameters are biased downward (see the literature beginning with Andrews, 1993a). In Andrews' results for first-order AR processes, for the sample sizes explored below over 150 observations, the bias even for fairly high estimated values of the AR parameter is relatively small, on the order of 0.02 for estimated coefficients as high as 0.95 (see Andrews, 1993a, Table II). A Monte Carlo exercise that estimates the dominant root for a fourth-order univariate AR finds that the (downward) bias in estimating the root from least-squares estimates of the AR parameters is between 0.001 and 0.05 for the sample sizes employed below. The true coefficients in the Monte Carlo exercise are calibrated to those estimated on the CPI, PCE, GDP, CPI-X and PCE-X data used in this chapter. The "true" dominant roots in these cases generally lie between 0.75 and 0.95. The largest bias arises for the CPI-X measure; the smallest arises for the CPI measure, which is nearly zero. This suggests that the least-squares bias is of modest concern in this application. Note that the standard errors of the estimated roots in Table 4 are generally quite a bit larger than the bias. Nonetheless, bias adjustments as estimated from the Monte Carlo exercise are applied to the estimated roots and coefficient sums in Table 4.

²⁵ Bias adjustments are applied to the estimated dominant roots; see end note 24.

²⁶ Standard errors are computed from a Monte Carlo simulation that draws 100,000 permutations of the estimated residuals for each inflation measure and sample, creating a new inflation series for each such residual permutation, given the original estimated AR coefficients and re-estimating the AR and the dominant root for each permutation.

Table 4 Dominant root of autoregressive process for inflation

Measure	Sample	Dominant root ^a (Std. Error)	Sum of AR coeff.'s (c(1)) ^a (Std. Error)
CPI	66–08	0.94 (0.032)	0.91 (0.055)
	66–84	0.92 (0.034)	0.92 (0.068)
	85–08	0.70 (0.12)	0.41 (0.18)
	95–08	0.64★ (0.11)	–0.039 (0.32)
PCE	66–08	0.96 (0.029)	0.94 (0.046)
	66–84	0.91 (0.034)	0.92 (0.065)
	85–08	0.82 (0.094)	0.61 (0.14)
	95–08	0.68★ (0.11)	0.22 (0.26)
GDP	66–08	0.97 (0.027)	0.95 (0.039)
	66–84	0.91 (0.033)	0.90 (0.075)
	85–08	0.90 (0.074)	0.72 (0.11)
	95–08	0.79 (0.12)	0.64 (0.17)
CPI-X	66–08	0.98 (0.027)	0.92 (0.042)
	66–84	0.92 (0.041)	0.85 (0.081)
	85–08	0.99 (0.039)	0.96 (0.054)
	95–08	0.75 (0.0641)	0.62 (0.16)

Continued

Table 4 Dominant root of autoregressive process for inflation—cont'd

Measure	Sample	Dominant root (Std. Error)	Sum of AR coeff.'s ($c(1)$) (Std. Error)
PCE-X	66–08	0.96 (0.023)	0.98 (0.033)
	66–84	0.91 (0.032)	0.90 (0.066)
	85–08	0.99 (0.044)	0.97 (0.061)
	95–08	0.72 (0.097)	0.49 (0.18)

^aBias adjusted

*Denotes complex roots

from the sum of the c_i 's to the dominant root is not uniform, particularly when the coefficients imply a complex pair of dominant roots (indicated by the *), as is the case for the CPI and the PCE in the latter subsample. In this case, even though the sum of AR coefficients may be small, the oscillatory behavior implied by the complex pair of roots may die out only slowly, implying significant persistence. For this reason, one should exercise caution in inferring persistence from the sum of the AR coefficients.

Because persistent relative price shifts can influence the persistence of measures of overall inflation, particularly in shorter samples, the bottom panel of Table 4 presents the same results for the core CPI and PCE inflation measures. Because these series are less noisy, the standard errors for the estimated dominant roots are almost uniformly smaller than those for the overall inflation measures in the top panel of the table.

The core-based measures provide less evidence of a decline in persistence in recent years. Neither CPI nor PCE inflation measures show a decline in the dominant root or the sum of the AR coefficients for the past 25 years. The dominant root estimate declines modestly for the period since 1995, but the standard error of this estimate is correspondingly larger. For the more recent and relatively short subsamples, these results suggest that large movements in key relative prices may well distort the extent to which underlying inflation persistence has changed, particularly if persistence is measured over a relatively short sample.

2.3 Evidence of changing reduced-form persistence in the United States

Recognizing the reduced-form nature of inflation persistence, a number of authors have looked for evidence that changes in the underlying determinants of inflation may have given rise to a change in reduced-form persistence. The leading cause of a change in inflation behavior is thought to be a change in the systematic behavior of the central bank. To illustrate potential underlying sources of changes in reduced-form persistence, consider the stylized, backward-looking model of inflation below

$$\begin{aligned}
\pi_t &= \pi_{t-1} + a x_t \\
x_t &= -b f_t \\
f_t &= c \pi_t
\end{aligned} \tag{12}$$

The first equation is a skeletal Phillips curve, in which the change in inflation is positively related to a variable, x_t , which we will take here to be the output gap. The output gap in turn depends negatively on the short-term policy rate, f_t (for federal funds rate), and the policy rate is a positive function of inflation (with an implicit target inflation rate of 0). The solution for inflation is

$$\begin{aligned}
\pi_t &= \alpha \pi_{t-1} \\
\alpha &\equiv \frac{1}{1 + abc}.
\end{aligned} \tag{13}$$

The solution for inflation is a first-order AR, which will exhibit less persistence — the coefficient α will be smaller — the larger the policy response to inflation (c), the more responsive the output gap is to the policy rate (b), and the more responsive inflation is to the output gap (a). In this simple framework, a central bank that behaves more aggressively in moving inflation toward its target will reduce the persistence of inflation more than one that moves less aggressively. A central bank that does not respond at all to inflation will render the inflation rate nonstationary: In this case, $\alpha = 1$, inflation follows a random walk, and the economy loses its nominal anchor. The intuition from this skeletal model generalizes to a number of more sophisticated models that include rational expectations and a richer description of the key elements sketched above.²⁷

2.3.1 Unknown breakpoint tests for univariate ARs

Several of the tests above are based on univariate ARs for inflation, using predetermined breakpoints to obtain a qualitative sense of the persistence of inflation across different subsamples. A more formal test for changes in the time series properties of inflation employs the unknown breakpoint testing methodology of Andrews (1993b) and Bai and Perron (1998). To isolate changes in the dynamic response of inflation to shocks around its mean from changes in the mean of inflation, we remove a time-varying mean from inflation by subtracting a very slow-moving Hodrick-Prescott filtered trend from the series. This simple methodology comes very close to the “inflation gap” methodology of Cogley and Sbordone (2008). The raw series and HP-filter-derived trend for the core CPI appear in Figure 8. The HP-filtered series all exhibit a zero mean and no evidence of a trend.²⁸

²⁷ A separate line of research examines the contribution of a time-varying inflation target to measures of inflation persistence. The chapter will return to this topic in more detail in Section 3.5.

²⁸ The HP filter parameter is set to 160,000. A filter parameter that is an order of magnitude larger or smaller does not qualitatively affect the results. In simple regressions of the detrended series on a constant, time trend, and trend squared over the sample 1966:Q1–2008:Q4, all of the series are estimated to have mean and trend terms that are insignificantly different from zero, and a corrected R^2 that is typically negative.

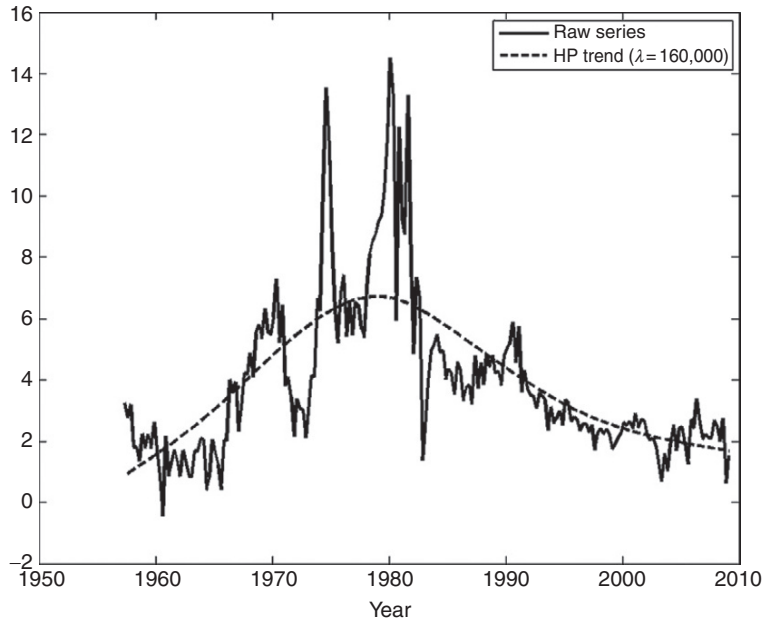


Figure 8 Inflation and its time-varying mean estimate (core CPI).

The results of the unknown breakpoint test are displayed in Table 5. The test regression is a four-quarter AR for the variable indicated, and it includes an intercept. All of the inflation series have the HP-filtered trend removed. As Table 5 suggests, the evidence of breakpoints in the AR dynamics of the inflation series is mixed. Inflation shows one significant break in the mid- to late-1970s for the core CPI and core PCE measures, no break at all for the GDP measure, and a single break somewhere in the last decade of data for the overall CPI and PCE measures. While this may be too stringent a test for changes in inflation dynamics, the results generally confirm the previous results and suggest caution in accepting findings that inflation dynamics have changed dramatically in the wake of arguably improved monetary policy. Section 4 discusses potential hazards in inferring changes in inflation dynamics in short samples, such as the post-1999 samples identified as possible breakpoints for the overall CPI and PCE measures in this unknown breakpoint test.²⁹

2.3.2 Pre-war inflation persistence under the gold standard

If persistence is expected to change as a result of changes in the monetary regime, it would be natural to compare inflation persistence across markedly different monetary regimes in economic history. The contrast in the United States between the post-

²⁹ The test uses the 5% critical value for rejecting $n + 1$ breaks in favor of n . Fifteen percent of the sample at either end is excluded from the set of potential breakpoints. The results are not sensitive to either assumption.

Table 5 Test for unknown breakpoints in univariate inflation process

Measure	Number of breakpoints	Dates
CPI	1	1999:Q1
PCE	1	1999:Q1
GDP	0	—
CPI-X	1	1980:Q2
PCE-X	1	1976:Q2

World War II fiat money system and the pre-World War I gold standard provides such a natural experiment. Barsky (1987) found that while U.S. inflation persistence was very high from 1960 to 1979, it was virtually nonexistent prior to World War I — indeed, Barsky’s ARIMA modeling of inflation during this period suggests that inflation is white noise. While some caveats apply to the results in Barsky’s analysis, his paper provides an excellent example of a more general point: the persistence of inflation (and more generally, its dynamic properties) must surely be influenced by the nature of the monetary regime in effect. A gold standard that aims to stabilize the price level must bear different implications from a fiat money standard that aims to stabilize the growth in the price level, which in turn must bear different implications from a fiat money standard that does not forcefully act to stabilize inflation around any particular goal. We reinforce this key link between monetary policy and inflation behavior in a variety of ways throughout this chapter.³⁰

2.3.3 A parameterized characterization of reduced-form persistence

Stock and Watson (2007) posited a relatively straightforward time series model of inflation that captures many of the features of reduced-form persistence discussed so far. The model can be expressed as an integrated moving-average process of order one, or IMA(1,1)

$$\Delta\pi_t = a_t - \Theta a_{t-1} \quad (14)$$

or equivalently as an unobserved components model with stochastic trend and stationary components³¹

³⁰ Barsky (1987) uses wholesale price data prior to World War I. To the extent that such data reflect movements of commodity prices rather than consumer goods and services prices, this result may be confounded by an inherent difference between commodities and final goods prices, rather than a difference in persistence across the monetary regimes. Note that Barsky (1987) also makes the link between persistence and forecastability, a link exploited in Cogley, Primiceri, and Sargent (2010), discussed in Section 2.3.4.

³¹ One can see the equivalence by substituting the definition for $\tau_t = \frac{e_t}{(1-L)}$ into the equation for inflation to obtain $\pi_t(1-L) = e_t + \eta_t(1-L)$, which, after rearrangement, yields an equation like (14), with inflation an integrated process with a moving average error term.

$$\pi_t = \tau_t + \eta_t \quad (15)$$

$$\tau_t = \tau_{t-1} + \varepsilon_t, \quad (16)$$

where η_t and ε_t are uncorrelated with one another, and are serially uncorrelated with mean zero and variances σ_η^2 and σ_ε^2 , respectively. One can think of τ_t as reflecting the “permanent” or trend component of inflation, and η_t as capturing the stationary component.

Interestingly, Stock and Watson (2007) found that the decline in the variance of inflation in recent decades is largely due to a marked decline in the variance of the permanent component, that is, a reduction in σ_ε^2 . It is still not possible, in their methodology, to reject the null of a unit root in inflation, but the variance of the shock that drives τ_t is estimated to be currently at historic lows.

2.3.4 Multivariate evidence of changes in reduced-form inflation persistence

For the most part, the measures of persistence discussed so far are univariate measures; that is, they use only the information in an inflation time series to draw inferences about its persistence. However, some authors have argued that one can draw more accurate inferences using a multivariate approach. In part, the intuition behind this claim is connected to the “trend inflation” models discussed in Section 3.5. In those models, the slow-moving or trend component of inflation accounts for much of the persistence in inflation. That trend in turn is most commonly associated with the central bank’s target rate of inflation. As a consequence, including variables that reflect the central bank’s inflation-targeting behavior, such as short-term policy rates, may help to identify both trend inflation and its persistence, and thus the persistence of inflation.

Cogley et al. (2010) used a time-varying VAR to estimate the trend component of inflation, which they associate with the central bank’s inflation target. They found continued persistence in inflation, but they associated it strongly with trend inflation. This implies that the Federal Reserve’s implicit target for inflation continues to have a unit root (or near-unit root), although Cogley et al. estimated the variance of that component to have declined, consistent with the findings in Stock and Watson (2007). The persistence of the “inflation gap” — the difference between actual inflation and its trend — appears to have declined in recent years.

Methodologically, Cogley et al. introduced a new measure of persistence that is related to the predictability of near-term movements in the variable of interest. Formally, persistence is calibrated by the R^2 of the j -step-ahead forecast of the variable. The higher the R^2 , the more predictable it is, and thus the more persistent, precisely because past shocks have a persistent influence on future inflation. They examine the R^2 s for 1-, 4-, and 8-quarter-ahead forecasts.³²

³² The intuition behind this measure of persistence is akin to the simple derivation of shock persistence in Eq. (9).

They find economically and statistically significant changes in the j -step-ahead R^2 s for the inflation gap from 1960 to 2006, with the 1-quarter ahead R^2 peaking at over 90% in the 1970s and early 1980s, falling to about 50% by the mid-1980s and through the end of their sample. The 4-quarter ahead R^2 s peak at 50 to 75% during the Great Inflation and decline to about 15% more recently; the 8-quarter ahead R^2 s peak at 20 to 35% in the same period, falling to 10% more recently. All of these changes appear to be quite significant statistically, judged by the estimated joint posterior distribution of the R^2 s in earlier and later periods.

2.4 International evidence of changing reduced-form persistence

A number of authors have developed empirical evidence on changes in the reduced-form persistence of inflation for samples of developed countries. Benati (2008) surveyed the evidence from a broad array of developed countries over long samples. His empirical work focuses on differences in estimated persistence across different monetary regimes, and his key hypothesis is that regimes that clearly anchor inflation (or the price level, as in the gold standard) induce less persistence in the inflation rate. Benati (2008) examined a number of European countries pre- and post-EMU; the UK, Canada, and Australia pre- and post-inflation targeting; and the United States pre- and post-Volcker disinflation.

A key finding in his paper, summarized in Table 6, reproduced from Benati (2008), is that reduced-form persistence has declined in recent years for all of the aforementioned countries that have adopted an inflation-targeting regime, and that this lower persistence is statistically quite different from the persistence exhibited prior to inflation targeting.³³ The conclusion from Benati's results is that countries that adopt a formal inflation-targeting regime are very likely to enjoy a decline in inflation persistence. The United States and Japan (not displayed in Table 6) are the obvious standouts, and they are also the countries that have not adopted a formal inflation-targeting regime. Benati (2008) also reported similar results using the degree of "indexation" in a structural New Keynesian model of the economy, that is, the estimates of μ in Eq. (23).³⁴

Levin and Piger (2004) also examined inflation persistence across a number of countries, focusing on the possibility that the reduced-form process for inflation has changed in recent years, perhaps owing to changes in the central bank's inflation objective. Their results, which employ unknown breakpoint methods in both the classical and Bayesian traditions, show that simply allowing for a change in the mean of inflation appears to reduce estimated reduced-form persistence for many of the countries in their sample. Other international evidence develops mixed results about changing inflation persistence. Ravenna (2000) documented a large post-1990 drop in Canadian inflation

³³ See Benati (2008), Tables I–VIII.

³⁴ Benati (2009) explored the stability of parameters that reflect intrinsic persistence across developed countries. In general, he found that these parameters, whether explicit or implicit, are not stable across changes in monetary regime.

Table 6 Estimates of reduced-form inflation persistence

Country	Early sample	Late sample	Test of difference
UK (RPI)	Bretton Woods to inflation targeting 0.95	Inflation targeting −0.07	0.00
Canada (CPI)	1971 to inflation targeting 0.90	Inflation targeting −0.33	0.00
Euro Area (GDP defl.)	Bretton Woods to EMU 1.01	EMU 0.35	0.00
U.S. (CPI)	Great Inflation 0.77	Post-Volcker 0.49	0.046
(PCE)	0.74	0.81	0.59

From Benati, 2008. With permission.

persistence. O'Reilly and Whelan (2005) employed methods very similar to those in Levin and Piger (2004), but found that for the Euro Area price indexes (as compared with Levin and Piger's individual-country price indexes) there has been no discernible change in inflation persistence.³⁵

2.5 Conclusions from the reduced-form evidence

From both theoretical and empirical perspectives, it seems likely that the contribution to inflation from its unit root component has diminished significantly in recent decades. In most macroeconomic models, inflation would contain an important unit root (in terms of contribution to variance) if the central bank was not acting to keep inflation low and stable, consistent with either an implicit or explicit inflation target, or if its target varied considerably over time, adding a significant “trend inflation” component to inflation that might imply a unit root. From a practical perspective, this suggests that macroeconomists can now think of inflation as a stationary series that will (in normal times) return to the central bank's inflation goal within a modest period, and that the central bank's inflation goal, while not written in stone (or anywhere else at present in the United States), is unlikely to vary significantly over time. Minor time variation in the inflation goal could add a small unit root component to inflation, but its contribution to the variance of inflation would likely be small.

With regard to the specific autocorrelation properties of a stationary inflation rate, the picture is considerably murkier. All authors agree that in the United States and many other developed countries, inflation exhibited considerable persistence from the 1960s through the mid-1980s. After that time, the statistical evidence is mixed.

³⁵ Note that their sample extends only through the fourth quarter of 2002. In communication with Luca Benati, he notes that his results would show the same stability in persistence if his sample stopped in 2002:Q4. Thus, evidence of a decline in European inflation persistence appears to arise primarily from the years from 2003 and forward.

For both the United States and other countries, studies fall on both sides of the argument about the possibility of declining reduced-form persistence. On a methodological note, for the United States, the evidence on changing persistence from so-called “core” measures appears to differ substantially from the evidence from so-called “headline” or total inflation measures. As a rule, the evidence of a change in persistence from core measures is less compelling than the evidence from headline measures.

Weighing all of the evidence, it seems reasonable to conclude that the persistence of inflation has declined somewhat in recent years. But how much it has declined and whether in the extreme case inflation is now a nonpersistent series, remain issues for further study. At the time of writing, we are in the midst of an economic environment characterized by large relative price swings and significant changes in common estimates of the output gap and marginal cost, factors that are thought to influence inflation. These conditions should provide data that will help economists test a number of aspects of inflation dynamics, including its persistence.

3. STRUCTURAL SOURCES OF PERSISTENCE

3.1 Inherited and intrinsic persistence

While establishing the degree of reduced-form persistence in inflation is an important first step, knowledge of the degree of reduced-form persistence is of limited use to a policymaker unless she can understand the underlying sources. As the policymaker contemplates potential policy actions, she must be able to determine whether or not the persistence is structural and thus may be taken as a stable feature of the economic landscape or instead is the reduced-form outcome of her own actions and the structure of the economy. In order to know this, she may find it useful to parse the sources of persistence into three types: (1) persistence that is generated by and thus inherited from the driving process, that is, by the behavior of the output gap or marginal cost; (2) persistence that is a part of the inflation process that is “intrinsic” to inflation (i.e., persistence that is imparted to inflation independent of the driving process); and (3) persistence that is induced by her own actions, which will often be reflected in the behavior of the driving process but more generally may be subsumed in the concept of inherited persistence. With respect to the last source, the research by Benati (2008) cited above suggests that central banks that are more explicit about their inflation goal — and act in accordance with that commitment — may enjoy less persistence in their nations’ inflation rates.

Disentangling these sources of inflation persistence is extraordinarily difficult in relatively short aggregate time series. This chapter will return to this issue later. To begin, it is important to distinguish theoretically among the potential sources of persistence in inflation. Significant differences in theoretical models will imply somewhat different ways of dissecting inflation persistence. We begin in Section 3.3 with the most widely used model of price-setting.

3.2 An alternative decomposition of persistence: Disinflations and supply shocks

For some economists, the decomposition of inflation persistence into intrinsic and inherited persistence may not map naturally into the defining macroeconomic events — the disinflation of the 1980s and the supply shocks of the 1970s — that they have in mind in considering the behavior of inflation in the postwar period. Thus, an alternative decomposition of persistence might focus instead on (1) persistence associated with reducing inflation from an inherited level (or perhaps a previous target level), as in the discussion of the sacrifice ratio above; (2) persistence associated with the response of inflation to a shock that moves inflation temporarily away from its inherited or desired level; and (3) persistence of the inherited or desired level, apart from the response to temporary shocks. These three concepts map well into the inherited/intrinsic taxonomy of the preceding subsection, and they also map well into the macroeconomic models that are discussed in Section 3.6.

The amount of persistence associated with an intentional disinflation from an inherited level will depend on several factors. If inflation exhibits intrinsic persistence, then a central bank that wishes to disinflate will need to take more vigorous policy actions (larger movements of its policy instrument) to offset the intrinsic inertia of inflation than if inflation does not exhibit intrinsic persistence. Of course, the more vigorously the central bank pursues its disinflation policy, the less persistence inflation will exhibit during the disinflation. The degree of vigor will be reflected in larger movements in the policy instrument, correspondingly larger movements in output and costs, and thus a quicker decline in inflation. Because the central bank's effects on inflation work through output and costs, this effect would be part of the inherited persistence component of inflation in the earlier taxonomy. Finally, the more fully credible the central bank's desire to disinflate, the more rapid the expectations incorporate a lower future inflation rate, and the less persistent the inflation will be in its decline to the new level.

As for the second, the response of inflation to temporary shocks that move inflation away from its inherited level will also depend on several factors. In the first instance, a one-time transitory shock to inflation will be manifested in a persistent deviation of inflation from its inherited level if inflation exhibits intrinsic persistence, perhaps because price-setters index current prices to past inflation. Equation (9) demonstrates how a series with higher intrinsic persistence (which in this simplified case simply means a larger AR coefficient) propagates temporary shocks further into the future. But in a more fully articulated model, the response of monetary policy to temporary shocks will also influence the persistence of inflation in response to such shocks. As shown next, if the central bank does not respond at least one-for-one to deviations of inflation from its desired level, inflation can wander indefinitely in response to temporary shock — it will contain a unit root that contributes importantly to the variance of inflation, and thus inflation will be extremely persistent. In sum, the combination of the central bank's response to inflation shocks and the intrinsic persistence of inflation jointly determine the persistence of inflation in response to such temporary shocks.

Finally, the persistence of the inherited or trend level of inflation, which recently has been taken to correspond to movements over time in the central bank's inflation target, is the focus of much of the literature beginning with the contributions of Cogley and Sbordone (2008). The topic is discussed in more detail in Sections 3.6 and 3.5. For the most part, this source of persistence can be seen as another kind of inherited persistence, as it derives fundamentally from the behavior of the central bank, not from the price-setting behavior of firms.

In what follows, this chapter adheres primarily to the first decomposition, because intrinsic and inherited persistence map neatly into standard and widely used modeling constructs, therefore clearly having testable implications (e.g., particular parameters will be expected to be large or small). The second decomposition, while intuitively appealing, reflects the response of inflation to particular sets of economic conditions. As a consequence, the empirical implications of the second decomposition for specific models of inflation are somewhat less clear.

3.3 Persistence in the Calvo/Rotemberg model

Many modern models of inflation derive from the seminal contributions of Calvo (1983) and Rotemberg (1982, 1983) and typically imply a Euler equation for inflation π_t that takes the form

$$\pi_t = \beta E_t \pi_{t+1} + \gamma x_t + \varepsilon_t. \quad (17)$$

E_t denotes the mathematical expectation using information available in period t , and β denotes the discount rate. The variable x_t represents a measure of the output gap or marginal cost, depending on details of the model. The role of the shock term, denoted ε_t , will be explored in greater detail later. The parameter γ is a function of the underlying frequency of price adjustment and the discount factor.³⁶ By iterating expectations forward and assuming that the expectation at time t of future shocks $\varepsilon_{t+i} = 0$, Eq. (17) can be expressed as

$$\pi_t = \gamma \sum_{i=0}^{\infty} \beta^i E_t x_{t+i} + \varepsilon_t. \quad (18)$$

As this rendering makes clear, inflation is the sum of two components, the discounted sum of expected marginal cost (say) and a shock that is by assumption *iid*, but that can in principle be serially correlated. This formulation clarifies the motivation for this inflation specification. In a Calvo world in which prices are expected to be fixed for some time, price-setters who can reset their prices set them equal to a markup over the discounted average of marginal cost that is expected to prevail over the expected life of the contract price.

³⁶ As Galí and Gertler (1999) demonstrated, $\gamma = \frac{(1-\theta)(1-\beta\theta)}{\theta}$. The more frequent the price adjustment, the larger the θ , and the smaller the coefficient on the driving process.

3.4 The analytics of inflation persistence: “Inherited” and “intrinsic” persistence

Expression (18) for inflation affords a natural taxonomy of the sources of inflation’s persistence. First, Eq. (18) implies that the inflation rate directly “inherits” the persistence in the variable x_t . If the output gap is a persistent series in the sense defined in Section 2, then other things equal, inflation will inherit some of that persistence. The reasons for persistence in the output gap or marginal cost; that is, the source of so-called “real rigidities,” is the subject of a number of papers (see, e.g., Blanchard & Galí, 2007, who focus on real wage rigidities that imply persistence in marginal cost, and Fuhrer, 2000 who focuses on rigidities in consumer spending arising from habit formation, which implies persistence in the output gap). Section 3.4.6 provides some empirical results on the persistence of widely used driving processes for the canonical inflation models.

3.4.1 The baseline case

In the simplest case, inflation is given by Eq. (17), and the process for x_t is a univariate first-order AR with autoregressive parameter ρ :

$$\begin{aligned}\pi_t &= \beta E_t \pi_{t+1} + \gamma x_t \\ x_t &= \rho x_{t-1} + u_t \\ \text{Var}(u_t) &= \sigma_u.\end{aligned}\tag{19}$$

For simplicity, no shock perturbs the inflation Euler equation ($\varepsilon_t = 0 \forall t$). In this case, one can show that the autocorrelation function for inflation is³⁷

$$A_i = \rho^i.$$

That is, inflation inherits exactly the autocorrelations of the first-order autoregressive process describing x_t . In this simple version of the model, the effects of monetary policy, real rigidities in consumption, or real wages — in short, the behavior of inflation arising from any aspect of the economy — must enter through their effects on x_t .³⁸

3.4.2 More complex cases

The autocorrelations of inflation become more complex when one allows for

- Nonzero shocks to the Euler equation ($\varepsilon_t \neq 0$)
- Variation in the size of γ , given nonzero shocks
- The possibility of some “backward-looking” element to inflation, as in Buiter and Jewitt (1981), Fuhrer and Moore (1995), or Christiano, Eichenbaum, and Evans (2005).

³⁷ See Fuhrer (2006) for derivations.

³⁸ Models in which price changes are state- rather than time-dependent allow for inherited persistence as well. Dotsey, King, and Wolman (1999) and Burstein (2006) examined cases in which variations in the size and persistence of money shocks result in more or less persistent inflation responses to money shocks.

These added complexities make the identification of underlying sources of persistence correspondingly complex, both theoretically and empirically. The following subsections derive the analytical results for these cases.

3.4.3 Shocks to the Euler equation

The augmentation of the Euler equation with an *iid* shock changes the interpretation of inflation persistence quite significantly, and in a way that is not well recognized in much of the literature on this subject. Modifying Eq. (19) to include this disturbance

$$\begin{aligned}\pi_t &= \beta E_t \pi_{t+1} + \gamma x_t + \varepsilon_t \\ x_t &= \rho x_{t-1} + u_t \\ \text{Var}(\varepsilon_t, u_t) &= \Sigma \equiv \begin{bmatrix} \sigma_\varepsilon^2 & 0 \\ 0 & \sigma_u^2 \end{bmatrix}\end{aligned}\quad (20)$$

makes a subtle but important difference in the autocorrelation function for inflation. The presence of the *iid* shock ε_t now makes the behavior of inflation a mixture of its inherited persistence from current and expected future x_t , with weight γ , and the nonpersistent shock process (with implicit weight of one). The larger the variance of the shock process, the more inflation looks like white noise with zero persistence. The smaller the γ , the smaller the importance of x_t in determining the autocorrelation of inflation will be.

Normalizing the variance of the shock to the x_t process to 1 for convenience, one can express the inflation autocorrelations in this case as

$$\begin{aligned}A_i &= \frac{\rho^i \gamma^2}{a \sigma_\varepsilon^2 + \gamma^2} \\ a &= (1 - \rho^2)(1 - \rho\beta)^2\end{aligned}\quad (21)$$

Note that Eq (21) indicates that the autocorrelations decay at rate ρ (the expression is pre-multiplied by ρ^i , and this is the only expression that varies with i). More generally, the expression suggests that inflation will be more autocorrelated the

- Higher ρ is; that is, the greater the persistence of the real driving variable x_t ;
- Higher γ is the larger the coefficient on the driving variable x_t , and thus the more of x_t 's persistence is inherited by inflation
- Smaller the variance of the shock ε_t is that disturbs inflation from the Euler equation.³⁹

Table 7 provides the first autocorrelation of inflation for various values of the key parameters in the simple inflation model.⁴⁰ For values of γ that correspond to those estimated in the literature (generally below 0.1), a modest relative variance for ε_t can imply a fairly low first autocorrelation. For example, if γ is estimated to be 0.05, and the variances

³⁹ Because we have normalized σ_u to 1, this should be interpreted as the smaller is σ_ε relative to σ_u .

⁴⁰ The full derivation for the results in Table 7 is provided in Fuhrer (2006).

Table 7 Value of A_1 for selected values of σ_e^2 and γ

σ_e^2	γ				
	.01	.03	.05	0.1	0.2
$\rho = 0.9$					
0	0.90	0.90	0.90	0.90	0.90
0.1	0.25	0.70	0.81	0.88	0.89
0.3	0.10	0.48	0.68	0.83	0.88
0.5	0.06	0.36	0.59	0.79	0.87
1	0.03	0.23	0.44	0.71	0.84
3	0.01	0.09	0.22	0.50	0.75
5	0.01	0.06	0.14	0.39	0.68
$\rho = 0.95$					
0.5	0.29	0.76	0.87	0.93	0.94
3	0.06	0.37	0.61	0.83	0.92
$\rho = 0.99$					
0.5	0.91	0.98	0.99	0.99	0.99
3	0.65	0.93	0.97	0.98	0.99

of the two shocks are the same, the first autocorrelation is 0.44. As the autocorrelation of the driving process approaches 1, the persistence of the driving process begins to dominate the white noise of the error term, as shown in the bottom panels of the table. Thus, even in this very simple model, it is clear that a persistent driving process does not need to impart any persistence to inflation, depending on the sizes of γ and σ_e .

3.4.4 The pivotal role of the coefficient on x_t

As the analytical results of the preceding subsection suggest, the influence of x_t on inflation – the size of the parameter γ – is pivotal both in interpreting the sources of inflation persistence and in identifying Eq. (20) as a Phillips curve or aggregate supply relation. If $\gamma = 0$, then (a) the Euler equation can no longer be interpreted as a Phillips curve; (b) inflation becomes decoupled from marginal cost or the output gap, and thus from monetary policy; (c) in most models, this decoupling will lead to indeterminacy for inflation, as monetary policy can no longer determine the steady-state value of inflation; and (d) inflation no longer inherits any persistence from x_t .⁴¹

⁴¹ To see point (c), consider the simplified model in Eqs. (12), and its solution in Eq. (13). When the Phillips parameter $a \rightarrow 0$, the solution for π_t becomes $\pi_t = \pi_{t-1}$. Inflation is a random walk, and thus fails to converge to any value in particular. This logic transfers to more complex specifications with explicit expectations.

Given the centrality of the parameter γ , it is of interest to determine how well identified this parameter is in the data. The answer varies from study to study, but a brief empirical exercise may help to illuminate the potential problems. Consider the generalized method of moments (GMM) estimates of a Euler equation that follow the format of Galí and Gertler (1999) in allowing for “rule-of-thumb” price-setters in addition to the Calvo price setters:

$$\pi_t = \lambda_b \pi_{t-1} + \lambda_f E_t \pi_{t+1} + \gamma x_t + \varepsilon_t. \quad (22)$$

The parameters λ_b and λ_f calibrate the amount of backward- and forward-looking price-setting behavior that influences inflation. Following Galí and Gertler (1999), we employ an instrument set that consists of four lags each of inflation, real marginal cost, a measure of the output gap, wage inflation, the spread of the 10-year Treasury constant maturity rate over the federal funds rate, and oil prices. Table 8 summarizes the results.

Only when allowing for 12th-order correlation in the weighting matrix does the coefficient on marginal cost enter with the correct sign and significantly at the 5% level. These results are provided as suggestive of the difficulties in identifying γ ; they are broadly consistent with the aggregated results found in Galí and Gertler (1999) and Rudd and Whelan (2006).⁴²

For comparison, the lower panel of the table provides maximum likelihood and Bayesian estimates of the same model, augmenting Eq. (22) with VAR equations for the variables employed as instruments above.⁴³

The VAR coefficients are taken as fixed from OLS estimates over the same sample, and the ML estimates of the λ 's (the weights on lagged and expected inflation in the hybrid model) and γ are presented in the table, along with BHHH standard errors. Once again, the estimates of γ are quite small and not significantly different from zero. Note that the ML estimate of the backward-looking component is somewhat larger than the GMM estimate, and is quite precisely estimated. The likelihood-ratio test for the restriction that λ_b and λ_f take the GMM values (with γ freely estimated) has p -value 0.0000. As we will see in the section on “trend inflation” models (Section 3.5), a pattern is emerging in which more tightly constrained models provide larger estimates of intrinsic inflation persistence.

The conclusion is that identification of the slope parameter in the New Keynesian Phillips curve poses a significant econometric challenge. Nonetheless, it is quite obviously central to the theory of inflation, both in the sense that the Euler equation is nearly meaningless without it, and in the sense that the value of γ bears important implications for the transmission of monetary policy and the persistence of inflation.

⁴² See Mavroeidis (2005) for a careful treatment of the difficulties in identifying New Keynesian Phillips curves.

⁴³ The Bayesian priors for the three parameters are conventional, with generalized beta densities for the λ 's, and a gamma density for γ . The prior distributions for the three parameters are centered on [0.5, 0.5, 0.05], respectively, with standard deviations of [0.2, 0.2, 0.02], respectively. The posterior distributions are estimated using a Markov-Chain Monte Carlo algorithm, with four simulation blocks of 200,000 draws each.

Table 8 Estimates of parameters in equation 20 Sample period: 1960–1997 GMM estimates (HAC Standard errors in parentheses)

Number of terms in weight matrix	λ_b	λ_f	γ
ma=4	—	0.99	−0.00067
	—	(0.012)	(0.0060)
	0.34	0.65	0.0042
	(0.044)	(0.045)	(0.0042)
ma = 12	—	0.99	0.0020
	—	(0.0088)	(0.0046)
	0.36	0.63	0.0065*
	(0.024)	(0.025)	(0.0033)
ML estimates (BHHH Standard errors in parentheses)			
	λ_b	λ_f	γ
	0.51	0.48	0.0047
	(0.027)	(0.031)	(0.0028)
Bayesian estimates (Max. of posterior, estimated posterior sd's in parentheses)			
	λ_b	λ_f	γ
	0.51	0.48	0.0184
	(0.044)	(0.044)	(0.014)

*significance level

3.4.5 Hybrid models of inflation and “intrinsic” persistence: Including lagged inflation

The debate over the empirical success of the basic specification summarized in Eq. (17) continues, with the ability of the specification to replicate the reduced-form persistence of inflation an important focus. A number of authors have proposed rationales for the presence of a lagged inflation term in their aggregate supply relation (aka intrinsic persistence), through indexation of price contracts (see Christiano et al., 2005), “rule-of-thumb” behavior (see Galí & Gertler, 1999), alternative contract assumptions (see Fuhrer & Moore, 1995), alterations to the Calvo framework that assume a rising, rather than a constant hazard for the ability to reset prices (see Mash, 2004; Sheedy, 2007), or alternatives to rational expectations (see Section 3.9, Orphanides & Williams, 2004;

Roberts, 1997). Woodford (2007) provided a very helpful summary of the state of modeling intrinsic inflation persistence.⁴⁴

An augmented Phillips curve specification that allows for the influence of lagged inflation takes the form

$$\pi_t = \mu\pi_{t-1} + (\beta - \mu)E_t\pi_{t+1} + \gamma x_t + \varepsilon_t, \quad (23)$$

with the rest of the specification as detailed in Eq. (20). In a sense that is central to this chapter, the presence of lagged inflation provides an augmented channel for what might be called “intrinsic” inflation persistence; that is, persistence that is not inherited from the driving process x_t . In the simpler model of Eq. (20) the *iid* shock ε_t provided a trivial source of intrinsic persistence; depending on the relative variance of that shock and the coefficient on x_t , the persistence inherent in x_t would be more or less inherited by inflation.

With a lag of inflation added in Eq. (23) any shock to the Euler equation will persist for longer, other things equal, independent of the evolution of the driving variable. A shock to inflation becomes part of the history of inflation, independent of shocks to x_t . In addition, the forward-looking component of the model incorporates the direct dependence on history, augmenting the direct effect on persistence. One can think of the model as comprising both rule-of-thumb and sophisticated forward-looking price-setters. The forward-looking price-setters, in forming an expectation for future inflation, must take into account the behavior of the rule-of-thumb price-setters, who set current prices based on lagged inflation. Thus, the sophisticated price-setters’ behavior reinforces the behavior of the rule-of-thumb price-setters.

To see this algebraically and graphically, consider the analytic expression for the autocorrelations of the model augmented with a lag:

$$A_1 = \frac{a}{b\sigma_e^2 - c\rho\mu} d + \lambda_s, \quad (24)$$

where $[a, b, c, d]$ are functions of the stable root λ_s (in turn a function of β and μ) and the other underlying structural parameters of the model. Fuhrer (2006) showed that the additive term in λ_s dominates A_1 , and that λ_s and thus A_1 rise monotonically with μ . Figure 9 plots the stable root and the first autocorrelation of inflation as a function of μ .⁴⁵ In generating Figure 9, we set $\beta = 0.98$, $\gamma = 0.05$, $\rho = 0.9$, and $\sigma_e^2 = 1$. It shows that the first autocorrelation rises rapidly from about 0.4 to above 0.9 as μ increases

⁴⁴ Disaggregated price data provide limited support for many of these theoretical rationales for intrinsic persistence. The data examined in Bils and Klenow (2004) and Nakamura and Steinsson (2008), for example, provide little evidence that prices rise at a roughly constant rate between more significant resets, as might be implied by firms following a rule-of-thumb, or indexation.

⁴⁵ Taken from Fuhrer (2006), Figure 2. Note that in this case, while the algebra is a bit messier, it can also be shown that the autocorrelations decay at rate ρ , so ρ and the first autocorrelation are sufficient statistics for the autocorrelation function.

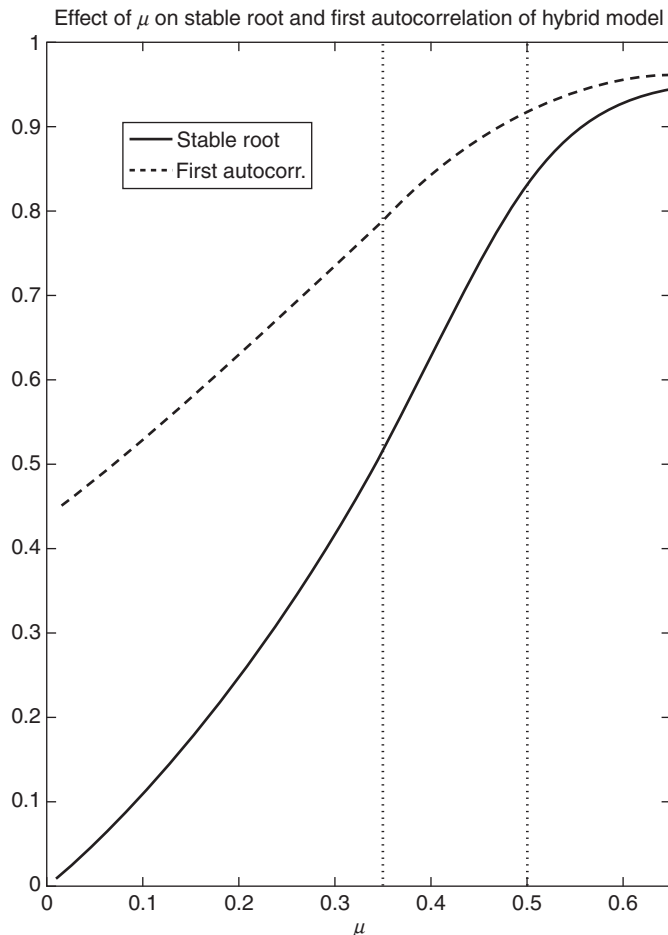


Figure 9 Dependence of stable root and first autocorrelation on μ .

from 0 to 0.6. It further emphasizes the role that forward-looking behavior plays in augmenting the direct effect of lagged inflation in the model.

Table 9 provides the first autocorrelation for this model for a variety of parameter settings. In particular, the size of μ varies from 0.1 to 0.9, and the relative variance of ε varies from 0 to 5. In addition, Table 9 displays the sensitivity of A_1 to the value of ρ .

As shown in the top panel of Table 9, for relatively high values of σ_ε^2 that significantly lower the first autocorrelation in the purely forward-looking model (the first column of Table 9, $\mu = 0$), a modest lag coefficient dramatically raises the persistence of inflation. The lower panel suggests that even with very little inherited persistence — $\rho = 0.5$ — a moderate value of μ implies a high degree of inflation persistence. An important implication of Table 9 is that it will be difficult to distinguish among

Table 9 Value of A_1 for selected values of σ_e^2 and μ $\gamma = 0.05, \beta = 0.98, \rho = 0.9$

σ_e^2	μ					
	0.0	0.1	0.3	0.5	0.7	0.9
0	0.90	0.92	0.96	0.99	1.00	1.00
0.3	0.68	0.74	0.86	0.96	0.98	0.99
0.5	0.59	0.66	0.82	0.94	0.97	0.98
1	0.44	0.53	0.74	0.92	0.97	0.98
2	0.29	0.39	0.64	0.89	0.96	0.98
3	0.22	0.32	0.59	0.88	0.96	0.98
5	0.14	0.25	0.54	0.86	0.96	0.98

$\rho = 0.5$						
σ_e^2	μ					
	0.0	0.1	0.3	0.5	0.7	0.9
0	0.5	0.58	0.76	0.94	0.99	0.99
0.3	0.02	0.13	0.44	0.84	0.96	0.98
0.5	0.012	0.12	0.43	0.84	0.96	0.98
1	0.0063	0.12	0.43	0.84	0.96	0.98
2	0.0032	0.11	0.42	0.83	0.96	0.98
3	0.0021	0.11	0.42	0.83	0.96	0.98
5	0.0013	0.11	0.42	0.83	0.96	0.98

sources of inflation persistence without imposing some structure on the inflation process, because the reduced-form implications of an inflation process that inherits a highly persistent driving process can be nearly the same as those of an inflation process that inherits little persistence from the driving process, but has a modest amount of intrinsic persistence.

3.4.6 The persistence of the driving process

Most researchers will agree that the observed persistence of inflation is determined at least in part by the inherited persistence of the driving process. Thus, in thinking about potential structural causes of a change in reduced-form persistence, a natural (but so far largely unexplored) question is to what extent there have been changes in the persistence of the

driving process.⁴⁶ In this section, we employ many of the same persistence measures that were previously used for inflation. We consider three candidates for the driving process: a measure of real marginal cost, proxied by the labor share (or equivalently, real unit labor costs) of the nonfarm business sector, and two measures of the output gap, the first a Hodrick-Prescott detrended log GDP gap, and the second the log deviation between real GDP and the Congressional Budget Office's estimate of potential GDP. We look at several subsamples. The first autocorrelation, the sum of the AR coefficients, and the dominant root of the AR process are displayed for each driving process and each subsample. The results are summarized in Table 10. As Table 10 indicates, there is remarkable stability in persistence measures across subsamples in all three proxies for the driving process, and for all three measures of persistence. It suggests little evidence of a change in persistence for the driving variables most common with inflation. For many of the measures and inflation variables, persistence appears to have *increased* in the more recent samples.⁴⁷ The results suggest that the stronger hypothesis that inflation has lost all its autocorrelation — that is, that inflation is an *iid* time series — is hard to justify. Unless the driving process is utterly decoupled from inflation, inflation must inherit some of its persistence. If inflation were completely decoupled from output or marginal cost, we would need an entirely new theory of inflation that steps outside the historical tradition of Phillips, Gordon, Calvo, and Rotemberg.⁴⁸

This decoupling between the (still ambiguous) evidence of declining reduced-form inflation persistence and a relatively stable and persistent driving process may help guide the search for structural interpretations of possibly changing inflation persistence. One simple interpretation is that the evidence for changes in reduced-form persistence is weak, and the stable, high persistence of the driving variables is consistent with that observation. Another is that while the inherited persistence from the driving process may not have changed much, the importance of the lagged inflation term in Phillips curves has diminished, leading to diminished intrinsic persistence in the face of unchanged inherited persistence. It may also be that the Phillips slope parameter and relative error variances have changed, which would affect the extent to which the unchanged persistence in the driving process is inherited by inflation. The next sections examine a number of these possibilities from the perspective of a standard dynamic stochastic general equilibrium (DSGE) model. Finally, it may be that inflation persistence

⁴⁶ As demonstrated earlier, a change in the coefficient on the driving process or in the relative variances of the inflation shock and the shock to the driving process may also affect the persistence of inflation.

⁴⁷ Of course, a number of authors have suggested that this standard proxy for real marginal cost is imperfect, and others have derived model-based measures of the output gap that can differ significantly from the simple measures used here. The results presented above are suggestive, and further research is warranted.

⁴⁸ This logic applies to “inflation gap” models of inflation as well, see Cogley and Sbordone (2008) and Cogley, Primiceri, and Sargent (2010). In this case, the deviation of inflation from trend inflation still depends on the expected and possibly lagged values of marginal cost. Thus, inflation will inherit the persistence of the driving process in these models as well.

Table 10 Estimated persistence of driving variables

Driving variable	1966–2008	66–83	84–08	95–08
First autocorrelation				
Real mc	0.92	0.80	0.92	0.90
HP gap	0.85	0.86	0.84	0.80
CBO gap	0.92	0.91	0.92	0.90
Sum of the AR coefficients ^a				
Real mc	0.97	0.78	0.99	0.99
HP gap	0.79	0.79	0.82	0.84
CBO gap	0.92	0.92	0.97	0.99
Dominant root of the AR process ^a				
Real mc	0.98	0.69	0.99	0.99
HP gap	0.81	0.82	0.79	0.79
CBO gap	0.78	0.78	0.82	0.87

^aBias adjusted.

has changed because the persistence of the trend component of inflation (Cogley & Sbordone, 2008) has declined. Evidence regarding this hypothesis is discussed in Section 3.5.

3.5 Persistence in models of “trend inflation”

A series of papers beginning with Cogley et al. (2010) and Cogley and Sbordone (2008) emphasizes the importance in modeling inflation of recognizing the slowly moving component of inflation that they dub “trend inflation.” The Cogley-Sbordone paper introduces two innovations. First, because long-run inflation is not a constant, the typical simplifications that give rise to the log-linearized Calvo model of Eq. (17) no longer apply. The standard log-linearization depends on the constancy of the long-run value of inflation.⁴⁹ Cogley and Sbordone derived the log-linear approximation that is appropriate when the long-run value of inflation has a trend. Hatted variables in the next equation denote deviations from steady-state values; for inflation, this implies the deviation of inflation from trend inflation

⁴⁹ See, for example, Woodford (2003) for the derivation.

$$\begin{aligned}\hat{\pi}_t = & \rho_t^\pi (\hat{\pi}_{t-1} - \hat{g}_t^\pi) + \zeta_t \hat{m}c_t + b_{1t} \tilde{E}_t \hat{\pi}_{t+1} + b_{2t} \tilde{E}_t \sum_{j=2}^{\infty} \varphi_{1t}^{j-1} \hat{\pi}_{t+j} \\ & + b_{3t} \tilde{E}_t \sum_{j=0}^{\infty} \varphi_{1t}^j (\hat{Q}_{t+j,t+j+1} + \hat{g}_{t+j+1}^y) + u_t\end{aligned}\quad (25)$$

where $\hat{g}_t^\pi$ and $\hat{g}_t^y$ are the innovation to trend inflation and the growth rate of real output, respectively; $\hat{m}c$ is the deviation of real marginal cost from its steady state; and $\hat{Q}_{t+j,t+j+1}$ is the one-period discount factor between periods $t+j$ and $t+j+1$. A key parameter in this specification is ρ_t^π , which calibrates the degree to which a lagged inflation term is required to match the AR properties of inflation, once trend inflation is accounted for. Regardless of the estimate of ρ_π , this specification is useful for researchers who wish to allow for time variation in the steady-state value of inflation. The most likely source of such trend variation is variation in the central bank's inflation target.

Second, Cogley and Sbordone found that the point estimate for ρ_t^π , the coefficient on lagged inflation in this specification, centers on zero. That is, once the model has accounted for the slow-moving variation in trend inflation, there is no need for a lag of inflation to account for the reduced-form persistence of inflation. While this empirical finding is controversial, the concept of trend inflation, which the authors associate with the central bank's time-varying inflation goal, is an important contribution to the inflation literature.

3.5.1 Cogley and Sbordone's measure of trend inflation

Table 11 displays the first autocorrelation for actual and detrended inflation, using Cogley and Sbordone's measure of trend inflation, for the subsamples in table 1 of their paper.⁵⁰ The table shows that the autocorrelations of detrended inflation are somewhat lower than those of the raw inflation data. However, in contrast to the table in Cogley and Sbordone (2008), the differences are small for the first two samples. For the most recent sample, the autocorrelation declines *both* for the detrended and for the raw series. Most of the decline in the detrended data's autocorrelation can be explained by a corresponding decline in that of the raw data. This suggests that there are other, equally important factors at work in explaining the persistence of inflation, and in explaining changes in persistence over time.

While not presented in Cogley and Sbordone (2008), it is of interest to examine their model's implication for the first autocorrelation of inflation. Using values of their key parameters that center on the median estimates over time presented in Figure 4 of their paper; that is, $\rho_t^\pi \approx 0$, $\zeta_t = 0.03$, $b_1 = 0.9$, $b_2 = 0.02$, and $b_3 = 0$, and assuming a

⁵⁰ Inflation is measured as four times the log change in the GDP deflator, as in Cogley and Sbordone (2008). The measure of trend inflation was kindly provided by the authors and replicates that in Figure 1 of their paper.

Table 11 First autocorrelation of inflation

Sample	Detrended π	Raw data
60:Q3–03:Q4	0.81	0.89
60:Q3–83:Q4	0.83	0.88
84:Q1–03:Q4	0.36	0.56

diagonal covariance matrix with variances estimated from a VAR over their sample period, I obtain a first autocorrelation for inflation of 0.22.⁵¹ This estimate differs markedly from their data’s first autocorrelation over the same sample, 1960–2003. Matching the autocorrelation of inflation from the data requires a significantly different set of parameters; for example, setting ρ_π near unity raises the first autocorrelation toward 0.8.

A recent paper by Barnes, Gumbau-Brisa, Lie, and Olivei (2009) examined the robustness of the finding that the detrended inflation model implies a value for ρ_t^π of 0. Details of Cogley and Sbordone’s estimation procedure make a significant difference to the estimation results. Barnes et al. (2009) showed that simply changing the form of the Euler equation, which implicitly imposes an additional constraint that is implied by the model, completely reverses the finding on ρ_t^π . They developed a precise estimate of this key “intrinsic persistence” parameter of about 0.8. This finding suggests caution in interpreting the rather striking implications of trend inflation models for inflation persistence.

3.6 Using a DSGE model to interpret structural sources of persistence

The skeletal model summarized in Eqs. (20) and (23) provides important insights into some of the structural sources of inflation persistence. However, the model leaves implicit the determination of real output and the role of monetary policy in influencing output and inflation. In theory, both the systematic component of monetary policy and the nature of the transmission of policy through the real side can have significant effects on the dynamic properties of inflation. In this section, we explore the quantitative effects on inflation of various aspects of an articulated DSGE model.

⁵¹ Note that setting b_1 to 1.05 as in Figure 4 of Cogley and Sbordone (2008) implies multiple solutions. I reduce the value of b_1 to 0.9 to keep it as high as possible, while still obtaining a unique solution to the model. The VAR employed in the exercise includes four lags each of the inflation rate from the GDP deflator, real marginal cost defined as in Cogley–Sbordone, the federal funds rate, and an output gap defined as the log-difference between real GDP and Hodrick–Prescott filtered real GDP. The VAR equations for marginal cost, the funds rate, and the output gap are used in conjunction with Eq. (25) to compute stability conditions, and to compute autocorrelations given the estimated covariance matrix of the shocks.

The model remains relatively simple. It comprises the “hybrid” inflation model discussed earlier, extended to allow for the presence of “trend inflation” as detailed in Cogley and Sbordone (2008) and in Section 3.5: the possibility of serially correlated markup shocks to the inflation Euler equation, perhaps motivated by sizable and persistent shocks to goods in a flexible price sector, such as energy or other imported goods (see de Walque, Smets, & Wouters, 2005); an optimizing IS curve that links real output to expected short-term real interest rates, allowing for real rigidity in the form of a lagged output term that can be motivated by the presence of habits in the consumer’s utility function;⁵² and a canonical policy or Taylor (1993) rule that makes the short-term policy interest rate a function of deviations of inflation and output from their desired levels. The last possibility also allows for any interest-rate smoothing.

The model can be summarized in the following set of equations⁵³

$$\begin{aligned}
 \pi_t - \bar{\pi}_t &= \rho_\pi(\pi_{t-1} - \bar{\pi}_{t-1} - (\bar{\pi}_t - \bar{\pi}_{t-1})) + b_1 E_t(\pi_{t+1} - \bar{\pi}_{t+1}) + \gamma \tilde{y}_t + \varepsilon_t \\
 \bar{\pi}_t &= \bar{\pi}_{t-1} + \zeta_t \\
 \varepsilon_t &= \rho_m \varepsilon_{t-1} + m_t \\
 \tilde{y}_t &= \mu_y \tilde{y}_{t-1} + (1 - \mu_y) E_t \tilde{y}_{t+1} - \gamma_\rho (r_t - E_t \pi_{t+1}) + u_t \\
 r_t &= \rho_r r_{t-1} + (1 - \rho_r) [a_\pi \pi_t + a_y \tilde{y}_t].
 \end{aligned} \tag{26}$$

The first equation expands the deviations notation of Eq.(25). As in Eq. 25, $\bar{\pi}_t$ is the t -period value of trend inflation, which is assumed to follow a random walk with shock ζ_t and variance σ_ζ^2 . The markup shock ε_t is assumed to follow a first-order AR process with AR(1) parameter ρ_m and shock m_t with variance σ_m^2 . The three shocks to the system, ζ_t , m_t , and u_t , are assumed to be independent and *iid* with diagonal covariance matrix

$$\begin{bmatrix} \sigma_\zeta^2 & 0 & 0 \\ 0 & \sigma_m^2 & 0 \\ 0 & 0 & \sigma_u^2 \end{bmatrix}.$$

The baseline parameters for the model are displayed in the second column of Table 12. The baseline value for the variance of trend inflation is taken from Cogley and Sbordone’s (2008) estimated trend inflation series. A univariate process is estimated for the series, and the conditional variance is used as an estimate of σ_ζ^2 .⁵⁴

⁵² See, for example, Fuhrer (2000).

⁵³ While the model affords a more structural decomposition of inflation persistence than is possible using the model in Eqs. (20) and (23), it still abstracts from some potentially important influences on inflation dynamics. This model uses the output gap, rather than the more common marginal cost measure, and thus abstracts from the role of wages and productivity in the inflation process.

⁵⁴ Consistent with the model in Eqs. (26), the coefficient on the lagged trend inflation series takes a value that is statistically indistinguishable from one.

Table 12 Baseline and alternative parameter sets for DSGE model

Parameter	Baseline value	Alternative values
ρ_π	0.6	(0.85, 0)
b_1	0.39	(0.1, 1.05)
π_y	0.1	0.025
ρ_m	0	0.9
μ_y	0.50000	0
γ_ρ	0.1	
ρ	0.8	
a_π	1.5	3
a_y	0.5	
σ_ε^2	0.05	0
σ_m^2	0.5	
σ_u^2	0.5	

We vary the values of the key parameters in the model to gauge the effect on the autocorrelations of inflation of changes in the systematic behavior of monetary policy, the variance of the shock to trend inflation, and changes in the price-setting and output sectors. The goal is to determine the extent to which the persistence of inflation — whether fixed or changing over time — can plausibly be attributed to specific underlying structural features, or to changes in those features.

Figures 10 and 11 display inflation's autocorrelation function for a variety of parameter configurations. The autocorrelation function corresponding to the baseline parameters in Table 12 is displayed in the solid line in all of the panels. It mirrors the properties of the autocorrelation functions displayed in Figure 7: The autocorrelations are high for the first several quarters, decaying gradually toward zero and turning negative for several quarters thereafter.

Figure 10 displays the autocorrelation function when the variance of trend inflation is at its baseline value. The top panel displays the change in inflation's autocorrelation function when the parameters governing the central bank's behavior and the response of inflation to the real side are altered. A sizable shift in the emphasis on inflation, a_π — from the conventional 1.5 to 3.0 — reduces the autocorrelations of inflation noticeably (the heavy dashed line in Figure 10).⁵⁵ The first autocorrelation of inflation is reduced

⁵⁵ A similar result, not shown, is obtained for a large shift in the emphasis on output in the policy rule.

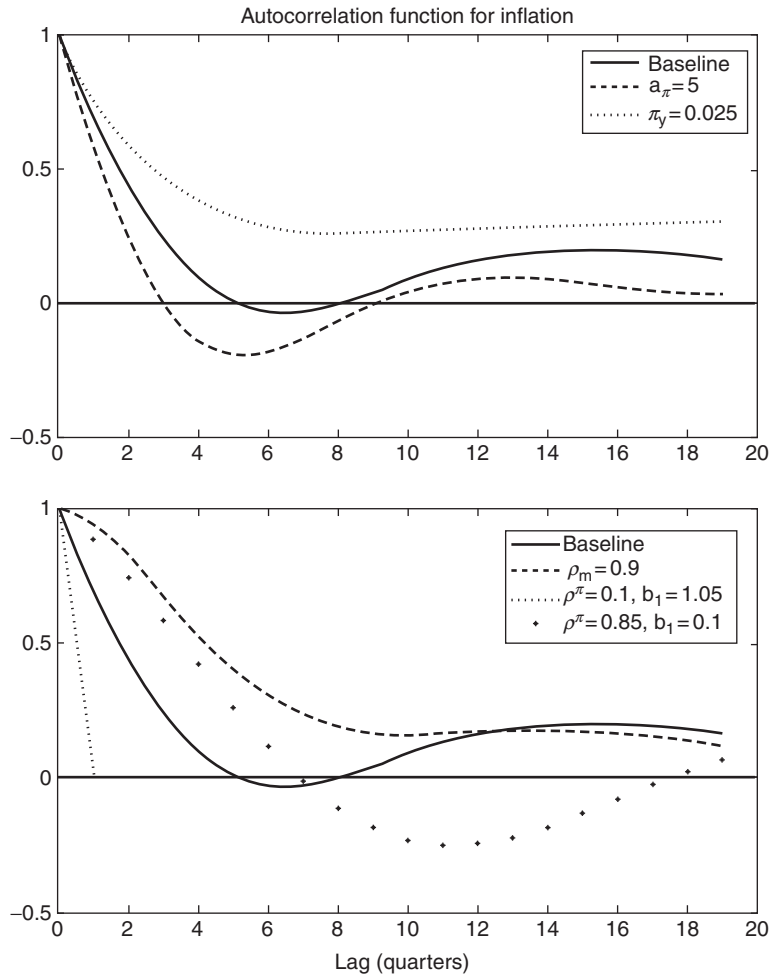


Figure 10 Effect of key structural parameters on inflation persistence, DSGE model.

from about 0.69 to 0.57, and subsequent autocorrelations drop a bit more rapidly toward zero. But the difference arising from more aggressive inflation-fighting is not dramatic, especially given the doubling in the policy response to inflation. Thus, while a dramatic change in the aggressiveness of monetary policy in achieving its inflation goal might well account for some of the reduction in reduced-form inflation persistence, one would not want to overstate this structural source of changes in inflation persistence.

A similarly modest change is induced by reducing the response of inflation to the output gap ($\pi_y = 0.025$), the light dashed line in the top panel of the figure.) Now the autocorrelations decay more gradually toward zero. The first autocorrelation of

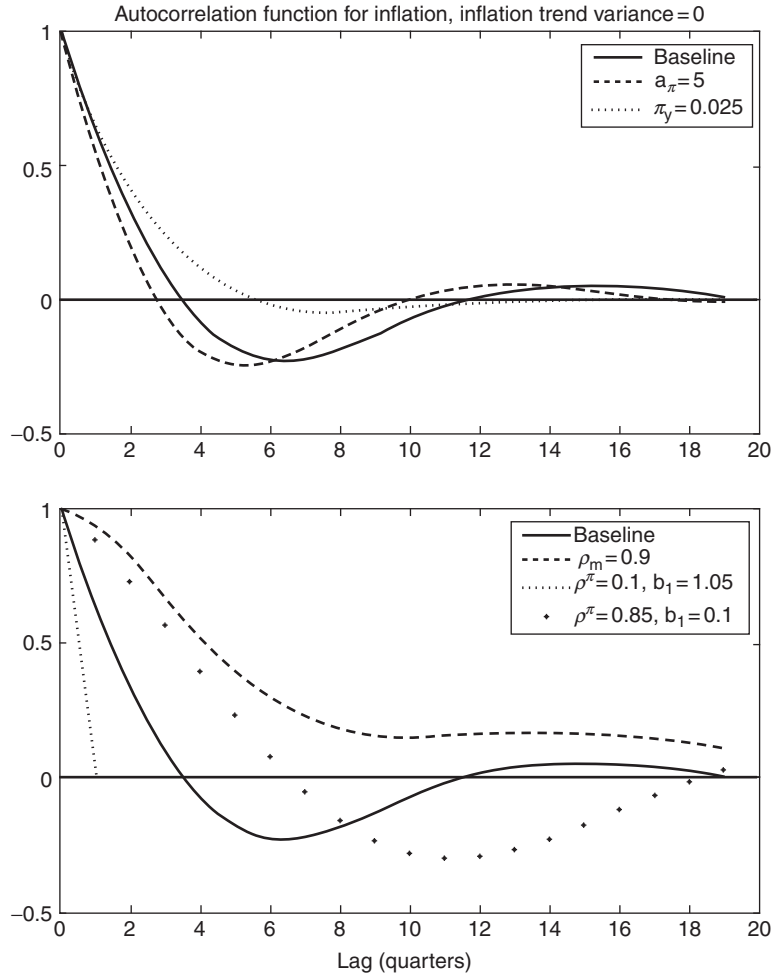


Figure 11 Effect of key structural parameters on inflation persistence, DSGE model.

inflation at 0.75 is modestly above that of the baseline case. This suggests that a smaller Phillips slope, which has been suggested by some recent empirical estimates (see Fuhrer, Olivei, & Tootell, (2009)), may have *increased* the persistence of inflation in recent years. Of course, this is a partial equilibrium effect. To the extent that the central bank recognizes this diminution in the strength of its transmission channel, it can alter its policy so as to offset this effect on persistence.

The bottom panel of Figure 10 displays inflation autocorrelations when other key aspects of the model economy are altered from the baseline. Here the changes are more dramatic. The most striking change occurs when the backward-looking component of the inflation equation is reduced, as shown in the light dashed line. With a primarily

forward-looking inflation equation ($b_1 = 1.05$, $\rho^\pi = 0.1$), the inflation autocorrelations jump quickly to zero (reminiscent of the disinflation simulations in Section 1.2), and are pinned at zero for all periods after the first. Similarly, increasing the weight on lagged inflation in the Euler equation yields a noticeable increase in inflation persistence, as shown in the dotted line. As emphasized in Section 3.5, the values of these key parameters are of considerable importance, and remain under considerable debate.

Of equal importance in determining the persistence of inflation is the persistence attributed to the markup shock, calibrated by the parameter ρ_m . As the heavy dashed line in the bottom panel suggests, a very persistent markup shock adds significantly to the autocorrelation of inflation. This result echoes early results on lagged dependent variables and serially correlated shocks in simple regression models. Consider the simple regression model with serially correlated shock

$$\begin{aligned} y_t &= \beta x_t + u_t \\ u_t &= \rho u_{t-1} + \varepsilon_t \\ &= \frac{\varepsilon_t}{1 - \rho L} \end{aligned}$$

where L is the lag operator. A serially correlated error implies that the regression equation can be rewritten in quasi-differenced form: Pre-multiply both sides of the equation above by $1 - \rho L$ to obtain

$$y_t = \rho y_{t-1} + \beta(x_t - \rho x_{t-1}) + \varepsilon_t.$$

In the context of the inflation Euler equation, this equation suggests that the presence of a highly correlated markup shock may be difficult to distinguish from the presence of significant indexation. Both can be interpreted as suggesting an important role for lagged inflation.⁵⁶

Figure 11 repeats these computations for the case in which the variance of trend inflation (σ_ζ^2) is set to zero – the inflation goal is constant across time. The simple conclusion from this exercise is that the results mirror those for the case in which the trend inflation variance is not zero. The effects of more aggressive monetary policy and a smaller Phillips curve slope are less pronounced. As in the baseline exercise, the largest effects on persistence arise from changes in the forward- and backward-looking components of the Euler equation, and to the degree of autocorrelation in the markup shock. Thus, for this calibration, particularly for the variance estimated for Cogley and Sbordone's trend inflation variable, the absence or presence of trend inflation is not of first-order importance in determining the persistence of inflation. This result contrasts sharply with those in Cogley and Sbordone (2008) and Cogley et al. (2010).

⁵⁶ Note that in the case of a serially correlated markup shock, a common factor restriction across the quasi-differenced y_t and x_t variables may help in identification. See Rudebusch (2006) for an application of the distinction between serially correlated errors and the lagged dependent variable in the Taylor (1993) rule.

Overall, these simulations suggest that while all of the aspects of the economy captured in the DSGE model contribute to the persistence of inflation, the most potent effects arise from two sources: the relative importance of indexation and expectations in the inflation Euler equation, and the autocorrelation of markup shocks. More inflation-responsive monetary policy has some effect, as does the slope of the Phillips curve (π_y), but these effects are noticeably smaller than the effect of eliminating the intrinsic persistence in the Phillips curve.⁵⁷ While these conclusions may vary somewhat depending on details of the DSGE specification, the results suggest that it may be inaccurate to attribute a great deal of the persistence — or a great deal of the change in persistence — to the behavior of monetary policy, including time variation in the inflation target. From this analysis, one would conclude that key parameters of the aggregate supply equation have the dominant effects on inflation persistence.

3.7 Persistence in state-dependent models of inflation

Most of the models in this chapter employ a time-dependent pricing convention; that is, the probability that a firm will adjust its price is solely a function of time, not of economic conditions. An important and appealing alternative is that the timing of price adjustment is endogenously chosen by firms in response to economic conditions. In general, the literature on state-dependent pricing (SDP) has focused relatively little attention on the issue of inflation persistence. While the early models of Caplin and Leahy (1991) and Dotsey, King, and Wolman (1999) focused on the search for “persistence mechanisms,” that focus centered largely on the difficulty in developing a non-zero and persistent output response to a one-time monetary shock. The intuition behind this difficulty is straightforward. In a classic (S,s) model of SDP motivated by a fixed cost to changing prices, modest to large monetary shocks could push all firms to their (S,s) boundary, and consequently all prices would adjust one-for-one with money: Money is neutral. For smaller monetary shocks, a smaller fraction of firms would adjust prices, so money could have a small aggregate effect. Thus, it can be difficult to avoid monetary neutrality in these models; persistent effects of monetary shocks on prices or inflation are even harder to come by.

Burstein (2006) suggested a variant of the SDP paradigm in which firms choose a price *path*, rather than a fixed price level, when they hit their (S,s) boundary. Equivalently, the firm faces a fixed cost to adjusting its price path rather than the level of prices.⁵⁸ Thus, a monetary shock that forces a firm to its boundary will lead the firm to set a sequence of price changes that in turn implies a persistent response of inflation and output to a monetary shock. Burstein’s model is capable of producing inflation

⁵⁷ See Rudebusch (2005) for a related set of results.

⁵⁸ Calvo, Celasun, and Kumhoff (2002) developed a related, time-dependent pricing model in which firms also choose price paths when they are allowed to reset prices.

responses to changes in the money growth rate that exhibit the hump shape typically found in the VAR literature.⁵⁹

A recent paper by Bakhshi, Khan, and Rudolf (2007) developed a Phillips curve from the Dotsey et al. SDP framework, and examined the dynamics of inflation. Interestingly, that Phillips curve includes lagged inflation terms, reflecting the fact that optimal relative prices set in period t depend on optimal relative prices set in previous price vintages (see their equations 9–11). They find that the persistence of inflation implied by the SDP-based Phillips curve is significantly lower than that implied by the time-dependent New-Keynesian Phillips curve, largely because the persistence induced by lagged inflation in the SDP Phillips curve is more than offset by the number of price-setters who reset following a shock.

While recent theoretical developments in SDP pricing models such as those in Burstein (2006) and Bakhshi et al. are promising, the empirical literature based on state-dependent models is considerably less well developed. As a consequence, there are few empirical results on inflation persistence that map well into the reduced-form and structural persistence concepts developed in this chapter. Overall, it seems fair to conclude that the theoretical results suggest that SDP models are likely to provide a less compelling structural explanation of reduced-form persistence than hybrid versions of the time-dependent models.

3.8 Persistence in sticky-information models

Mankiw and Reis (2002) proposed a model in which information, rather than prices, is sticky. In essence, the model applies the Calvo machinery to the updating of information rather than prices. In the Mankiw and Reis (2002) model, a fraction of price-setters get to update their information in each period in a manner analogous to that of the Calvo model. As a consequence, the age or vintage of price-setters' information sets will be described by a geometric distribution, as is the case for the duration of price contracts in the Calvo setting.⁶⁰

The model implies a Phillips curve that links inflation to output and a geometric weighted average of lagged expectations of inflation and output (see Mankiw & Reis, 2002, p. 1300):

⁵⁹ A fuller empirical examination of Burstein's specification would be of interest. A key question in this regard is how the price paths set by resetting firms are allowed to respond to different shocks — markup shocks, shocks to the central bank's inflation target, and so on. In the model of Calvo et al. (2002), it can be difficult to obtain data-consistent impulse responses to all shocks without making different assumptions about when price paths can and cannot respond to shocks. For more on this issue, see Fuhrer (2008).

⁶⁰ Chapter 5 in this *Handbook* outlines a number of issues that arise in a model in which price-setters have imperfect information.

$$\pi_t = \left[\frac{\alpha\lambda}{1-\lambda} \right] \gamma_t + \lambda \sum_{j=0}^{\infty} (1-\lambda)^j E_{t-1-j}(\pi_t + \alpha\Delta\gamma_t). \quad (27)$$

As the authors emphasize, in this Phillips curve it is past expectations of current conditions, rather than current expectations of future conditions, that determine inflation. The Mankiw-Reis Phillips curve is thus a close cousin to those of Fischer (1977) and Koenig (1996). The Mankiw-Reis paper documents a number of desirable features of their model, based on simulations in response to a permanent change in the level of demand, the growth rate of demand (a stylized “disinflation”), and an anticipated drop in the growth rate of demand.

In this model, one can see by inspection (and the authors verify) that inflation will inherit the persistence of the output process. The authors compute the autocorrelations of inflation under the assumption that a simple quantity equation holds, trivially linking money to output, and that money growth follows an autoregressive process

$$\begin{aligned} m_t &= p_t + \gamma_t \\ \Delta m_t &= 0.5\Delta m_{t-1} + \varepsilon_t. \end{aligned} \quad (28)$$

Under these assumptions, the autocorrelations for inflation are indeed quite high. Figure 12 displays the autocorrelations taken from Table I of the paper by Mankiw and Reis (the solid line). The dashed line in the figure displays the autocorrelations when one allows for supply shocks (shocks to the Phillips curve). As discussed in Section 3.4.3, the autocorrelation in the presence of supply shocks depends on the variance of supply shocks relative to the shocks to the driving process. Here, we have set the ratio of supply shocks to money growth shocks at 0.25. As the figure indicates, supply shocks with relatively modest variance dramatically alter the implications for inflation’s autocorrelations in this model, as in the Calvo/Rotemberg model.

3.9 Persistence in learning models

When the agents in the models discussed earlier know the structure and parameters of the model, they can use that knowledge to form expectations by taking the mathematical expectation of the variable of interest. That is the essence of rational expectations, and up to this point the structural models in this chapter have assumed rational expectations — although the preceding section discusses the ramifications of imperfect information, albeit in a rational expectations setting. Nonetheless, one might view the literature on imperfect information as providing a motivation for considering models in which agents learn about their economic environment.⁶¹

⁶¹ Chapters 4 and 5 in this *Handbook* consider a variety of issues that arise in models in which agents must learn about aspects of their economic environment.

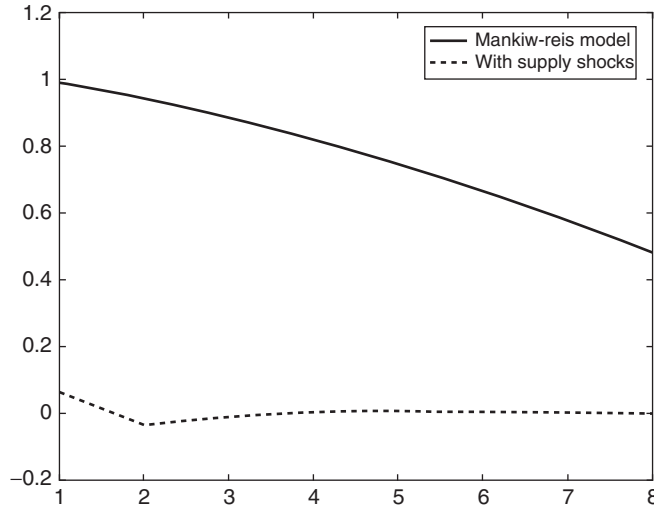


Figure 12 Autocorrelations of inflation, Mankiw-Reis model.

A long tradition dating back at least to Bray (1982) and Marcet and Sargent (1989) examines the dynamics of macroeconomic variables when the agents in the model lack perfect knowledge of the model and consequently must learn about their environment. Orphanides and Williams (2004), Williams (2006), Adam (2005), and Slobodyan and Wouters (2007) explored monetary economies in which agents must learn about their economic environment. Learning can significantly alter the dynamics of an otherwise standard macroeconomic model. Consider the simple case in which inflation is governed by a two-equation model comprising a Calvo-like Phillips curve and a reduced-form equation for output, as in Eq. (19) in Section 3.4.1:

$$\pi_t = F_{t-1}\pi_{t+1} + \gamma x_t. \quad (29)$$

$$x_t = \rho x_{t-1} + \mu_t. \quad (30)$$

The key difference is that expectations of inflation are not the mathematical expectation, denoted by the E_t operator, but are reasonable forecasts given the information available to the agents at time $t - 1$, denoted by F_{t-1} . One plausible way to formalize learning, posited in Williams (2006), is for agents to estimate the reduced-form of the model using an adaptive estimation rule. The unique and stable reduced-form solution of this model under rational expectations is given by

$$\begin{bmatrix} \pi_t \\ x_t \end{bmatrix} = \begin{bmatrix} 0 & \rho \frac{\gamma}{1 - \rho} \\ 0 & \rho \end{bmatrix} \begin{bmatrix} \pi_{t-1} \\ x_{t-1} \end{bmatrix}. \quad (31)$$

As discussed earlier, the model implies no dependence on lagged inflation, although inflation will indeed exhibit persistence to the extent that x_t does. But if agents are not endowed with knowledge of the solution coefficients under rational expectations, they may instead attempt to estimate a reduced-form equation for inflation and x_t such as

$$\begin{bmatrix} \pi_t \\ x_t \end{bmatrix} = \hat{A}_t \begin{bmatrix} \pi_{t-1} \\ x_{t-1} \end{bmatrix} + e_t. \quad (32)$$

If their initial estimate for $\hat{A}_t$ coincides with the solution in Eq. (31), then the model will behave as under rational expectations. But in general, the model will exhibit different dynamics, governed in part by agents' current estimate $\hat{A}_t$ in Eq. (32). If e_t is unforecastable from period t , then Eq. (32) implies that

$$F_{t-1}\pi_{t+1} \equiv e_\pi \hat{A}^2 \begin{bmatrix} \pi_{t-1} \\ x_{t-1} \end{bmatrix}, \quad (33)$$

where e_π selects the row of $\hat{A}$ corresponding to the inflation equation in the reduced form. Substituting this expectation for inflation in Eq. (29), we obtain⁶²

$$\pi_t = \hat{a}_{11}\pi_{t-1} + \hat{a}_{12}x_{t-1} + \gamma x_t. \quad (34)$$

Depending on the current estimated values of $\hat{a}_{ij}$, inflation may now exhibit some intrinsic persistence, and will inherit — more or less, depending on the value of $\hat{a}_{12}$ — the persistence of x_t .⁶³

This very stylized example makes it clear that learning can add another layer of dynamics to inflation. If agents use a forecasting rule that differs from the rational expectations solution, they can add intrinsic persistence to an otherwise forward-looking model that implies no such persistence.

4. INFERENCE ABOUT PERSISTENCE IN SMALL SAMPLES: “ANCHORED EXPECTATIONS” AND THEIR IMPLICATIONS FOR INFLATION PERSISTENCE

Implicit in the analysis of Benati (2008) and explicit in a paper by Williams (2006) is the suggestion that inflation expectations that are “well anchored” by the central bank's explicit commitment to an inflation target may have altered the persistence and, more generally, the overall dynamics of inflation. Theoretically, this can be true to a degree. Consider once again the simple model of Eqs. 12, modified slightly to make the central bank's inflation target explicit:

⁶² The time subscripts on the elements of $\hat{A}_t$ are dropped for notational convenience.

⁶³ Additional dynamics would, in principle, be added as the agents' estimates of the $\hat{a}_{ij}$ evolve over time. The specifics vary depending on the estimation rule assumed, and a detailed investigation of this issue lies outside the scope of this chapter.

$$\begin{aligned}
\pi_t &= \pi_{t-1} + ax_t \\
x_t &= -b(f_t - \bar{\pi}) \\
f_t &= \bar{\pi} + c(\pi_t - \bar{\pi}).
\end{aligned} \tag{35}$$

As long as $c \neq 0$, the steady state for this model is given by⁶⁴

$$\begin{aligned}
\pi_t &= \bar{\pi} \\
x_t &= 0 \\
f_t &= \bar{\pi}.
\end{aligned} \tag{36}$$

But when $c = 0$, the central bank does nothing to move inflation toward a particular target, the steady state for inflation becomes indeterminate, and the solution for inflation becomes

$$\pi_t = \pi_{t-1}. \tag{37}$$

In contrast to the stable AR process in Eq. 13, the inflation rate in this case follows a random walk. This simple model demonstrates the importance in this class of models of the central bank systematically pursuing an inflation target. When central banks do so, inflation is stationary. When they do not, it is not. In most models with explicit expectations, a similar proposition holds, as is well known. The intuition for a model with rational expectations is relatively straightforward. Consider a simple model based on the Calvo (1983) specification of inflation

$$\begin{aligned}
\pi_t &= \beta E_t \pi_{t+1} + \gamma x_t \\
x_t &= -a(r_t - E_t \pi_{t+1}) \\
r_t &= \bar{\pi} + b(\pi_t - \bar{\pi}).
\end{aligned} \tag{38}$$

As demonstrated in Eq. 18, the solution for π_t is the weighted sum of future x_t , which in turn depends on the future short-term real interest rates, $r_t - E_t \pi_{t+1}$. The central bank sets the short-term nominal interest rate r_t . Well-anchored inflation expectations require two things from the central bank. First, the central bank must have an inflation goal that is known to the private agents in the economy, and second, the central bank must move its policy rate in a way that systematically pushes the inflation rate toward that goal. Put differently, the “Taylor principle” (Taylor, 1999) operates in this model: As long as the central bank moves the policy rate by more than the inflation rate, so that it is increasing the short-term real rate when inflation is above its target and conversely when it is below, then the expected path of x_t will be consistent with returning inflation to its target under arbitrary initial conditions. In this sense, inflation expectations will be well anchored, and inflation will have a determinate solution and be stationary. As is discussed in the Introduction and in more detail later, in a purely

⁶⁴ The equilibrium real rate of interest is set to zero for convenience.

forward-looking model such as this one, well-anchored expectations not only imply determinacy and stationarity, but an inflation rate that follows a white noise process.⁶⁵

Williams (2006) suggested that in recent years expectations may have become so well anchored that inflation may be well characterized by random deviations around a constant

$$\pi_t = c + \varepsilon_t.$$

As a description of the data, this simple model does reasonably well in recent years, as Figure 13 suggests. It shows the errors made by a model that forecasts the four-quarter log change in prices as equal to its sample mean, versus one that sets the forecast equal to the previous four-quarter change.⁶⁶ A version of the random walk model has been advocated by Atkeson and Ohanian (2001) as an alternative to poorly performing Phillips curves. The errors over Williams's sample are smaller for the constant-based forecast, with a root-mean-squared error of 0.29 versus 0.37 for the random walk model.⁶⁷

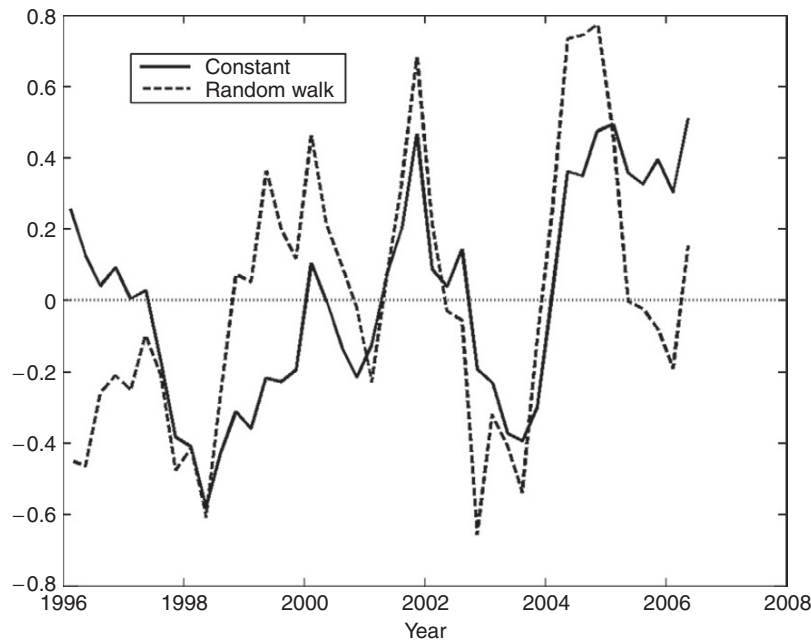


Figure 13 Errors for Williams (2006) and random walk inflation models in recent years.

⁶⁵ This implication also relies on the assumption that all the shocks disturbing the equations above are *iid*.

⁶⁶ Of course, the sample mean forecast uses information not available in real time to the forecaster.

⁶⁷ Adding the last three years to the sample diminishes the advantage of the constant-based model. The RMSEs for the constant-based and random-walk models for the extended sample are 0.32 and 0.35, respectively. The constant-based model, with an estimated mean of about 1.8, consistently underforecasts inflation over the last three years of the sample.

The diversity of results on reduced-form inflation persistence presented in the preceding subsections together should suggest some caution in arriving at conclusions about a possible change in inflation persistence in the past decade or two. We examine a simple Monte Carlo exercise to highlight the difficulty in inferring changes in inflation dynamics and their implications for persistence in relatively short samples.

Inflation in the exercise is generated by three simple equations: a backward-looking Phillips curve with lagged coefficients that sum to 1, an output equation that makes output a function of its own lag and the real interest rate, and a simple policy rule in which the policy rate responds only to the gap between inflation and the inflation target.

$$\begin{aligned}\pi_t &= \sum_{i=1}^4 a_i \pi_{t-i} + b \tilde{y}_t + \varepsilon_t \\ \tilde{y}_t &= c \tilde{y}_{t-1} - d(r_{t-1} - \pi_t) + u_t \\ r_t &= e r_{t-1} + (1 - e)(f(\pi_t - \bar{\pi})).\end{aligned}\tag{39}$$

The variances of the two shocks are calibrated from VAR equations estimated from 1966 to 2006, or split at 1984 to allow for changes in the estimated variance due to the so-called Great Moderation.⁶⁸ The baseline values for b and d are 0.1, for c 0.8, for e 0.8, and for f 1.5. The inflation target is set to 2.0. Fifty thousand samples of length 40 quarters are created using draws of the shocks to the inflation and output equations, assuming the individual shocks are random normal draws scaled by the estimated variances of ε and u . Initial conditions are randomized for each draw, rather than using a “burn-in” period.⁶⁹

For each simulated sample, the following simple regression models are estimated. The first is designed as a simple test that inflation exhibits no persistence, perhaps because inflation expectations are well anchored at the central bank’s target, and can thus be well modeled as white noise around a constant c . The null for this model is that $\alpha = 0$, which would replicate the result of Williams (2006). The second allows for the influence of output. Both are clearly misspecified, but the issue is whether with a relatively short sample one could be misled into believing that the properties of inflation had changed when in fact the underlying inflation process still exhibits persistence and a correlation with output.

$$\pi_t = \alpha \pi_{t-1} + c\tag{40}$$

⁶⁸ The equations estimate a bivariate VAR in HP-filtered real GDP and the annualized rate of inflation in the core consumer price index. The three samples are considered at 1966:Q1–2006:Q4, 1966:Q1–1984:Q4, and 1985:Q1–2006:Q4.

⁶⁹ Tests for convergence of the distribution of estimated parameters suggest that 50,000 samples are more than adequate to assure convergence for this exercise.

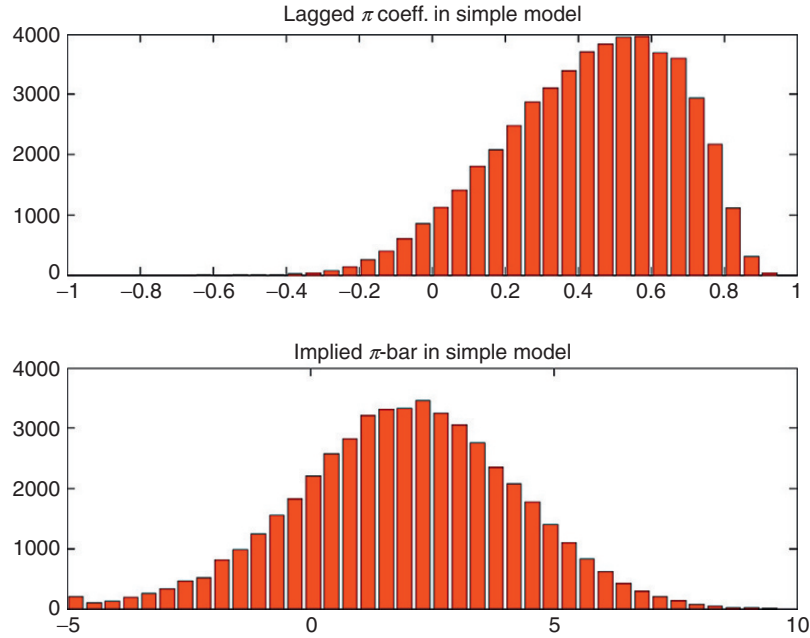


Figure 14 Distribution of estimated coefficients.

$$\pi_t = \alpha\pi_{t-1} + by_{t-1} + c \quad (41)$$

The resulting regression estimates are displayed in the summary table below and the histograms in the panels of Figure 14. For the first model, Figure 14 and Table 13 show that the standard error on the estimated lag coefficient is large; the median t -statistic for the hypothesis that $\alpha = 0$ has a p -value above 0.05. The estimate of c implies a mean for π_t of almost exactly 2.⁷⁰ One could be forgiven for inferring from these estimates that inflation had little or no persistence, and was well anchored around the central bank's inflation target.⁷¹

The estimates for the second regression model, displayed in the bottom panel of Table 13, suggest that matters are even worse when one tests for the presence of a Phillips-like relationship in the misspecified model for a small sample. Now, the median estimate of α is smaller still, it is imprecisely estimated (the median p -value for the test of $\alpha = 0$ is greater than 0.10), and the effect of output is biased downward and very

⁷⁰ The standard error of the intercept, at 0.16, is quite small. The distribution of the mean of inflation, which is a nonlinear function of α and c ($\bar{\pi} = \frac{c}{1-\alpha}$), implies a relatively large standard error for the mean of 2.3.

⁷¹ A good portion of the downward bias evident in the estimates in Figure 14 arises from the truncation of the lags of inflation compared with the lags in the data-generating process. Still, in large samples, the regression should retrieve the sum of the underlying AR coefficients in the single AR coefficient. As the sample size in the previous exercise increases from 10 years to 25 to 50, the median estimate of α in Model 1 increases from 0.47 to 0.70 to 0.84.

Table 13 Distribution of estimated coefficients
Model 1

Coefficient	Median	Std. deviation
α	0.47	0.24
c	1.06	0.16
Model 2		
Coefficient	Median	Std. deviation
α	0.41	0.25
b	0.069	0.33
c	1.2	0.17

imprecisely estimated. Once again, the mean of the estimated inflation target matches the true value quite precisely. One might conclude from simple regressions such as these that inflation has little or no persistence and that the Phillips correlation is absent. These small samples allow one to recover the average value of inflation over the period, but the rest of inflation dynamics may be very poorly estimated.

5. MICROECONOMIC EVIDENCE ON PERSISTENCE

5.1 Persistence in micro data: U.S. evidence

Up to this point, we have focused on macroeconomic, aggregate evidence bearing on inflation persistence. Yet the dominant models in the literature aim to provide microeconomic foundations for inflation, based on the price-setting decisions of individual firms. In this regard, it is striking that the now large literature that examines micro-price data has emerged relatively recently, led by the work of Bils and Klenow (2004). Bils and Klenow employed an unpublished data set of about 350 CPI expenditure categories that account for about 70% of consumer expenditures. In examining the persistence and volatility of inflation, they used a subset of 123 spending categories that account for about 63% of spending.

For each of the CPI components that they examined, p_i , they estimate a simple AR(1) process to assess the degree of persistence and volatility for its inflation rate, dp_i (where dp_i is the log change in price series p_i)

$$dp_{i,t} = \rho dp_{i,t-1} + \varepsilon_{i,t}. \quad (42)$$

They found that the average persistence of inflation, which they defined as the arithmetic average of the ρ_i 's, is slightly negative (-0.05) for their shorter sample (1995–2000) and only modestly positive (0.26) for their longer sample (1959–2000).

Interestingly, for the shorter sample, they found that the degree of persistence across price categories is *positively* correlated with the frequency of price change, clearly contrary to predictions from the Calvo/Rotemberg and Taylor models. Over the longer sample period, the correlation is positive but not statistically different from zero.

The authors estimate that adjusting for the presence of “temporary sales” — explicitly transitory reductions in prices that revert within days or weeks and would likely lead to an understatement of the persistence of nonsale prices — would have only a small impact on their estimates of persistence. Of course, the fact that temporary sales occur regularly casts some doubt on the underlying assumption of time-dependent pricing models, in which prices are simply fixed for long periods. It also raises the question as to whether the relevant price measures should exclude or include temporary sales. This issue is discussed in more detail in Nakamura and Steinsson (2008), who employed a data set that allowed them to study prices at a more disaggregated level. They found that temporary sales have a larger effect on the measured frequency of price changes than in the Bils and Klenow data; excluding temporary sales reduces the frequency of price changes by about one-half. They did not estimate the effect of temporary sales on inflation persistence in their paper.

The results from these seminal papers suggest that the micro data exhibit behavior that is at odds with the prevailing time-dependent pricing models’ description of underlying price behavior. Nakamura and Steinsson’s (2008) paper suggested that several key features are also at odds with a menu-cost model. But whether the micro data are consistent with the estimates of persistence in aggregate data is less clear, for two reasons. First, aggregate price series may exhibit quite different properties from the individual price series. Second, what one observes in individual price changes likely reflects the combined influence of firm- or industry-specific shocks and macro shocks. The two points are related, and the relative importance of the two is an empirical question, but if individual prices respond differently to macro versus micro shocks, then it will be important to sort out these influences in evaluating the relevance of micro-data evidence for aggregate inflation persistence.

Section 5.3 discusses some recent work bearing on the first point. Boivin, Giannoni, and Mihov (2009) provided results bearing on the second point. They estimated a small number of common factors (principal components) from a large number of macroeconomic variables. They then related individual price changes to these common factors to decompose individual inflation rates into idiosyncratic, sector-specific fluctuations and macroeconomic fluctuations. Denoting the matrix of common factors by C_t and the log change in the individual price series p_{it} as dp_{it} , they estimated regressions

$$dp_{it} = \lambda_i C_t + \varepsilon_{it}. \quad (43)$$

The R^2 of each regression indicates the fraction of individual price changes that may be attributed to the common macroeconomic factors; one minus this R^2 is the fraction of variation attributable to sector-specific sources. Their baseline results suggest that about 85% of the variation in the individual price changes may be attributed to sector-specific shocks.

Using the decomposed individual inflation series, Boivin et al. (2009) estimated simple ARs for each of the series and their two components, measuring persistence by the sum of the AR coefficients.⁷² They find, like Bils and Klenow (2004), that individual inflation series exhibit relatively little persistence. The idiosyncratic components of the individual inflation series, ε_{it} , exhibit essentially no persistence, while the common components ($\lambda_i C_{it}$) vary in persistence from negative for some series to above 0.95 for some health-care components and tenant room and board. These findings imply that the aggregate inflation measures, which are quite persistent in their data, inherit their persistence from the common macroeconomic components of the individual price series, particularly from those that exhibit the highest persistence. The non-persistent idiosyncratic components essentially wash out in aggregation.

5.2 Persistence in micro data: Euro Area evidence

Altissimo, Ehrmann, and Smets (2006) examined both aggregate and disaggregated data to explore the properties of Euro Area inflation. Their conclusions on aggregate data echoed those of others — inflation has been persistent; its reduced-form persistence has declined in recent years so that it now exhibits moderate persistence, although how moderate depends on how it is estimated; and its decline may be attributable to a stable and well-focused monetary regime that anchors long-run inflation expectations. Their disaggregated sectoral data suggested that individual price series exhibit less persistence on average than their corresponding aggregates.

The studies of Angeloni et al. (2006) and Alvarez et al. (2006) found ample evidence of infrequent price changes at the micro level. The former estimates that while there is substantial heterogeneity across sectors, prices on average are quite sticky, exhibiting a four to six quarter duration. Seen through the lens of the Calvo model, these estimates suggest that the “Calvo parameter,” which indexes the frequency of price changes, is large, implying a small effect of marginal cost on inflation. From Eq. (21), the smaller is the effect of marginal cost on inflation, the less of the persistence in marginal cost is inherited by inflation. Neither of these studies examines the persistence of disaggregated price series, nor do the studies explore the complications in aggregating disaggregated series with heterogeneous dynamics. But to the extent that they find relatively infrequent price changes, and to the extent that one feels comfortable mapping these into aggregate Calvo parameters, the studies imply less inherited persistence, other things equal.

5.3 More on aggregation and persistence

Bils and Klenow (2004) found quite a lot of difference between the persistence properties of inflation based on individual price series and their expenditure-share-weighted aggregate inflation measure. For their longer sample, they estimated an AR coefficient

⁷² The data of Boivin et al. (2009) are observed at the monthly frequency, and they estimate ARs with 13 lags.

of 0.63, with a standard error of 0.03, in contrast to the much smaller average of the AR coefficients for the individual price series. Boivin et al. (2009) found that the low-persistence, idiosyncratic components of disaggregated inflation series appear to wash out in aggregate inflation measures, leaving the persistent common macroeconomic components to dominate the persistence of aggregate inflation. These observations suggest that aggregation of price series may play an important role in determining the degree of aggregate inflation persistence.⁷³

Several recent papers examine the role of aggregation in inflation persistence. Mumtaz, Zabczyk, and Ellis (2009) employed a methodology that draws importantly on Boivin et al. (2009) studying disaggregated price data for the UK. They also found that the persistence of the aggregate inflation measure is biased upwards relative to the persistence of the underlying price series and that the bias is driven by the macro components of those series. Altissimo, Mojon, and Zaffaroni (2009) delved more deeply into the aggregation process, using existing results on aggregation of time series (see, e.g., Granger, 1980 for parametric results and Zaffaroni, 2004 for nonparametric results that are more closely related to the work in Altissimo et al., 2009). Similarly to Boivin et al. (2009), they assumed that individual price series may be characterized by an unobserved components model that makes each price change series, $dp_{i,t}$, a function of its own idiosyncratic persistence, a common shock, u_t , and an idiosyncratic shock, $\varepsilon_{i,t}$:

$$dp_{i,t} = \alpha_i dp_{i,t-1} + u_t + \varepsilon_{i,t}. \quad (44)$$

By assuming a particular form for the distribution of the persistence parameters $f(\alpha)$, as in Granger (1980), one can derive results for the persistence of the simple aggregate of the individual price changes, $dP_t = (1/n) \sum_{i=1}^n dp_{i,t}$. The relative contributions of the common and idiosyncratic shocks will be important in understanding the relationship between individual and aggregate prices, but so too will be the differences in the way in which the common shock is propagated in the individual prices; that is, the differences in the α_i 's. Price series that perpetuate the common shock through a larger α_i will have a larger effect on the persistence of the aggregate than series with a smaller α_i .

Altissimo, Mojon, and Zaffaroni (2009) find that disaggregated price changes exhibit significantly less persistence than the aggregate and that the preponderance of the variance of the individual price series is accounted for by idiosyncratic volatility, in agreement with all of the studies cited above. Like Boivin et al., (2009), they found that a single principal component of the individual price series dominates in explaining the low-frequency variation in the individual price changes. It follows that this factor must account for the common persistence among the disaggregated series. Estimation

⁷³ On a note of theoretical counterpoint, Carvalho (2006) developed a multisector version of the Calvo model with heterogeneity in price stickiness that implies an aggregate inflation rate with *less* persistence than that of the standard Calvo model.

of a model like eq. (44) reveals in addition that the propagation of the common shock in individual series is indeed quite varied. The high persistence of the common shock in services prices, in particular, combined with the relatively high weight of services in the aggregate price index, accounts for a substantial part of the persistence in aggregate inflation.

6. CONCLUSIONS

It may be early to draw firm conclusions about the structural sources of inflation persistence, or about the extent to which these sources have changed and manifested themselves in changes in reduced-form inflation persistence. In the first case, it may be premature because there is not yet widespread agreement about the appropriate mapping between micro data or reduced-form aggregate data and our structural models. In the second case, we have a fairly short sample from which to draw inferences about potential changes (see Section 4 for more details).

Still, the research to date allows one to draw some conclusions. First, to the extent that reduced-form persistence has changed, policymakers need to gain clarity about the source of the change. This chapter discussed a number of structural channels through which persistence may have changed. It suggested that one may decompose the structural sources of persistence embodied in inflation into two components: “Intrinsic” persistence — in essence, the influence of lagged inflation in structural Phillips curves — and inherited persistence, which is persistence inherited from the driving process for inflation. In conventional theories of inflation, inherited persistence can change because the persistence of the driving process has changed, or because the coefficient on the driving process has changed, or because the relative variances of the shocks to inflation and the driving process have changed.

The analysis presented in this chapter suggested that it is unlikely that any change in persistence has arisen from a change in the persistence of the driving process, as this has remained remarkably stable throughout the period. In addition, a DSGE model-based analysis suggests that while changes in the systematic component of monetary policy likely have led to less-persistent inflation, the largest changes in persistence are most likely due to changes in the so-called intrinsic sources of inflation persistence — whether those arise from indexation, rule-of-thumb price-setters, or a rising price reset hazard. Finally, the models that depart from the standard Calvo framework suggest that other aspects of the economy that impinge upon inflation persistence may be responsible for changes in inflation persistence. These may include smaller or less-frequent changes in “trend inflation” or a smaller role for learning, as central bank transparency about its goals has increased.

Second, we have now accumulated an impressive and growing body of evidence on the behavior of price- (and wage-) setting at the disaggregated level. This evidence

strongly suggests that some of the inferences drawn from micro data about the frequency of price changes, as well as the degree of inflation persistence, may pertain largely to price responses to industry- or firm-specific shocks. The response to aggregate shocks by the aggregate component common to the individual price series may well have quite different properties from the responses of individual firms to idiosyncratic shocks. Integrating this evidence into our structural models, perhaps along the lines of models of “rational inattention” (see Gorodnichenko, 2008; Maćkowiak & Wiederholt, 2009; Sims, 2003) seems a promising avenue for research.

Finally, we are currently accumulating additional evidence that should allow us to take a firmer stance on whether reduced-form persistence has changed, and to discern the structural sources of any such changes. The upheaval created by the 2007–2009 financial crisis and recession, with the concomitant prospect of a prolonged period of elevated unemployment and depressed marginal cost, suggests that over the next decade we will have accumulated evidence that will allow us to test more fully the hypothesis of a decline in reduced-form inflation persistence and to test competing theories that attribute the structural sources of persistence.

REFERENCES

- Adam, K., 2005. Learning to forecast and cyclical behavior of output and inflation. *Macroecon. Dyn.* 9, 1–27.
- Altissimo, F., Ehrmann, M., Smets, F., 2006. Inflation persistence and price-setting behaviour in the Euro Area: A summary of the IPN evidence. National Bank of Belgium working papers No. 95 October.
- Altissimo, F., Mojon, B., Zaffaroni, P., 2009. Can aggregation explain the persistence of inflation? *J. Monetary Econ.* 56, 231–241.
- Alvarez, L., Dhyne, E., Hoeberichts, M., Kwapil, C., Le Bihan, H., Lunnemann, P., et al., 2006. Sticky prices in the euro area: a summary of new micro-evidence. *J. Eur. Econ. Assoc.* 4, 575–584.
- Andrews, D., 1993a. Exactly median-unbiased estimation of first-order autoregressive/unit-root models. *Econometrica* 61, 139–165.
- Andrews, D., 1993b. Testing for structural instability and structural change with unknown change point. *Econometrica* 61, 821–856.
- Angeloni, I., Aucremanne, L., Ehrmann, M., Galí, J., Levin, A., Smets, F., 2006. New evidence on inflation persistence and price stickiness in the Euro area: Implications for macro modeling. *J. Eur. Econ. Assoc.* 4, 562–574.
- Atkeson, A., Ohanian, L., 2001. Are Phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis Quarterly Review* 25 (1), 2–11 (Winter).
- Bai, J., Perron, P., 1998. Estimating and testing for multiple structural changes in linear models. *Econometrica* 66, 47–78.
- Bakhshi, H., Khan, H., Rudolf, B., 2007. The Phillips curve under state-dependent pricing. *J. Monetary Econ.* 54, 2321–2345.
- Ball, L., 1994. Credible disinflation with staggered price-setting. *Am. Econ. Rev.* 84 (1), 282–289.
- Ball, L., Cecchetti, S.G., 1990. Inflation and uncertainty at short and long horizons. *Brookings Pap. Econ. Act.* 21, 215–254.
- Barnes, M., Gumbau-Brisa, F., Lie, D., Olivei, G., 2009. Closed-form estimates of the New Keynesian Phillips curve with time-varying trend inflation. Federal Reserve Bank of Boston, Working paper.
- Barsky, R.B., 1987. The Fisher hypothesis and the forecastability and persistence of inflation. *J. Monetary Econ.* 19, 3–24.

- Benati, L., 2008. Investigating inflation persistence across monetary regimes. *Q. J. Econ.* 123 (3), 1005–1060.
- Benati, L., 2009. Are intrinsic inflation persistence models structural in the sense of Lucas (1976)? March European Central Bank, Working paper series No. 1038.
- Bils, M., Klenow, P., 2004. Some evidence on the importance of sticky prices. *J. Polit. Econ.* 112 (5), 947–985.
- Blanchard, O., Galí, J., 2007. Real wage rigidities and the new Keynesian model. *J. Money Credit Bank.* 39 (1), 35–65, Supplement to.
- Boivin, J.P., Giannoni, M., Mihov, I., 2009. Sticky prices and monetary policy: Evidence from disaggregated U.S. data. *Am. Econ. Rev.* 99 (1), 350–384.
- Bray, M., 1982. Learning, estimation, and the stability of rational expectations. *J. Econ. Theory* 26, 318–339.
- Buiter, W., Jewitt, I., 1981. Staggered wage setting with real wage rigidities: Variations on a theme of Taylor. *The Manchester School* 49, 211–228.
- Burstein, A.T., 2006. Inflation and output dynamics with state-dependent pricing decisions. *J. Monetary Econ.* 53, 1235–1257.
- Calvo, G.A., 1983. Staggered prices in a utility-maximizing framework. *J. Monetary Econ.* 12 (3), 383–398.
- Calvo, G.A., Celasun, O., Kumhoff, M., 2002. A theory of rational inflationary inertia. In: Aghion, P., Frydman, R., Stiglitz, J., Woodford, M. (Eds.), *Knowledge, information and expectations in modern macroeconomics: In honor of Edmund S. Phelps*. Princeton University Press, Princeton, NJ.
- Caplin, A., Leahy, J., 1991. State-dependent pricing and the dynamics of money and output. *Q. J. Econ.* 106 (3), 683–708.
- Carvalho, C., 2006. Heterogeneity in price stickiness and the real effects of monetary shocks. *Frontiers of Macroeconomics* 2 (1) Article 1.
- Christiano, L., Eichenbaum, M., Evans, C., 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. *J. Polit. Econ.* 113 (1), 1–45.
- Cogley, T., Sbordone, A., 2008. Trend inflation, indexation, and inflation persistence in the New Keynesian Philips curve. *Am. Econ. Rev.* 98 (5), 2102–2126.
- Cogley, T., Primiceri, G.E., Sargent, T.J., 2010. Inflation-gap persistence in the U.S. *American Economic Journal: Macroeconomics* 2 (1), 43–69.
- de Walque, G., Smets, F., Wouters, R., 2005. Price setting in general equilibrium: Alternative specifications. *Computing in Economics and Finance* 370.
- Dornbusch, R., 1976. Expectations and exchange rate dynamics. *J. Polit. Econ.* 84 (6), 1161–1176.
- Dotsey, M., King, R., Wolman, A., 1999. State-dependent pricing and the general equilibrium dynamics of money and output. *Q. J. Econ.* 114, 655–690.
- Estrella, A., Fuhrer, J., 2002. Dynamic inconsistencies: Counterfactual implications of a class of rational expectations models. *Am. Econ. Rev.* 92 (4), 1013–1028.
- Fischer, S., 1977. Long-term contracts, rational expectations, and the optimal money supply rule. *J. Polit. Econ.* 85 (1), 191–205.
- Friedman, M., 1968. The role of monetary policy. *Am. Econ. Rev.* 58 (1), 1–17.
- Fuhrer, J., 2000. Habit formation in consumption and its implications for monetary-policy models. *Am. Econ. Rev.* 90 (3), 367–390.
- Fuhrer, J., 2006. Intrinsic and inherited inflation persistence. *International Journal of Central Banking* 2 (3), 49–86.
- Fuhrer, J., 2008. Special issue comment on optimal price setting and inflation inertia in a rational expectations model. *J. Econ. Dyn. Control* 32, 2536–2542.
- Fuhrer, J., Moore, G., 1992. Monetary policy rules and the indicator properties of asset prices. *J. Monet. Econ.* 29, 303–336.
- Fuhrer, J., Moore, G., 1995. Inflation persistence. *Q. J. Econ.* 110 (1), 127–159.
- Fuhrer, J., Olivei, G., Tootell, G.M.B., 2009. Empirical estimates of changing inflation dynamics. Federal Reserve Bank of Boston, Working paper.

- Galí, J., Gertler, M., 1999. Inflation dynamics: A structural econometric analysis. *J. Monet. Econ.* 44, 195–222.
- Gordon, R.J., 1982. Price inertia and policy ineffectiveness in the United States, 1890–1980. *J. Polit. Econ.* 90 (6), 1087–1117.
- Gordon, R.J., King, S.R., Modigliani, F., 1982. The output cost of disinflation in traditional and vector autoregressive models. *Brookings Pap. Econ. Act.* 1982 (1), 205–244.
- Gorodnichenko, Y., 2008. Endogenous information, menu costs and inflation persistence. NBER Working Paper 14183.
- Granger, C., 1980. Long memory relationships and the aggregation of dynamic models. *J. Econom.* 14, 227–238.
- Gray, J., 1977. Wage indexation: A macroeconomic approach. *J. Monet. Econ.* 2 (2), 221–235.
- Ireland, P., 2007. Changes in the Federal Reserve's inflation target: Causes and consequences. *J. Money Credit Bank.* 39 (8), 1851–1882.
- Koenig, E., 1996. Aggregate price adjustment: The Fischerian alternative. Federal Reserve Bank of Dallas, Working Paper No. 9615.
- Levin, A., Piger, J., 2004. Is inflation persistence intrinsic in industrial economies?. European Central Bank, Working paper No. 334.
- Lucas Jr., R.E., 1972. Expectations and the neutrality of money. *J. Econ. Theory* 4, 103–124.
- Mackowiak, B., Wiederholt, M., 2009. Optimal sticky prices under rational inattention. *Am. Econ. Rev.* 99 (3), 769–803.
- Mankiw, N.G., Reis, R., 2002. Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve. *Q. J. Econ.* 117 (4), 1295–1328.
- Marcet, A., Sargent, T.J., 1989. Convergence of least-squares learning in environments with hidden state variables and private information. *J. Polit. Econ.* 97 (6), 1306–1322, December.
- Mash, R., 2004. Optimizing microfoundations for inflation persistence. Oxford University, Oxford, UK Oxford University Department of Economics, Discussion Paper No. 183.
- Mavroeidis, S., 2005. Identification issues in forward-looking models estimated by GMM, with an application to the Phillips curve. *J. Money Credit Bank.* 37 (3), 421–448.
- Mumtaz, H., Zabczyk, P., Ellis, C., 2009. What lies beneath: What can disaggregated data tell us about the behaviour of prices?. Bank of England, Working Paper No. 364.
- Muth, J., 1961. Rational expectations and the theory of price movements. *Econometrica* 29 (3), 315–335.
- Nakamura, E., Steinsson, J., 2008. Five facts about prices: A reevaluation of menu cost models. *Q. J. Econ.* 112, 1415–1464.
- Okun, A.M., 1977. Efficient disinflationary policies. *Am. Econ. Rev.* 68, 348–352; (May 1978, Papers and Proceedings 1977).
- O'Reilly, G., Whelan, K., 2005. Has Euro-area inflation persistence changed over time? *Rev. Econ. Stat.* 87 (4), 709–720.
- Orphanides, A., Williams, J.C., 2004. Imperfect knowledge, inflation expectations, and monetary policy. In: Bernanke, B., Woodford, M. (Eds.), *The inflation targeting debate*, University of Chicago Press.
- Phelps, E.S., 1968. Money-wage dynamics and labor-market equilibrium. *J. Polit. Econ.* 76 (4), 678–711 Part 2.
- Pivetta, F., Reis, R., 2007. The persistence of inflation in the United States. *J. Econ. Dyn. Control* 31, 1326–1358.
- Ravenna, F., 2000. The impact of inflation targeting in Canada: A structural analysis. New York University; Manuscript.
- Roberts, J., 1997. Is inflation sticky? *J. Monet. Econ.* 39 (2), 173–196.
- Roberts, J., 2006. Monetary policy and inflation dynamics. *International Journal of Central Banking* 2, (3), September.
- Rotemberg, J.J., 1982. Sticky prices in the United States. *J. Polit. Econ.* 90 (6), 1187–1211.
- Rotemberg, J.J., 1983. Aggregate consequences of fixed costs of price adjustment. *Am. Econ. Rev.* 73 (3), 433–436.
- Rudd, J., Whelan, K., 2006. Can rational expectations sticky-price models explain inflation dynamics? *Am. Econ. Rev.* 96 (1), 303–320.

- Rudebusch, G., 2005. Assessing the Lucas Critique in monetary policy models. *J. Money Credit Bank.* 37 (2), 245–272.
- Rudebusch, G., 2006. Monetary policy inertia: Fact or fiction? *International Journal of Central Banking* 2 (4), 85–135.
- Sargent, T.J., Wallace, N., 1975. “Rational” expectations, the optimal monetary instrument, and the optimal money supply rule. *J. Polit. Econ.* 83 (2), 241–254.
- Sheedy, K.D., 2007. Intrinsic inflation persistence. Centre for Economic Performance Discussion Paper No. 837.
- Sims, C.A., 2003. Implications of rational inattention. *J. Monet. Econ.* 50, 665–690.
- Slobodyan, S., Wouters, R., 2007. Learning in an estimated medium scale DSGE model. National Bank of Belgium, Working paper.
- Stock, J.H., Watson, M., 2007. Has inflation become harder to forecast? *J. Money Credit Bank.* 39, 3–34.
- Taylor, J., 1980. Aggregate dynamics and staggered contracts. *J. Polit. Econ.* 88 (1), 1–23.
- Taylor, J., 1993. Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy* 39, 195–214.
- Taylor, J., 1999. A historical analysis of monetary policy rules. In: Taylor, J.B. (Ed.), *Monetary policy rules*. University of Chicago Press, Chicago, pp. 319–341.
- Williams, J., 2006. The Phillips curve in an era of well-anchored inflation expectations. Federal Reserve Bank of San Francisco, Unpublished working paper.
- Woodford, M., 2003. *Interest and prices: Foundations of a theory of monetary policy*. Princeton University Press, Princeton, NJ.
- Woodford, M., 2007. Interpreting inflation persistence: Comments on the conference on Quantitative Evidence on Price Determination. *J. Money Credit Bank.* 39 (1), 203–210, Supplement to.
- Zaffaroni, P., 2004. Contemporaneous aggregation of linear dynamic models in large economies. *J. Econom.* 120, 75–102.

Monetary Policy and Unemployment[☆]

Jordi Galí

CREI, Universitat Pompeu Fabra, and Barcelona GSE

Contents

1. Introduction	488
2. Evidence on the Cyclical Behavior of Labor Market Variables and Inflation	491
3. A Model with Nominal Rigidities and Labor Market Frictions	495
3.1 Households	495
3.2 Firms	497
3.2.1 <i>Final goods</i>	497
3.2.2 <i>Intermediate goods</i>	499
3.2.3 <i>A brief detour: Labor market frictions and inflation dynamics</i>	501
3.3 Monetary policy	502
3.4 Labor market frictions and wage determination	503
3.4.1 <i>The case of flexible wages</i>	503
3.4.2 <i>The case of sticky wages</i>	506
3.4.3 <i>Relation to the New Keynesian wage inflation equation</i>	513
3.5 Aggregate demand and output	514
4. Equilibrium Dynamics: The Effects of Monetary Policy and Technology Shocks	515
4.1 Steady state and calibration	515
4.2 The effects of monetary policy and technology shocks	517
4.3 The role of labor market frictions	520
4.4 The role of price stickiness	523
4.5 The role of wage stickiness	526
5. Labor Market Frictions, Nominal Rigidities and Monetary Policy Design	528
5.1 The social planner's problem	528
5.1.1 <i>The efficient steady state</i>	529
5.2 Optimal monetary policy	529
5.2.1 <i>The case of flexible wages</i>	530
5.2.2 <i>The case of sticky wages</i>	530
6. Possible Extensions	535
6.1 Real wage rigidities and wage indexation	535
6.2 Greater wage flexibility for new hires	535

[☆] Many of the insights contained in this chapter are based on earlier joint work with Olivier Blanchard, who sparked my interest in the subject. I also thank the editors, Jan Eeckhout, Chris Pissarides, Carlos Thomas, and participants at the CREI Faculty Lunch and the Conference on “Key Development in Monetary Economics,” for helpful comments at different stages of this project. Tomaz Cajner and Lien Laureys provided excellent research assistance. I acknowledge financial support from the European Research Council, the Ministerio de Ciencia e Innovación and the Government of Catalonia.

6.3 Smaller wealth effects	536
6.4 Other demand shocks	537
7. Conclusions	537
References	543

Abstract

Much recent research has focused on the development and analysis of extensions of the New Keynesian framework that model labor market frictions and unemployment explicitly. This chapter describes some of the essential ingredients and properties of those models, and their implications for monetary policy.

JEL classification: E32

Keywords

Nominal Rigidities

Labor Market Frictions

Wage Rigidities

1. INTRODUCTION

The existence of involuntary unemployment has long been recognized as one the main ills of modern industrialized economies. And the rise in unemployment that invariably accompanies all economic downturns is, arguably, one of the main reasons why cyclical fluctuations are generally viewed as undesirable.

Despite the central role of unemployment in the policy debate, that variable has been — at least until recently — conspicuously absent from the new generation of models that have become the workhorse for the analysis of monetary policy, inflation and the business cycle, and which are generally referred to as New Keynesian.¹ That absence may be justified on the grounds that explaining unemployment and its variations has never been the focus of that literature, so there was no need to model that phenomenon explicitly. But this could be interpreted as suggesting that there is no independent role for unemployment — as distinguished, say, from measures of output or employment — as a determinant of inflation (or other macro variables) or as a variable that central banks should be concerned about and even respond to in a systematic way. In other words, under the previous view, unemployment and the frictions

¹ The reader can find a textbook exposition of the New Keynesian model in Walsh (2003a), Woodford (2003), and Galí (2008). An early version and analysis of the baseline New Keynesian model can be found in Yun (1996), who used a discrete-time version of the staggered price-setting model originally developed in Calvo (1983). King and Wolman (1996) provided a detailed analysis of the steady state and dynamic properties of the model. Goodfriend and King (1997); Rotemberg and Woodford (1999); and Clarida, Galí, and Gertler (1999) were among the first to conduct a normative policy analysis using that framework.

underlying it are not essential for understanding fluctuations in nominal and real variables, nor a key ingredient in the design of monetary policy.²

On the other hand, understanding the determinants of unemployment and the nature of its fluctuations has been at the heart of a parallel literature, one that has built on the search and matching models in the Diamond–Mortensen–Pissarides tradition.³ Since the influential work of Hall (2005) and Shimer (2005), pointing to the difficulties of a calibrated version of such a model to account for the size of observed fluctuations in unemployment and other labor market variables, that literature has taken a more quantitative turn and sparked the interest of mainstream macroeconomists. Yet, and at least until recently, the models used in that literature have been purely *real*, and hence they had nothing to say about the role of monetary policy, either as a source of unemployment fluctuations, or as a tool to stabilize those fluctuations.⁴

Over the past few years, however, a growing number of researchers have turned their attention toward the development and analysis of frameworks that combine elements from the two traditions described earlier. The typical framework in this literature combines the nominal rigidities and consequent monetary non-neutralities of New Keynesian models with the real frictions in labor markets that are characteristic of the search and matching models. To the extent of my knowledge, Chéron and Langot (2000) were the first to bring together nominal rigidities and labor market frictions, showing how the resulting framework could generate both a Beveridge curve (a negative correlation between vacancies and unemployment) and a Phillips curve (a negative correlation between inflation and unemployment) in the presence of both technology and monetary shocks. Subsequently, Walsh (2003b, 2005) and Trigari (2009) analyzed the impact of embedding labor market frictions into the basic New Keynesian model with sticky prices but flexible wages, with a focus on the size and persistence of the effects of monetary policy shocks.

More recent contributions have extended that work in two dimensions. First, they have relaxed the assumption of flexible wages, and introduced different forms of nominal and real wage rigidity. The work of Trigari (2006) and Christoffel and Linzert (2005) falls into that category. Secondly, the focus of analysis has gradually turned to normative issues, and more specifically, to the implications of labor market frictions and unemployment for the design of monetary policy. Thus, the work of Blanchard and Galí (2010; in a model with real wage rigidities) and Thomas (2008a; under nominal wage rigidities) provides an explicit analysis of the optimal monetary policy in the

² The term “unemployment” cannot be found in the index of Walsh (2003a) or Woodford (2003), two textbooks providing a modern treatment of monetary economics. In Galí (2008) I briefly mention “unemployment” in the concluding chapter, but only in reference to the recent extensions of the New Keynesian model discussed in this chapter.

³ Early contributions to the current vintage of search and matching models include Diamond (1982a,b), Mortensen (1982a,b), and Pissarides (1984). See Pissarides (2000) for a comprehensive exposition of the search and matching approach.

⁴ Incidentally, it is worth pointing out that standard RBC models share the shortcomings of both paradigms: they neither can explain involuntary unemployment nor have any role for monetary policy.

context of a simple New Keynesian model with labor market frictions.⁵ As argued later, and perhaps not surprisingly, those two extensions are not unrelated: the presence of wage rigidities has important implications, not only for the macroeconomic effects of different shocks, but also for the relative desirability of alternative policies.

While still in its infancy, the above-mentioned literature has already provided some insights of interest and has laid the ground for a possible “evolution” of the estimated DSGE models currently used for policy analysis, one that would introduce labor market frictions and unemployment explicitly in the full-fledged monetary models of the kind originally developed by Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2003, 2007). The recent work of Gertler, Sala, and Trigari (2008) and Christiano, Trabandt, and Walentin (2010) provides an excellent illustration of the progress being made in that direction.

The objective of this chapter is twofold. First, to describe some of the essential ingredients of a model that combines labor market frictions and nominal rigidities. And, secondly, to illustrate how such a model can be used to address questions of interest pertaining to the interaction between labor market frictions and nominal rigidities. Two broad questions are emphasized in the analysis below:

- What is the role of labor market frictions in shaping the economy’s response to aggregate shocks?
- What are the implications of those frictions for the design of monetary policy? In particular, should central banks pay attention to unemployment when setting interest rates?

To address those questions, I develop an extension of the New Keynesian model that allows for labor market frictions and unemployment. The model is highly stylized, combining elements found in existing papers, but abstracting from ingredients that (in my view) are not essential given the purpose at hand. Relative to the relevant literature, the main novelty of the framework developed here lies in the introduction of variable labor market participation. That feature is meant to overcome the surprising contrast between the importance given by the New Keynesian literature to the elasticity of labor supply (e.g., as a determinant of the persistence of the effects of monetary policy shocks) and the assumption of a fully inelastic labor supply found almost invariably in existing models with labor market frictions. In the latter, changes in unemployment match one-for-one those in employment (with the opposite sign), so there is no information contained in measures of unemployment that is not revealed by observing employment.

Several lessons emerge from the analysis, which are summarized next in the form of bullet points.

- Quantitatively realistic labor market frictions are likely to have, by themselves, a limited effect on the economy’s equilibrium dynamics. Instead, their main role is “to make room” for wage rigidities, with the latter leading to inefficient responses to shocks and significant trade-offs for monetary policy.

⁵ See also the analysis in Arseneau and Chugh (2008) in a model with flexible prices and quadratic costs of nominal wage adjustment.

- When combined with a realistic Taylor-type rule, the introduction of price rigidities in a model with labor market frictions has a limited impact on the economy's equilibrium response to real shocks (although it is sufficient to make monetary policy non-neutral).
- If the conditions that guarantee the efficiency of the steady state are assumed, the optimal policy under flexible wages (i.e., wages subject to period-by-period Nash bargaining) is one of strict inflation targeting, which requires that the price level be stabilized at all times. If, instead, nominal wages are bargained over and readjusted infrequently, the optimal policy involves moderate deviations from price stability and can be approximated well by a simple interest rate rule that responds to price inflation with a coefficient of about 1.5.
- Deviations in the unemployment rate from its efficient level are generally a source of welfare losses above and beyond those generated by fluctuations in the output or employment gaps. An optimized simple interest rate rule calls for a systematic (although relatively weak) stabilizing policy response to inefficient fluctuations in unemployment.

The chapter is organized as follows. Section 2 presents some evidence on the cyclical behavior of labor market variables and inflation, as well as a simple structural interpretation of their fluctuations. Section 3 develops a baseline model with labor market frictions and price rigidities, allowing for two alternative wage-setting environments (flexible and sticky wages). Section 4 discusses the properties of a calibrated version of the model, focusing on the implied responses to monetary and technology shocks. Section 5 presents the welfare criterion associated with the model under the assumption of an efficient steady state, and discusses the responses to a technology shock under the optimal monetary policy and the optimal simple rule. Section 6 discusses possible model extensions. Section 7 presents conclusions.

2. EVIDENCE ON THE CYCLICAL BEHAVIOR OF LABOR MARKET VARIABLES AND INFLATION

This section summarizes the cyclical properties of employment, the labor force, the unemployment rate, the real wage and inflation in the post-war U.S. economy. I use quarterly data corresponding to the sample period 1948Q1–2008Q4 and drawn from the HAVER database. GDP is taken to be the benchmark cyclical indicator. As a wage measure I used hourly compensation in the nonfarm business sector. The GDP deflator is the price level used to compute inflation and the real wage. Employment, the labor force, and GDP are normalized by working age population and, together with the real wage, are expressed in natural logarithms. All variables are detrended using a band-pass filter that seeks to preserve fluctuations with a periodicity between 6 and 32 quarters.

The first panel of Table 1 reports two key unconditional second moments for the cyclical component of each variable: its standard deviation relative to GDP and its correlation with GDP. Many of the facts reported in the table are well known but are summarized here as a reminder. Thus, note that employment is substantially more

Table 1 Cyclical Properties

	Unconditional		Demand		Technology	
	$\frac{\sigma(x)}{\sigma(y)}$	$\rho(x, y)$	$\frac{\sigma(x)}{\sigma(y)}$	$\rho(x, y)$	$\frac{\sigma(x)}{\sigma(y)}$	$\rho(x, y)$
Employment	0.60	0.83	0.59	0.92	0.90	0.51
Labor force	0.20	0.30	0.20	0.31	0.39	0.02
Unemployment rate	0.49	−0.90	0.50	−0.93	0.62	−0.76
Real wage	0.44	0.07	0.32	−0.78	0.27	0.27
Inflation	0.19	0.27	0.18	0.37	0.27	0.60

volatile than the labor force, with unemployment lying somewhere in between. The real wage is also shown to be substantially less volatile than GDP. Turning to the correlation with GDP, we see that both employment and the labor force are procyclical, although the latter only moderately so (their respective correlations are 0.83 and 0.30). The unemployment rate is highly countercyclical, with a correlation with GDP close to −0.9. Price inflation is mildly procyclical, but the real wage is essentially acyclical.

In addition to the unconditional statistics just summarized, Table 1 also reports conditional statistics based on a decomposition of each variable into “technology-driven” and “demand-driven” components. The decomposition is based on a partially identified VAR with five variables: (log) labor productivity, (log) employment, the unemployment rate, price inflation, and the average price markup. The latter is computed as the difference between (log) labor productivity and the (log) real wage.⁶ Following the strategy proposed in Galí (1999), I identified technology shocks as the only source of the unit root in labor productivity. The structural VAR contains four additional shocks that are left unidentified, and referred to loosely as “demand” shocks. I define the “demand” component of each variable of interest as the sum of its components associated with each of those four shocks.⁷

The second and third panels in Table 1 report some statistics of interest for the demand and technology components of a number of variables, computed after detrending the estimated components with a band-pass filter analogous to the one applied earlier to the raw data. Note that the conditional second moments associated with the demand-driven component are very similar to the unconditional second moments. This is not surprising once one realizes that nontechnology shocks account for the bulk of the volatility of the cyclical component of all variables (statistics not shown here). The only exception lies in the strong negative correlation between the real wage and

⁶ The baseline results discussed next are based on a specification of the VAR with (log) employment in first differences and the unemployment rate detrended using a second-order polynomial of time. The main findings are robust to an alternative specification with employment detrended in log-levels.

⁷ The reader is referred to Galí (1999) for a detailed description of the econometric approach.

GDP conditional on demand shocks, which contrasts with the near zero unconditional correlation between the same variables.

The conditional statistics associated with the technology-driven components are shown in the third panel of Table 1. Note that the labor force is now largely acyclical and the real wage mildly procyclical, both of which contrast with the corresponding unconditional statistics. Also, while the technology components of employment and the unemployment rate are shown to be procyclical and countercyclical, as measured by the corresponding correlation with GDP, a look at the estimated dynamic responses of those variables to a technology shock reveals a more complex pattern. Figure 1 displays the estimated responses to a favorable technology shock; that is, one that is shown to increase output and labor productivity permanently. Note that output hardly changes in the short run, with its response building up only gradually over time. On the other hand, employment declines on impact in response to that shock, and only gradually reverts back to its initial level. A similar result can be found in Galí (1999); Basu, Fernald, and Kimball (2006); Francis and Ramey (2005); and Galí and Rabanal (2004), among others, using alternative VAR specifications (and with a focus on hours rather than employment).⁸ The previous authors have argued that such estimated responses to a technology shock are at odds with the predictions of a standard calibrated real business cycle model, which would call for a simultaneous upward adjustment of output and employment in response to a technology improvement. The existence of short-run demand constraints, possibly resulting from the interaction of nominal rigidities and a not-fully-accommodating monetary policy, has been posited as an explanation for that evidence.

Figure 1 also provides evidence on the response of variables other than output and employment to a positive technology shock. In particular we see that the labor force declines slightly but permanently after that shock. That decline in the labor force can only offset partially the larger fall in employment, thus leading to a persistent increase in the unemployment rate, which is only reverted after six quarters. Similar evidence of a short-run rise in unemployment in response to a positive supply shock can also be found in Blanchard and Quah (1989) and, more recently, in Barnichon (2008). The latter author argues that such evidence implies a rejection of a central prediction of the standard search and matching model, although it can be accounted for once that model is extended to allow for nominal rigidities and a suitable monetary policy rule.

Next I explore whether a model that combines nominal rigidities and labor market frictions can account for different aspects of the evidence just described.

⁸ The previous evidence is not uncontroversial. For a critical perspective on that evidence see Christiano, Eichenbaum, and Vigfusson (2003) and Chari, Kehoe, and McGrattan (2008).

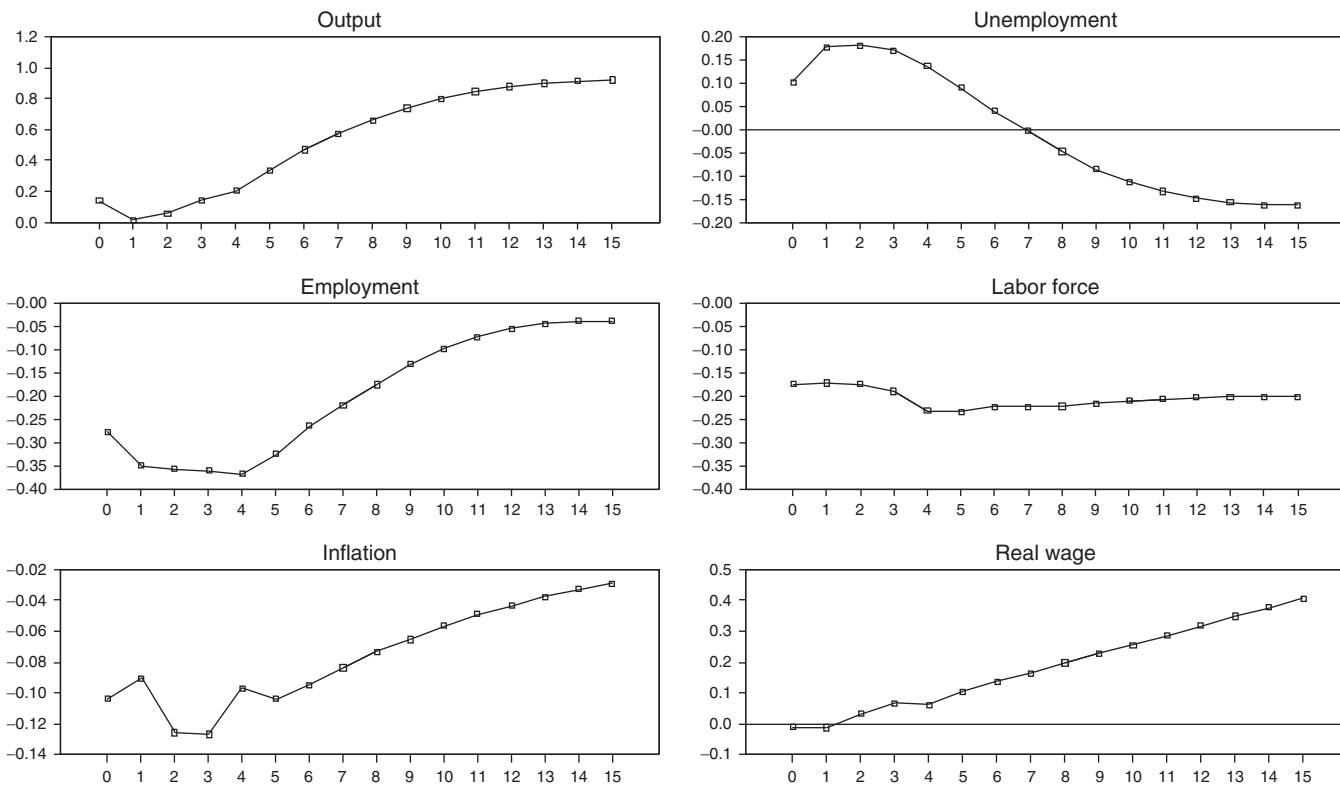


Figure 1 Estimated effects of technology shocks.

3. A MODEL WITH NOMINAL RIGIDITIES AND LABOR MARKET FRICTIONS

3.1 Households

I assume a large number of identical households. Each household is made up of a continuum of members represented by the unit interval. There is assumed to be full consumption risk sharing within each household.⁹ The household seeks to maximize the objective function

$$E_0 \sum_{t=0}^{\infty} \beta^t U(C_t, L_t) \quad (1)$$

where $\beta \in [0, 1]$ is the discount factor, $C_t \equiv \left(\int_0^1 C_t(i)^{1-\frac{1}{\epsilon}} di \right)^{\frac{\epsilon}{\epsilon-1}}$ is an index of the quantities consumed of the different types of final goods, and L_t is an index of the total effort or time that household members allocate to labor market activities. More specifically, I define L_t as

$$L_t = N_t + \psi U_t \quad (2)$$

where N_t and U_t denote, respectively, the fraction of household members who are employed and unemployed (and looking for a job).¹⁰ Parameter $\psi \in [0, 1]$ represents the marginal disutility generated by an unemployed member relative to an employed one. Nonparticipation in the labor market generates no disutility to the household. Note that the labor force (or participation rate) is given by $N_t + U_t \equiv F_t$. The following constraints must be satisfied for all t : $C_t(i) \geq 0$, all $i \in [0, 1]$, $0 \leq N_t + U_t \leq 1$, $U_t \geq 0$ and $N_t \geq 0$.

The household's period utility is assumed to take the form

$$U(C_t, L_t) \equiv \log C_t - \frac{\chi}{1+\varphi} L_t^{1+\varphi} \quad (3)$$

and where the disutility implied by labor market activities can be interpreted as resulting from foregone leisure and/or consumption of home produced goods. Note that by setting $\psi = 0$ the resulting utility function specializes to one commonly used in monetary models of the business cycle. That specification is consistent with a balanced growth path and involves a direct parametrization of the Frisch labor supply elasticity, which is given by $1/\varphi$. On the other hand, if $\varphi = 0$ is assumed, we can interpret the term $\chi N_t + \chi \psi U_t$ as the sum of the disutilities of labor market activities of all household

⁹ Merz (1995) was the first to adopt the assumption of a representative "large" household with a conventional utility function in the context of a search model.

¹⁰ I focus on variations in labor input at the extensive margin, and abstract from possible variations over time in hours per worker (or effort per worker). Even though the latter displays nontrivial cyclical movements in the data, its introduction seems unnecessary to convey the basic points made below. See Trigari (2009) and Thomas (2008), among others, for examples of related models that allow for variation in (disutility-generating) hours per worker.

members, with work and unemployment generating, respectively, individual disutilities of χ and $\chi\psi$ (with no disutility generated by nonparticipation).¹¹ Note also that the chosen specification differs from the one generally used in the search and matching literature, where the marginal rate of substitution is assumed to be constant, thus implying a fully inelastic labor supply above a certain threshold wage.

Employment evolves over time according to

$$N_t = (1 - \delta)N_{t-1} + x_t U_t^0 \quad (4)$$

where δ is a constant separation rate, x_t is the job finding rate, and U_t^0 is the fraction of household members who are unemployed (and looking for a job) at the beginning of period t . Note that $U_t = (1 - x_t) U_t^0$.¹²

The household faces a sequence of budget constraints given by

$$\int_0^1 P_t(i) C_t(i) di + Q_t B_t \leq B_{t-1} + \int_0^1 W_t(j) N_t(j) dj + \Pi_t$$

where $P_t(i)$ is the price of good i , $W_t(j)$ is the nominal wage paid by firm j , B_t represents purchases of one-period bonds (at a price Q_t), and Π_t is a lump-sum component of income (which may include, among other items, dividends from ownership of firms or lump-sum taxes). The above sequence of period budget constraints is supplemented with a solvency condition which prevents the household from engaging in Ponzi schemes.

Optimal demand for each good takes the familiar form:

$$C_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\epsilon} C_t \quad (5)$$

where $P_t \equiv \left(\int_0^1 P_t(i)^{1-\epsilon} di \right)^{\frac{1}{1-\epsilon}}$ denotes the price index for final goods. Note also that Eq. (5) implies that total consumption expenditures can be written as $\int_0^1 P_t(i) C_t(i) di = P_t C_t$.

The intertemporal optimality condition is given by

$$Q_t = \beta E_t \left\{ \frac{C_t}{C_{t+1}} \frac{P_t}{P_{t+1}} \right\} \quad (6)$$

In the model with frictionless, perfectly competitive labor markets the household would determine how much labor to supply, taking as given the (single) market wage.

¹¹ See, for example, Shimer (2009).

¹² Note that Eq. (4) implies that current hires become productive in the same period. This is the timing assumed in Blanchard and Galí (2010) and consistent with the bulk of the business cycle literature, where employment is assumed to be a non-predetermined variable. In contrast, most search and matching models assume it takes one period for a new hire to become productive, thus making employment predetermined, and preventing it from responding contemporaneously to shocks.

The wage would adjust so that all the labor supplied is employed, implying the absence of involuntary unemployment. Thus, we would have $L_t = N_t$ for all t , and under the assumed preferences, an intratemporal optimality condition would hold, equating the real wage to the marginal rate of substitution, $W_t/P_t = \chi C_t N_t^\varphi$, and implicitly determining the quantity of labor supplied. The present model departs from that Walrasian benchmark in an important respect: the wage does not “automatically” adjust to guarantee that all the labor supplied is employed. Instead, the wage is bargained bilaterally between individual workers and firms to split the surplus generated by existing employment relations. Employment is then the result of the aggregation of firms’ hiring decisions, given the wage protocol. In other words, employment is demand determined, with the households’ participation decision influencing employment only indirectly, through its impact on wages and on hiring costs.

3.2 Firms

As in much of the literature on nominal rigidities and labor market frictions, I assume a model with a two-sector structure. Firms in the final goods sector do not use labor as an input, but are subject to nominal rigidities in the form of restrictions to the frequency of their price-setting decisions. On the other hand, firms in the intermediate goods sector take the price of the good they produce as given, use labor as an input (subject to hiring costs), and engage in wage bargaining with its workers. That modeling strategy, originally proposed in Walsh (2005), has the advantage of getting around the difficulties associated with having price-setting decisions and wage bargaining concentrated in the same firms.¹³

3.2.1 Final goods

I assume a continuum of monopolistically competitive firms indexed by $i \in [0, 1]$, each producing a differentiated final good. All firms have access to an identical technology

$$Y_t(i) = X_t(i)$$

where $X_t(i)$ is the quantity of the (single) intermediate good used by firm i as an input.

Under *flexible prices* each firm would set the price of its good optimally each period, subject to a demand schedule with constant price elasticity ϵ .¹⁴ Profit maximization thus implies the familiar price-setting condition:

$$P_t(i) = \mathcal{M}^p(1 - \tau)P_t^f$$

¹³ See Kuester (2007) and Thomas (2008b) for an analysis of a version of the model where price-setters are subject to labor market frictions.

¹⁴ As discussed later, this requires that the demand of final goods coming from intermediate goods firms (to pay for their hiring costs), has the same price elasticity as the demand originating in households.

where P_t^I is the price of the intermediate good, $\mathcal{M}^p \equiv \frac{\epsilon}{\epsilon-1}$ is the optimal or desired (gross) markup and τ is a subsidy on the purchases of intermediate goods. Note that $(1 - \tau)P_t^I$ is the nominal marginal cost facing the final goods firm. Since all firms choose the same price it follows that

$$P_t = \mathcal{M}^p(1 - \tau)P_t^I$$

for all t .

Instead of flexible prices, I assume in much of what follows a price-setting environment as in Calvo (1983), with each firm being able to adjust its price each period only with probability $1 - \theta_p$. That probability is independent across firms and independent of the time elapsed since the last price adjustment. Thus, parameter $\theta_p \in [0, 1]$ also represents the fraction of firms that keep their prices unchanged in any given period and can thus be interpreted as an index of price rigidities.

All firms adjusting their price in any given period choose the same price, denoted by P_t^* , since they face an identical problem. The (log-linearized) optimal price setting condition in this environment is given by¹⁵

$$P_t^* = \mu^p + (1 - \beta\theta_p) \sum_{k=0}^{\infty} (\beta\theta_p)^k (E_t\{p_{t+k}^I\} - \tau) \quad (7)$$

where lower case letters denote the logs of the original variables, and $\mu^p \equiv \log \mathcal{M}^p$. Thus, firms that adjust their price in any given period, choose a (log) price that is equal to the desired (log) markup over a weighted average of current and (expected) future (log) marginal costs, with the weights being a function of both the discount factor β and the Calvo parameter θ_p .

By combining Eq. (7) with the (log-linearized) law of motion for the aggregate price level given by¹⁶

$$p_t = \theta_p p_{t-1} + (1 - \theta_p) p_t^* \quad (8)$$

one can derive the inflation equation

$$\pi_t^p = \beta E_t\{\pi_{t+1}^p\} - \lambda_p \hat{\mu}_t^p \quad (9)$$

where $\pi_t^p \equiv p_t - p_{t-1}$ is price inflation, $\hat{\mu}_t^p \equiv \mu_t^p - \mu^p = p_t - (p_t^I - \tau) - \mu^p$ denotes the deviation of the (log) average price markup from its desired (and steady state) value, and $\lambda_p \equiv \frac{(1-\theta_p)(1-\beta\theta_p)}{\theta_p}$. Equation (9) makes clear that whatever is the influence of labor market frictions and wage-setting practices on the dynamics of price inflation, it must

¹⁵ See, for example, Galí (2008, Chapter 3), for details of the derivation.

¹⁶ Equation (8) can be derived by log-linearizing the expression for the aggregate price level P_t around a zero inflation steady state, and using the fact that a fraction $1 - \theta_p$ of firms set the same price P_t^* , while the price index for the remaining fraction that keep their price unchanged is P_{t-1} , since they are drawn randomly from the universe of firms.

necessarily work through their impact on firms' markups, since variations in price inflation are the result of misalignments between current and desired price markups.

3.2.2 Intermediate goods

The intermediate good is produced by a continuum of identical, perfectly competitive firms, represented by the unit interval and indexed by $j \in [0, 1]$. All such firms have access to a production function

$$Y_t^I(j) = A_t N_t(j)^{1-\alpha}$$

Variable A_t represents the state of technology, which is assumed to be common across firms and to vary exogenously over time. More precisely, I assume that $a_t \equiv \log A_t$ follows an AR(1) process with autoregressive coefficient ρ_a and variance σ_a^2 .

Employment at firm j evolves according to

$$N_t(j) = (1 - \delta)N_{t-1}(j) + H_t(j) \quad (10)$$

where $\delta \in (0, 1)$ is an exogenous separation rate, and $H_t(j)$ represents the measure of workers hired by firm j in period t . Note that new hires start working in the period they are hired. That timing assumption, which follows Blanchard and Galí (2010), deviates from the standard one in the search and matching literature (which requires a one period lag before a hired worker becomes productive), but is consistent with conventional business cycle models, where employment is not a predetermined variable.

3.2.2.1 Labor market frictions

Following Blanchard and Galí (2010), I introduce labor market frictions in the form of a cost per hire, represented by G_t and defined in terms of the bundle of final goods. That cost is assumed to be exogenous to each individual firm.

Though G_t is taken as given by each individual firm, it is natural to think of it as depending on aggregate factors. One natural such determinant is the degree of tightness in the labor market, which can be approximated by the job finding rate $x_t \equiv H_t/U_t^0$; that is, the ratio of aggregate hires, $H_t \equiv \int_0^1 H_t(j) dj$, to the size of the unemployment pool at the beginning of the period, U_t^0 . More specifically, I assume¹⁷

$$\begin{aligned} G_t &= G(x_t) \\ &= \Gamma x_t^\gamma \end{aligned}$$

¹⁷ Instead, Blanchard and Galí (2010) assumed a hiring cost of the form $A_t \Gamma x_t^\gamma$. At the possible cost of less realism, that formulation has the advantage of preserving the homogeneity of the efficiency conditions with respect to the technology shock A_t , leading to a constrained-efficient allocation with a constant employment, which is a convenient benchmark.

Relation to the matching function approach. The above formulation is equivalent to the matching function approach adopted by the search literature. Under the latter, firms and workers match according to a function $M(V_t, U_t^0)$ where V_t represents the number of aggregate vacancies, and where a firm can post vacancies at a unit cost Γ . Under the assumption of homogeneity of degree one in the matching function, the fraction of posted vacancies that get filled within the period is given by $M(V_t, U_t^0)/V_t \equiv q(V_t/U_t^0)$, where $q' < 0$. On the other hand, the job finding rate is given by $x_t = M(V_t, U_t^0)/U_t^0 \equiv p(V_t/U_t^0)$ where $p' > 0$. It follows that a fraction $q(p^{-1}(x_t))$ of vacancies posted are filled with the resulting cost per hire being given by $G_t = \Gamma/q(p^{-1}(x_t))$, which is increasing in x_t . In particular, under the assumption of a Cobb-Douglas matching function $M(V_t, U_t^0) = V_t^\varsigma U_t^{01-\varsigma}$ we have $G_t = \Gamma x_t^{\frac{1-\varsigma}{\varsigma}}$, which coincides with the above specification of the cost function, for $\gamma \equiv \frac{1-\varsigma}{\varsigma}$.

In the presence of labor market frictions, wages (and, as a result, employment) may differ across firms, since they cannot be automatically arbitrated out by workers switching from low to high wage firms. I make this explicit by using the subindex j to refer to the wage and other variables that are potentially firm-specific. Given a wage $W_t(j)$, the optimal hiring policy of firm j is described by the condition

$$MRPN_t(j) = \frac{W_t(j)}{P_t} + G_t - (1 - \delta)E_t\{\Lambda_{t,t+1} G_{t+1}\} \quad (11)$$

where $MRPN_t(j) \equiv (P_t^I/P_t)(1 - \alpha) A_t N_t(j)^{-\alpha}$ is the marginal revenue product of labor (expressed in terms of final goods) and $\Lambda_{t,t+k} \equiv \beta^k (C_t/C_{t+k})$ is the stochastic discount factor for k -period ahead (real) payoffs.¹⁸ In other words, each period the firm hires workers up to the point where the marginal revenue product of labor equals the cost of a marginal worker. The latter, represented by the right-hand side of Eq. (11), has three components: (i) the real wage $W_t(j)/P_t$, (ii) the hiring cost G_t , and (iii) the discounted savings in future hiring costs that result from having to hire $(1 - \delta)$ fewer workers the following period. Equivalently, and solving Eq. (11) forward, we have:

$$G_t = E_t \left\{ \sum_{k=0}^{\infty} \Lambda_{t,t+k} (1 - \delta)^k \left(MRPN_{t+k}(j) - \frac{W_{t+k}(j)}{P_{t+k}} \right) \right\}$$

that is, the hiring cost must equate the (expected) surplus generated by the (marginal) worker.¹⁹

For notational convenience it is useful to define the *net* hiring cost as $B_t \equiv G_t - (1 - \delta)E_t\{\Lambda_{t,t+1} G_{t+1}\}$. Thus, one can rewrite Eq. (11) more compactly as:

¹⁸ Note that intermediate good firms are perfectly competitive and thus take the price P_t^I as given.

¹⁹ Implicitly it is assumed that the firm is always doing some positive hiring. This will be the case if exogenous separations are large enough and shocks are small enough.

$$MRPN_t(j) = \frac{W_t(j)}{P_t} + B_t \quad (12)$$

The previous optimality condition can be used to derive an expression for the (log) average price markup in the final goods sector, which was previously shown to be the driving force of inflation. Using $n_t \simeq \int_0^1 n_t(j) dj$ and $w_t \simeq \int_0^1 w_t(j) dj$ as approximate measures of (log) aggregate employment and the (log) average nominal wage around a symmetric steady state, log-linearization of Eq. (12) and subsequent integration over all firms yields the following expression for the average markup in the final goods sector:²⁰

$$\hat{\mu}_t^p = (a_t - \alpha \hat{n}_t) - [(1 - \Phi)\hat{\omega}_t + \Phi \hat{b}_t] \quad (13)$$

where $\omega_t \equiv w_t - p_t$ is the average (log) real wage, and $\Phi \equiv \frac{B}{(W/P)+B}$ measures the importance of (nonwage) hiring costs relative to the wage. Also, note for future reference that

$$\hat{b}_t = \frac{1}{1 - \beta(1 - \delta)} \hat{g}_t - \frac{\beta(1 - \delta)}{1 - \beta(1 - \delta)} (E_t\{\hat{g}_{t+1}\} - \hat{r}_t) \quad (14)$$

where $\hat{g}_t = \gamma \hat{x}_t$ and where r_t denotes the real return on a riskless one-period bond.²¹

Finally, note that Eq. (12) also implies

$$\alpha(n_t(j) - n_t) = -(1 - \Phi)(\omega_t(j) - \omega_t) \quad (15)$$

that is, the relative demand for labor by any given firm depends exclusively on its relative wage, with the corresponding elasticity being given by $-(1 - \Phi)/\alpha$. Note that this is a consequence of the hiring cost being common to all firms and independent of each firm's hiring and employment levels.²²

3.2.3 A brief detour: Labor market frictions and inflation dynamics

Empirical assessments of the price-setting block of the New Keynesian model have often focused on inflation Eq. (9) and made use of the fact that, in the absence of labor market frictions, the average price markup (or, equivalently, the real marginal cost, with the sign reversed) is given by

²⁰ Under the assumption that $\frac{p_t}{P_t}$, N_t , $\frac{W/P}{A}$ and $\frac{B}{A}$ have well-defined steady states, the previous equation will also hold in log-levels (with an added constant term), and hence will be consistent with nonstationary technology.

²¹ The price of a one-period riskless real bond is given by $\exp\{-r_t\} = E_t\{\Lambda_{t,t+1}\}$. Log-linearizing around a steady state we have

$$\hat{r}_t \equiv r_t - \rho \simeq -E_t\{\hat{\lambda}_{t,t+1}\}$$

where $\rho \equiv -\log \beta$ and $\lambda_{t,t+1} \equiv \log \Lambda_{t,t+1}$.

²² The assumption of a decreasing returns technology is required for wage differentials across firm to be consistent with equilibrium, given the assumption of price-taking behavior (otherwise only the firm with the lowest wage would not be priced out of the market). As an alternative, Thomas (2008a) assumed a constant returns technology, but combined it with the assumption of firm-specific convex vacancy posting costs, in the form of management utility losses.

$$\begin{aligned}\hat{\mu}_t^p &= (a_t - \alpha \hat{n}_t) - \hat{\omega}_t \\ &= -\hat{s}_t^n\end{aligned}$$

where $\hat{s}_t^n \equiv \hat{\omega}_t - (\hat{y}_t - \hat{n}_t)$ is the (log) labor income share, expressed as a deviation from its mean. The latter variable is readily available for most industrialized countries and can thus be used to construct a measure of the average markup, which can in turn serve as the basis for any empirical evaluation of Eq. (9).²³

The analysis above implies that in the presence of labor market frictions

$$\begin{aligned}\hat{\mu}_t^p &= (a_t - \alpha \hat{n}_t) - [(1 - \Phi)\hat{\omega}_t + \Phi\hat{b}_t] \\ &= -\hat{s}_t^n - \Phi(\hat{b}_t - \hat{\omega}_t)\end{aligned}$$

Thus, the resulting empirical inflation equation may be written as

$$\pi_t^p = \beta E_t \{ \pi_{t+1}^p \} + \lambda_p (\hat{s}_t^n + \Theta(\hat{b}_t - \hat{\omega}_t)) \quad (16)$$

Given Eq. (11) and the fact that $\hat{g}_t = \gamma \hat{x}_t$ it follows that in the presence of labor market frictions the measure of the average markup takes the form of a “corrected” labor income share, where the correction involves information on the current and future job finding rate.

In a recent paper, Krause, López-Salido, and Lubik (2008) revisited the empirical evidence on inflation dynamics using an equation similar to Eq. (16), together with data on the job finding rate to construct a modified markup series. They concluded that the impact of labor market frictions on the driving variable of inflation is rather limited. To some extent this is something one could anticipate for, as discussed later, under a realistic calibration of hiring costs, $\frac{B}{W/P} = (0.045) (1 - \beta (1 - \delta)) \simeq 0.006$, implying too small a coefficient Φ to make a significant difference in the markup measure, at least in the absence of implausibly large fluctuations in net hiring costs relative to wages.

3.3 Monetary policy

Under the model’s baseline specification, monetary policy is assumed to be described by a simple Taylor-type interest rate rule represented by

$$i_t = \rho + \phi_\pi \pi_t^p + \phi_\gamma \hat{y}_t + \nu_t \quad (17)$$

where $i_t \equiv -\log Q_t$ is the yield on a one-period nominally riskless bond, $\rho \equiv -\log \beta$ is the household’s discount rate, and ν_t is an exogenous policy shifter, which is assumed to follow an AR(1) process with AR coefficient ρ_ν and variance σ_ν^2 .

²³ See Galí and Gertler (1999); Galí, Gertler, and López-Salido (2001); and Sbordone (2002) for early applications of that approach.

Following Taylor (1993, 1999b), I take a properly calibrated version of the previous rule as a rough approximation to actual monetary policy in the United States. Much of the recent literature on nominal rigidities and labor market frictions has also adopted an interest rate rule similar to Eq. (17), even though some details may differ across papers.²⁴ Even though Eq. (17) is used as a baseline specification of monetary policy, I also consider alternative specifications of the policy rule when I turn to the normative analysis in Section 6.

Next I turn to a description of wage determination.

3.4 Labor market frictions and wage determination

I consider two alternative assumptions regarding wage setting: flexible wages and sticky wages. Under flexible wages, all wages are renegotiated and (potentially) adjusted every period. Under sticky wages only a constant fraction of firms can adjust their nominal wages in any given period. In both cases, the wage is determined according to a Nash bargaining protocol, with constant shares of the total surplus associated with each existing employment relation accruing to the worker (or his household) and the firm, respectively.

In contrast with the existing monetary models with labor market frictions, the following framework incorporates an explicit (albeit stylized) modeling of the participation decision. This is possible through the introduction of a (utility) cost to labor market participation, which the household must trade-off against the probability and benefits resulting from becoming employed.²⁵

Next I show, for both the flexible and sticky wage environments, how the surplus is split between households and firms as a function of the wage. In all cases, workers are assumed to act in a way consistent with maximization of the utility of their household, as specified in Eqs.(1) and (3) (as opposed to maximization of their hypothetical “individual” utility).

3.4.1 The case of flexible wages

Under this scenario each firm negotiates every period with its workers over their individual compensation. The value accruing to the representative household from a member employed at firm j , expressed in terms of final goods, is given by:

$$\mathcal{V}_t^N(j) = \frac{W_t(j)}{P_t} - MRS_t + E_t\{\Lambda_{t,t+1}((1 - \delta)\mathcal{V}_{t+1}^N(j) + \delta\mathcal{V}_{t+1}^U)\}$$

²⁴ Thus, Walsh (2005), Faia (2008), and Trigari (2009) include the lagged nominal rate in the rule as a source of inertia, but impose that the shock be serially uncorrelated. In addition, Walsh (2005) also assumed no systematic response to output, whereas Faia (2008) also included unemployment as an argument of the rule. Chéron and Langot (2000) and Walsh (2003b) are an exception in that they assume an exogenous process for the money supply, a less appealing specification from the point of view of realism.

²⁵ My approach generalizes the one used by Shimer (2010) in the context of a real search and matching model.

where $MRS_t \equiv \chi C_t L_t^\varphi$ is the household's marginal rate of substitution between consumption and labor market effort (or, equivalently, the marginal disutility of labor market effort, expressed in terms of the final goods bundle), and $\mathcal{V}_t^U$ is the value generated by a member who is unemployed at the beginning of period t .²⁶ The latter is given by

$$\mathcal{V}_t^U = x_t \int_0^1 \frac{H_t(z)}{H_t} \mathcal{V}_t^N(z) dz + (1 - x_t)(-\psi MRS_t + E_t\{\Lambda_{t,t+1} \mathcal{V}_{t+1}^U\})$$

The value associated with nonparticipation is normalized to zero. Under the assumption of an interior allocation with positive nonparticipation, the household must be indifferent between sending an additional member to the labor market or not. Thus, it must be the case that $\mathcal{V}_t^U = 0$ for all t . The latter condition in turn implies:

$$\psi MRS_t = \frac{x_t}{1 - x_t} \int_0^1 \frac{H_t(z)}{H_t} \mathcal{S}_t^H(z) dz \quad (18)$$

where $\mathcal{S}_t^H(j) \equiv \mathcal{V}_t^N(j) - \mathcal{V}_t^U(j) = \mathcal{V}_t^N(j)$ denotes the surplus accruing to the household from an established employment relation at firm j .²⁷

Thus we have:

$$\mathcal{S}_t^H(j) = \frac{W_t(j)}{P_t} - MRS_t + (1 - \delta) E_t\{\Lambda_{t,t+1} \mathcal{S}_{t+1}^H(j)\} \quad (19)$$

On the other hand, the surplus from an existing employment relation accruing to firm j is given by

$$\mathcal{S}_t^F(j) = MRP N_t(j) - \frac{W_t(j)}{P_t} + (1 - \delta) E_t\{\Lambda_{t,t+1} \mathcal{S}_{t+1}^F(j)\} \quad (20)$$

Note that under the maintained assumption that the firm is maximizing profits, it follows from Eqs. (11) and (20) that $\mathcal{S}_t^F(j) = G_t$ for all $j \in [0,1]$ and t . In other words, the surplus that a profit maximizing firm gets from an existing employment relation equals the hiring cost (which is also the cost of replacing a current worker by a new one, and thus what a firm “saves” from maintaining an existing relation).

The reservation wage for a worker employed at firm j is the minimum wage consistent with a non-negative surplus. It is given by

²⁶ Note that in defining the surplus relative to the value of an unemployed person at the beginning of the period, I am implicitly assuming that if no wage agreement is reached the worker always has a chance to join the pool of the unemployed and look for a job in the same period.

²⁷ Note that under the assumption that $\psi = 0$, there would be no cost associated with remaining unemployed so, to the extent the surplus from employment $\mathcal{S}_t^H(j)$ was positive, there would be full participation, so that $U_t = 1 - N_t$ for all t .

$$\Omega_t^H(j) = MRS_t - (1 - \delta) E_t \{ \Lambda_{t,t+1} \mathcal{S}_{t+1}^H(j) \}$$

The corresponding reservation wage for the firm, that is, the wage consistent with a non-negative surplus for the firm is

$$\Omega_t^F(j) = MRPN_t + (1 - \delta) E_t \{ \Lambda_{t,t+1} \mathcal{S}_{t+1}^F(j) \}$$

The *bargaining set* at firm j in period t is defined by the range of wage levels consistent with a non-negative surplus for both the firm and the worker, and thus corresponds to the interval $[\Omega_t^H(j), \Omega_t^F(j)]$. Note that the size of the bargaining set is given by

$$\begin{aligned} \Omega_t^F(j) - \Omega_t^H(j) &= \mathcal{S}_t^F(j) + \mathcal{S}_t^H(j) \\ &\geq G_t \end{aligned}$$

In other words, the presence of labor market frictions in the form of hiring costs guarantees the existence, in equilibrium, of a nontrivial bargaining set and, as a consequence, room for bargaining between firms and workers. As emphasized by Hall (2005), any wage that lies within the bargaining set is consistent with a privately efficient employment relation; that is, one that neither the worker nor the firm has an incentive to terminate.

Until the work of Hall (2005) and Shimer (2005), the search and matching literature has generally relied on the assumption of *period-by-period Nash bargaining* between workers and firms as a “selection rule” to determine the prevailing wage. This has also been the case for the more recent vintage of models with sticky prices, when no wage rigidities are assumed (see, e.g., Walsh, 2003b, 2005 and Trigari, 2009). In what follows, I take the assumption of period-by-period Nash bargaining as the one defining the *flexible wage* economy, leaving a discussion of an alternative for the next subsection.

Period-by-period Nash bargaining implies that the firm and each of its workers determine the wage in period t by solving the problem

$$\max_{W_t(j)} \mathcal{S}_t^H(j)^{1-\xi} \mathcal{S}_t^F(j)^\xi$$

subject to Eqs. (19) and (20), and where $\xi \in (0, 1)$ denotes the relative bargaining power of firms vis à vis workers.

The solution to that problem implies the following constant share rule:

$$\xi \mathcal{S}_t^H(j) = (1 - \xi) \mathcal{S}_t^F(j)$$

The associated (Nash) wage is thus given by

$$\begin{aligned} \frac{W_t(j)}{P_t} &= \xi \Omega_t^H(j) + (1 - \xi) \Omega_t^F(j) \\ &= \xi MRS_t + (1 - \xi) MRPN_t(j) \end{aligned} \tag{21}$$

Using Eq. (12) to substitute for $MRPN_{it}(j)$ we confirm that the wage is common to all firms and, as a result, so will be employment, the hiring rate, and the marginal revenue product. Thus, we can henceforth omit the j index in what follows and write the Nash wage as

$$\frac{W_t}{P_t} = \xi MRS_t + (1 - \xi)MRPN_t \quad (22)$$

which combined with Eq. (11) (evaluated at the symmetric equilibrium) implies

$$G_t - (1 - \delta) E_t \{ \Lambda_{t,t+1} G_{t+1} \} = \xi (MRPN_t - MRS_t) \quad (23)$$

Finally, note that under Nash bargaining the participation condition Eq. (18) can be rewritten as²⁸

$$\xi \psi MRS_t = (1 - \xi) \frac{x_t}{1 - x_t} G_t \quad (24)$$

3.4.2 The case of sticky wages

The flexibility of wages implied by the assumption of period-by-period Nash bargaining made in the previous subsection stands in conflict with the empirical evidence. More specifically, Eq. (22) implies that the nominal wage of all workers should experience continuous adjustments in response to changes in the price level, consumption, employment, productivity and any other variable that may affect the marginal rate of substitution or the marginal revenue product of firms. By contrast, the evidence based on observation of individual wages point to substantial nominal wage rigidities. Thus, Taylor's (1999a) survey of the evidence concluded that the average frequency of wage changes is about one year. Evidence of similar (and even stronger) nominal wage rigidities can be found in more recent studies using U.S. micro data (e.g., Barattieri, Basu, & Gottschalk, 2009) as well as micro data and surveys from many European countries (European Central Bank, 2009).

Motivated by that evidence, and by the difficulties of calibrated search and matching models with *flexible wages* to account for the observed volatility of unemployment or the “excess smoothness” of the real wage relative to labor productivity and GDP, many researchers have introduced different forms of wage rigidities in models with labor market frictions. As argued by Hall (2005), those frictions “make room” for such rigid wages, since they imply a nontrivial wage bargaining set consistent with privately efficient employment relations. In Hall's words, that property “...provides a full answer to the condemnation of sticky wage models in Robert Barro (1977), for invoking an inefficiency that intelligent actors could easily avoid.”

²⁸ As before, Eq. (24) is only needed when $\psi > 0$, so that $N_t \neq L_t$.

Perhaps not surprisingly given the indeterminacy inherent to the existence of a bargaining set, the range of proposals to model wage rigidities in the literature is broad. Thus, some authors introduce real wage rigidities (in either real or monetary models) by postulating an “ad hoc” real wage schedule, which implies (potentially) continuous adjustment of all wages, although one that is smoother than that implied by period-by-period Nash bargaining (see, e.g., Hall, 2005; Blanchard and Galí, 2007, 2010; Christoffel and Linzert, 2005). An alternative approach to modeling wage rigidities assumes staggered wage setting, so that only a fraction of workers are allowed to bargain over and adjust their wage in any given period. In that case, each individual wage remains unchanged for several periods, either in real terms (Gertler & Trigari, 2009) or, more realistically, in nominal terms (as in Bodart et al., 2006; Gertler, Sala, & Trigari, 2008; and Thomas, 2008a).

Here I follow the last group of authors and introduce wage rigidities in the form of staggered nominal wage setting à la Calvo. More specifically, I assume that the *nominal* wages paid by a given firm to its employees are renegotiated (and likely reset) with probability $1 - \theta_w$ each period, independently of the time elapsed since the last adjustment at that firm. The newly set wage is determined through Nash bargaining between each individual worker and the firm. Once the nominal wage is set, it remains unchanged until a new opportunity for resetting the wage arises. As a result, in any given period the wage (both real and nominal) will generally deviate from the flexible Nash wage derived in the previous subsection. Yet, and to the extent that shocks are not too large, the wage will remain within the relevant bargaining set and will thus be privately efficient to maintain the corresponding employment relation.

Most important, I assume that workers hired between renegotiation periods are paid the *average* wage prevailing at the firm. Thus, the average wage will have an influence on the firm’s hiring and employment levels. Yet, I assume that the number of workers is large enough that neither the firm nor the worker bargaining over the wage internalize the impact that their choice will have on the average wage. In a symmetric equilibrium all workers will get the same wage, which ex post will be equal to the average.²⁹ It is important to stress that the previous assumption is not an innocuous one. If new hires could negotiate their wage freely at the time of being hired, the existence of long spells with unchanged nominal wages for incumbent workers would have no direct impact on the hiring decisions and, as a result, on output and employment, as emphasized by Pissarides (2009). The empirical evidence on the relevance of wage stickiness for new hires remains controversial. Some authors have provided evidence pointing to greater wage flexibility for new hires (see, e.g., Haefke, Sontag, & van Rens, 2008, and the references in Pissarides, 2009), while others reject the existence

²⁹ This assumption simplifies the subsequent analysis considerably.

of any significant differences between new hires and incumbent workers (e.g., Gertler & Trigari, 2009, and Galuscak et al., 2008).³⁰

An immediate consequence of the staggering assumption is that wages will generally differ across firms, and so will employment and output. That dispersion in the allocation of workers across otherwise identical firms, coupled with the assumption of decreasing returns, is inefficient from a social viewpoint, a point further discussed below in the context of the normative analysis of the model.³¹

Next, I derive the basic equations describing the surpluses accruing to households and firms from existing employment relations, as a preliminary step to the analysis of wage determination as the outcome of a Nash bargain.

Let $\mathcal{V}_{t+k|t}^N$ denote the value accruing to a household in period $t + k$ from the employment of a member at a firm that last reset its wage in period t . Under the previous assumption we have:

$$\begin{aligned} \mathcal{V}_{t+k|t}^N = & \frac{W_t^*}{P_{t+k}} - MRS_{t+k} \\ & + E_{t+k} \left\{ \Lambda_{t+k,t+k+1} \left[(1 - \delta) \left(\theta_w \mathcal{V}_{t+k+1|t}^N + (1 - \theta_w) \mathcal{V}_{t+k+1|t+k+1}^N \right) + \delta \mathcal{V}_{t+k+1}^U \right] \right\} \end{aligned} \quad (25)$$

for $k = 0, 1, 2, 3 \dots$ where W_t^* denotes the nominal wage newly set in period t .³² Note that the last term on the right-hand side of Eq. (25) reflects the fact that the continuation value depends on whether wages are readjusted or not in the following period.

On the other hand, the value accruing to a household in period t from a member who is unemployed (but part of the labor force) at the beginning of period t is given by:

$$\mathcal{V}_t^U = x_t \int_0^1 \left(\frac{H_t(z)}{H_t} \right) \mathcal{V}_t^N(z) dz + (1 - x_t) (-\psi MRS_t + E_t \{ \Lambda_{t,t+1} V_{t+1}^U \})$$

³⁰ See Section 6 for a brief discussion of an extension by Bodart et al. (2006) allowing for differential flexibility between incumbents and new hires.

³¹ The inefficiencies resulting from staggered nominal wage-setting were already stressed in Erceg et al. (2000), in the context of a model without labor market frictions. Wage-staggering in Thomas (2008a) leads to an aggregate inefficiency as a result of the convexity of vacancy posting costs at the level of each firm. Here the inefficiency results from the presence of decreasing returns to labor.

³² Note that even though newly set wages can in principle differ across workers and firms, ex post all individual wages set in any given period will be identical. That justifies the omission of firm or worker indexes in W_t^* .

Again, optimal participation implies $\mathcal{V}_t^U = 0$ for all t . As a result

$$\mathcal{S}_{t+k|t}^H = \frac{W_t^*}{P_{t+k}} - MRS_{t+k} + (1 - \delta)E_{t+k} \left\{ \Lambda_{t+k,t+k+1} (\theta_w \mathcal{S}_{t+k+1|t}^H + (1 - \theta_w) \mathcal{S}_{t+k+1|t+k+1}^H) \right\} \quad (26)$$

and

$$\psi MRS_t = \frac{x_t}{1 - x_t} \int_0^1 \left(\frac{H_t(z)}{H_t} \right) \mathcal{S}_t^H(z) dz \quad (27)$$

Iterating Eq. (26) forward and evaluating the resulting expression at $k = 0$, one can determine the household surplus from an employment relation at a firm whose wages are currently being reset:

$$\begin{aligned} \mathcal{S}_{t|t}^H = E_t \left\{ \sum_{k=0}^{\infty} ((1 - \delta)\theta_w)^k \Lambda_{t,t+k} \left(\frac{W_t^*}{P_{t+k}} - MRS_{t+k} \right) \right\} \\ + (1 - \theta_w)(1 - \delta)E_t \left\{ \sum_{k=0}^{\infty} ((1 - \delta)\theta_w)^k \Lambda_{t,t+k+1} \mathcal{S}_{t+k+1|t+k+1}^H \right\} \end{aligned} \quad (28)$$

On the other hand, the period $t + k$ surplus accruing to a firm that last renegotiated its wages in period t , resulting from a marginal employment relation, is given by

$$\mathcal{S}_{t+k|t}^F = MRPN_{t+k|t} - \frac{W_t^*}{P_{t+k}} + (1 - \delta)E_{t+k} \left\{ \Lambda_{t+k,t+k+1} (\theta_w \mathcal{S}_{t+k+1|t}^F + (1 - \theta_w) \mathcal{S}_{t+k+1|t+k+1}^F) \right\} \quad (29)$$

for $k = 0, 1, 2, 3, \dots$, where $MRPN_{t+k|t} \equiv \frac{P_{t+k}^f}{P_{t+k}} (1 - \alpha) A_{t+k} N_{t+k|t}^{-\alpha}$ is the firm's marginal revenue product of labor, and $N_{t+k|t}$ its employment level.

Note, for future reference, that when combined with the optimal choice of employment by the firm at each point in time (as described by Eq. 11), Eq. (29) implies:

$$\mathcal{S}_{t+k|t}^F = G_{t+k}$$

for all t and k . In other words, the surplus accruing to the firm is always equal to the current hiring cost, independently of how long the wage has remained unchanged.

Iterating Eq. (29) forward and evaluating the resulting expression at $k = 0$ yields

$$\begin{aligned} \mathcal{S}_{t|t}^F = E_t \left\{ \sum_{k=0}^{\infty} ((1 - \delta)\theta_w)^k \Lambda_{t,t+k} \left(MRPN_{t+k|t} - \frac{W_t^*}{P_{t+k}} \right) \right\} \\ + (1 - \theta_w)(1 - \delta)E_t \left\{ \sum_{k=0}^{\infty} ((1 - \delta)\theta_w)^k \Lambda_{t,t+k+1} \mathcal{S}_{t+k+1|t+k+1}^F \right\} \end{aligned} \quad (30)$$

In the present environment, the Nash bargained wage at a firm that resets nominal wages in period t is given by the solution to

$$\max_{W_t^*} (\mathcal{S}_{t|t}^H)^{1-\xi} (\mathcal{S}_{t|t}^F)^\xi$$

subject to Eqs. (28) and (30). The implied sharing rule is given by

$$\xi \mathcal{S}_{t|t}^H = (1 - \xi) \mathcal{S}_{t|t}^F \quad (31)$$

which, combined with Eqs. (28) and (30), requires that the nominal wage newly set in period t satisfy the condition:

$$E_t \left\{ \sum_{k=0}^{\infty} ((1 - \delta)\theta_w)^k \Lambda_{t,t+k} \left(\frac{W_t^*}{P_{t+k}} - \Omega_{t+k|t}^{tar} \right) \right\} = 0 \quad (32)$$

where

$$\Omega_{t+k|t}^{tar} \equiv \xi MRS_{t+k} + (1 - \xi) MRP N_{t+k|t} \quad (33)$$

can be interpreted as the k -period ahead *target* real wage. Note that the expression for the latter corresponds to that of the relevant Nash wage under *flexible wages*, as derived in the previous subsection (see Eq. 21).

Log-linearizing the wage setting rule (Eq. 32) around a zero inflation steady state we obtain:

$$w_t^* = (1 - \beta(1 - \delta)\theta_w) E_t \sum_{k=0}^{\infty} (\beta(1 - \delta)\theta_w)^k E_t \left\{ \omega_{t+k|t}^{tar} + p_{t+k} \right\} \quad (34)$$

where $\omega_{t+k|t}^{tar} \equiv \log \Omega_{t+k|t}^{tar}$. In other words, the nominal wage set through Nash bargaining corresponds to a weighted average of the current and expected future *target* nominal wages relevant to the firm that is resetting wages. The weights decline geometrically with the horizon, at a rate that is a function of the degree of wage stickiness and the separation rate, since both those factors determine the expected duration of the newly set wage.

Next, I rewrite the above expression in terms of *average* target wages. Log-linearizing Eq. (33) around a symmetric steady state we have

$$\hat{\omega}_{t+k|t}^{tar} = \Upsilon(\hat{c}_{t+k} + \varphi \hat{l}_{t+k}) + (1 - \Upsilon)(-\hat{\mu}_{t+k}^p + a_{t+k} - \alpha \hat{n}_{t+k|t}) \quad (35)$$

where $\Upsilon \equiv \frac{\xi MRS}{W/P}$. Let ω_t^{tar} denote the (log) *average* target wage, defined as the current target wage for a (hypothetical) firm whose employment matched average employment. Formally,

$$\hat{\omega}_t^{tar} \equiv \Upsilon(\hat{c}_t + \varphi \hat{l}_t) + (1 - \Upsilon)(-\hat{\mu}_t^p + a_t - \alpha \hat{n}_t) \quad (36)$$

Note that one can interpret $\hat{\omega}_t^{tar}$ as the Nash bargained wage that would be observed in a flexible wage environment, conditional on the levels of consumption and (average) marginal revenue product generated by the equilibrium allocation under sticky wages.

Combining Eqs. (35) and (36) with Eq. (15)

$$\hat{\omega}_{t+k|t}^{tar} = \hat{\omega}_{t+k}^{tar} + (1 - \Upsilon)(1 - \Phi)(w_t^* - w_{t+k}) \quad (37)$$

Substituting Eq. (37) into Eq. (34), and after some algebraic manipulation we can derive the difference equation

$$w_t^* = \beta(1 - \delta)\theta_w E_t\{w_{t+1}^*\} - \frac{1 - \beta(1 - \delta)\theta_w}{1 - (1 - \Upsilon)(1 - \phi)}(\hat{w}_t - \hat{w}_t^{tar}) + (1 - \beta(1 - \delta)\theta_w)w_t \quad (38)$$

The law of motion for the (log) average wage $w_t \equiv \int_0^1 w_t(j) dj$ is given by

$$w_t = \theta_w w_{t-1} + (1 - \theta_w)w_t^* \quad (39)$$

Combining Eqs. (38) and (39), one can derive the following wage inflation equation:

$$\pi_t^w = \beta(1 - \delta)E_t\{\pi_{t+1}^w\}\lambda_w(\hat{w}_t - \hat{w}_t^{tar}) \quad (40)$$

where $\lambda_w \equiv \frac{(1 - \beta(1 - \delta)\theta_w)(1 - \theta_w)}{\theta_w(1 - (1 - \Upsilon)(1 - \Phi))}$. Note that the driving variable behind fluctuations in wage inflation is *the wage gap* $\omega_t - \omega_t^{tar}$, defined as the deviation between the average wage and the average *target* wage.³³

Finally, and as shown in Appendix 4 in this chapter, the optimal participation condition (Eq. 27) can be approximated around the zero inflation steady state as follows:

$$\hat{c}_t + \varphi \hat{l}_t = \frac{1}{1 - x} \hat{x}_t + \hat{g}_t - \Xi \pi_t^w \quad (41)$$

where $\Xi \equiv \frac{\xi(W/P)}{(1 - \xi)G} \frac{\theta_w}{(1 - \theta_w)(1 - \beta(1 - \delta)\theta_w)}$. Note that under flexible wages $\theta_w = 0$, implying $\Xi = 0$. The left-hand side of Eq. (41) measures the cost of labor market participation (through joining the pool of unemployed at the beginning of the period), while the right-hand side is the expected reward from that participation, both expressed as log deviations from their steady-state values. That reward is increasing in the job finding rate and in the size of current hiring costs (since workers with newly set wages will generate a surplus proportional to that variable), and decreasing in wage inflation (since

³³ Thomas (2008a) derived a similar representation for wage inflation — in the context of a slightly different model with efficient hours choice — convex vacancy posting costs, and constant returns.

the latter is positively related to the gap between the newly set wage and the average wage, with the latter being the one that is relevant to the participation decision).

3.4.2.1 Sustainability of the fixed wage

Both the firm and the worker will find it efficient to maintain an existing employment relation as long as their respective surpluses are positive. Thus, for a worker and firm that last reset the wage in period t , this will be the case as long as the nominal wage W_t^* remains within the bargaining set bounded by the reservation wages of the firm and the worker.

Formally, we require

$$W_t^* \in [\underline{W}_{t+k|t}, \overline{W}_{t+k|t}]$$

where

$$\underline{W}_{t+k|t} \equiv P_{t+k} \left(MRS_{t+k} - (1 - \delta) E_{t+k} \left\{ \Lambda_{t+k,t+k+1} \left(\theta_w S_{t+k+1|t}^H + (1 - \theta_w) S_{t+k+1|t+k+1}^H \right) \right\} \right)$$

and

$$\overline{W}_{t+k|t} \equiv P_{t+k} (MRPN_{t+k|t} + (1 - \delta) E_{t+k} \{ \Lambda_{t+k,t+k+1} G_{t+k+1} \})$$

Note that in the zero inflation steady state we have $W^* = P(\xi \underline{W} + (1 - \xi) \overline{W})$, so that the newly set wage lies within the bargaining set. Thus, the probability that the wage of any firm remains within that set outside the steady state will be larger the more stable the prices and consumption, employment, unemployment, and technology (the variables underlying MRS_t and $MRPN_{t+k|t}$). This will be the case, in turn, if shocks are “sufficiently small,” an assumption that I maintain in what follows. Notice, however, that given the Calvo structure, which implies that there are *some* wages that remained unchanged for arbitrarily long periods, it will be unavoidable that a small fraction of firms violate that condition in finite time (which would call for terminating the relationship or, more plausibly, violating the exogenous Calvo constraint on the timing of wage adjustments). Gertler and Trigari (2009) and Thomas (2008a) conduct simulations of related models and conclude that, for plausible calibrations of the wage rigidity parameter and shocks of empirically plausible size, the typical wage has a very small probability of falling outside the bargaining set before it gets to be readjusted. On those grounds, and following the literature, in my analysis I ignore that possibility, thus assuming that no wage ever hits the boundaries of the bargaining set.³⁴

³⁴ See Galí and van Rens (2009) for a model in which wages are adjusted only when they hit the boundaries of the bargaining set.

3.4.3 Relation to the New Keynesian wage inflation equation

Equation (40) has a structure analogous to the wage inflation equation that arises in the New Keynesian model with staggered nominal wage setting, as originally developed by Erceg, Henderson, and Levin (2000; EHL, henceforth). In the latter, each household is specialized in supplying a differentiated type of labor service, whose demand has a constant elasticity ϵ_w . In any given period it is allowed to reset the corresponding nominal wage unilaterally with a constant probability $1 - \theta_w$. The implied (log-linearized) optimal wage setting rule in the EHL model takes the form

$$w_t^* = \mu^w + (1 - \beta\theta_w)E_t \sum_{k=0}^{\infty} (\beta\theta_w)^k E_t \{mrs_{t+k|t} + p_{t+k}\} \quad (42)$$

where $\mu^w \equiv \log \frac{\epsilon_w}{\epsilon_w - 1}$ is the desired (log) wage markup of the real wage over the marginal rate of substitution (i.e., the one prevailing in the absence of wage rigidities). The previous optimal wage-setting rule can be contrasted with Eq. (34), the one prevailing under staggered wage setting with Nash bargaining.

The wage inflation equation that results from combining the log-linearized optimal wage setting rule (Eq. 42) with a law of motion for the average wage identical to Eq. (39) can be written as

$$\pi_t^w = \beta E_t \{\pi_{t+1}^w\} - \lambda_{ehl}(\hat{\omega} - \widehat{mrs}_t) \quad (43)$$

where mrs_t is the average (log) marginal rate of substitution between consumption and hours, and λ_{ehl} is a coefficient that is inversely related to the degree of wage stickiness θ_w . In particular, under the specification of preferences used in the model above with $\psi = 0$, we have $\widehat{mrs}_t = \hat{c}_t + \varphi \hat{n}_t$ and $\lambda_{ehl} \equiv (1 - \beta\theta_w)(1 - \theta_w)/(\theta_w(1 + \epsilon_w\varphi))$.³⁵

Three main differences with respect to Eq. (40) are worth pointing out.

First, the “effective” discount factor is smaller in the model with frictions, since it incorporates the probability of termination of each relationship (and thus of the associated wage), whereas in the EHL model the wage applies to the same group of workers throughout its duration, not to a specific relation that may be subject to termination. Secondly, the implicit target wage in the EHL model is given by the average marginal rate of substitution (augmented with a constant desired wage markup), whereas in the model with frictions the target wage is also a function of the marginal revenue product of labor, since that variable also influences the total surplus to be split through the wage negotiation. Finally, the difference in the coefficient on the wage gap between the two formulations captures the different adjustments needed to express the wage inflation equation in terms of average variables: the average marginal rate of substitution in the EHL model, and the average marginal revenue product of labor in the

³⁵ See Galí (2010) for a discussion of the relation between the New Keynesian Wage inflation equation and the original Phillips curve.

present model. Note that under the special parameter configuration $\delta = 0$ and $\xi = 1$, the form of the wage inflation equation of the present model matches exactly that of the EHL model.

3.5 Aggregate demand and output

Under the assumption that hiring costs take the form of a bundle of final goods given by the same CES function as the one defining the consumption index, the demand for each final good will be given by $Y_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon} (C_t + G_t H_t)$, where $H_t \equiv \int_0^1 H_t(j) dj$ denotes aggregate hires. Thus, the implied constancy of the price elasticity of demand justifying the constant desired markup $\mathcal{M}^p \equiv \frac{\epsilon}{\epsilon-1}$ assumed earlier.

Letting aggregate output be given by $Y_t \equiv \left(\int_0^1 Y_t(i)^{1-\frac{1}{\epsilon}} di\right)^{\frac{\epsilon}{\epsilon-1}}$ it can be easily checked that the aggregate goods market clearing condition may be written as

$$Y_t = C_t + G_t H_t \quad (44)$$

Hence, aggregate demand has two components. The first component is consumption, which evolves according to the Euler equation (6). The second component is the demand for final goods originating in firms' hiring activities.

Turning to the supply side, one can derive the following aggregate relation between final goods and intermediate input

$$\begin{aligned} X_t &\equiv \int_0^1 X_t(i) dj \\ &= Y_t \int_0^1 \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon} di \end{aligned} \quad (45)$$

where the term $D_t^p \equiv \int_0^1 \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon} di \geq 1$ captures the inefficiency resulting from dispersion in the quantities produced and consumed of the different final goods, which is a consequence of the price dispersion caused by staggered price setting.

On the other hand, the total supply of intermediate goods is given by

$$\begin{aligned} X_t &= \int_0^1 Y_t^I(j) dj \\ &= A_t N_t^{1-\alpha} \int_0^1 \left(\frac{N_t(j)}{N_t}\right)^{1-\alpha} dj \end{aligned} \quad (46)$$

where the term $D_t^w \equiv 1 / \int_0^1 \left(\frac{N_t(j)}{N_t}\right)^{1-\alpha} dj \geq 1$ captures the inefficiency resulting from dispersion in the allocation of labor across firms due to the staggering of wages, combined with the assumption of decreasing returns ($\alpha > 0$).

As shown in Appendix 1 in this chapter, in a neighborhood of the zero inflation steady state we have $D_t^p \simeq 1$ and $D_t^w \simeq 1$ up to a first-order approximation. Thus, combining Eqs. (45) and (46) we obtain the approximate aggregate production relation:

$$Y_t = A_t N_t^{1-\alpha} \quad (47)$$

For the sake of convenience, Appendix 3 collects all the model's (log) linearized equilibrium conditions, as derived in the previous sections. Next, I use those equilibrium conditions to characterize the behavior of a calibrated version of my model economy.

4. EQUILIBRIUM DYNAMICS: THE EFFECTS OF MONETARY POLICY AND TECHNOLOGY SHOCKS

This section presents the equilibrium responses of several variables of interest to the model's exogenous shocks — monetary policy and technology — and discusses how those responses are affected by nominal rigidities and labor market frictions. As a preliminary step I discuss the model's steady state, which is partly the basis for the calibration.

4.1 Steady state and calibration

The model's steady state is independent of the degree of price and wage rigidities, and of the monetary policy rule. For simplicity, I assume a steady state with zero inflation and no secular growth. I normalize the level of technology in the steady state to be $A = 1$. Notice that in steady state there are no relative price distortions so $D^p = D^w = 1$. Thus, the goods market clearing condition, evaluated at the steady state, can be written as

$$N^{1-\alpha} = C + \delta N \Gamma x^\gamma \quad (48)$$

Evaluating Eq. (23) at the steady state we have

$$(1 - \beta(1 - \delta)) \Gamma x^\gamma = \xi \left(\frac{1 - \alpha}{\mathcal{M}^p(1 - \tau)} N^{-\alpha} - \chi CL^\varphi \right) \quad (49)$$

Finally, the steady state participation condition requires

$$(1 - x) \xi \psi \chi CL^\varphi = (1 - \xi) \Gamma x^{1+\gamma} \quad (50)$$

The remaining steady state conditions include:

$$xU = (1 - x) \delta N \quad (51)$$

$$L = N + \psi U \quad (52)$$

To calibrate the model I adopt the following strategy. First, I pin down the steady-state employment rate, participation rate, and job finding rate using observed average values in the post-war U.S. economy. This leads to the choice of $N = 0.59$ and

$F = N + U = 0.62$, which in turn imply $U = 0.03$. Note that the implied unemployment rate as a fraction of the labor force — the conventional definition — is then close to 5% ($0.03/0.62 = 0.048$). Following Blanchard and Galí (2010), I set the steady-state value for the (quarterly) job finding rate x to 0.7. The implied separation rate is thus $\delta = (x/1 - x)U/N = 0.12$. Following convention I set $\alpha = 1/3$ and $\beta = 0.99$. Parameter φ is the inverse of the Frisch labor supply elasticity, a more controversial parameter due to the conflict between micro and macro evidence. I set $\varphi = 5$ in the baseline calibration, which corresponds to a Frisch elasticity of 0.2.

The baseline values for the parameters determining the degree of price and wage stickiness are set to imply average durations of one year in both cases, i.e. $\theta_p = 0.75$ and $\theta_w = 0.75$. This is roughly consistent with microeconomic evidence on wage and price setting.³⁶

Using the equivalence with the matching function approach discussed earlier and using estimates of the latter, I set $\gamma = 1$. I also assume $\mathcal{M}^p(1 - \tau) = 1$, so that the subsidy fully offsets the distortionary effects of the market power of final goods firms, which is one of the conditions for an efficient steady state. Following Hagedorn and Manovskii (2008) and Shimer (2009), who rely on the evidence reported in Silva and Toledo (2009), I take the average cost of hiring a worker to be 4.5% of the quarterly wage; that is, $G = 0.045 (W/P)$. Accordingly, the share of hiring costs in GDP is $\Theta = \delta NG/Y = (0.045)\delta S''$, where S'' is the labor income share. Setting the latter to 2/3 we have $\Theta = 0.0014$; that is, slightly above one-tenth of a percentage point of GDP. It follows that $\Gamma = G/x^\gamma = \Theta/(N^\gamma x^\gamma \delta) = 0.02$.

This leaves me with three free (although related) parameters, the firm's share in the Nash bargain (ξ), the weight of unemployment in the disutility of labor market effort (ψ), and the parameter scaling that disutility (χ). Given the value for one of these parameters, I can determine the remaining two by combining Eqs. (48), (49), and (50). Given the earlier choice of $\gamma = 1$, perhaps a natural benchmark setting for ξ is 0.5, which — as shown next — would be the value consistent with an efficient steady state and is often assumed in the literature. Yet, that configuration implies $\psi = 0.041$, a weight on unemployment that is arguably unrealistically small if one takes into consideration not only the time allocated to job search activities by the unemployed, but also the psychological costs of unemployment.³⁷ Thus, and as an alternative parameter

³⁶ See, for example, Nakamura and Steinsson (2008) and Basu and Gottschalk (2009) for recent U.S. micro evidence on price and wage rigidities, respectively.

³⁷ Thus, if the disutility of the unemployed (relative to the nonparticipant) results exclusively from the time allocated to job search activities and we take the standard work week for the employed to be of 40 hours, that calibration would that the unemployed 1.6 hours a week are allocated to job search activities. This is somewhat below the 2.5 hours per week of job search observed in time use surveys, as discussed in Krueger and Mueller (2008). The latter paper also provides survey-based evidence of subjective well-being, showing that unemployed individuals in the United States report considerably lower life satisfaction than the employed. Under literal interpretation of the model that evidence would call for a value above unity.

configuration I choose $\xi = 0.05$, which is associated with $\psi = 0.82$, possibly a more plausible value. As discussed next, the choice of a value in that range has significantly different, and more plausible, implications for the economy's response to a monetary policy shock. The implied settings for χ corresponding to the two calibrations are 15.5 and 12.3, respectively.

Finally, I calibrate the coefficients in the interest rate rule in a way consistent with the specification in Taylor (1993); that is, I set $\phi_\pi = 1.5$ and $\phi_y = 0.5/4 = 0.0125$ (the latter adjustment justified by Taylor's use of annualized inflation rate vs. quarter-to-quarter inflation here). That calibration is generally viewed as a reasonable approximation to monetary policy in United States, at least over the past three decades.

4.2 The effects of monetary policy and technology shocks

Figure 2A displays the dynamic responses of six macro variables (output, unemployment, employment, labor force, inflation, and the real wage) to an exogenous monetary policy shock, under the baseline assumption of $\xi = 0.5$, which is consistent with an efficient steady state. More specifically, disturbance v_t in the interest rate rule is assumed to rise by 0.25 percentage points, and to die out gradually according to an AR(1) process with an AR coefficient $\rho_v = 0.5$. Note that, in the absence of an endogenous component in the rule, such an experiment would be associated with a one percentage point increase in the (annualized) interest rate.

Although the estimated VAR model discussed in Section 2 did not specifically seek to identify monetary shocks, to the extent that those shocks and other demand shocks generate similar patterns among the variables considered, we can use the estimated conditional moments associated with demand shocks as a rough benchmark when evaluating the model's response to a monetary policy shock.

Figure 2A shows that both output and employment decline in response to the tightening of monetary policy, due to the contraction in consumption (not shown) resulting from the interest rate hike. Note also that the labor force increases by nearly 5%, driving up the unemployment rate by about 5 percentage points. In light of the evidence presented in Section 2, both responses seem implausibly large and, in the case of the labor force, it appears to go in the wrong direction. Note also that price inflation is procyclical in a way consistent with the evidence. The procyclical response of the real wage is, on the other hand, at odds with the estimated negative correlation with output conditional on demand shocks.

Figure 2B displays the corresponding responses to a technology shock. The latter takes the form of a one percent increase in a_t , which dies out gradually according to an AR(1) process with an AR coefficient of 0.9. Note that, in a way consistent with the estimated impulse responses shown in Figure 1, output rises and inflation declines, as one would expect from a positive technology shock. Note also that the real wage

risers gradually in the short run, a natural consequence of the existence of nominal wage rigidities. Furthermore, and in contrast with the standard search and matching model, employment declines and unemployment increases in response to the same positive technology shock. This is consistent with the evidence presented in Section 2 and in the literature referred therein. As was the case with monetary shocks, however, the rise in unemployment is largely driven by the increase in the labor force, which is far more volatile than employment and comoves negatively with the latter variable. This is in contrast with an estimated correlation (conditional on demand shocks) between the labor force and employment of 0.85.

A possible reason for the unrealistically large fluctuations in the labor force and unemployment shown in Figure 2A and B is the low value of parameter ψ (about 0.04) associated with the calibration underlying those figures. Such a low value implies a small penalty on fluctuations in those variables, given employment. Figure 3A and B

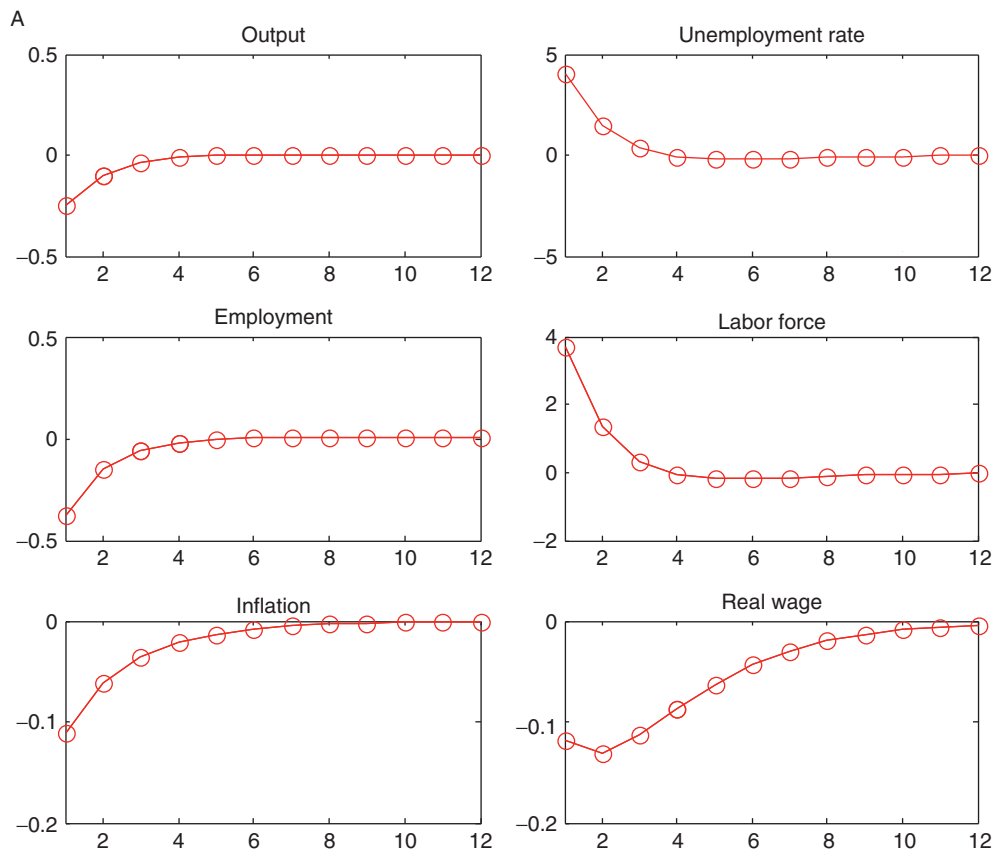


Figure 2A The effects monetary policy shocks: sticky wages ($\xi=0.5$).

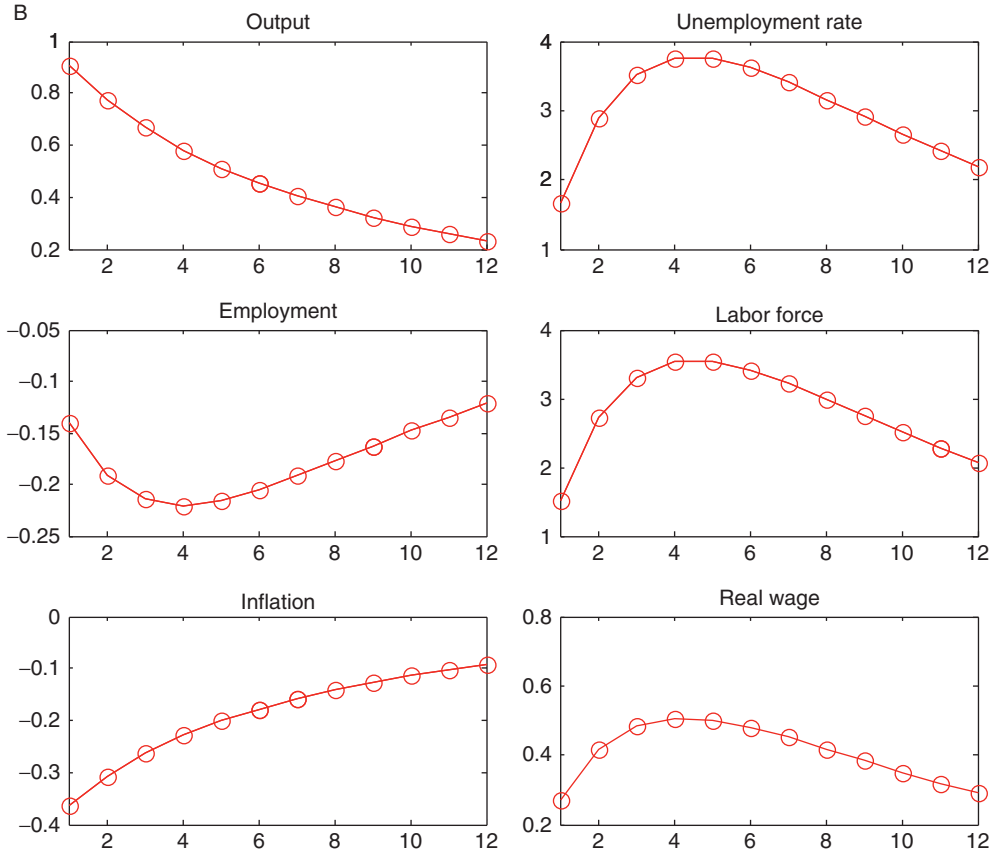


Figure 2B The effects of technology shocks: sticky wages ($\xi=0.5$).

shows the model's implied responses to monetary and technology shocks under the alternative calibration, with $\psi = 0.82$ and $\xi = 0.05$. As the figures make clear, now the labor force experiences much smaller variations, and comoves positively with employment. The latter's movements are the dominant force behind the variations in unemployment, in a way consistent with the evidence. The response of the remaining variables is not qualitatively affected. Thus, the only variable whose response is at odds with the evidence in Section 2 is the real wage, which responds procyclically to a monetary shock in the model, while displaying a negative correlation with output conditional on "demand" shocks in the data. That discrepancy could be due, however, to the presence of shocks other than technology shocks or monetary shocks (e.g., fiscal policy or labor supply shocks) that may be responsible for the negative correlation picked up by the partially identified VAR discussed in Section 2. Given the previous findings, and unless otherwise noted, I stick to this alternative calibration in the remainder of the paper.

4.3 The role of labor market frictions

To ascertain the role played by the presence of labor market frictions in shaping the economy's response to different shocks, I compare the model's implied responses to those shocks in the presence or not of such frictions. A perfectly competitive labor market is assumed in the case of no frictions. In both cases I maintain the assumption of flexible wages.

Figure 4A and B displays the economy's response to a monetary policy and a technology shock, respectively. Note that, in most cases the difference is quantitatively very small. Qualitatively, the only significant difference lies in the nonzero unemployment response to either shock in the presence of frictions, whereas in their absence a perfectly competitive labor market guarantees that there is no unemployment, implying

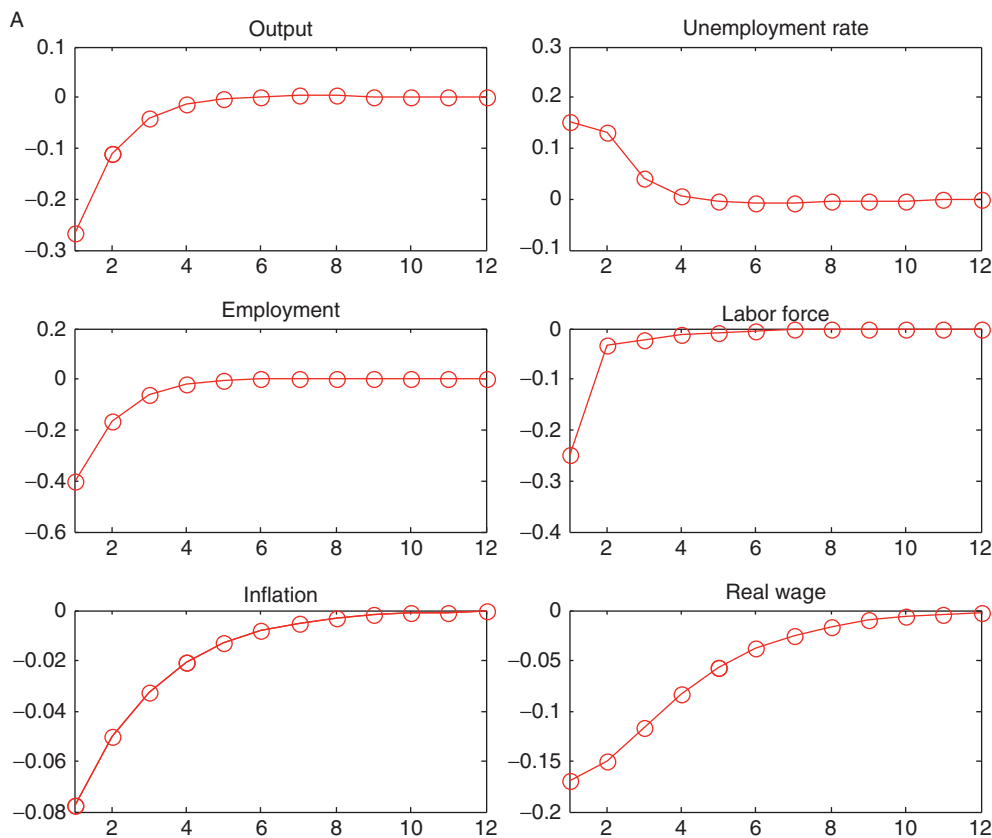


Figure 3A The effects of monetary policy shocks: sticky wages ($\xi=0.05$).

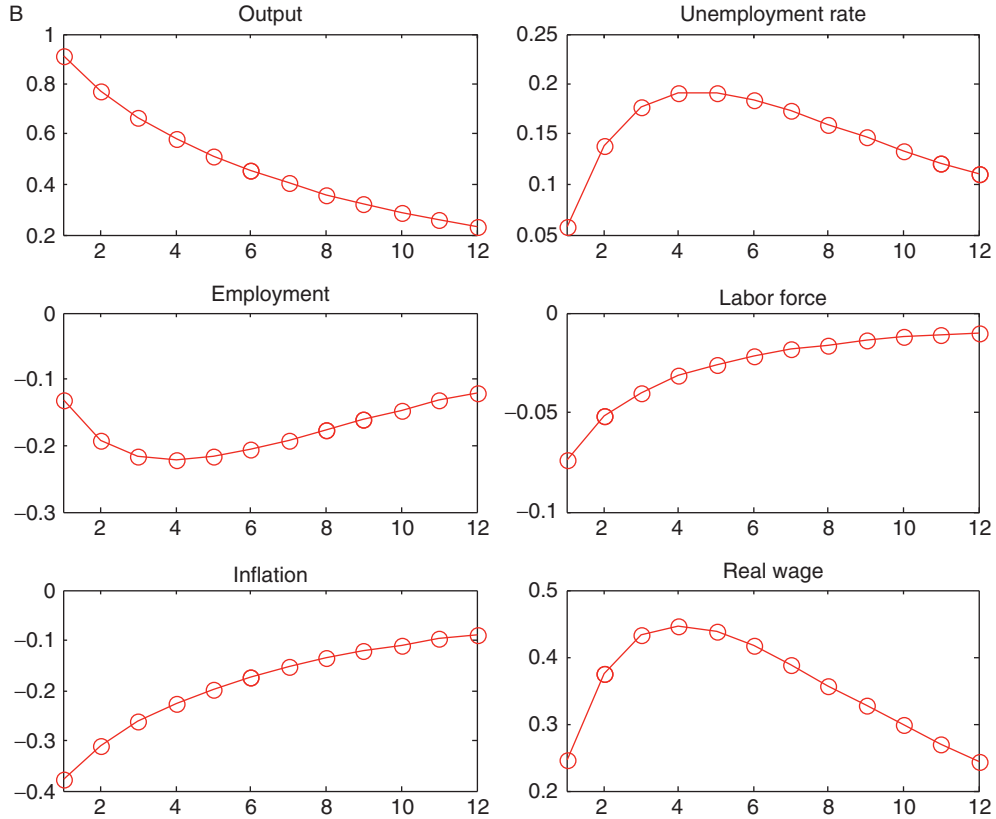


Figure 3B The effects of technology shocks: sticky wages ($\xi=0.05$).

that its response to shocks is flat at zero, as shown in Figure 4A and B. The variations in unemployment generated by the introduction of frictions are, however, very small for both shocks. This result is reminiscent of the so-called Shimer puzzle; i.e., the finding of too small a volatility of unemployment implied by a calibrated (real) search and matching framework with flexible wages and driven by technology shocks (Shimer, 2005).

The finding of a small role of labor market frictions in the response to monetary policy shocks contrasts somewhat with the conclusions from a related analysis in Walsh (2005). More precisely, Walsh showed that the introduction of labor market frictions has consequences on the pattern of the response of output and inflation to a monetary policy shock roughly equivalent to a substantial increase in the degree of price rigidities³⁸ in an otherwise standard New Keynesian model with Walrasian labor markets. In practice, it leads to

³⁸ Corresponding to an increase in the Calvo parameter θ_p from 0.5 to 0.85, which is equivalent to raising the average duration of prices from two to more than six quarters.

a significantly more sluggish response of inflation and a larger and more persistent response of output. A possible explanation for the discrepancy between Walsh's results and those found here lies in the fact that his model with labor market frictions assumes a constant marginal disutility from work, whereas his New Keynesian model introduces (with no apparent justification) a different utility function with an increasing marginal disutility of work. The latter feature will generally make wages and hence marginal costs more sensitive to variations in activity, thus leading to a larger response of prices in the short run, and a more dampened output response.³⁹

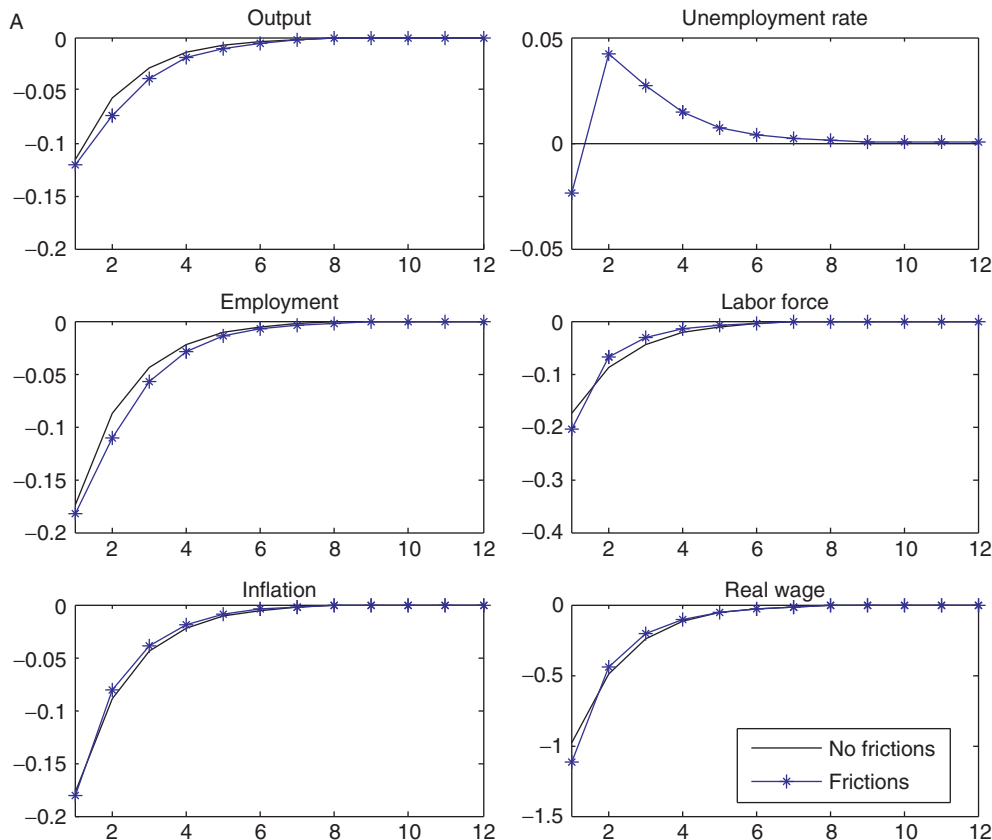


Figure 4A The role of labor market frictions: *flexible wages, monetary policy shock.*

³⁹ A similar discrepancy arises vis à vis Trigari (2009) in her comparison of the responses of a search model and a New Keynesian model to a monetary policy shock. Thus, in Trigari's search model labor adjustment takes place along two margins, hours per worker and employment, whereas in her New Keynesian model only hours per worker are allowed to vary. As argued by Trigari herself, that difference makes the elasticity of marginal cost to output larger in the New Keynesian model, which accounts for the weaker and less persistent response of output in the latter case.

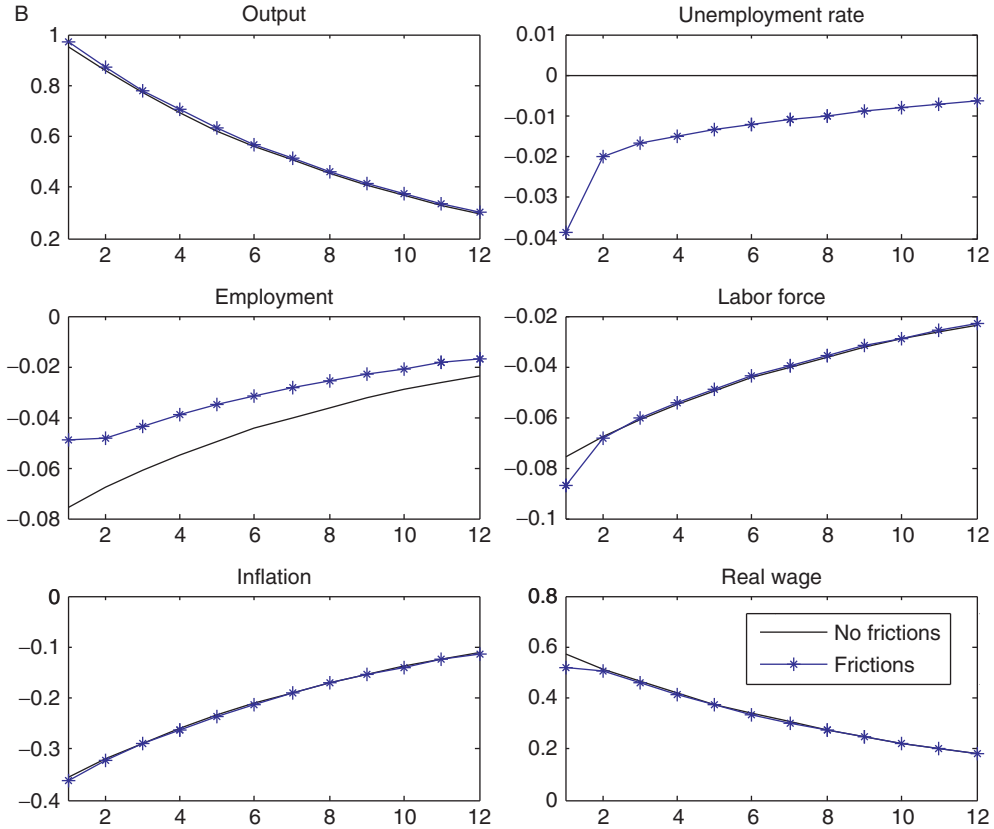


Figure 4B The role of labor market frictions: *flexible wages, technology shock*.

4.4 The role of price stickiness

How does the introduction of sticky prices affect, qualitatively and quantitatively, the response of unemployment and other variables to aggregate shocks? To address this question I analyze the response to monetary and technology shocks of two versions of the model economy developed earlier, with the presence or not of staggered price setting in the final goods sector as the only different among them. In both cases I maintain the assumption of full wage flexibility.

Figure 5A and B displays the corresponding impulse response functions. First, and not surprisingly, we see that the introduction of price stickiness has a significant impact on the economy's response to a monetary policy shock (Figure 5A). Thus, under flexible prices no real variable is affected by the shock, and only inflation declines in response to the tightening of policy. In contrast, once a realistic degree of price

stickiness is allowed for, the model implies a decline in output, employment, and the labor force, with a rise in the unemployment rate (after a tiny one period decline). Inflation and the real wage also decline, as expected.

The impact of price stickiness on the response to a positive technology shock (Figure 5B) appears to be much more limited. In particular, the effect on the size of the output response — more muted under sticky prices — is hardly discernible. The difference is sufficient, however, to account for a sign reversal in the response of employment, from positive to negative, although quantitatively the size of the employment adjustment is very small in both cases. Combined with a small influence (in the

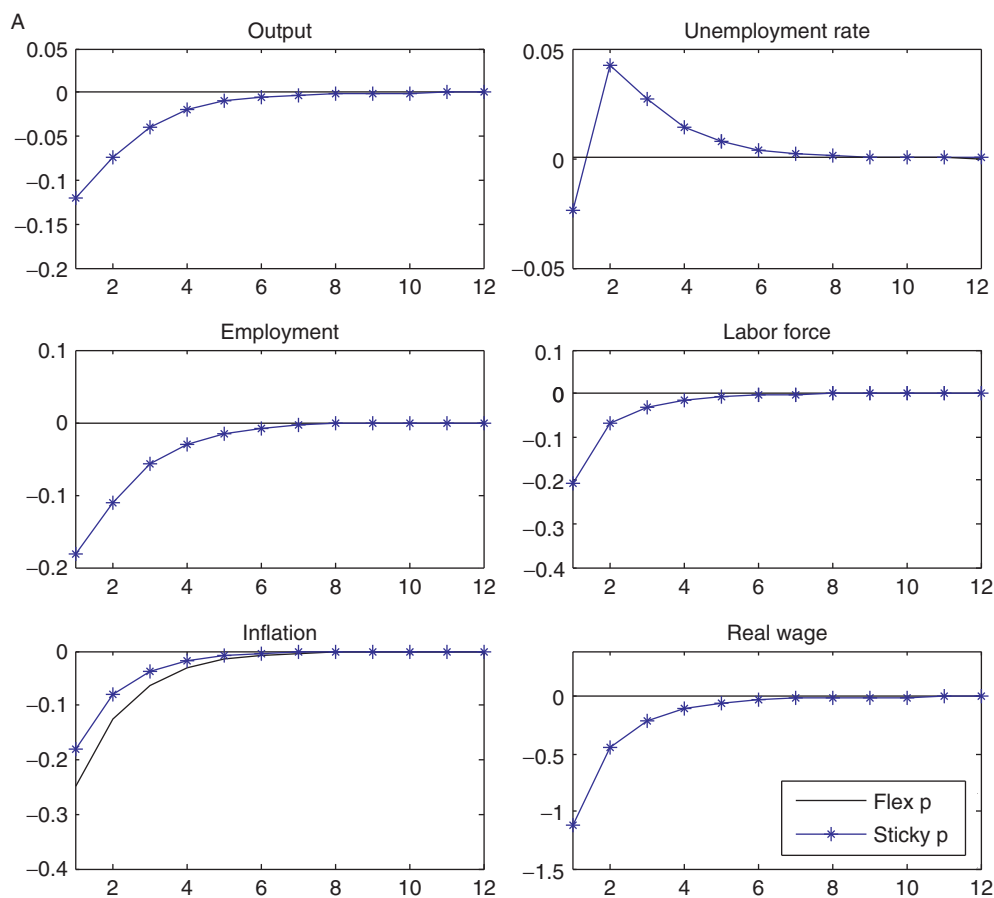


Figure 5A The role of price stickiness: *flexible wages, monetary policy shock.*

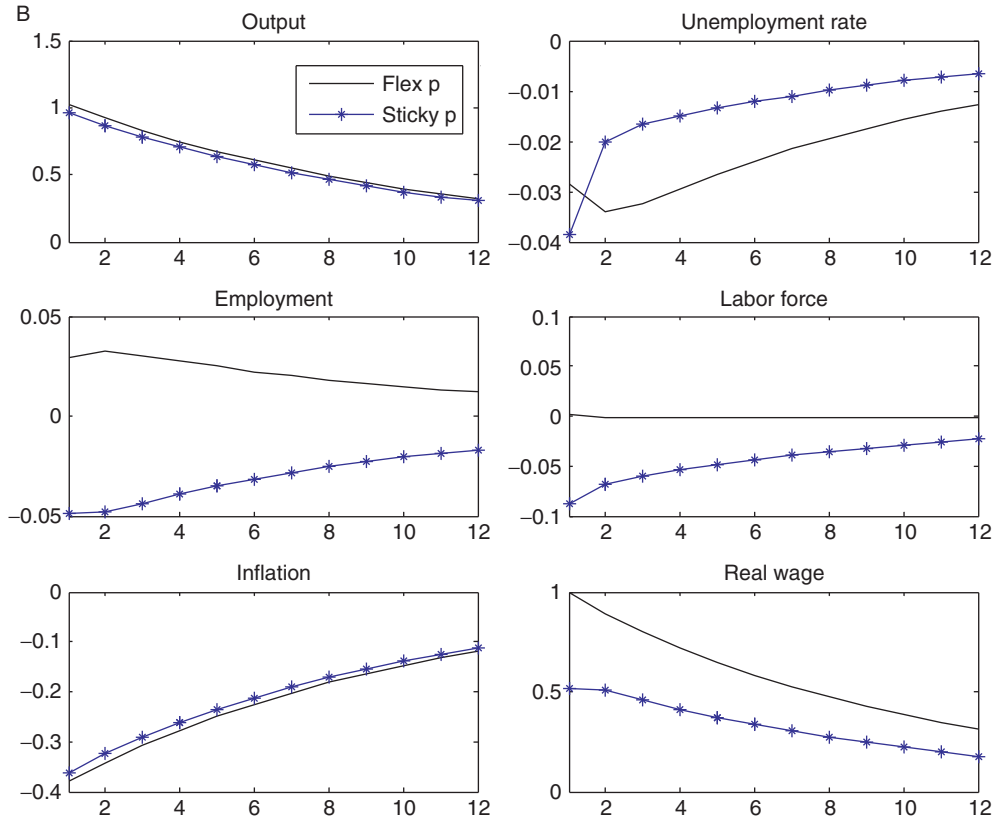


Figure 5B The role of price stickiness: *flexible wages, technology shock*.

same direction) on the response of the labor force, the impact of price stickiness on the response of unemployment to the technology shock is almost negligible.⁴⁰ The only sizable impact of price stickiness appears to be on the response of the real wage, which declines considerably as a result of the large rise in the markup of final goods firms that results from their failure to lower prices to match the decline in the price of intermediate goods. This is reflected in a muted rise in the marginal revenue product of intermediate goods firms and, as a result, on the wage.

⁴⁰ See Andrés, Domenech, and Ferri (2006) for a similar exercise in a model with endogenous capital accumulation, price indexation, and endogenous match destruction. Their findings point to a stronger role for price rigidities in accounting for the volatility of vacancies relative to unemployment, but not so much for the volatility of unemployment, which goes down slightly when stronger price rigidities are assumed.

4.5 The role of wage stickiness

Finally, I turn to an examination of the role played by wage stickiness in shaping the responses of the economy with labor market frictions to monetary and technology shocks. Figure 6A and B displays, respectively, the simulated responses to those shocks. For each type of shock, responses under two alternative calibrations are displayed. The solid line corresponds to an economy with flexible prices ($\theta_w = 0$), whereas the starred line assumes $\theta_w = 0.75$, implying an average duration of wages of one year. In both cases prices are assumed to be sticky.

As Figure 6A makes clear, the presence of sticky wages strengthens substantially the effects of a monetary policy shock on economic activity. In particular, the decline in output and employment is roughly twice as large as in the case of flexible wages. Since the response of the labor force is hardly affected, the resulting increase in unemployment is also much larger.

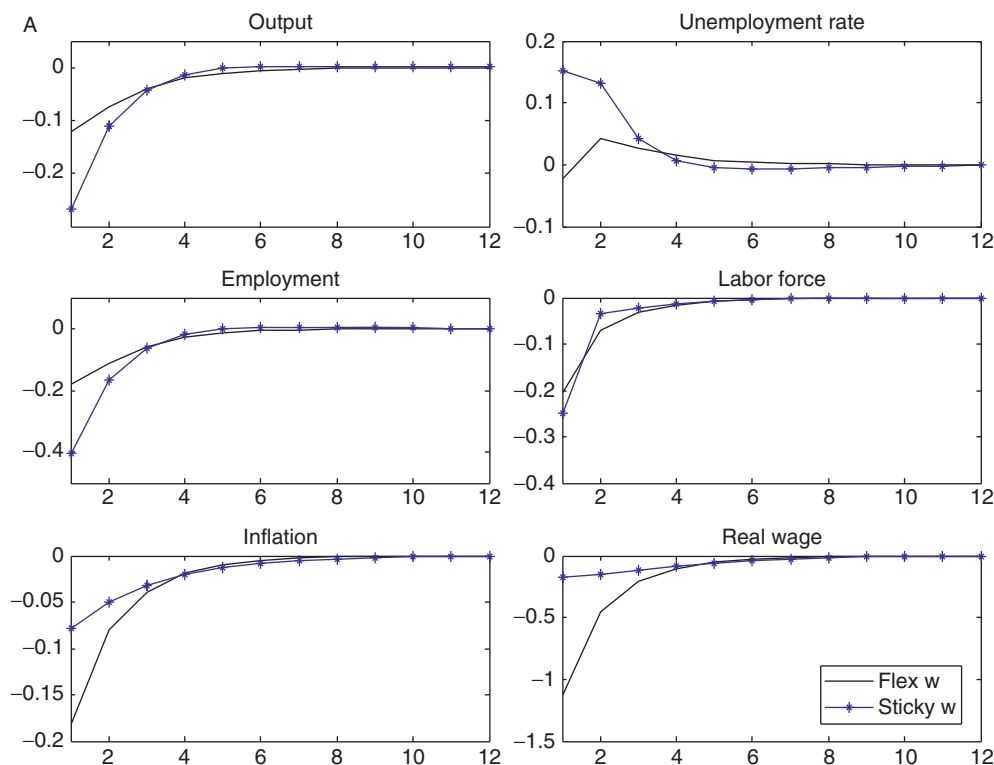


Figure 6A The role of wage stickiness: sticky prices, monetary policy shock.

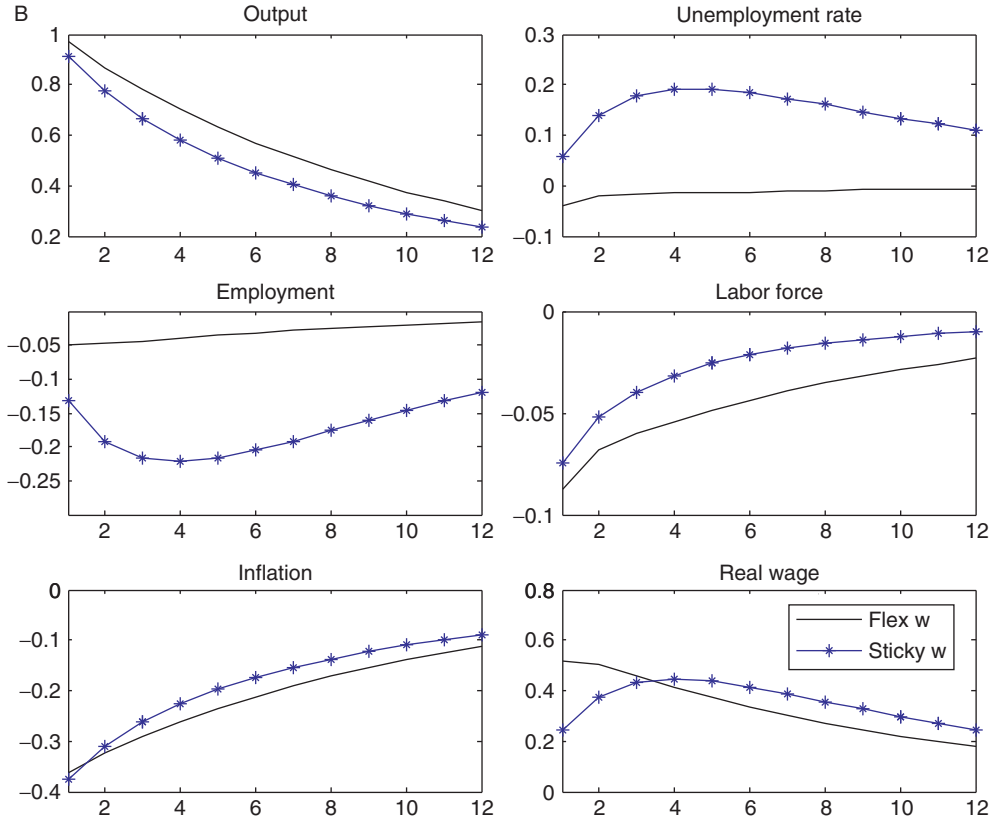


Figure 6B The role of wage stickiness: *sticky prices, technology shock*.

In addition, and not surprisingly, we see how the average real wage shows a much smoother response in the presence of staggered contracts, leading to less downward pressure on marginal costs and, as a result, a smaller decline in inflation.

The impact of wage stickiness on the responses to a technology shock is also substantial, as shown in Figure 6B. In particular, the negative response of employment is now larger, and that of the labor force (slightly) smaller. This is sufficient for the response of the unemployment rate to switch its sign, and thus to rise in response to a positive technology shock. Once again, that implication contrasts with the prediction of real models with labor market frictions (e.g., Shimer, 2005), but is consistent with the evidence presented in Section 2.

Note also that the introduction of sticky wages dampens the response of the real wage even further in the short run, driving closer to the near-zero short-run response uncovered by the empirical evidence in Section 2.

As discussed previously, the presence of labor market frictions does not appear to have much impact on the economy's response to shocks. The indirect impact is, however, more substantial to the extent that it justifies the presence of sticky wages in equilibrium.

Having looked at some of the positive predictions of the model under alternative sets of assumptions, I turn next to its normative implications.

5. LABOR MARKET FRICTIONS, NOMINAL RIGIDITIES AND MONETARY POLICY DESIGN

I start this section by describing the constrained-efficient allocation, and then turn my attention to the optimal design of monetary policy in the presence of labor market frictions and nominal rigidities. Ultimately, the purpose of the analysis is to shed light on how the existence of unemployment and wage rigidities should influence the conduct of monetary policy.

5.1 The social planner's problem

The social planner maximizes the representative household's utility

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(\log C_t - \frac{\chi}{1+\varphi} L_t^{1+\varphi} \right)$$

subject to the resource constraint

$$C_t + \Gamma x_t^\gamma H_t = A_t N_t^{1-\alpha}$$

and the definitions

$$\begin{aligned} L_t &= N_t + \psi U_t \\ H_t &= N_t - (1 - \delta)N_{t-1} \\ x_t &= \frac{H_t}{U_t/(1 - x_t)} \end{aligned}$$

In contrast with firms and households, the social planner internalizes the impact of its hiring and participation decisions on the job finding rate x_t and, hence, of the hiring cost. The optimality conditions characterizing the resulting constrained-efficient allocation are given by

$$MRS_t = MPN_t - (1 + \gamma)(G_t - (1 - \delta)E_t\{\Lambda_{t,t+1}G_{t+1}\}) \quad (53)$$

and

$$\psi MRS_t = \gamma \frac{x_t}{1 - x_t} G_t \quad (54)$$

where $MPN_t \equiv (1 - \alpha)A_tN_t^{-\alpha}$ is the marginal product of labor and, as above, $MRS_t \equiv \chi C_t L_t^\varphi$ is the marginal disutility of labor market effort, expressed in terms of the final consumption bundle.

5.1.1 The efficient steady state

Evaluated at the steady state, the previous two efficiency conditions take the form:

$$(1 + \gamma)(1 - \beta(1 - \delta))\Gamma x^\gamma = (1 - \alpha)N^{-\alpha} - \chi CL^\varphi \quad (55)$$

$$(1 - x)\psi\chi CL^\varphi = \gamma\Gamma x^{1+\gamma} \quad (56)$$

By comparing Eqs. (55) and (56) with the corresponding steady-state conditions of the decentralized economy Eqs. (49) and (50), it is easy to see that the latter's steady state will be efficient whenever

$$\mathcal{M}^p(1 - \tau) \quad (57)$$

and

$$\xi(1 + \gamma) = 1 \quad (58)$$

In words, condition (57) requires that the subsidy on the purchases of intermediate goods should exactly offset the impact of firms' market power, as reflected in the desired gross markup $\mathcal{M}^p$. Condition (58) is a version of the Hosios condition similar to the one derived in Blanchard and Galí (2010). It involves an inverse relation between firms' relative bargaining power, ξ , and the elasticity of hiring costs, γ . That inverse relation captures the negative externality (in the form of larger hiring costs) caused by firms' hiring decisions, and the positive externality resulting from higher participation (in the form of reduced hiring costs). The stronger these externalities (corresponding to a larger γ) are, the lower the relative bargaining power of firms (the smaller ξ), which is consistent with an efficient allocation, since the implied higher wages would induce fewer hires and more participation.

5.2 Optimal monetary policy

For simplicity, and throughout this section, I maintain the assumption of a constrained-efficient steady state; that is, conditions (57) and (58) are assumed to hold. The assumption of an efficient steady state is often made in the literature on optimal monetary policy, for in that case the latter focuses exclusively on offsetting (or at least alleviating) the consequences of inefficient fluctuations in response to shocks.⁴¹ Like before, I consider the two scenarios of flexible and sticky wages in turn.

⁴¹ See Woodford (2003) and Galí (2008) for a discussion of these issues in the context of the New Keynesian model without frictions.

5.2.1 The case of flexible wages

Under the assumption of period-by-period Nash bargaining of wages analyzed in Section 4.1, it is easy to check that the optimal monetary policy corresponds to a strategy of strict inflation targeting, that is, full stabilization of the price level. To see this, note from Eq. (9) that under that policy the markup of final goods firms will remain constant and equal to the desired level; that is, $P_t/P_t^I = \mathcal{M}^p(1 - \tau)$, for all t . Combined with assumption (57), it follows that $MRPN_t = MPN_t = (1 - \alpha)A_tN_t^{-\alpha}$ for all t . Thus, and imposing Eq. (58), one can easily check that equilibrium conditions (23) and (24) match exactly the efficiency conditions (53) and (54). In other words, the resulting equilibrium allocation is efficient.

Intuitively, under assumptions (57) and (58), the equilibrium of an economy in which both prices and (Nash bargained) wages are flexible involves a constrained-efficient allocation. Under flexible wages, a monetary policy that succeeds in fully stabilizing the price level replicates that natural allocation, and is thus optimal. That policy can be implemented with the assumed interest rate rule by choosing an arbitrarily large coefficient ϕ_π . That environment is thus characterized by what Blanchard and Galí (2007) referred to as “the divine coincidence,” or the absence of a trade-off between inflation stabilization and the attainment of an efficient allocation — one implies the other.

The previous finding hinges on the efficiency of the flexible price equilibrium allocation, guaranteed by assumptions (57) and (58). Faia (2009) analyzed the optimal policy in a related model (i.e., one with labor market frictions, sticky prices, and flexible wages), while relaxing the assumption of efficiency of the flexible price allocation. She shows that in that case it is optimal for the central bank to deviate from a policy of strict inflation targeting, although the size of the deviations implied by her calibrated model are quantitatively small.

5.2.2 The case of sticky wages

As is well known from the analysis of Erceg et al. (2000) and others, when both prices and wages are sticky it is generally impossible for the central bank to replicate the constrained-efficient equilibrium allocation, which under assumptions (57) and (58) corresponds to the equilibrium allocation in the absence of nominal rigidities (the *natural* allocation, for short), as previously discussed. The intuition behind that result is straightforward: in response to real shocks the real wage will generally adjust in the equilibrium with flexible prices and wages, and that adjustment will be necessary to support the resulting (constrained-efficient) allocation. Any adjustment of the real wage requires some variation in either the price level or the nominal wage. But in the presence of sticky prices and wages such variations will occur only in response to deviations of average price markups and/or average real wages from their natural

counterparts (see Eqs. 9 and 40), from which it follows that the natural (and efficient, under my assumptions) allocation will not be attainable.

To determine the optimal policy in that context I start by deriving a second-order approximation to the representative household's utility losses caused by deviations from the constrained efficient allocation due to the presence of nominal rigidities. In so doing I restrict myself to the case of small fluctuations around the efficient steady state. As derived in Appendix 4, the loss function takes the following form (expressed in terms of the consumption-equivalent loss, as a fraction of GDP):

$$\begin{aligned} \mathbb{L} \equiv & \frac{1}{2} E_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{\epsilon}{\lambda_p} (\pi_t^p)^2 + \frac{(1 + \Phi)^2 (1 - \alpha)}{\alpha \lambda_w^*} (\pi_t^w)^2 \right. \\ & \left. + \frac{(1 + \varphi)(1 - \Omega)N}{(1 - \alpha)L} \left(\tilde{y}_t + \frac{(1 - \alpha)\psi U}{N} \tilde{u}_t \right)^2 \right) \end{aligned} \quad (59)$$

where $\tilde{y}_t \equiv y_t - y_t^n$ and $\tilde{u}_t \equiv u_t - u_t^n$ are, respectively, the output and unemployment gaps relative to their natural counterparts (where the latter are defined as their equilibrium values under flexible prices and wages); $\lambda_w^* \equiv (1 - \theta_w)(1 - \beta\theta_w)/\theta_w$ is inversely related to the degree of wage rigidities θ_w ; and $1 - \Omega \equiv \frac{MRS}{MPN} = 1 - \frac{B(1+\gamma)}{MPN}$ is the steady-state gap between the marginal rate of substitution and the marginal product of labor resulting from the existence of labor market frictions. Note that in the absence of labor market frictions and under flexible wages $\lambda_w^* \rightarrow \infty$, $\Omega = 0$, $U = 0$, so the previous loss function collapses to the one familiar from the basic New Keynesian model.⁴²

The presence of labor market frictions has two implications for the welfare criterion. First, to the extent that they are accompanied by staggered nominal wage-setting, fluctuations in wage inflation will generate welfare losses due to the implied dispersion in wages and the resulting losses from an inefficient allocation of labor across firms.⁴³ Note that here the size of the welfare losses resulting from any given departure from wage stability is (i) increasing in $1 - \Phi$ (which measures the weight of wages in the total cost of employing a new worker), (ii) decreasing in the degree of diminishing returns to labor α (for the latter dampens the extent of employment dispersion caused by any given level of wage dispersion), and (iii) increasing in the degree of wage stickiness θ_w (which determines the degree of wage dispersion caused by a given deviation from zero wage inflation).

⁴² See the expression in Galí (2008, p. 81), under $\alpha = 1$.

⁴³ By contrast, in the monopoly union model of Erceg et al. (2000) the welfare losses from wage inflation are a consequence of the distorted allocation of employment across labor types *within each firm*, resulting from dispersion in their wages caused by staggered wage setting.

Secondly, and to the extent that $\psi > 0$, the welfare criterion above points to a specific role for unemployment gap fluctuations as a source of welfare losses, beyond that associated with variations in the output gap (or the employment gap, which by construction is proportional to output gap). That role is related to the fact that unemployment is a component of effective labor market effort, and that fluctuations in the latter (relative to its efficient benchmark) generate disutility. The importance of unemployment fluctuations is thus increasing in ψ and U , which determine the weight of unemployment in the total disutility from market effort.

The equilibrium allocation under the optimal monetary policy can be determined by minimizing Eq. (59) subject to the log-linearized equilibrium conditions listed in Appendix 2 (excluding the Taylor rule). Figure 7 displays the equilibrium responses to a technology shock of the same variables considered earlier, under the optimal policy. For the sake of comparison it also displays the corresponding responses under the Taylor rule used previously. The simulation is based on a calibration with stickiness in both prices and wages. Note that the optimal response implies some deviation from price stability. In particular it requires a temporary decline in inflation, which makes it possible for the real wage to adjust upward with a smaller upward adjustment of nominal wages.⁴⁴ It also allows for a stronger accommodation of the increase in productivity, as reflected in the larger positive response of output. In accordance, employment is allowed to rise, and unemployment to decline. Note also that the optimal policy is associated with a smaller decline in inflation than the Taylor rule. Despite the greater price stability, the cumulative response of the real wage is stronger under the optimal policy, which requires positive wage inflation (not shown) in contrast with the wage deflation associated with the equilibrium under the Taylor rule.

Is there a simple interest rate rule that the central bank could follow that would improve on the assumed Taylor rule? To answer that question I compute the optimal rule among the class of interest rate rules of the form:

$$i_t = \rho + \phi_p \pi_t^p + \phi_y \hat{y}_t + \phi_w \pi_t^w + \phi_u u_t$$

where I have added wage inflation and the unemployment rate as arguments, relative to the conventional Taylor rule. The coefficients that minimize the household's welfare loss, determined by iterating over all possible configurations, are $\phi_p = 1.51$, $\phi_y = -0.10$, $\phi_w = 0.01$, and $\phi_u = -0.025$. Figure 8 summarizes the dynamic response of the economy under that optimal simple rule, and compares it to the corresponding responses under the fully optimal policy, and makes clear the differences between the two are practically negligible. Note that relative to the standard Taylor rule, the optimized simple rule calls for further accommodation of supply-driven output variation

⁴⁴ See Thomas (2008a) for a related result in the context of a similar model.

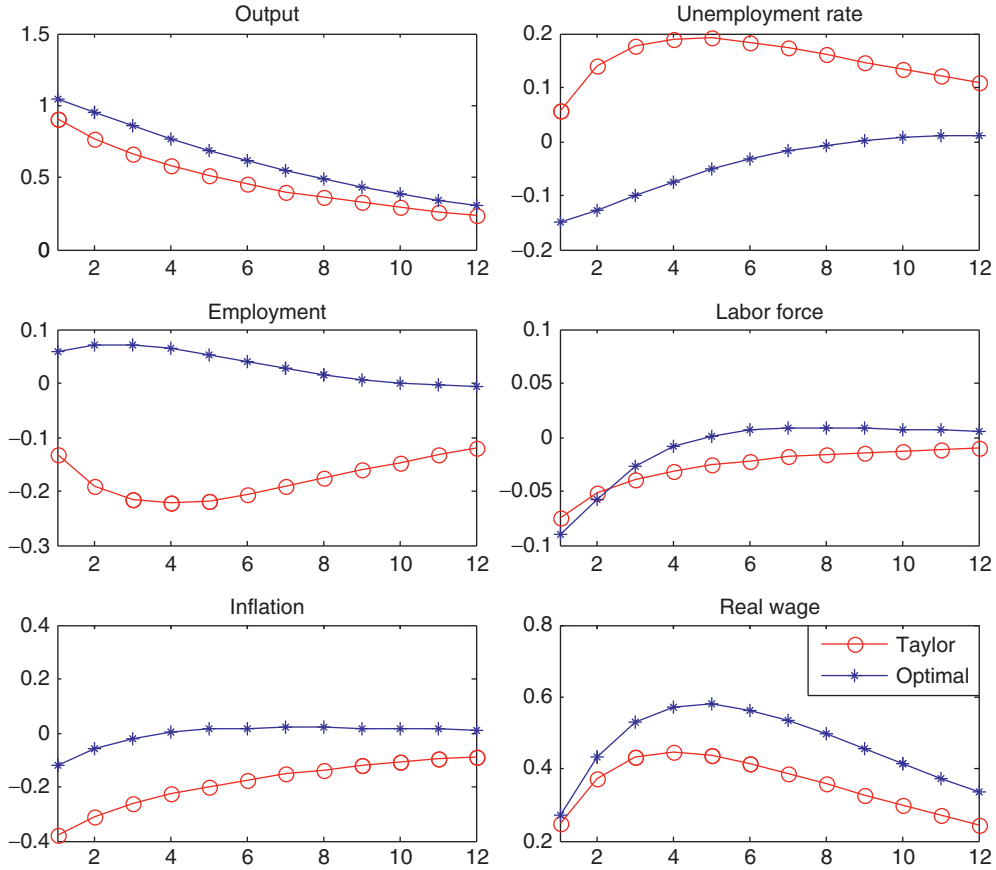


Figure 7 Monetary policy design: Optimal versus Taylor: *sticky prices and wages, technology shock.*

and also puts some weight on stabilization of unemployment. Interestingly, the optimal coefficient on price inflation is very close to 1.5, the value often assumed in standard calibrations of the Taylor rule (following Taylor, 1993). Perhaps more surprisingly, the weight on wage inflation is close to zero. This is in contrast with the findings in Erceg et al. (2000), where stabilization of wage inflation emerges as a highly desirable policy from a welfare viewpoint.⁴⁵ On the other hand, the desirability of a systematic policy response to unemployment fluctuations is in line with the findings on optimal simple rules in Blanchard and Galí (2010) and Faia (2009).

⁴⁵ The structure of the present model and the associated inefficiencies resulting from wage dispersion lead to a coefficient on wage volatility in the loss function that is about one-third the size of the coefficient on price inflation. That ranking is reversed for standard calibrations of the Erceg et al. (2000) model.

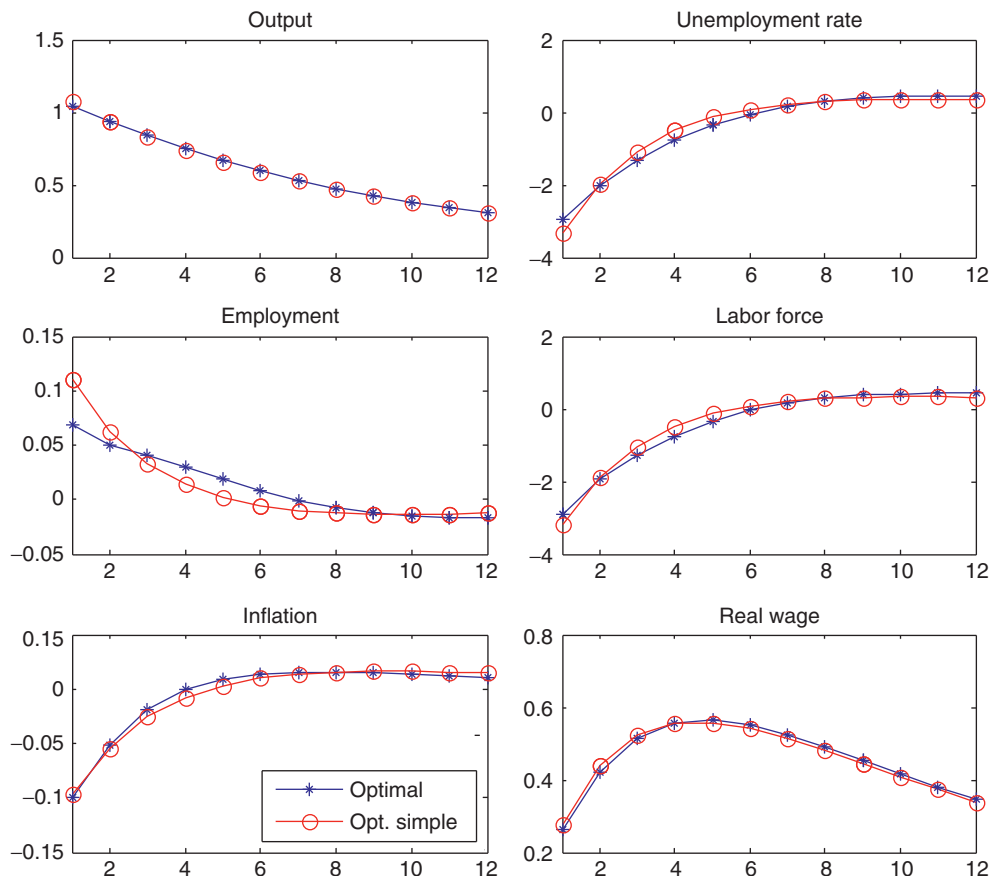


Figure 8 Monetary policy design: Optimal versus optimal simple: *sticky prices and wages, technology shock*.

Given the relatively small values of the coefficients on variables other than price inflation in the optimized interest rate rule, a rule of the form $i_t = \rho + 1.5\pi_t^p$ leads to technology shock responses (not shown) that are similar to those generated by the optimized one. That rule can be interpreted as capturing the notion of *flexible* inflation targeting, whereby central banks seek to attain a prespecified inflation target only gradually (“in the medium term,” using the language of the European Central Bank), as opposed to the *strict* inflation targeting that is optimal in environments in which price stickiness is the only nominal distortion.

The previous findings are consistent, at least in a qualitative sense, with the existing literature on optimal monetary policy in environments with labor market frictions and wage rigidities, despite the differences in modeling details. This is the case, in

particular, for Blanchard and Galí (2010; in a model with real wage rigidities) and Thomas (2008; in a model with staggered nominal wage setting like the present one).

6. POSSIBLE EXTENSIONS

As argued in the Introduction, it is not the goal of this chapter to offer an exhaustive analysis of existing models of monetary policy and unemployment. Instead, I have developed and analyzed a relatively streamlined model, but one which in my view contains the key ingredients to illustrate the consequences of the coexistence of nominal rigidities and labor market frictions. The model is, however, sufficiently flexible to be able to accommodate many extensions that can already be found in the literature. A list of some of those extensions, with a brief description of ways to introduce them, but without any further analysis, is next.

6.1 Real wage rigidities and wage indexation

As emphasized by Blanchard and Galí (2007, 2010) the presence of real wage rigidities may have implications for the optimal design of monetary policy that are likely to differ from the ones generated by a model with nominal wage rigidities only (like the one emphasized here). Among other things, in the presence of real wage rigidities, the policymaker cannot use price inflation to facilitate the adjustment of real wages. A simple way to introduce real wage rigidities would be to allow for (possibly partial) wage indexation to contemporaneous wage inflation between wage renegotiations. Formally, one can assume:

$$W_{t+k|t} = W_{t+k-1|t} (P_{t+k}/P_{t+k-1})^\zeta$$

for $k = 1, 2, 3, \dots$ and $W_{t|t} = W_t^*$, and where $W_{t+k|t}$ is the nominal wage in period $t + k$ for an employment relationship whose wage was last renegotiated in period t . Note that parameter $\zeta \in [0, 1]$ measures the degree of indexation. An alternative specification, often used in the New Keynesian literature (e.g., Smets & Wouters, 2007) and adopted by Gertler et al. (2008), assumes instead indexation to past inflation. Formally,

$$W_{t+k|t} = W_{t+k-1|t} (P_{t+k-1}/P_{t+k-2})^\zeta$$

for $k = 1, 2, 3, \dots$. In the latter case, even with full indexation, price inflation can still be used to speed up the adjustment of real wage to shocks that warrant such an adjustment, due to the lags in indexation.

6.2 Greater wage flexibility for new hires

As previously discussed, a number of authors (Carneiro, Guimaraes, & Portugal, 2008; Haefke et al., 2008; Pissarides, 2009) have argued that while the wages of incumbent workers display some clear rigidities, the latter may not have allocative consequences

(to the extent they remain within the bargaining set) since the wage that determines hiring decision is the wage of new hires, which is likely to be more flexible, according to some evidence. Even though that evidence remains controversial and has been disputed in some quarters (see references earlier in this paragraph), it may be of interest to see how such differential flexibility can be introduced in the model, and to explore its positive and normative implications. A tractable and flexible way of introducing that feature, proposed in Bodart et al. (2006), involves the assumption that new hires at a firm are paid either the average wage (with probability η) or a freely negotiated wage (with probability $1 - \eta$). Parameter η is thus an index of the degree of relative wage flexibility for new hires. That assumption would require a change in the equation describing the value of unemployment, since the probability of bargaining over wage at the time of being hired would now be $1 - \theta_w \eta$, instead of $1 - \theta_w$. One could then quantify the extent to which the responses to shocks and the optimal policy vary with η .

6.3 Smaller wealth effects

The earlier analysis relied on a specification of utility with wealth effects of labor supply that are likely to be implausibly large. That could explain the unusual unrealistic behavior of the labor force under some of the calibrations previously discussed. One way to get around that problem is to assume the following alternative specification of the utility function, originally proposed in Galí (2010):⁴⁶

$$U(C_t, L_t) \equiv \Theta_t \log C_t - \frac{\chi}{1 + \varphi} L_t^{1 + \varphi}$$

where $\Theta_t \equiv \overline{C}_t / Z_t$, $\overline{C}_t$ is aggregate consumption (taken as given by each individual household), and

$$Z_t = Z_{t-1}^{\vartheta} \overline{C}_t^{1 - \vartheta}$$

and $\vartheta \in [0, 1]$. In that case the marginal rate of substitution between consumption and market effort is given (in logs) by

$$mrs_t = z_t + \varphi l_t$$

where $z_t = (1 - \vartheta)c_t + \vartheta z_{t-1}$. Thus, changes in consumption will have an arbitrarily small effect on the short-run supply for market effort, if ϕ is close to unity. Given that the gap between z_t and c_t is stationary (even when c_t displays a linear trend or a unit root), the previous specification of utility will still be consistent with a balanced growth path.

⁴⁶ See Jaimovich and Rebelo (2009) for an alternative specification of utility in the same spirit.

6.4 Other demand shocks

The analysis of optimal monetary policy above assumes the economy faces only a technology shock (naturally, the monetary policy shock is turned off for the purposes of that exercise). How the policy implications may vary once a shock other than technology is introduced seems worthy of investigation. In particular, it may be the case that in that scenario the optimal policy will attach a greater weight to output stabilization.⁴⁷

7. CONCLUSIONS

Over the past few years a growing number of researchers have turned their attention toward the development and analysis of extensions of the New Keynesian framework that model unemployment explicitly. This chapter has described some of the essential ingredients and properties of those models, and their implications for monetary policy.

The analysis of a calibrated version of the model developed here suggests that labor market frictions are unlikely — either by themselves or through their interaction with sticky prices — to have large effects on the equilibrium response to shocks, in an economy with nominal rigidities and a monetary policy described by a simple Taylor-type rule. In that respect, perhaps the most important contribution of those frictions lies in their ability to reconcile the presence of wage rigidities with privately efficient employment relations. The presence of those nominal wage rigidities has, on the other hand, important consequences for the economy's response to shocks as well as for the optimal design of monetary policy. Thus, in the model developed earlier, the optimal policy allows for significant deviations from price stability to facilitate the adjustment of real wages to real shocks. Furthermore, the outcome of that policy can be approximated by means of a simple interest rate rule that responds to both price inflation and the unemployment rate.

APPENDIX 1

Proof of Lemma

From the definition of the price index:

$$\begin{aligned}
 1 &= \int_0^1 \left(\frac{P_t(i)}{P_t} \right)^{1-\epsilon} di \\
 &= \int_0^1 \exp\{(1-\epsilon)(p_t(i) - p_t)\} di \\
 &\simeq 1 + (1-\epsilon) \int_0^1 (p_t(i) - p_t) di + \frac{(1-\epsilon)^2}{2} \int_0^1 (p_t(i) - p_t)^2 di
 \end{aligned}$$

⁴⁷ Sveen and Weinke (2008) made a forceful case for the importance of demand shocks in accounting for labor market dynamics.

where the approximation results from a second-order Taylor expansion around the zero inflation steady state. Thus, and up to second order, we have

$$p_t \simeq E_i\{p_t(i)\} + \frac{(1-\epsilon)}{2} \int_0^1 (p_t(i) - p_t)^2 di$$

where $E_i\{p_t(i)\} \equiv \int_0^1 p_t(i) di$ is the cross-sectional mean of (log) prices. In addition,

$$\begin{aligned} \int_0^1 \left(\frac{P_t(i)}{P_t} \right)^{-\epsilon} di &= \int_0^1 \exp\{-\epsilon(p_t(i) - p_t)\} di \\ &\simeq 1 - \epsilon \int_0^1 (p_t(i) - p_t) di + \frac{\epsilon^2}{2} \int_0^1 (p_t(i) - p_t)^2 di \\ &\simeq 1 + \frac{\epsilon}{2} \int_0^1 (p_t(i) - p_t)^2 di \\ &\simeq 1 + \frac{\epsilon}{2} \text{var}_i\{p_t(i)\} \geq 1 \end{aligned}$$

where the last equality follows from the observation that, up to second order,

$$\begin{aligned} \int_0^1 (p_t(i) - p_t)^2 di &\simeq \int_0^1 (p_t(i) - E_i\{p_t(i)\})^2 di \\ &\equiv \text{var}_i\{p_t(i)\} \end{aligned}$$

Finally, using the definition of d_t^p we obtain

$$d_t^p \simeq \frac{\epsilon}{2} \text{var}_i\{p_t(i)\} \geq 0$$

On the other hand,

$$\begin{aligned} \int_0^1 \left(\frac{N_t(j)}{N_t} \right)^{1-\alpha} dj &= \int_0^1 \exp\{(1-\alpha)(n_t(j) - n_t)\} dj \\ &\simeq 1 + (1-\alpha) \int_0^1 (n_t(j) - n_t) dj + \frac{(1-\alpha)^2}{2} \int_0^1 (n_t(j) - n_t)^2 dj \\ &\simeq 1 - \frac{\alpha(1-\alpha)}{2} \int_0^1 (n_t(j) - n_t)^2 dj \leq 1 \end{aligned}$$

where the third equality follows from the fact that $\int_0^1 (n_t(j) - n_t) dj \simeq -\frac{1}{2} \int_0^1 (n_t(j) - n_t)^2 dj$ (using a second-order approximation of the identity $1 \equiv \int_0^1 \frac{N_t(j)}{N_t} dj$).

Log-linearizing the optimal hiring condition (11) around a symmetric equilibrium we have

$$n_t(j) - n_t \simeq -\frac{1 - \Phi}{\alpha} (w_t(j) - w_t)$$

thus

$$\int_0^1 \left(\frac{N_t(j)}{N_t} \right)^{1-\alpha} dj \simeq 1 - \frac{(1 - \Phi)^2 (1 - \alpha)}{2\alpha} \int_0^1 (w_t(j) - w_t)^2 dj$$

implying

$$d_t^w \equiv -\log \int_0^1 \left(\frac{N_t(j)}{N_t} \right)^{1-\alpha} \simeq \frac{(1 - \Phi)^2 (1 - \alpha)}{2\alpha} \text{var}_j \{w_t(j)\} \geq 0$$

APPENDIX 2

Linearization of participation condition

Lemma. Define $Q_t \equiv \int_0^1 \left(\frac{H_t(z)}{H_t} \right) \mathcal{S}_t^H(z) dz$. Then, around a zero inflation deterministic steady state we have

$$\hat{q}_t \simeq \hat{g}_t - \Xi \pi_t^w$$

$$\text{where } \Xi \equiv \frac{\xi(W/P)}{(1-\xi)G} \frac{\theta_w}{(1-\theta_w)(1-\beta(1-\delta)\theta_w)}$$

Proof of Lemma:

$$\begin{aligned} Q_t &\simeq \int_0^1 \mathcal{S}_t^H(z) dz \\ &= (1 - \theta_w) \sum_{q=0}^{\infty} \theta_w^q \mathcal{S}_{t|t-q}^H \\ &= (1 - \theta_w) \sum_{q=0}^{\infty} \theta_w^q (\mathcal{S}_{t|t}^H + \mathcal{S}_{t|t-q}^H - \mathcal{S}_{t|t}^H) \end{aligned}$$

where the first equality holds up to a first order approximation in a neighborhood of a symmetric steady state.

Using the Nash bargaining condition (31) we have:

$$\xi Q_t = (1 - \xi) G_t + \xi (1 - \theta_w) \sum_{q=0}^{\infty} \theta_w^q (\mathcal{S}_{t|t-q}^H - \mathcal{S}_{t|t}^H)$$

Note, however, that

$$\begin{aligned}\mathcal{S}_{i|t-q}^H - \mathcal{S}_{i|t}^H &= E_t \left\{ \sum_{k=0}^{\infty} ((1-\delta)\theta_w)^k \Lambda_{t,t+k} \left(\frac{W_{t-q}^*}{P_{t+k}} - \frac{W_t^*}{P_{t+k}} \right) \right\} \\ &= \left(\frac{W_{t-q}^* - W_t^*}{P_t} \right) E_t \left\{ \sum_{k=0}^{\infty} ((1-\delta)\theta_w)^k \Lambda_{t,t+k} \left(\frac{P_t}{P_{t+k}} \right) \right\}\end{aligned}$$

Using the law of motion for the aggregate wage,

$$\begin{aligned}(1-\theta_w) \sum_{q=0}^{\infty} \theta_w^q (\mathcal{S}_{i|t-q}^H - \mathcal{S}_{i|t}^H) &= \left(\frac{W_t - W_t^*}{P_t} \right) E_t \left\{ \sum_{k=0}^{\infty} ((1-\delta)\theta_w)^k \Lambda_{t,t+k} \left(\frac{P_t}{P_{t+k}} \right) \right\} \\ &= -\pi_t^w \left(\frac{\theta_w}{1-\theta_w} \right) \frac{W_{t-1}}{P_t} E_t \left\{ \sum_{k=0}^{\infty} ((1-\delta)\theta_w)^k \Lambda_{t,t+k} \left(\frac{P_t}{P_{t+k}} \right) \right\} \\ &\simeq -\pi_t^w \left(\frac{\theta_w}{(1-\theta_w)(1-\beta(1-\delta)\theta_w)} \right) \left(\frac{W}{P} \right)\end{aligned}$$

where the approximation holds in a neighborhood of the zero inflation steady state. It follows that

$$\xi Q_t \simeq (1-\xi) G_t - \xi \left(\frac{\theta_w}{(1-\theta_w)(1-\beta(1-\delta)\theta_w)} \right) \left(\frac{W}{P} \right) \pi_t^w$$

or, equivalently, in (log) deviations from steady state values:

$$\hat{q}_t \simeq \hat{g}_t - \Xi \pi_t^w$$

where $\Xi \equiv \frac{\xi(W/P)}{(1-\xi)G} \frac{\theta_w}{(1-\theta_w)(1-\beta(1-\delta)\theta_w)}$.

APPENDIX 3

Log-linearized equilibrium conditions

- *Technology, Resource Constraints and Miscellaneous Identities*
- Goods market clearing (44)

$$\hat{y}_t = (1-\Theta)\hat{c}_t + \Theta(\hat{g}_t + \hat{h}_t)$$

- where $\Theta \equiv \frac{\delta NG}{Y}$.
- Aggregate production function

$$\hat{y}_t = a_t + (1-\alpha)\hat{n}_t$$

- Aggregate hiring and employment

$$\delta \hat{h}_t = \hat{n}_t - (1-\delta)\hat{n}_{t-1}$$

- *Hiring cost*

$$\hat{g}_t = \gamma \hat{x}_t$$

- *Job finding rate*

$$\hat{x}_t = \hat{h}_t - \hat{u}_t^o$$

- *Effective market effort*

$$\hat{l}_t = \left(\frac{N}{L}\right) \hat{n}_t + \left(\frac{\psi U}{L}\right) \hat{u}_t$$

- *Labor force*

$$\hat{f}_t = \left(\frac{N}{F}\right) \hat{n}_t + \left(\frac{U}{F}\right) \hat{u}_t$$

- *Unemployment:*

$$\hat{u}_t = \hat{u}_t^o - \frac{x}{1-x} \hat{x}_t$$

- *Unemployment rate*

$$\hat{ur}_t = \hat{f}_t - \hat{n}_t$$

- *Decentralized Economy: Other Equilibrium Conditions*
- *Euler equation*

$$\hat{c}_t = E_t\{\hat{c}_{t+1}\} - \hat{r}_t$$

- *Fisherian equation*

$$\hat{r}_t = \hat{i}_t - E_t\{\pi_{t+1}\}$$

- *Inflation equation*

$$\pi_t = \beta E_t\{\pi_{t+1}\} - \lambda_p \hat{\mu}_t^p$$

- *Optimal hiring condition*

$$\begin{aligned} \alpha \hat{n}_t &= a_t - [(1 - \Phi) \hat{\omega}_t + \Phi \hat{b}_t] - \hat{\mu}_t^p \\ \hat{b}_t &= \frac{1}{1 - \beta(1 - \delta)} \hat{g}_t - \frac{\beta(1 - \delta)}{1 - \beta(1 - \delta)} (E_t\{\hat{g}_{t+1}\} - \hat{r}_t) \end{aligned}$$

- *Optimal participation condition (only when $\psi > 0$)*

$$\hat{c}_t + \varphi \hat{l}_t = \frac{1}{1-x} \hat{x}_t + \hat{g}_t - \Xi \pi_t^w$$

- where $\Xi \equiv \frac{\xi(W/P)}{(1-\xi)G} \frac{\theta_w}{(1-\theta_w)(1-\beta(1-\delta)\theta_w)}$ (note $\Xi = 0$ under flexible wages). When $\psi = 0$, $\hat{l}_t = \hat{n}_t$ and $\hat{f}_t = 0$ hold instead.
- *Interest rate rule*

$$\hat{i}_t = \phi_\pi \pi_t + \phi_y \hat{y}_t + v_t$$

- Wage Setting Block: Flexible Wages
- Nash wage equation

$$\hat{\omega}_t = (1 - \Upsilon)(\hat{c}_t + \varphi \hat{l}_t) + \Upsilon(-\hat{\mu}_t^p + a_t - \alpha \hat{n}_t)$$

where $\Upsilon \equiv \frac{(1-\xi)MRPN}{W/P}$

- Wage Setting Block: Sticky Wages

$$\begin{aligned}\hat{\omega}_t &= \hat{\omega}_{t-1} + \pi_t^w - \pi_t^p \\ \pi_t^w &= \beta(1 - \delta)E_t\{\pi_{t+1}^w\} - \lambda_w(\hat{\omega}_t - \hat{\omega}_t^{tar}) \\ \hat{\omega}_t^{tar} &= (1 - \Upsilon)(\hat{c}_t + \varphi \hat{l}_t) + \Upsilon(-\hat{\mu}_t^p + a_t - \alpha \hat{n}_t)\end{aligned}$$

- Social Planner's Problem: Efficiency Conditions

$$\begin{aligned}a_t - \alpha \hat{n}_t &= (1 - \Omega)(\hat{c}_t + \varphi \hat{l}_t) + \Omega \hat{b}_t \\ \hat{c}_t + \varphi \hat{l}_t &= \frac{1}{1 - x} \hat{x}_t + \hat{g}_t\end{aligned}$$

where $\Omega \equiv \frac{(1+\gamma)B}{MPN}$.

APPENDIX 4

Sketch of the derivation of loss function

Combining a second-order expansion of the utility of the representative household and the resource constraint around the constrained-efficient allocation yields

$$E_0 \sum_{t=0}^{\infty} \beta^t \tilde{U}_t \simeq -E_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{1}{1 - \Theta} (d_t^p + d_t^w) + \frac{1}{2} (1 + \varphi) \chi L^{1+\varphi} \hat{l}_t^2 \right)$$

As shown in Appendix $d_t^p \simeq \frac{\epsilon}{2} \text{var}_i(p_t(i))$, and $d_t^w \simeq \frac{(1-\Phi)^2(1-\alpha)}{2\alpha} \text{var}_j(w_t(j))$. I make use of the following property of the Calvo price and wage setting environment:

Lemma:

$$\begin{aligned}\sum_{t=0}^{\infty} \beta^t \text{var}_i\{p_t(i)\} &= \frac{\theta_p}{(1 - \theta_p)(1 - \beta\theta_p)} \sum_{t=0}^{\infty} \beta^t (\pi_t^p)^2 \\ \sum_{t=0}^{\infty} \beta^t \text{var}_j\{w_t(j)\} &= \frac{\theta_w}{(1 - \theta_w)(1 - \beta\theta_w)} \sum_{t=0}^{\infty} \beta^t (\pi_t^w)^2\end{aligned}$$

Proof: Woodford (2003, Chapter 6).

Combining the previous results and letting $\mathbb{L} \equiv -E_0 \sum_{t=0}^{\infty} \beta^t \tilde{U}_t(C/Y)$ denote the utility losses expressed as a share of steady state GDP we can write

$$\mathbb{L} \equiv \frac{1}{2} E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{\epsilon}{\lambda_p} (\pi_t^p)^2 + \frac{(1 + \Phi)^2 (1 - \alpha)}{\alpha \lambda_w^*} (\pi_t^w)^2 + (1 + \varphi) (\chi CL^{1+\varphi}/Y) \tilde{l}_t^2 \right]$$

where $\lambda_w^* \equiv (1 - \theta_w)(1 - \beta\theta_w)/\theta_w$.

Next note that, up to first order,

$$\begin{aligned} \tilde{l}_t &= \left(\frac{N}{L(1 - \alpha)} \right) \tilde{y}_t + \left(\frac{\psi U}{L} \right) \tilde{u}_t \\ &= \left(\frac{N}{L(1 - \alpha)} \right) \left(\tilde{y}_t + \frac{(1 - \alpha)\psi U}{N} \tilde{u}_t \right) \end{aligned}$$

Thus we have:

$$\mathbb{L} \equiv \frac{1}{2} E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{\epsilon}{\lambda_p} (\pi_t^p)^2 + \frac{(1 + \Phi)^2 (1 - \alpha)}{\alpha \lambda_w^*} (\pi_t^w)^2 + \frac{(1 + \varphi)(1 - \Omega)N}{(1 - \alpha)L} \left(\tilde{y}_t + \frac{(1 - \alpha)\psi U}{N} \tilde{u}_t \right)^2 \right]$$

where $1 - \Omega \equiv \frac{MRS}{MPN} = 1 - \frac{B(1+\gamma)}{MPN}$ is the steady state gap between the marginal rate of substitution and the marginal product of labor resulting from the existence of labor market frictions.

REFERENCES

- Andrés, J., Domenech, R., Ferri, J., 2006. Price rigidity and the volatility of vacancies and unemployment. Universidad de Valencia, Mimeo.
- Arseneau, D.M., Chugh, S.K., 2008. Optimal fiscal and monetary policy with costly wage bargaining. *J. Monet. Econ.* 55 (8), 1401–1414.
- Barattieri, A., Basu, S., Gottschalk, P., 2009. Some evidence on the importance of sticky wages. Boston College, Mimeo.
- Barnichon, R., 2008. Productivity, aggregate demand and unemployment fluctuations. Finance and Economics Discussion Series 2008–47. Federal Reserve Board.
- Barro, R.J., 1977. Long term contracting, sticky prices and monetary policy. *J. Monet. Econ.* 3 (3), 305–316.
- Basu, S., Fernald, J., Kimball, M., 2006. Are technology improvements contractionary? *Am. Econ. Rev.* 96 (5), 1418–1448.
- Blanchard, O.J., Galí, J., 2007. Real wage rigidities and the New Keynesian model. *J. Money Credit Bank.* 39 (1), 35–66 supplement to volume.
- Blanchard, O.J., Galí, J., 2010. Labor markets and monetary policy: A New Keynesian model with unemployment. *Am. Econ. J: Macroeconomics.* 2 (2), 1–3.
- Blanchard, O.J., Quah, D., 1989. The dynamic effects of aggregate demand and supply disturbances. *Am. Econ. Rev.* 79 (4), 655–673.

- Bodart, V., de Walque, G., Pierrard, O., Sneessens, H., Wouters, R., 2006. Nominal wage rigidities in a New Keynesian model with frictional unemployment. Mimeo. Unpublished manuscript.
- Calvo, G., 1983. Staggered prices in a utility maximizing framework. *J. Monet. Econ.* 12, 383–398.
- Carneiro, A., Guimaraes, P., Portugal, P., 2008. Real wages and the business cycle: Accounting for worker and firm heterogeneity. Mimeo. Unpublished manuscript.
- Chari, V.V., Kehoe, P.J., McGrattan, E., 2008. Are structural VARs with long-run restrictions useful in developing business cycle theory? *J. Monet. Econ.* 55 (8), 1337–1352.
- Chéron, A., Langot, F., 2000. The Phillips and Beveridge Curves Revisited. *Econ. Lett.* 69, 371–376.
- Christiano, L.J., Eichenbaum, M., Vigfusson, R., 2003. What happens after a technology shock?. NBER WP# 9819.
- Christiano, L.J., Eichenbaum, M., Evans, C.L., 2005. Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *J. Pol. Econ.* 113 (1), 1–45.
- Christiano, L.J., Trabandt, M., Walentin, K., 2010. Involuntary Unemployment and the Business Cycle. Unpublished manuscript.
- Christoffel, K., Linzert, T., 2005. The role of real wage rigidities and labor market frictions for unemployment and inflation dynamics. European Central Bank. Discussion Paper 556.
- Clarida, R., Galí, J., Gertler, M., 1999. The science of monetary policy: A New Keynesian perspective. *J. Econ. Lit.* 37, 1661–1707.
- Diamond, P.A., 1982a. Aggregate demand management in search equilibrium. *J. Polit. Econ.* 90, 881–894.
- Diamond, P.A., 1982b. Wage determination and efficiency in search equilibrium. *Rev. Econ. Stud.* 49, 217–227.
- Erceg, C.J., Henderson, D.W., Levin, A.T., 2000. Optimal monetary policy with staggered wage and price contracts. *J. Monet. Econ.* 46 (2), 281–314.
- European Central Bank, 2009. Wage dynamics in Europe: Final report of the wage dynamics network. http://www.ecb.int/home/html/researcher_wdn.en.html.
- Faia, E., 2008. Optimal monetary policy rules in a model with labor market frictions. *J. Econ. Dyn. Control* 32 (5), 1600–1621.
- Faia, E., 2009. Ramsey monetary policy with labor market frictions. *J. Monet. Econ.* 56, 570–581.
- Francis, N., Ramey, V., 2005. Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited. *J. Monet. Econ.* 52 (8), 1379–1399.
- Galí, J., 1999. Technology, employment, and the business cycle: Do technology shocks explain aggregate fluctuations? *Am. Econ. Rev.* 89 (1), 249–271.
- Galí, J., 2008. Monetary policy, inflation, and the business cycle. An introduction to the New Keynesian framework. Princeton University Press, Princeton, NJ.
- Galí, J., 2010. The return of the wage Phillips curve. Unpublished manuscript.
- Galí, J., Gertler, M., 1999. Inflation dynamics: A structural econometric analysis. *J. Monet. Econ.* 44 (2), 195–222.
- Galí, J., Gertler, M., López-Salido, D., 2001. European inflation dynamics. *Eur. Econ. Rev.* 45 (7), 1237–1270.
- Galí, J., Rabanal, P., 2004. Technology shocks and aggregate fluctuations: How well does the RBC model fit postwar U.S. data? NBER Macroeconomics Annual 2004, 225–288.
- Galí, J., van Rens, T., 2009. The vanishing procyclicality of labor productivity. Unpublished manuscript.
- Galuscak, K., Murphy, A., Nicolitsas, D., Smets, F., Strzelecki, P., Vodopivec, M., et al., 2008. The determination of wages of newly hired workers: Survey evidence on internal vs. external factors. Unpublished manuscript.
- Gertler, M., Sala, L., Trigari, A., 2008. An estimated monetary DSGE model with unemployment and staggered nominal wage setting. *J. Money Credit Bank.* 40 (8), 1713–1764.
- Gertler, M., Trigari, A., 2009. Unemployment fluctuations with staggered Nash wage bargaining. *J. Polit. Econ.* 117 (1), 38–86.
- Goodfriend, M., King, R.G., 1997. The new neoclassical synthesis and the role of monetary policy. NBER Macroeconomics Annual 231–282.
- Haefke, C., Sonntag, M., van Rens, T., 2008. Wage rigidity and job creation. Unpublished manuscript.

- Hagedorn, M., Manovskii, I., 2008. The cyclical behavior of equilibrium unemployment and vacancies revisited. *Am. Econ. Rev.* 98 (4), 1692–1706.
- Hall, R., 2005. Employment fluctuations with equilibrium wage stickiness. *Am. Econ. Rev.* 95 (1), 50–64.
- Jaimovich, N., Rebelo, S., 2009. Can news about the future drive the business cycle? *Am. Econ. Rev.* 99 (4), 1097–1118.
- King, R.G., Wolman, A.L., 1996. Inflation Targeting in a St. Louis Model of the 21st Century, Federal Reserve Bank of St. Louis Review, 78 (3).
- Krause, M., López-Salido, D., Lubik, T.A., 2008. Inflation dynamics with search frictions: A structural econometric analysis. *J. Monet. Econ.* 55 (5), 892–916.
- Krueger, A.B., Mueller, A., 2008. The lot of the unemployed: A time use perspective. IZA Discussion Paper no. 3490.
- Kuester, K., 2007. Real price and wage rigidities in a model with matching frictions. European Central Bank. Working Paper Series no. 720.
- Merz, M., 1995. Search in the labor market and the real business cycle. *J. Monet. Econ.* 36, 269–300.
- Mortensen, D.T., 1982a. The matching process as a noncooperative/bargaining game. In: McCall, J. (Ed.), *The economics of information and uncertainty*. University of Chicago Press, Chicago, pp. 233–254.
- Mortensen, D.T., 1982b. Property rights and efficiency in mating, racing and related games. *Am. Econ. Rev.* 72, 968–979.
- Nakamura, E., Steinsson, J., 2008. Five facts about prices: A reevaluation of menu cost models. *Q. J. Econ.* 123 (4), 1415–1464.
- Pissarides, C., 1984. Search intensity, job advertising and efficiency. *J. Labor Econ.* 2, 128–143.
- Pissarides, C., 2000. *Equilibrium unemployment theory*. MIT Press, Cambridge, MA.
- Pissarides, C., 2009. The unemployment volatility puzzle: is wage stickiness the answer? *Econometrica* 77 (5), 1339–1369.
- Rotemberg, J., Woodford, M., 1999. Interest rate rules in an estimated sticky price model. In: Taylor, J.B. (Ed.), *Monetary policy rules*. University of Chicago Press, Chicago.
- Sbordone, A., 2002. Prices and unit labor costs: Testing models of pricing behavior. *J. Monet. Econ.* 45 (2), 265–292.
- Smets, F., Wouters, R., 2003. An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area. *J. Europ. Eco. Assoc.* 1 (5), 1123–1175.
- Smets, F., Wouters, R., 2007. Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *Am. Econ. Rev.* 97 (3), 586–606.
- Shimer, R., 2005. The cyclical behavior of equilibrium unemployment and vacancies. *Am. Econ. Rev.* 95 (1), 25–49.
- Shimer, R., 2010. *Labor markets and business cycles*. Princeton University Press, Princeton, NJ. in press.
- Silva, J., Toledo, M., 2009. Labor turnover costs and the cyclical behavior of vacancies and unemployment. *Macroecon. Dyn.* 13 (Suppl. 1), 76–96.
- Sveen, T., Weinke, L., 2008. New Keynesian perspectives on labor market dynamics. *J. Monet. Econ.* 55 (5), 921–930.
- Taylor, J.B., 1993. Discretion versus policy rules in practice. *Carnegie-Rochester Series on Public Policy* 39, 195–214.
- Taylor, J.B., 1999a. Staggered price and wage setting in macroeconomics. In: Taylor, J.B., Woodford, M. (Eds.), *Handbook of macroeconomics*. Elsevier, New York, pp. 1341–1397 (Chapter 15).
- Taylor, J.B., 1999b. An historical analysis of monetary policy rules. In: Taylor, J.B. (Ed.), *Monetary policy rules*. University of Chicago Press, Chicago.
- Thomas, C., 2008a. Search and matching frictions and optimal monetary policy. *J. Monet. Econ.* 55 (5), 936–956.
- Thomas, C., 2008b. Search frictions, real rigidities and inflation dynamics. Banco de España. Working paper 2008-06.
- Trigari, A., 2006. The role of search frictions and bargaining in inflation dynamics Unpublished manuscript, Bocconi University.

- Trigari, A., 2009. Equilibrium unemployment, job flows, and inflation dynamics. *J. Money Credit Bank.* 41 (1), 1–33.
- Walsh, C., 2003a. *Monetary theory and policy*. MIT Press, Cambridge, MA.
- Walsh, C., 2003b. Labor market search and monetary shocks. In: Altug, S., Chadha, J., Nolan, C. (Eds.), *Elements of dynamic macroeconomic analysis*. Cambridge University Press, Cambridge, UK, pp. 451–486.
- Walsh, C., 2005. Labor market search, sticky prices, and interest rate rules. *Rev. Econ. Dyn.* 8, 829–849.
- Woodford, M., 2003. *Interest and prices: Foundations of a theory of monetary policy*. Princeton University Press, Princeton, NJ.
- Yun, T., 1996. Nominal price rigidity, money supply endogeneity, and business cycles. *J. Monet. Econ.* 37, 345–370.

Financial Intermediation and Credit Policy in Business Cycle Analysis[☆]

Mark Gertler and Nobuhiro Kiyotaki

NYU and Princeton

Contents

1. Introduction	548
2. A Canonical Model of Financial Intermediation and Business Fluctuations	551
2.1 Physical setup	552
2.2 Households	554
2.3 Banks	555
2.3.1 Case 1: Frictionless wholesale financial market ($\omega = 1$)	559
2.3.2 Case 2: Symmetric frictions in wholesale and retail financial markets ($\omega = 0$)	561
2.4 Evolution of bank net worth	563
2.5 Nonfinancial firms	564
2.5.1 Goods producer	564
2.5.2 Capital goods producers	564
2.6 Equilibrium	565
3. Credit Policies	566
3.1 Lending facilities (direct lending)	567
3.2 Liquidity facilities (discount window lending)	569
3.3 Equity injections	571
3.4 Government expenditures and budget constraint	574
4. Crisis Simulations and Policy Experiments	574
4.1 Calibration	575
4.2 Crisis experiment	576
4.2.1 No policy response	576
4.2.2 Credit policy response	579
5. Issues and Extensions	581
5.1 Tightening margins	582
5.2 Regulatory arbitrage and securitized lending	584
5.3 Outside equity, externalities, and moral hazard	586
6. Concluding Remarks	589
References	597

[☆] Thanks to Michael Woodford, David Andolfatto, Larry Christiano, Harris Dellas, Ian Dew-Becker, Giovanni Di Bartolomeo, Chris Erceg, Simon Gilchrist, Arvind Krishnamurthy, Ramon Marimon and Shinichi Nishiyama for helpful comments. Thanks also to Albert Queralto Olive for excellent research assistance.

Abstract

We develop a canonical framework to think about credit market frictions and aggregate economic activity in the context of the current crisis. We use the framework to address two issues in particular: first, how disruptions in financial intermediation can induce a crisis that affects real activity; and second, how various credit market interventions by the central bank and the Treasury of the type we have seen recently, might work to mitigate the crisis. We make use of earlier literature to develop our framework and characterize how very recent literature is incorporating insights from the crisis.

JEL classification: E30, E44, E50.

Keywords

Asset Prices
Credit Policy
Financial Intermediation
Moral Hazard
Net Worth
Spreads

1. INTRODUCTION

To motivate interest in a paper on financial factors in business fluctuations it used to be necessary to appeal either to the Great Depression or to the experiences of many emerging market economies. This is no longer necessary. Over the past few years the United States and much of the industrialized world have experienced the worst post-war financial crisis, and the global recession that has followed also appears to have been the most severe of this era. But there is evidence that the financial sector has stabilized and the real economy has stopped contracting and output growth has resumed. The path to full recovery, however, remains highly uncertain.

The timing of recent events poses a challenge for writing a *Handbook* chapter on credit market frictions and aggregate economic activity. It is true that over the last several decades there has been a robust literature in this area. Bernanke, Gertler, and Gilchrist (BGG; 1999) surveyed much of the earlier work a decade ago in the *Handbook of Macroeconomics*. Since the time of that survey, the literature has continued to grow. While much of this work is relevant to the current situation, it obviously did not anticipate all the key empirical phenomena that have played out during the current crisis. A new literature that builds on the earlier work is rapidly emerging to address these issues. Most of these papers are in preliminary working paper form.

Our plan in this chapter is to look both forward and backward. We look forward in the sense that we offer a canonical framework to think about credit market frictions and aggregate economic activity in the context of the current crisis. The framework is not meant as a comprehensive description of recent events but rather as a first pass

at characterizing some of the key aspects and at laying out issues for future research. We looked backward by making use of earlier literature to develop the particular framework we offer. In doing so, we address how this literature may be relevant to the new issues that have arisen. We also, as best we can, characterize how very recent literature is incorporating insights from the crisis.

From our vantage, there are two broad aspects of the crisis that have not been fully captured in work on financial factors in business cycles. First, by all accounts, the current crisis has featured a significant disruption of financial intermediation.¹ Much of the earlier macroeconomics literature with financial frictions emphasized credit market constraints on nonfinancial borrowers and treated intermediaries largely as a veil (see, e.g., BGG). Second, to combat the crisis, both the monetary and fiscal authorities in many countries, including the United States, have employed various unconventional policy measures that involve some form of direct lending in credit markets.

From the standpoint of the Federal Reserve, these “credit” policies represent a significant break from tradition. In the post-war era, the Fed scrupulously avoided any exposure to private sector credit risk. However, in the current crisis the central bank has acted to offset the disruption of intermediation by making imperfectly secured loans to financial institutions and by lending directly to high-grade, nonfinancial borrowers. In addition, the fiscal authority acting in conjunction with the central bank injected equity into the major banks with the objective of improving credit flows. Although the issue is not without considerable controversy, many observers argue that these interventions helped stabilized financial markets and, consequently, helped limit the decline of real activity. Since these policies are relatively new, much of the existing literature is silent about them.

With this background in mind, we begin in the next section by developing a baseline model that incorporates financial intermediation into an otherwise frictionless business cycle framework. Our goal is twofold: first to illustrate how disruptions in financial intermediation can induce a crisis that affects real activity; and second, to illustrate how various credit market interventions by the central bank and the Treasury of the type we have seen recently might work to mitigate the crisis.

As in Bernanke and Gertler (1989), Kiyotaki and Moore (1997) and others, we endogenize financial market frictions by introducing an agency problem between borrowers and lenders.² The agency problem works to introduce a wedge between the cost of external finance and the opportunity cost of internal finance, which adds to the overall cost of credit that a borrower faces. The size of the external finance premium, further, depends on the condition of borrower balance sheets. Roughly

¹ For a description of the disruption of financial intermediation during the current recession, see Brunnermeier (2009), Gorton (2010), and Bernanke (2009). For a more general description of financial crisis over the last several hundred years, see Reinhart and Rogoff (2009).

² A partial of other macro models with financial frictions in this vein includes Williamson (1987), Kehoe and Levine (1993), Holmstrom and Tirole (1998), Carlstrom and Fuerst (1997), Caballero and Krishnamurthy (2001), Krishnamurthy (2003), Christiano et al. (2005), Lorenzoni (2008), Fostel and Geanakoplos (2008), and Brunnermeier and Sannikov (2009).

speaking, as a borrower's percentage stake in the outcome of an investment project increases, the incentive to deviate from the interests of lenders' declines. The external finance premium then declines as a result.

In general equilibrium, a "financial accelerator" emerges. As balance sheets strengthen with improved economics conditions, the external finance problem declines, which works to enhance borrower spending, thus enhancing the boom. Along the way, there is mutual feedback between the financial and real sectors. In this framework, a crisis is a situation where balance sheets of borrowers deteriorate sharply, possibly associated with a sharp deterioration in asset prices, causing the external finance premium to jump. The impact of the financial distress on the cost of credit then depresses real activity.³

Bernanke and Gertler (1989), Kiyotaki and Moore (1997) and others focused on credit constraints faced by nonfinancial borrowers.⁴ As we noted earlier, however, the evidence suggests that disruption of financial intermediation is a key feature of both recent and historical crises. Thus we focus our attention on financial intermediation.

We begin by supposing that financial intermediaries have skills in evaluating and monitoring borrowers, which makes it efficient for credit to flow from lenders to non-financial borrowers through intermediaries. In particular, we assume that households deposit funds in financial intermediaries that in turn lend funds to nonfinancial firms. We then introduce an agency problem that potentially constrains the ability of intermediaries to obtain funds from depositors. When the constraint is binding (or there is some chance it may bind), the intermediary's balance sheet limits its ability to obtain deposits. In this instance, the constraint effectively introduces a wedge between the loan and deposit rates. During a crisis, this spread widens substantially, which in turn sharply raises the cost of credit that nonfinancial borrowers face.

As recent events suggest, however, in a crisis, financial institutions face difficulty not only in obtaining depositor funds in retail financial markets but also in obtaining funds from one another in wholesale ("interbank") markets. Indeed, the first signals of a crisis are often strains in the interbank market. We capture this phenomenon by subjecting financial institutions to idiosyncratic "liquidity" shocks, which have the effect of creating surplus and deficits of funds across financial institutions. If the interbank market works perfectly, then funds flow smoothly from institutions with surplus funds to those in need. In this case, loan rates are thus equalized across different financial institutions. Aggregate behavior in this instance resembles the case of homogeneous intermediaries.

However, to the extent that the agency problem that limits an intermediary's ability to obtain funds from depositors also limits its ability to obtain funds from other financial

³ Most of the models focus on the impact of borrower constraints on producer durable spending. See Monacelli (2009) and Iacoviello (2005) for extensions to consumer durables and housing. Jermann and Quadrini (2009), among others, focused on borrowing constraints on employment.

⁴ An exception is Holmstrom and Tirole (1997). More recent work includes He and Krishnamurthy (2009) and Angeloni and Faia (2009).

institutions and to the extent that nonfinancial firms can obtain funds only from a limited set of financial intermediaries, disruptions of interbank markets are possible that can affect real activity. In this instance, intermediaries with deficit funds offer higher loan rates to nonfinancial firms than intermediaries with surplus funds. In a crisis this gap widens. Financial markets effectively become segmented and sclerotic. As we show, the inefficient allocation of funds across intermediaries can further depress aggregate activity.

In Section 3 we incorporate credit policies within the formal framework. In practice the central bank employed three broad types of policies. The first, which was introduced early in the crisis, was to permit discount window lending to banks secured by private credit. The second, introduced in the wake of the Lehman default was to lend directly in relatively high-grade credit markets, including markets in commercial paper, agency debt, and mortgage-backed securities. The third (and most controversial) involved direct assistance to large financial institutions, including the equity injections and debt guarantees under the Troubled Assets Relief Program (TARP) as well as the emergency loans to JP Morgan Chase (who took over Bear Stearns) and AIG.

We stress that within our framework, the net benefits from these various credit market interventions are increasing in the severity of the crisis. This helps account for why it makes sense to employ them only in crisis situations.

In Section 4, we use the model to simulate numerically a crisis that has some key features of the current crisis. Absent credit market frictions, the disturbance initiating the crisis induces only a mild recession. With credit frictions (especially those in interbank market), however, an endogenous disruption of financial intermediation works to magnify the downturn. We then explore how various credit policies can help mitigate the situation.

Our baseline model is quite parsimonious and meant mainly to exposit the key issues. In Section 5, we discuss a number of questions and possible extensions. In some cases, we discuss a relevant literature, stressing the implications of this literature for going forward.

2. A CANONICAL MODEL OF FINANCIAL INTERMEDIATION AND BUSINESS FLUCTUATIONS

Overall, the specific business cycle model is a hybrid of Gertler and Karadi's (2009) framework that allows for financial intermediation and Kiyotaki and Moore's (2008) framework that allows for liquidity risk. We keep the core macro model simple in order to clearly see the role of intermediation and liquidity. On the other hand, we also allow for some features prevalent in conventional quantitative macro models (such as Christiano, Eichenbaum, & Evans, 2005; Smets & Wouters, 2007) to get a rough sense of the importance of the factors we introduce.⁵

⁵ Some recent monetary DSGE models that incorporate financial factors include Christiano et al. (2003, 2010) and Gilchrist, Yankov, and Zakrešjek (2009).

For simplicity we restrict attention to a purely real model and only credit policies, as opposed to conventional monetary models. Extending the model to allow for nominal rigidities is straightforward (see, e.g., Gertler & Karadi, 2009), and permits studying conventional monetary policy along with unconventional policies. However, because much of the insight into how credit market frictions may affect real activity and how various credit policies may work can be obtained from studying a purely real model, we abstract from nominal frictions.⁶

2.1 Physical setup

Before describing our economy with financial frictions, we present the physical environment.

There are a continuum of firms of mass unity located on a continuum of islands. Each firm produces output using an identical constant returns to scale Cobb–Douglas production function with capital and labor as inputs. Capital is not mobile, but labor is perfectly mobile across firms and islands. Because labor is perfectly mobile, we can express aggregate output Y_t as a function of aggregate capital K_t and aggregate labor hours L_t as:

$$Y_t = A_t K_t^\alpha L_t^{1-\alpha}, \quad 0 < \alpha < 1, \quad (1)$$

where A_t is aggregate productivity which follows a Markov process.

Each period investment opportunity arrives randomly to a fraction π^i of islands. On a fraction $\pi^n = 1 - \pi^i$ of islands, there are no investment opportunities. Only firms on islands with investment opportunities can acquire new capital. The arrival of investment opportunities is i.i.d. across time and across islands. The structure of this idiosyncratic risk provides a simple way to introduce liquidity needs by firms, following Kiyotaki and Moore (2008). Let I_t denote aggregate investment, δ the rate of physical depreciation and ψ_{t+1} a shock to the quality of capital. Then the law of motion for capital is given by:

$$\begin{aligned} K_{t+1} &= \psi_{t+1} [I_t + \pi^i (1 - \delta) K_t] + \psi_{t+1} \pi^n (1 - \delta) K_t \\ &= \psi_{t+1} [I_t + (1 - \delta) K_t]. \end{aligned} \quad (2)$$

The first term of the right reflects capital accumulated by firms on investing islands and the second is capital that remains on noninvesting islands, after depreciation. Summing across islands yields a conventional aggregate relation for the evolution of capital, except for the presence of the disturbance ψ_{t+1} , which we refer to as a capital quality shock. Following the finance literature (e.g., Merton, 1973), we introduce the capital quality shock as a simple way to introduce an exogenous source of variation in the

⁶ There are several insights that monetary models add. First, if the zero lower bound on the nominal interest is binding, the financial market disruptions will have a larger effect than otherwise. This is because the central bank is not free to further reduce the nominal rate to offset the crisis. Second, to the extent there are nominal price and/or wage rigidities that induce countercyclical markups, the effect of the credit market disruption and aggregate activity is amplified. See, for example, Gertler and Karadi (2009) and Del Negro et al. (2010) for an illustration of both of these points.

value of capital. As will become clear later, the market price of capital will be endogenous within our framework. In this regard, the capital quality shock will serve as an exogenous trigger of asset price dynamics. The random variable ψ_{t+1} is best thought of as capturing some form of economic obsolescence, as opposed to physical depreciation.⁷ We assume the capital quality shock ψ_{t+1} also follows a Markov process.⁸

Firms on investing islands acquire capital from capital goods producers who operate in a national market. There are convex adjustment costs in the gross rate of change in investment for capital goods producers. Aggregate output is divided between household consumption C_t , investment expenditures, and government consumption G_t ,

$$Y_t = C_t + \left[1 + f\left(\frac{I_t}{I_{t-1}}\right) \right] I_t + G_t \quad (3)$$

where $f\left(\frac{I_t}{I_{t-1}}\right)I_t$ reflects physical adjustment costs, with $f(1) = f'(1) = 0$ and $f''(I_t/I_{t-1}) > 0$. Thus the aggregate production function of capital goods producers is decreasing returns to scale in the short run and is constant returns to scale in the long run.

Next we turn to preferences:

$$E_t \sum_{i=0}^{\infty} \beta^i \left[\ln C_{t+i} - \gamma C_{t+i-1} \right] - \frac{\chi}{1+\varepsilon} L_{t+i}^{1+\varepsilon} \quad (4)$$

where E_t is the expectation operator conditional on date t information and $\gamma \in (0,1)$. We abstract from many frictions in the conventional DSGE framework (e.g., nominal price and wage rigidities, variable capital utilization, etc.). However, we allow both habit formation of consumption and adjustment costs of investment because, as the DSGE literature has found, these features are helpful for reasonable quantitative performance and because they can be kept in the model at minimal cost of additional complexity.

If there were no financial frictions, the competitive equilibrium would correspond to a solution of the planner's problem that involves choosing aggregate quantities $(Y_t, L_t, C_t, I_t, K_{t+1})$ as a function of the aggregate state $(C_{t-1}, I_{t-1}, K_t, A_t, \psi_t)$ to maximize the expected discounted utility of the representative household subject to the resource constraints. This frictionless economy (a standard real business cycle model) will serve as a benchmark to which we may compare the implications of the financial frictions.

⁷ One way to motivate this disturbance is to assume that final output is a CES composite of a continuum of intermediate goods that are in turn produced by employing capital and labor in a Cobb-Douglas production technology. Suppose that, once capital is installed, capital is good-specific and that each period a random fraction of goods becomes obsolete and is replaced by new goods. The capital used to produce the obsolete goods is now worthless and the capital for the new goods is not fully on line. The aggregate capital stock will then evolve according to Eq. (2).

⁸ Other recent papers that make use of this kind of disturbance include Gertler and Karadi (2009), Brunnermeier and Sannikov (2009), and Gourio (2009).

In the following sections we will introduce banks that intermediate funds between households and nonfinancial firms in a retail financial market. In addition, we will allow for a wholesale interbank market, where banks with surplus funds on noninvestment islands lend to banks in need of funds on investing islands. We will also introduce financial frictions that may impede credit flows in both the retail and wholesale financial markets and then study the consequences for real activity.

2.2 Households

In our economy with credit frictions, households lend to nonfinancial firms via financial intermediaries. Following Gertler and Karadi (2009), we formulate the household sector in a way that permits maintaining the tractability of the representative agent approach.

In particular, there is a representative household with a continuum of members of measure unity. Within the household there are $1 - f$ “workers” and f “bankers.” Workers supply labor and return their wages to the household. Each banker manages a financial intermediary (which we will call a “bank”) and transfers non-negative dividends back to the household subject to its flow of fund constraint. Within the family there is perfect consumption insurance.

Households do not hold capital directly. Rather, they deposit funds in banks. (It may be best to think of them as depositing funds in banks other than the ones they own). In our model, bank deposits are riskless one period securities. Households may also hold riskless one period government debt, which is a perfect substitute for bank deposits.

Let W_t denote the wage rate, T_t lump-sum taxes, R_t the gross return on riskless debt from $t - 1$ to t , D_{ht} the quantity of riskless debt held, and Π_t net distributions from ownership of both banks and nonfinancial firms. Then the household chooses consumption, labor supply, and riskless debt (C_t , L_t , D_{ht+1}) to maximize expected discounted utility (Eq. 4) subject to the flow of funds constraint,

$$C_t = W_t L_t + \Pi_t - T_t + R_t D_{ht} - D_{ht+1}. \quad (5)$$

Let u_{Ct} denote the marginal utility of consumption and $\Lambda_{t,t+1}$ the household’s stochastic discount factor. Then the household’s first-order conditions for labor supply and consumption/saving are given by

$$E_t u_{Ct} W_t = \chi L_t^\varepsilon, \quad (6)$$

$$E_t \Lambda_{t,t+1} R_{t+1} = 1, \quad (7)$$

with

$$u_{Ct} \equiv (C_t - \gamma C_{t-1})^{-1} - \beta \gamma (C_{t+1} - \gamma C_t)^{-1} \text{ and} \\ \Lambda_{t,t+1} \equiv \beta \frac{u_{Ct+1}}{u_{Ct}}.$$

Because banks may be financially constrained, bankers will retain earnings to accumulate assets. Absent some motive for paying dividends, they may find it optimal to accumulate to the point where the financial constraint they face is no longer binding. To limit bankers' ability to save to overcome financial constraints, we allow for turnover between bankers and workers. In particular, we assume that with i.i.d. probability $1 - \sigma$, a banker exits next period, (which gives an average survival time $= \frac{1}{1-\sigma}$). Upon exiting, a banker transfers retained earnings to the household and becomes a worker. Note that the expected survival time may be quite long (in our baseline calibration it is ten years.) It is critical, however, that the expected horizon is finite, in order to motivate payouts while the financial constraints are still binding.

Each period, $(1 - \sigma)f$ workers randomly become bankers, keeping the number in each occupation constant. Finally, because in equilibrium bankers will not be able to operate without any financial resources, each new banker receives a "startup" transfer from the family as a small constant fraction of the total assets of entrepreneurs. Accordingly, Π_t is net funds transferred to the household; that is, funds transferred from exiting bankers minus the funds transferred to new bankers (aside from small profits of capital producers).

An alternative to our approach of having a consolidated family of workers and bankers would be to have the two groups as distinct sets of agents, without any consumption insurance between the two groups. It is unlikely, however, that the key results of our paper would change qualitatively. By sticking with complete consumption insurance, we are able to have lending and borrowing in equilibrium and still maintain tractability of the representative household approach.

2.3 Banks

To finance lending in each period, banks raise funds in a national financial market. Within the national financial market, there is a retail market (where banks obtain deposits from households) and a wholesale market (where banks borrow and lend amongst one and another).

At the beginning of the period each bank raises deposits d_t from households in the retail financial market at the deposit rate R_{t+1} . After the retail financial market closes, investment opportunities for nonfinancial firms arrive randomly to different islands. Banks can only make loans to nonfinancial firms located on the same island. As we stated earlier, for a fraction π^i of locations, new investment opportunities are available to finance as well as existing projects. Conversely, for a fraction $\pi^n = 1 - \pi^i$, no new investments are available to finance, only existing ones. On the interbank market, banks on islands with new lending opportunities will borrow funds from those on islands with no new project arrivals.⁹

⁹ Our model is thus one where liquidity problems emerge in part due to limited market participation, in the spirit of Allen and Gale (1994, 2007) and others. This is because within our framework (i) only banks of the same island can make loans to nonfinancial firms and (ii) banks on investing islands cannot raise additional funds in the retail financial market after they learn their customers have investment opportunities.

Financial frictions affect real activity in our framework via the impact on funds available to banks. For simplicity, however, there is no friction in transferring funds between a bank and nonfinancial firms in the same island. In particular, we suppose that the bank is efficient at evaluating and monitoring nonfinancial firms of the same island, and also at enforcing contractual obligations with these borrowers. We assume the costs to a bank of performing these activities are negligible. Accordingly, given its supply of available funds, a bank can lend frictionlessly to nonfinancial firms of the same island against their future profits. In this regard, firms are able to offer banks perfectly state-contingent debt. It is simplest to think of the bank's claim on nonfinancial firms as equity.

After learning about its lending opportunities, a bank decides the volume of loans s_t^h to make to nonfinancial firms and the volume of interbank borrowing b_t^h where the superscript $h = i, n$ denotes the island type (i for investing and n for noninvesting) on which the bank is located during the period. Let Q_t^h be the price of a loan (or "asset"); that is, the market price of the bank's claim on the future returns from one unit of present capital of nonfinancial firm at the end of period. We index the asset price by h because, owing to temporal market segmentation, Q_t^h may depend on the volume of opportunities that the bank faces.

For an individual bank, the flow-of-funds constraint implies the value of loans funded within a given period, $Q_t^h s_t^h$, must equal the sum of the bank net worth n_t^h , its borrowings on the interbank market b_t^h and deposits d_t :

$$Q_t^h s_t^h = n_t^h + b_t^h + d_t. \quad (8)$$

Note that d_t does not depend upon the volume of the lending opportunities, which is not realized at the time of obtaining deposits.

Let R_{bt} be the interbank interest rate from periods $t - 1$ to period t . Then net worth at t is the gross payoff from assets funded at $t - 1$, net borrowing costs, as follows:

$$n_t^h = [Z_t + (1 - \delta)Q_t^h]\psi_t s_{t-1} - R_{bt} b_{t-1} - R_t d_{t-1}, \quad (9)$$

where Z_t is the dividend payment at t on the loans the bank funds at $t - 1$. (Recall that ψ_t is an exogenous aggregate shock to the quality of capital). Observe that the gross payoff from assets depends on the location specific asset price Q_t^h , which is the reason n_t^h depends on the realization of the location specific shock at t .

Given that the bank pays dividends only when it exits (which occurs with a constant probability), the objective of the bank at the *end* of period t is the expected present value of future dividends, as follows

$$V_t = E_t \sum_{i=1}^{\infty} (1 - \sigma) \sigma^{i-1} \Lambda_{t,t+i} n_{t+i}^h, \quad (10)$$

where $\Lambda_{t,t+i}$ is the stochastic discount factor, which is equal to the marginal rate of substitution between consumption of date $t + i$ and date t of the representative household.

To maintain tractability, we make assumptions to ensure that we do not have to keep track of the distribution of net worth across islands.

In particular, we allow for arbitrage at the beginning of each period (before investment opportunities arrive) to ensure that ex ante expected rates of return to intermediation are equal across islands. In particular, we suppose that a fraction of banks on islands where expected returns are low can move to islands where they are high. Before they move, they sell their existing loans to nonfinancial firms to the other banks that remain on the island in exchange for interbank loans that the remaining banks have been holding in their portfolios. These transactions keep each existing loan to nonfinancial firms on the island it was initiated. At the same time, they permit arbitrage to equalize returns across markets ex ante.

As will become clear later, ex ante expected returns being equalized across islands requires that the ratio of total intermediary net worth to total capital on each island is the same at the beginning of each period.¹⁰ Thus, given this arbitrage activity and given that the liquidity shock is i.i.d., we do not have to keep track of the beginning of period distribution of net worth across islands.

To motivate an endogenous constraint on the bank's ability to obtain funds in either the retail or wholesale financial markets, we introduce the following simple agency problem: We assume that after a bank obtains funds, the banker managing the bank may transfer a fraction θ of "divertable" assets to his or her family. Divertable assets consist of total gross assets $Q_t^h S_t^h$ net a fraction ω of interbank borrowing b_t^h . If a bank diverts assets for its personal gain, it defaults on its debt and is shut down. The creditors may re-claim the remaining fraction $1 - \theta$ of funds. Because its creditors recognize the bank's incentive to divert funds, they will restrict the amount they lend. In this way a borrowing constraint may arise.

We allow for the possibility that a bank may be constrained not only in obtaining funds from depositors but also in obtaining funds from other banks, although we permit the tightness of the constraint faced in each market to differ. In particular, the parameter ω indexes (inversely) the relative degree of friction in the interbank market.

With $\omega = 1$, banks cannot divert assets financed by borrowing from other banks: lending banks are able to perfectly recover the assets that underlie the loans they make. In this case, the interbank market operates frictionlessly, and banks are not constrained in borrowing from one another. They may only be constrained in obtaining funds from depositors.

¹⁰ In turn, this requires a movement of net worth from low return to high return islands that is equal in total to the quantity of interbank loans issued in the previous period. The asset exchange between moving and staying banks described in the text accomplishes this arbitrage.

In contrast, with $\omega = 0$, lending banks are no more efficient than depositors in recovering assets from borrowing banks. In this case, the friction that constrains a bank's ability to obtain funds on the interbank market is the same as for the retail financial market. In general, we can allow parameter ω to differ for borrowing versus lending banks. However, maintaining symmetry simplifies the analysis without affecting the main results.

We assume that the banker's decision over whether to divert funds must be made at the end of the period after the realization of the idiosyncratic uncertainty that determines its type, but before the realization of aggregate uncertainty in the following period. Here the idea is that if the banker is going to divert funds, it takes time to position assets and this must be done between the periods (e.g., during the night). Let $V_t(s_t^h, b_t^h, d_t)$ be the maximized value of V_t , given an asset and liability configuration (s_t^h, b_t^h, d_t) at the end of period t . Then, to ensure the bank does not divert funds, the following incentive constraint must hold for each bank type:

$$V_t(s_t^h, b_t^h, d_t) \geq \theta(Q_t^h s_t^h - \omega b_t^h). \quad (11)$$

In general the value of the bank at the end of period $t - 1$ satisfies the Bellman equation

$$\begin{aligned} & V_{t-1}(s_{t-1}, b_{t-1}, d_{t-1}) \\ &= E_{t-1} \Lambda_{t-1,t} \sum_{h=i,n} \pi^h \left\{ (1 - \sigma) n_t^h + \sigma \underset{d_t}{\text{Max}} [\underset{s_t^h, b_t^h}{\text{Max}} V_t(s_t^h, b_t^h, d_t)] \right\}. \end{aligned} \quad (12)$$

Note that the loans and interbank borrowing are chosen after a shock to the loan opportunity is realized while deposits are chosen before.

To solve the decision problem, we first guess that the value function is linear:

$$V_t(s_t^h, b_t^h, d_t) = \mathcal{V}_{st} s_t^h - \mathcal{V}_{bt} b_t^h - \mathcal{V}_t d_t \quad (13)$$

where $\mathcal{V}_{st}$, $\mathcal{V}_{bt}$, and $\mathcal{V}_t$ are time-varying parameters, and verify this guess later. Note that $\mathcal{V}_{st}$ is the marginal value of assets at the end of period t , $\mathcal{V}_{bt}$ is the marginal cost of interbank debt, and $\mathcal{V}_t$ is the marginal cost of deposits.¹¹

Let λ_t^h be the Lagrangian multiplier for the incentive constraint (11) faced by bank of type h and $\bar{\lambda}_t \equiv \sum_{h=i,n} \pi^h \lambda_t^h$ be the average of this multiplier across states. Then given

the conjectured form of the value function, we may express the first order conditions for d_t , s_t^h , and λ_t^h , as:

$$(\mathcal{V}_{bt} - \mathcal{V}_t)(1 + \bar{\lambda}_t) = \theta \omega \bar{\lambda}_t, \quad (14)$$

¹¹ The parameters in the conjectured value function are independent of the individual bank's type, because the value function is measured after the bank finishes its transaction for the current period and because the shock to the loan opportunity is i.i.d. across periods.

$$\left(\frac{\mathcal{V}_{st}}{Q_t^h} - \mathcal{V}_{bt}\right)(1 + \lambda_t^h) = \lambda_t^h \theta (1 - \omega), \quad (15)$$

$$\left[\theta - \left(\frac{\mathcal{V}_{st}}{Q_t^h} - \mathcal{V}_t\right)\right] Q_t^h s_t^h - [\theta \omega - (\mathcal{V}_{bt} - \mathcal{V}_t)] b_t^h \leq \mathcal{V}_t n_t^h. \quad (16)$$

According to Eq. (14), the marginal cost of interbank borrowing exceeds the marginal cost of deposit if and only if the incentive constraint is expected to bind for some state ($\bar{\lambda}_t^h > 0$) and the interbank market operates more efficiently than the retail deposit market (i.e., $\omega > 0$, meaning that assets financed by interbank borrowing are harder to divert than those financed by deposits). Equation (15) states that the marginal value of assets in terms of goods $\frac{\mathcal{V}_{st}}{Q_t^h}$ exceeds the marginal cost of interbank borrowing by banks on type h island to the extent that the incentive constraint is binding ($\lambda_t^h > 0$) and there is a friction in interbank market ($\omega < 1$). Finally, Eq. (16) is the incentive constraint. It requires that the values of the bank's net worth (or equity capital), $\mathcal{V}_t n_t^h$, must be at least as large as weighted measure of assets $Q_t^h s_t^h$ net of interbank borrowing b_t^h that a bank holds. In this way, the agency problem introduces an endogenous balance sheet constraint on banks.

The model for the general case with $0 \leq \omega \leq 1$ is somewhat cumbersome to solve. There are, however, two interesting special cases that provide insight into the models workings. In case 1, there is a perfect interbank market, which arises when $\omega = 1$. In case 2, the frictions in the interbank market are of the same magnitude as in the retail financial market, which arises when $\omega = 0$. Next, we characterize each of the cases. The Appendix in this chapter provides a solution for the general case of an interbank friction with $\omega < 1$.

2.3.1 Case 1: Frictionless wholesale financial market ($\omega = 1$)

If banks cannot divert assets financed by interbank borrowing ($\omega = 1$), interbank lending is frictionless. As Eq. (15) suggests, perfect arbitrage in the interbank market equalizes the shadow values of assets in each market, implying $\frac{\mathcal{V}_{st}}{Q_t^b} = \frac{\mathcal{V}_{st}}{Q_t^l}$, which in turn implies $Q_t^b = Q_t^l = Q_t$. The perfect interbank market further implies that the marginal value of assets in terms of goods $\frac{\mathcal{V}_{st}}{Q_t}$ must equal the marginal cost of borrowing on the interbank market $\mathcal{V}_{bt}$,

$$\frac{\mathcal{V}_{st}}{Q_t} = \mathcal{V}_{bt}. \quad (17)$$

Because asset prices are equal across island types, we can drop the h superscript in this case. Accordingly, let μ_t denote the excess value of a unit of assets relative to deposits; that is, the marginal value of holding assets $\frac{\mathcal{V}_{st}}{Q_t}$ net the marginal cost of deposits $\mathcal{V}_t$.

Then, given that banks are constrained in the retail deposit market, Eqs. (14) and (15) imply that the

$$\mu_t \equiv \frac{\mathcal{V}_{st}}{Q_t} - \mathcal{V}_t > 0. \quad (18)$$

It follows that the incentive constraint (16) in this case may expressed as

$$Q_t s_t - b_t = \phi_t n_t \quad (19)$$

with

$$\phi_t = \frac{\mathcal{V}_t}{\theta - \mu_t}. \quad (20)$$

Note that since interbank borrowing is frictionless, the constraint applies to assets inter-mediated minus interbank borrowing. How tightly the constraint binds depends positively on the fraction of net assets the bank can divert and negatively on the excess value of bank assets, given by μ_t . The higher the excess value, the greater the franchise value of the bank and the less likely it is to divert funds.

Let Ω_{t+1} be the marginal value of net worth at date $t + 1$ and let R_{kt+1} be the gross rate of return on bank assets. Then, after combining the conjectured value function with the Bellman equation, we can verify that the value function is linear in (s_t^h, b_t^h, d_t) if μ_t and $\mathcal{V}_t$ satisfy:

$$\mathcal{V}_t = E_t \Lambda_{t,t+1} \Omega_{t+1} R_{t+1} \quad (21)$$

$$\mu_t = E_t \Lambda_{t,t+1} \Omega_{t+1} (R_{kt+1} - R_{t+1}) \quad (22)$$

with

$$\begin{aligned} \Omega_{t+1} &= 1 - \sigma + \sigma(\mathcal{V}_{t+1} + \phi_{t+1} \mu_{t+1}), \text{ and} \\ R_{kt+1} &= \psi_{t+1} \frac{Z_{t+1} + (1 - \delta) Q_{t+1}}{Q_t}. \end{aligned}$$

Let us define the “augmented stochastic discount factor” as the stochastic discount factor $\Lambda_{t,t+1}$ weighted by the (stochastic) marginal value of net worth Ω_{t+1} . (The marginal value of net worth is a weighted average of marginal values for exiting and for continuing banks. If a continuing bank has an additional net worth, it can save the cost of deposits and can increase assets by the leverage ratio ϕ_{t+1} , where assets have an excess value equal to μ_{t+1} per unit). According to Eq. (21), the cost of deposits per unit to the bank $\mathcal{V}_t$ is the expected product of the augmented stochastic discount factor and the deposit rate R_{t+1} . Similarly from Eq. (22), the excess value of assets per unit, μ_t , is the expected product of the augmented stochastic discount factor and the excess return $R_{kt+1} - R_{t+1}$.

Since the leverage ratio net of interbank borrowing, ϕ_t , is independent of both bank-specific factors and island-specific factors, we can sum across individual banks to obtain the relation for the demand for total bank assets $Q_t S_t$ as a function of total net worth N_t as:

$$Q_t S_t = \phi_t N_t \quad (23)$$

where ϕ_t is given by Eq. (20). Overall, a setting with a perfect interbank is isomorphic to one where banks do not face idiosyncratic liquidity risks. Aggregate bank lending is simply constrained by aggregate bank capital.

If the banks' balance sheet constraints are binding in the retail financial market, there will be excess returns on assets over deposits. However, a perfect interbank market leads to arbitrage in returns to assets across market as follows:

$$E_t \Lambda_{t,t+1} \Omega_{t+1} R_{kt+1} = E_t \Lambda_{t,t+1} \Omega_{t+1} R_{bt+1} > E_t \Lambda_{t,t+1} \Omega_{t+1} R_{t+1}. \quad (24)$$

As will become clear, a crisis in such an economy is associated with an increase in the excess return on assets for banks of all types.

2.3.2 Case 2: Symmetric frictions in wholesale and retail financial markets ($\omega = 0$)

In this instance the bank's ability to divert funds is independent of whether the funds are obtained in either the retail or wholesale financial markets. This effectively makes the borrowing constraint the bank faces symmetric in the two credit markets. As a consequence, interbank loans and deposits become perfect substitutes as sources of finance. Accordingly, Eq. (14) implies that the marginal cost of interbank borrowing is equal to the marginal cost of deposits

$$\mathcal{V}_{bt} = \mathcal{V}_t. \quad (25)$$

Here, even if banks on investing islands are financially constrained, banks on noninvesting islands may or may not be. Roughly speaking, if the constraint on interbank borrowing binds tightly, banks in noninvesting islands will be more inclined to use their funds to refinance existing investments rather than lend them to banks on investing islands. This raises the likelihood that banks on noninvesting islands will earn zero excess returns on their assets.

Because asset supply per unit of bank net worth is larger on investing islands than on noninvesting islands, the asset price is lower; that is, $Q_t^i < Q_t^n$. Intuitively, given that the leverage ratio constraint limits banks' ability to acquire assets, prices will clear at lower values on investing islands where supplies per unit of bank net worth are greater. In the previous case of a perfect interbank market, funds flow from noninvesting to investing islands to equalize asset prices. Here, frictions in the interbank market limit the degree of arbitrage, keeping Q_t^i below Q_t^n .

A lower asset price on the investing island, means a higher expected return. Let $\mu_t^h \equiv \frac{\mathcal{V}_t}{Q_t^h} - \mathcal{V}_t$ be the excess value of assets on a type h island. Then we have:

$$\mu_t^i > \mu_t^n \geq 0. \quad (26)$$

The positive excess return implies that banks in the investing islands are finance constrained. Thus the leverage ratios for banks on each island type are given by:

$$\frac{Q_t^i S_t^i}{n_t^i} = \phi_t^i = \frac{\mathcal{V}_t}{\theta - \mu_t^i} \quad (27)$$

$$\frac{Q_t^n S_t^n}{n_t^n} \leq \phi_t^n = \frac{\mathcal{V}_t}{\theta - \mu_t^n}, \text{ and } \left(\frac{Q_t^n S_t^n}{n_t^n} - \phi_t^n \right) \mu_t^n = 0. \quad (28)$$

In this case the method of undetermined coefficients yields

$$\mathcal{V}_t = E_t \Lambda_{t,t+1} \sum_{h'=i,n} \pi^{h'} \Omega_{t+1}^{h'} R_{t+1} = E_t \Lambda_{t,t+1} \Omega_{t+1}^{h'} R_{t+1} \quad (29)$$

$$\mu_t^h = E_t \Lambda_{t,t+1} \Omega_{t+1}^{h'} (R_{kt+1}^{hh'} - R_{t+1}) \quad (30)$$

with

$$\Omega_{t+1}^{h'} = 1 - \sigma - \sigma(\mathcal{V}_{t+1} + \phi_{t+1}^{h'} \mu_{t+1}^{h'}), \text{ and}$$

$$R_{kt+1}^{hh'} = \psi_{t+1} \frac{Z_{t+1} + (1 - \delta) Q_{t+1}^{h'}}{Q_t^h}$$

With an imperfect interbank market, both the marginal value of net worth $\Omega_{t+1}^{h'}$ and the return on assets $R_{kt+1}^{hh'}$ depend on which island type a bank enters in the subsequent period. Accordingly, we index each by h' and take expectations over h' conditional on date t information denoted as E_t .

Because leverage ratios differ across islands, we aggregate separately across bank types to obtain the aggregate relations:

$$Q_t^i S_t^i = \phi_t^i N_t^i \quad (31)$$

$$Q_t^n S_t^n \leq \phi_t^n N_t^n, \text{ and } (Q_t^n S_t^n - \phi_t^n N_t^n) \mu_t^n = 0, \quad (32)$$

where ϕ_t^i and ϕ_t^n are given by Eqs. (27) and (28). As we will see, in the general equilibrium, investment will depend on the price of capital on “investing” islands, Q_t^i . Accordingly, it is the aggregate balance sheet constraint on asset demand for banks on investing islands, given by Eq. (31) that becomes critical for interactions between financial conditions and production.

Next, from Eqs. (25), (26), (29), and (30), we learn that the returns obey

$$\begin{aligned} E_t \Lambda_{t,t+1} \Omega_{t+1}^{h'} R_{kt+1}^{ih'} &> E_t \Lambda_{t,t+1} \Omega_{t+1}^{h'} R_{kt+1}^{nh'} \\ &\geq E_t \Lambda_{t,t+1} \Omega_{t+1}^{h'} R_{bt+1} = E_t \Lambda_{t,t+1} \Omega_{t+1}^{h'} R_{t+1}. \end{aligned} \quad (33)$$

with $\geq$ holds with strict inequality iff $\mu_t^n > 0$ and holds with equality iff $\mu_t^n = 0$. With an imperfect interbank market, a crisis is associated with both a rise in the excess return for banks on investing islands and increase in the dispersion of returns between island types.

As we show in the Appendix in this chapter, for the case where the interbank market is imperfect but operates with less friction than the retail deposit market (i.e., $0 < \omega < 1$), the interbank rate will lie between the return on loans and the deposit rates. Intuitively, because a dollar interbank credit will tighten the incentive constraint by less than a dollar of deposits (since lending banks are able to recover a greater fraction of creditor assets than are depositors), the interbank rate exceeds the deposit rate. However, because lending banks are not able to perfectly recover assets $\omega < 1$, there is still imperfect arbitrage, which keeps the expected discounted interbank rate below the expected discounted rate of return to loans.

2.4 Evolution of bank net worth

Let total net worth for type h banks, N_t^h , equal the sum of the net worth of existing bankers N_{ot}^h (o for old) and of entering bankers N_{yt}^h (y for young):

$$N_t^h = N_{ot}^h + N_{yt}^h. \quad (34)$$

Net worth of existing bankers equals earnings on assets net debt payments made in the previous period, multiplied by the fraction that survive until the current period, σ :

$$N_{ot}^h = \sigma \pi^h \{ [Z_t + (1 - \delta) Q_t^h] \psi_t S_{t-1} - R_t D_{t-1} \}. \quad (35)$$

Because the arrival of investment opportunity is independent across time, the interbank loans are net out in the aggregate here. We assume that the family transfers to each new banker is the fraction $\xi/(1 - \sigma)$ of the total value assets of exiting bankers, implying:

$$N_{yt}^h = \xi \pi^h [Z_t + (1 - \delta) Q_t^h] \psi_t S_{t-1}. \quad (36)$$

Finally, by the balance sheet of the entire banking sector, deposits equal the difference between total assets and bank net worth as follows,

$$D_t = \sum_{h=i,n} (Q_t^h S_t^h - N_t^h). \quad (37)$$

Observe that the evolution of net worth depends on fluctuations in the return to assets. Further, the higher the leverage of the bank, the larger the percentage impact of return fluctuations on net worth will be. Note also that a deterioration of capital quality (a decline in ψ_t) directly reduces net worth. As we will show, there is also a second round effect, as the decline in net worth induces a fire sale of assets, depressing asset prices and thus further depressing bank net worth.

2.5 Nonfinancial firms

There are two types of nonfinancial firms: goods producers and capital goods producers.

2.5.1 Goods producer

Competitive goods producers on different islands operate a constant returns to scale technology with capital and labor inputs, given by Eq. (1). Since labor is perfectly mobile across islands, firms choose labor to satisfy

$$W_t = (1 - \alpha) \frac{Y_t}{L_t} \quad (38)$$

It follows that we may express gross profits per unit of capital Z_t as follows:

$$Z_t = \frac{Y_t - W_t L_t}{K_t} = \alpha A_t \left(\frac{L_t}{K_t} \right)^{1-\alpha}. \quad (39)$$

As we noted earlier, conditional on obtaining funds from a bank, a goods producer does not face any further financial frictions and can commit to pay all the future gross profits to the creditor bank. A goods producer with an opportunity to invest obtains funds from an intermediary by issuing new state-contingent securities (equity) at the price Q_t^i . The producer then uses the funds to buy new capital goods from capital goods producers. Each unit of equity is a state-contingent claim to the future returns from one unit of investment:

$$\psi_{t+1} Z_{t+1}, (1 - \delta) \psi_{t+1} \psi_{t+2} Z_{t+2}, (1 - \delta)^2 \psi_{t+1} \psi_{t+2} \psi_{t+3} Z_{t+3}, \dots$$

Through perfect competition, the price of new capital goods is equal to Q_t^i and goods producers earn zero profits state by state.

Note that given constant returns and perfect labor mobility, we do not have to keep track of the distribution of capital across islands. As in the standard competitive model with constant returns, the size distribution of firms is indeterminate.

2.5.2 Capital goods producers

Capital producers operate in a national market. They make new capital using input of final output and subject to adjustment costs, as described in Section 2.1. They sell new capital to firms on investing islands at the price Q_t^i . Given that households own capital producers, the objective of a capital producer is to choose I_t to solve:

$$\max E_t \sum_{\tau=t}^{\infty} \Lambda_{t,\tau} \left\{ Q_{\tau}^i I_{\tau} - \left[1 + f\left(\frac{I_{\tau}}{I_{\tau-1}}\right) \right] I_{\tau} \right\}$$

From profit maximization, the price of capital goods is equal to the marginal cost of investment goods production as follows:

$$Q_t^i = 1 + f\left(\frac{I_t}{I_{t-1}}\right) + \frac{I_t}{I_{t-1}} f'\left(\frac{I_t}{I_{t-1}}\right) - E_t \Lambda_{t,t+1} \left(\frac{I_{t+1}}{I_t}\right)^2 f'\left(\frac{I_{t+1}}{I_t}\right) \quad (40)$$

Profits (which arise only outside of steady state), are redistributed lump sum to households.

2.6 Equilibrium

To close the model (in the case without government policy), we require market clearing in both the market for securities and the labor market. Total securities issued on investing and noninvesting islands correspond to aggregate capital acquired by each type, as follows:

$$\begin{aligned} S_t^i &= I_t + (1 - \delta)\pi_i K_t \\ S_t^n &= (1 - \delta)\pi^n K_t. \end{aligned} \quad (41)$$

Note that demand for securities by banks is given by Eq. (23) in the case of a frictionless interbank market and by Eqs. (31) and (32) in the case of an imperfect interbank market. Observe first that the market price of capital on each island type will, in general, depend on the financial condition of the associated banks. Second, with an imperfect interbank market, the asset price will be generally lower (or, equivalently, state-contingent loans rates offered by banks will be generally greater) on investing islands than elsewhere.¹²

Finally, the condition that labor demand equals labor supply requires that

$$(1 - \alpha) \frac{Y_t}{L_t} \cdot E_t u_{C_t} = \chi L_t^{\varepsilon} \quad (42)$$

Because of Walras' Law, once the market for goods, labor, securities, and interbank loans is cleared, the market for riskless debt will be cleared automatically:

$$D_{ht} = D_t + D_{gt},$$

where D_{gt} is supply of government debt. This completes the description of the model.

Absent credit market frictions, the model reduces to a real business cycle framework modified with habit formation and flow investment adjustment costs. With the credit market frictions, however, balance sheet constraints on banks' ability to obtain funds in retail and wholesale markets may limit real investment spending, affecting aggregate

¹² This verifies the earlier conjecture in Section 2.3.2. For the more general case of imperfect interbank market, see Appendix 1.

real activity. As we will show, a crisis is possible where weakening of bank balance sheets significantly disrupts credit flows, depressing real activity.

As we have discussed, one example of a factor that could weaken bank balance sheets is a deterioration of the underlying quality of capital. A negative quality shock directly reduces the value of bank net worth, forcing banks to reduce asset holdings. A second round effect on bank net worth arises as the fire sale of assets reduces the market price of capital. Further, the overall impact on bank equity of the decline in asset values is proportionate to the amount of bank leverage. With highly leveraged banks, a substantial percentage drop in bank equity may arise, leading to a significant disruption of credit flows. We illustrate this point clearly in Section 4.

3. CREDIT POLICIES

During the crisis the various central banks, including the U.S. Federal Reserve, made use of their powers as a lender of last resort to facilitate credit flows. To justify such actions, the Federal Reserve appealed to Section 13.3 of the Federal Reserve Act, which permits it in “unusual and exigent circumstances” to make loans to the private sector, as long as the loans are judged to be of sufficiently high grade. The statute makes clear that in normal times the Federal Reserve is not permitted to take on private credit risk. In a crisis, however, it has the freedom to fulfill its responsibility as lender of last resort, provided that it does not absorb undue risk.

In practice, the Federal Reserve employed three general types of credit policies. First, early on it expanded discount window operations by permitting discount window loans to be collateralized by high-grade private securities and also by extending the availability of the window to nonbank financial institutions. Second, it lent directly in high-grade credit markets, funding assets that included commercial paper, agency debt, and mortgage-backed securities. Third, the Treasury, acting in concert with the Federal Reserve, injected equity in the banking system along with supplying bank debt guarantees (together with the Federal Deposit Insurance Corporation).

There is some evidence that these types of policies were effective in stabilizing the financial system. The expanded liquidity helped smooth the flow of funds between financial institutions effectively by dampening the turmoil-induced increases in the spread between the interbank lending rate (LIBOR) and the Treasury Bill rate. The enhanced financial distress following the Lehmann failure, however, proved to be too much for the liquidity facilities alone to handle. At this point, the Federal Reserve set up facilities to lend directly to the commercial paper market and a number of weeks later phased in programs to purchase agency debt and mortgage-backed securities. Credit spreads in each of these markets fell.

The equity injections also came soon after Lehmann. Although not without controversy, the equity injections appeared to reduce stress in banking markets. Upon the

initial injection of equity in mid-October 2008, credit default swap rates of the major banks fell dramatically. By this time, the receiving banks have paid back a considerable portion of the funds. Although risks remain, the government appears to have made money on many of these programs.

In the following subsections, we take a first pass at analyzing how these policies work, using our baseline model.¹³ As we showed in the previous section, within the context of our model, the financial market frictions open the possibility of periods of distress where excess returns on assets are abnormally high. Because they are balance sheet constrained, private financial intermediaries cannot immediately arbitrage these returns. One can see the point of the Federal Reserve's various credit programs as facilitating this arbitrage in times of crisis. In this regard, each of the various policies works somewhat differently, as we discuss next.

Before proceeding, we emphasize that, consistent with the Federal Reserve Act, these interventions are used only during crises and not during normal times. Indeed, within the logic of the model, the net benefits from credit policy are increasing in the distortion of credit markets that the crisis induces, as measured by the excess return on capital.

3.1 Lending facilities (direct lending)

We characterize lending broadly as the facilities the Federal Reserve set up for direct acquisition of high quality private securities.

Lending facilities work as follows: We suppose that the central bank has both an advantage and a disadvantage relative to private lenders. The advantage is that unlike private intermediaries, the central bank is not balance sheet constrained (at least in the same way). Private citizens do not have to worry about the central bank defaulting. The liabilities it issues are government debt and it can credibly commit to honoring this debt (aside from inflation). Thus, in periods of distress where private intermediaries are unable to obtain additional funds, the central bank can obtain funds and then channel them to markets with abnormal excess returns.¹⁴

In the current crisis, the Federal Reserve funded the initial expansion of its lending programs by issuing government debt (that it borrowed from the Treasury) and then later made use of interest bearing reserves. The latter are effectively government debt. It is true that the interest rate on reserves fell to zero as the federal funds rate reached its lower bound, giving these reserves the appearance of money. However, once the Federal Reserve moves the funds rate above zero, it will also raise the interest rate on reserves.

¹³ For related attempts at model credit policy, see Curdia and Woodford (2009a,b), Reis (2009), and Sargent and Wallace (1983).

¹⁴ Others have also emphasized how that special nature of government liabilities can give rise to a productive role for government financial intermediations. See, for example, Sargent and Wallace (1983), Kiyotaki and Moore (2008), Gertler and Karadi (2009), and Shleifer and Vishny (2010). As originally noted by Wallace (1981), unless there is something special about government liabilities, the Miller-Modigliani theorem applies to government finance.

In this regard, the Federal Reserve's unconventional policies should be thought of as expanded central intermediation as opposed to expanding the money supply. In the case of lending facilities, a key advantage of the central bank is that it is not constrained in its ability to access funds the same way private intermediaries may be in time of financial distress. Another equally important advantage is that the Federal Reserve can lend in many markets. By contrast, private banks face a limited market participation constraint; that is, they can only lend to nonfinancial firms of the same island.

At the same time, we suppose that the central bank is less efficient at intermediating funds. It faces an efficiency cost τ per unit, which may be thought of as a cost of evaluating and monitoring borrowers that is above and beyond what a private intermediary (who has specific knowledge of a particular market) would pay.¹⁵

To obtain funds, the central bank issues government debt to the private sector that is a perfect substitute for bank deposits, and pays the riskless real rate R_{t+1} . It lends the funds in market h at the private loan rate R_{kt+1}^{hh} which depends upon the state of the next period h' . Observe that the central banks are not offering the funds at a subsidized rate. However, by expanding the supply of funds available in the market, it will reduce equilibrium lending rates.

Let S_t^h be total securities of type h intermediated, S_{pt}^h total securities of type h intermediated by private banks, and S_{gt}^h total type h securities intermediated by the central bank. Then total intermediation of type h assets is given by:

$$Q_t^h S_t^h = Q_t^h (S_{pt}^h + S_{gt}^h) \quad (43)$$

We suppose the central bank chooses to intermediate the fraction φ_t^h of total credit in market h :

$$S_{gt}^h = \varphi_t^h S_t^h \quad (44)$$

where φ_t^h may be thought of as an instrument of central bank credit policy.

Assuming that banks' investing regions are constrained under a symmetric frictions in wholesale and retail financial markets ($\omega = 0$), lending facilities expand the total amount of assets intermediated in the market. Combining Eqs.(31), (43), and (44), yields

$$Q_t^i S_t^i = \frac{1}{1 - \varphi_t^i} \phi_t^i N_t^i \quad (45)$$

The effect on asset demand for noninvesting regions depends on whether or not banks in these regions are balance sheet constrained (i.e., on whether the excess return $\mu_t^n > 0$ is positive). If they are, then lending facilities affect asset demands similarly to the way

¹⁵ Other potential costs include the potential for politicization of credit flows. We abstract from this consideration, although we think it provides another important reason for why credit policies are more appropriate in crises than normal times.

they do in investing regions, only the superscript i is replaced by n in Eq. (45). One other hand, if banks in noninvesting regions are not constrained (i.e., $\mu_t^n = 0$), then central bank credit merely displaces private credit, leaving total asset demand in the sector unaffected. Let S_t^{n*} be total asset demand consistent with a zero excess return on assets on noninvesting islands in equilibrium. Then

$$Q_t^n S_t^{n*} = Q_t^n S_{pt}^n + \varphi_t^n Q_t^n S_t^{n*}, \text{ iff } \mu_t^n = 0. \quad (46)$$

Here an increase in central credit provision crowds out private intermediation one for one. Only when private intermediaries are financially constrained does central bank intermediation expand the overall supply of credit.

3.2 Liquidity facilities (discount window lending)

With liquidity facilities, the central bank uses the discount window to lend funds to banks that in turn lend them out to nonfinancial borrowers. Typically, liquidity facilities are used to offset disruption of interbank markets. Such was the case in the current crisis.

Another distinguishing feature of liquidity facilities is that central bank lending is typically done at a penalty rate. This prescription dates back to Bagehot (1873). The idea is that during a liquidity crises, it is the breakdown of markets for short-term funds that is responsible for many borrowers having limited credit access, as opposed to lack of credit worthiness of individual borrowers. Because excess returns for these borrowers are abnormally high during the crisis, they are more than willing to borrow at penalty rates. Offering the funds at a penalty rate further discourages inefficient use of central bank credit by the private sector.

In this section we use our model to illustrate how discount window lending may facilitate the flow of interbank lending during a crisis. To do so, we restrict attention to the case ($\omega = 0$), where borrowers in the interbank market face symmetric constraints on obtaining funds in both the wholesale and retail markets. In this instance, banks with surplus funds face the same risk as depositors that borrowing banks may divert a fraction of gross assets for their own purposes.

We suppose the central bank offers discount window credit at the noncontingent interest rate R_{mt+1} to banks who borrow on the interbank market. It funds this activity by issuing government debt that is a perfect substitute for household deposits. For discount window lending to expand the supply of funds in the interbank market, however, the central bank must have an advantage over private lenders in supplying funds to borrowing banks. Otherwise discount window lending will simply supplant private interbank lending.

Here we suppose that the central bank is better able to enforce repayment than private lenders. In particular for any unit of discount window credit supplied, a borrowing bank can divert only the fraction $\theta(1 - \omega_g)$ of assets, with $0 < \omega_g \leq 1$. Recall that for

credit supplied by a private lender, the borrowing bank can divert the fraction $\theta > \theta(1 - \omega_g)$. Here the idea is that the government may have additional means at its disposal (IRS records, access to credit records, legal punishments, etc.) to retrieve assets. We suppose, however, that after a certain level of discount window lending, the central bank's ability to retrieve assets more efficiently than the private sector disappears. Think of this as reflecting some capacity constraint on the central bank's ability to efficiently process discounted window loans secured by private credit.¹⁶

Let m_t^h be discount window borrowing for a bank of type h . The flow of funds constraint is now,

$$Q_t^h s_t^h = n_t^h + b_t^h + m_t^h + d_t. \quad (47)$$

with $m_t^h \geq 0$. Let $V_t(s_t^h, b_t^h, m_t^h, d_t)$ be the value of a bank who holds assets and liabilities $(s_t^h, b_t^h, m_t^h, d_t)$ at the end of period t . For the bank to continue operating this value must not fall below the gain from diverting assets, taking into account the central bank's advantage in retrieving assets. Accordingly, in this case the incentive constraint is given by:

$$V_t(s_t^h, b_t^h, m_t^h, d_t) \geq \theta(Q_t^h s_t^h - \omega_g m_t^h). \quad (48)$$

We defer the details of the bank's decision problem for this case to the Appendix at the end of the chapter. Accordingly, let μ_{mt} be the excess cost to a bank of discount window credit relative to deposits

$$\mu_{mt} = E_t \Lambda_{t,t+1} \Omega_{t+1}^{h'} (R_{mt+1} - R_{t+1}). \quad (49)$$

Next note that, because we are restricting attention to the case of symmetric frictions in private interbank and retail financial markets ($\omega = 0$), the interbank rate equals the deposit rate: $R_{bt+1} = R_{t+1}$. Then from the first-order conditions we learn that for both private interbank borrowing and discount window to be actively used, we need:

$$\mu_{mt} = \omega_g \mu_t^i \quad (50)$$

where μ_t^i is the excess value of assets on investing islands, given by Eq. (30).

According to Eq. (50), to make borrowers indifferent between discount window and private credit at the margin, the central bank should set R_{mt+1} to make the excess cost of discount window credit equal to the fraction ω_g of the excess value of assets. Intuitively, because a unit of discount window credit permits a borrowing bank to expand

¹⁶ Alternatively, if we had asset heterogeneity this constraint might reflect a limitation on the kind of bank assets that might be suitable collateral for discount window lending. For example, information-intensive commercial and industrial loans are not good collateral for discount window loans since they require expertise for monitoring and evaluation. On the other hand, agency debt or high-grade securitized mortgage might be suitable, but banks might only have a limited fraction in their portfolios.

assets by a greater amount than a unit private interbank credit, it is willing to pay a higher cost for this form of credit. In this way, the model generates an endogenously determined penalty rate for discount window lending.

Let M_t be the total supply of discount window credit offered to the market. Then one can show that the market demand for assets by investing banks is given by

$$Q_t^i S_{pt}^i = \phi_t^i N_t^i + \omega_g M_t. \quad (51)$$

Thus, as long as $\omega_g > 0$, discount window lending can expand the total level of assets intermediated by banks on investing regions.

Because the excess value of bank assets on noninvesting islands is less than that on investing islands; that is, $\mu_t^n < \mu_t^i$, banks on noninvesting islands will not borrow from the discount window. Given that the discount rate is set to satisfy Eq. (50), discount window lending will be too expensive for banks who do not have new investment to finance.

The question then arises as to why the central bank does not simply expand discount lending to drive excess values of assets to zero. As we noted earlier, it is reasonable to suppose that there are capacity constraints on the central bank's ability to adequately monitor the asset management activities of banks, (even though we do not formally incorporate it into our model). With a capacity constraint on discount window lending (secured by private credit), the central bank may need to use other tools such as direct lending or equity injections during crisis periods of high excess returns. While liquidity facilities may be useful for improving the flow of funds in interbank markets, in a major crisis other kinds of interventions may be necessary to stabilize financial markets.

3.3 Equity injections

With equity injections, the fiscal authority coordinates with the monetary authority to acquire ownership positions in banks. As with direct central bank lending we suppose that there are efficiency costs associated with government acquisition of equity. Let this cost be τ_e per unit of equity acquired. During a financial crisis, however, the net benefits from equity injections may be positive and significant.

The effect of equity injections depends on three factors: (i) the payout rule for government equity, (ii) the price at which the government acquires the equity relative to the market price, and (iii) the advantage the government might have relative to private creditors in addressing the agency problem with banks.

The government injects equity into banks who stay active (instead of exiting) at the beginning of the period before banks learn whether their customers have opportunities to invest or not. This is different from the direct lending and discount window lending activities of the central bank that are conducted after the arrival of investment opportunities. By this difference in timing, we try to capture a feature that the equity injections are slower than the direct lending and discount window lending. For simplicity we restrict attention to the case with a perfect interbank market in which banks cannot divert

assets that are financed by interbank borrowing. (See the Appendix for a general case). Then the asset price is equal across regions with different investment opportunity.

We suppose that a unit of government equity has the same payout stream as a unit of private equity. The government may hold the equity stake until the bank exits and then receive the liquidation value of its assets, equal to $Z_t + (1 - \delta)Q_t$ per unit of capital times the number of units of capital its shares are worth. Alternatively it may sell off its holding at this value before the bank exits, assuming the crisis has passed.

Accordingly, one can effectively divide the total number of securities held by the bank at time t between those privately owned, s_{pt} , and those publicly owned, s_{gt} :

$$s_t = s_{pt} + s_{gt} \quad (52)$$

Let n_{gt} be the market value of government equity. The bank's balance sheet identity then implies:

$$Q_t s_t = n_t + b_t + d_t + n_{gt} \quad (53)$$

where each security the government holds is valued at the market price Q_t , implying:

$$n_{gt} = Q_t s_{gt} \quad (54)$$

To acquire equity, the government may pay a price Q_{gt} that is above Q_t . One rationale for the government paying a premium is that the market price is below its normal value due to financial distress. For example, the government could pick Q_{gt} so that the excess return on government equity, μ_{gt} , equals zero, as follows:

$$\mu_{gt} = E_t \Lambda_{t,t+1} \Omega_{t+1} (R_{gkt+1} - R_{t+1}) \quad (55)$$

where R_{gkt+1} is the gross return on a unit of government equity injected at time t is

$$R_{gkt+1} = \psi_{t+1} \frac{Z_{t+1} + (1 - \delta)Q_{t+1}}{Q_{gt}} \quad (56)$$

Since the excess return of private equity is positive (see equation (22)), $Q_{gt} > Q_t$.

The premium the government pays for equity is effectively a transfer to the bank that shows up in its net worth as follows:

$$n_t = [Z_t + (1 - \delta)Q_t] \psi_t s_{pt-1} - R_{bt} b_{t-1} - R_{dt} d_{t-1} + (Q_{gt} - Q_t) [s_{gt} - (1 - \delta) \psi_t s_{gt-1}] \quad (57)$$

where $(Q_{gt} - Q_t) [s_{gt} - (1 - \delta) \psi_t s_{gt-1}]$ is the “gift” to the bank from new government equity purchases.

We suppose that the bank cannot divert assets financed by government equity. As with discount window lending, the government has an advantage relative to the private creditors in recovering assets. Accordingly, the incentive constraint becomes

$$V_t(s_t - s_{get}, b_t, d_t) \geq \theta(Q_t(s_t - s_{get}) - b_t).$$

where as before b_t is interbank borrowing (with $\omega = 1$).

Let N_{gt} be total government equity in the banking system and S_{gt} be total holdings of government equity. Then we can aggregate to obtain the following expressions for aggregate asset demand and for the evolution of net worth:

$$Q_t S_t = \phi_t N_t + N_{gt} \quad (58)$$

$$N_t = (\sigma + \xi)[Z_t + (1 - \delta)Q_t]\psi_t S_{pt-1} - \sigma R_t D_{t-1} + (Q_{gt} - Q_t)[S_{gt} - (1 - \delta)\psi_t S_{gt-1}] \quad (59)$$

where ϕ_t is the leverage ratio privately intermediated assets in the case of a perfect interbank market (see Eq. 20), and with $N_{gt} = Q_t S_{gt}$. Thus, in this case equity injections expand the value of assets intermediated one-for-one, as Eq. (58) suggests. In addition, to the extent the government paying pays a premium over the market price (which is depressed due to the financial crisis), the equity injection also expands private bank net worth, as Eq. (59) indicates. This in turn expands asset demand by a multiple equal to the leverage ratio ϕ_t .

One additional important effect of government equity injections is that they reduce the impact of unanticipated changes in asset values on private bank equity. Absent government equity, for example, the bank absorbs entirely the loss from an unanticipated decline in asset values, given that its obligations to outsiders are all in the form of non-contingent debt. With public equity, however, the government shares proportionately in the loss.

A key question now is what might determine the allocation of credit policy intervention between direct lending, discount window lending, and equity injections. We argued earlier that in the context of our model, it might be natural to think of capacity constraints on discount window lending secured by private credit. As long as the efficiency costs of direct central bank lending are not large, extensive use of direct lending makes sense. For high-grade instruments like commercial paper, agency debt, and mortgage-backed securities, it is reasonable to suppose the costs of central bank intermediation are not large. This might account for why direct central bank lending in the current crisis involved these kinds of assets. On the other hand, it is easy to imagine that other forms of bank lending, such as commercial and industrialized loans, which involve extensive evaluation and monitoring, would be quite costly for the central bank to intermediate. In this case, in a period of crisis, equity injections that enhance the ability of private banks to make these kinds of loans would seem desirable (if the efficiency cost of government equity injection is not too large). In our model, capital is homogeneous. Getting at this issue, accordingly, will involve extending our framework to allow for asset heterogeneity.

3.4 Government expenditures and budget constraint

Here government consumption G_t consists of “normal” government expenditures $\bar{G}$ and intermediation expenditures. Let S_{gt}^h be total securities of type $h = i, n$ acquired via direct central bank lending, and S_{get} securities acquired via equity injections. Then G_t is given by

$$G_t = \bar{G} + \tau_e S_{get} + \tau \sum_{h=i,n} S_{gt}^h \quad (60)$$

Putting together fiscal and monetary authority, government expenditures are financed by lump-sum taxes T_t and net earnings from credit market interventions as

$$\begin{aligned} G_t + Q_{gt}[S_{get} - (1 - \delta)\psi_t S_{get-1}] + \sum_{h=i,n} Q_t^h[S_{gt}^h - (1 - \delta)\psi_t S_{gt-1}^h] \\ = T_t + Z_t\psi_t(S_{gt-1} + S_{get-1}) + R_{mt}M_{t-1} - M_t + D_{gt} - R_t D_{gt-1} \end{aligned} \quad (61)$$

where M_t is total discount window lending and D_{gt} is government bond. As we discussed earlier, the price the government pays for equity, Q_{gt} , could exceed the market price.

Note that the during the crisis the government will earn extra returns on its portfolio, since excess private returns in the market are positive, but private intermediaries are constrained from exploiting this. On the other hand, the government may takes losses on its portfolio. Here we assume that lump-sum taxes adjust to finance the losses. It would be interesting to consider distortionary taxes to get a better sense of the costs faced in pursuing these policies.

4. CRISIS SIMULATIONS AND POLICY EXPERIMENTS

In this section we present some numerical experiments designed to illustrate how the model may capture some key features of a financial crisis and also how credit policy might work to mitigate the crisis. The analysis is meant only to be suggestive. In this regard, our aim is to show how vulnerability of the financial system might propagate the effects of a disturbance to asset values and aggregate production that might otherwise have a relatively modest effect on the economy. In addition to identifying the significance of balance sheet effects on intermediaries in the process, we also isolate the importance of an imperfect interbank market.

We start with the calibration and then turn to a “crisis” simulation. After examining how the crisis plays out in the absence of any kind of policy response, we analyze how credit policy might work to mitigate the crisis. We focus on direct lending since this policy is the simplest to present. Although, we do not report the results here, the other policies ultimately affect the economy in a similar fashion.

4.1 Calibration

There are eleven parameters for which we need to assign values. Seven are standard preference and technology parameters. These include the discount factor β , the habit parameter γ , the utility weight on labor χ , the inverse of the Frisch elasticity of labor supply ε , the capital share parameter α , the depreciation rate δ , and the elasticity of the price of capital with respect to investment η . For these parameters we use reasonably conventional values, as reported in Table 1. The one exception involves the labor supply elasticity: To compensate partly for the absence of labor market frictions, we use a Frisch labor elasticity of ten, which is well above the range found in the business cycle literature and typically lies between unity and three. We emphasize that this compensation is only partial: Had we instead incorporated the various key of quantitative DSGE models, including variable capital utilization and nominal price and wage

Table 1 Parameter Values for Baseline Model

Households

β	0.990	Discount rate
γ	0.500	Habit parameter
χ	5.584	Relative utility weight of labor
ε	0.100	Inverse Frisch elasticity of labor supply

Financial intermediaries

π^t	0.250	Probability of new investment opportunities
θ	0.383	Fraction of assets divertable: perfect interbank market
	0.129	Fraction of assets divertable: imperfect interbank market
ξ	0.003	Transfer to entering bankers: perfect interbank market
	0.002	Transfer to entering bankers: imperfect interbank market
σ	0.972	Survival rate of the bankers

Intermediate good firms

α	0.330	Effective capital share
δ	0.025	Steady-state depreciation rate

Capital producing firms

$I f' / f$	1.500	Inverse elasticity of net investment to the price of capital

Government

$\frac{G}{Y}$	0.200	Steady-state proportion of government expenditures

rigidities, employment volatility in our framework would be much greater, even with a conventional labor supply elasticity.

The four additional parameters are specific to our model. The first is the probability of an investment opportunity, π^i . The last three are the financial sector parameters: σ the quarterly survival probability of bankers, ξ the transfer parameter for new bankers, and θ the fraction of gross assets the banker can divert. We set π^i equal to 0.25, implying that new investment opportunities on an island arise once a year on average. We set $\sigma = 0.975$, implying that bankers survive for ten years on average.

Finally, we choose ξ and θ to hit the following two targets: an average credit spread of 100 basis points per year and an economy-wide leverage ratio of 4. The choice of a leverage ratio of four reflects a crude first pass attempt to average across sectors with vastly different financial structures. For example, before the beginning of the crisis, most housing finance was intermediated by financial institutions with leverage ratios between twenty (commercial banks) and thirty (investment banks). The total housing stock, however, was only about one-third of the overall capital stock. Leverage ratios are clearly smaller in other sectors of the economy. We base the steady-state target for the spread on the pre-2007 spreads as a rough average of the following spreads: mortgage rates versus government bond rates, BAA corporate bond rates versus government bonds, and commercial paper rates versus T-bill rates.

We consider both the case of a perfect interbank market ($\omega = 1$) and of an imperfect interbank market ($\omega = 0$). As we noted earlier, with a perfect interbank market, the model economy behaves as if banks were homogenous and did not face an idiosyncratic arrival of lending opportunities. Under our calibration, within a local region of the steady state, all banks are symmetrically constrained; that is, have similar excess returns on assets.

With an imperfect interbank market, under our calibration only banks on investing islands are constrained (within a local region of the steady state). Those on noninvesting islands have sufficient funds relative to lending opportunities to bid the price on assets to the point where the excess return over deposit costs is zero. They lend surplus funds to banks in investing regions. For reasonable variations of our calibration, banks remain unconstrained in noninvesting regions and remain constrained in investing regions.

Finally, we suppose that the capital quality shock obeys a first-order autoregressive process.

4.2 Crisis experiment

4.2.1 No policy response

We now turn to the crisis experiment. Broadly speaking, what triggered the recent financial crisis was a decline in real estate values that precipitated a wave of losses on mortgage-backed securities held by financial intermediaries. Our model is not

sufficiently rich to capture precisely this phenomenon, particularly since it does not include housing. The initiating feature of the current crisis that we can capture, however, is the deterioration in value of intermediary portfolios.

In particular, the initiating disturbance we consider is an exogenous decline in capital quality.¹⁷ What we are trying to capture in a simple way is an exogenous force that triggers a decline in the value of intermediary assets. Within the model economy, the initial exogenous decline is then magnified in two ways. First, because banks are leveraged, the effect of decline in assets values on bank net worth is enhanced by a factor equal to the leverage ratio. Second, the drop in net worth tightens banks' borrowing constraint inducing effectively a fire sale of assets that further depresses asset values. The crisis then feeds into real activity as the decline in asset values leads to a fall in investment.

The initiating disturbance is a 5% unanticipated decline in capital quality with an autoregressive factor of 0.66. We fix the size of the shock simply to produce downturn of roughly similar magnitude to the one observed over the 2008–2009 financial crisis.

We began by analyzing the performance of the model economy without credit policy and we start with the case of a perfect interbank market. Figure 1 reports the impulse responses of the key economic variables to a negative shock to capital quality. The dotted line is the model without financial frictions and the solid line is our baseline model with a perfect interbank market.

Note first that the negative disturbance produces only a modest downturn in the frictionless model. The loss of capital initially produces a drop in output and consumption. However, high returns to capital induce an increase in investment and employment. Therefore without financial frictions, the economy smoothly converges to a normal state as in a Cass-Koopmans optimal growth model with a smaller initial capital stock than the steady state.

With financial frictions the output decline at the trough is roughly twice as large as in the frictionless case. It is also significantly more protracted. The 5% decline in the quality of capital leads to a roughly 50% decline in bank net worth. The magnified effect is due to bank leverage and to the fall in the market price capital, arising from the fire sale of assets induced by the tightening of bank borrowing constraints. The contraction in asset prices induces a decline in investment that is nearly double the output decline. It is the enhanced decline in investment that is ultimately responsible for the magnified drop in output in the case with financial frictions. Finally, the employment drop, while nearly several percentage points larger than in the frictionless case,

¹⁷ What is critical for our crisis experiment is that the initiating disturbance leads to a decline in the market prices of intermediary assets. Another type of disturbance that could initiate a decline in asset values would be an unfavorable “news shock” about the future payoff to capital as in Gilchirst and Leahy (2002); Christiano, Motto, and Rostagno (2010); or Gourio (2009). Yet another possibility would be to introduce “noise” shocks, as in La’O (2010).

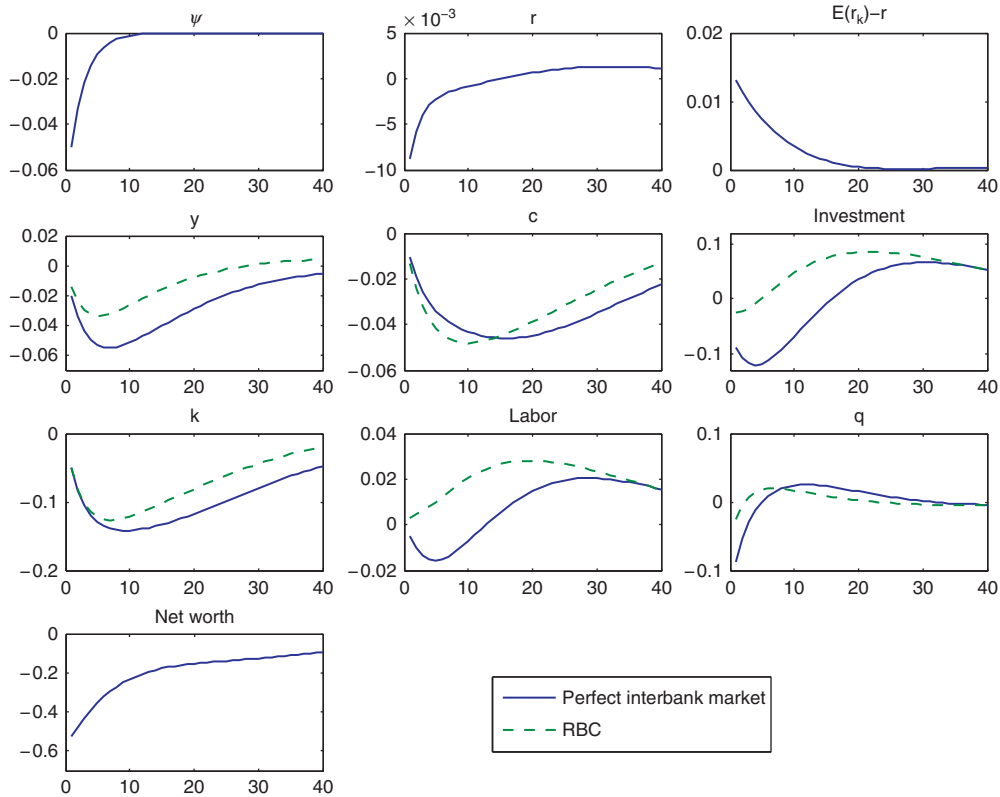


Figure 1 Crisis experiment: Perfect interbank market.

is relatively modest. This simply reflects the absence of various standard labor market frictions that would enhance the response.

That financial factors are at work during the crisis is reflected in the behavior of the spread between the expected return to capital and the riskless interest rate. In the frictionless model this spread does not move (to a first order.) In the case with financial frictions, the spread rises on impact as a product of the decline in bank net worth. The increase in the cost of capital is responsible for the magnified drop in investment and output.

Financial factors also contribute to the slow recovery back to trend. To reduce the spread between the expected return to capital and the riskless rate, bank net worth must increase. But this process takes time, as Figure 1 illustrates. As long as the spread is above trend, financial factors are a drag on the real economy. Note that throughout this convergence process, banks are effectively deleveraging since they are building up equity relative to debt. The model captures how the deleveraging process can slow down a recovery.

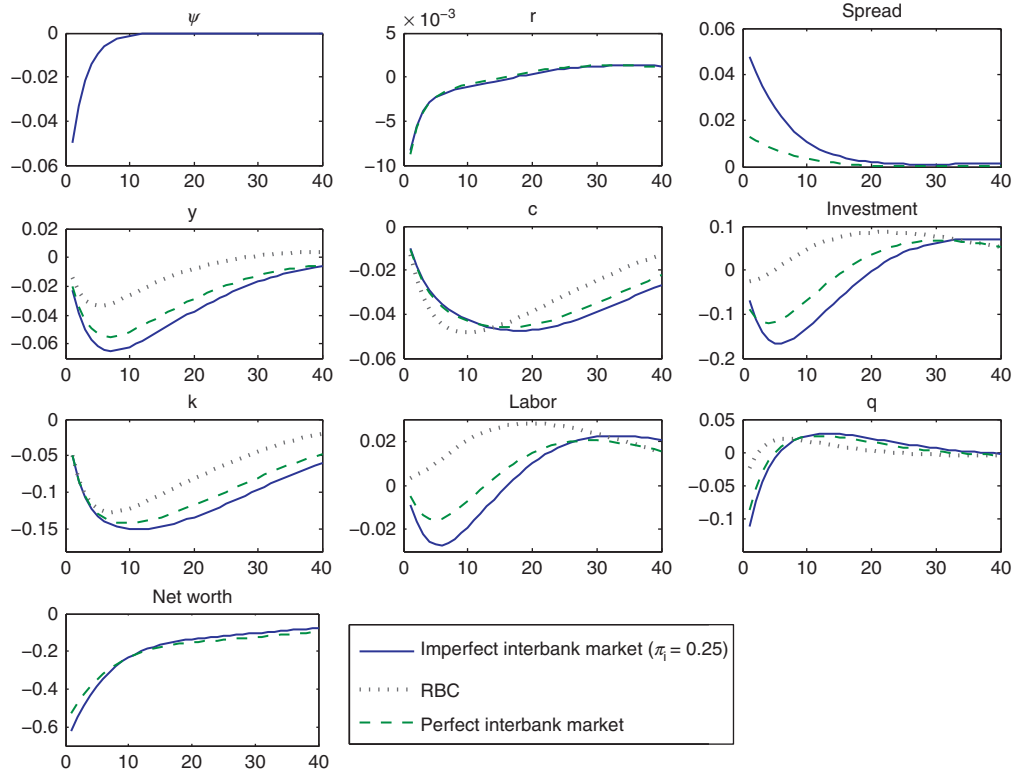


Figure 2 Crisis experiment: Imperfect interbank market.

Next we turn to the case with the imperfect interbank market in Figure 2. Observe that frictions in the interbank market magnify the overall decline. The overall decline in investment is roughly a third larger relative to the perfect interbank market case, the output decline 20% larger, and the employment decline nearly double. Intuitively, in this case investing banks are limited in their ability to obtain funds on the interbank market once the crisis hits. In addition, banks on investing islands have higher leverage than those on noninvesting islands because the asset price is lower in investing islands. Accordingly, asset prices in investing islands fall by more than they otherwise would, leading to an enhanced drop in overall investment. Symptomatic of the imperfect interbank market is the sharp rise in the spread between the return on capital and the riskless rate, which increases well above 5%, as compared to 1% in the case of a perfect interbank market.

4.2.2 Credit policy response

Here we analyze the impact of direct central bank lending as a means to mitigate the impact of the crisis. Symptomatic of the financial distress in the simulated crisis is a large

increase in the spread between the expected return on capital on investing islands and the riskless interest rate. In practice, it was the appearance of abnormally large credit spreads in various markets that induced the Federal Reserve to intervene with credit policy. Accordingly we suppose that the Federal Reserve adjusts the fraction of private credit it intermediates to the difference between spread on investing islands, $(E_t R_{kt+1}^{ih'} - R_{t+1})$, and its steady state value $(ER_k^{ih'} - R)$, as:

$$\varphi_t = v_g[(E_t R_{kt+1}^{ih'} - R_{t+1}) - (ER_k^{ih'} - R)] \quad (62)$$

To be clear, the rule applies only during a crisis; that is, during “unusual and exigent” circumstances.

We begin with the case a perfect interbank market. In this case the return on assets is equalized across islands. It does not matter to which locale the central bank supplies credit. If it intermediates funds on noninvesting islands, banks in these locations will lend any surplus funds to banks on investing islands to the point where the return on assets is equalized across locations.

We set the policy parameter v_g equal to 100. Figure 3 reports the impulses for this case. The policy intervention dampens the overall decline in output by nearly one-third. The increase in central bank credit significantly reduces the rise in the spread, which in turn reduces the overall drop in investment. At its peak, central bank credit increases to slightly over 10% of the capital stock.

With an imperfect interbank market the central bank acquires assets on investing islands. What we have in mind here is that the central bank is targeting assets with high excess returns; that is, assets that may be underfunded due to shortages of intermediary capital in the relevant market. Note that by charging the market rate to borrowers in these regions, the policy screens out borrowers on noninvesting islands who earn lower returns.

Figure 4 reports the results for this case. The credit policy similarly works to dampen the output decline by mitigating the increase in the spread. Interestingly, the policy is more effective at containing the crisis in this case. What matters are the leverage constraints on bank borrowing in investing locations, as opposed to leverage constraints economy-wide. By directly facilitating credit flows in investing regions, a given level of central bank intermediation can be more effective in relaxing financial constraints. Note in this case that at the peak, central bank credit intermediation is only about 5% of total assets intermediated, which is less than half of what it was in the economy of the frictionless interbank market. However, it is roughly 20% of assets intermediated in investing regions. The high percentage of central bank intermediation in this distressed region is what accounts for the effectiveness of the policy. This occurs even though total central bank intermediation is smaller than in the case of the perfect interbank market.

As we noted earlier, both discount window lending and equity injections work in a similar fashion to mitigate a crisis. It would be interesting to extend our framework to

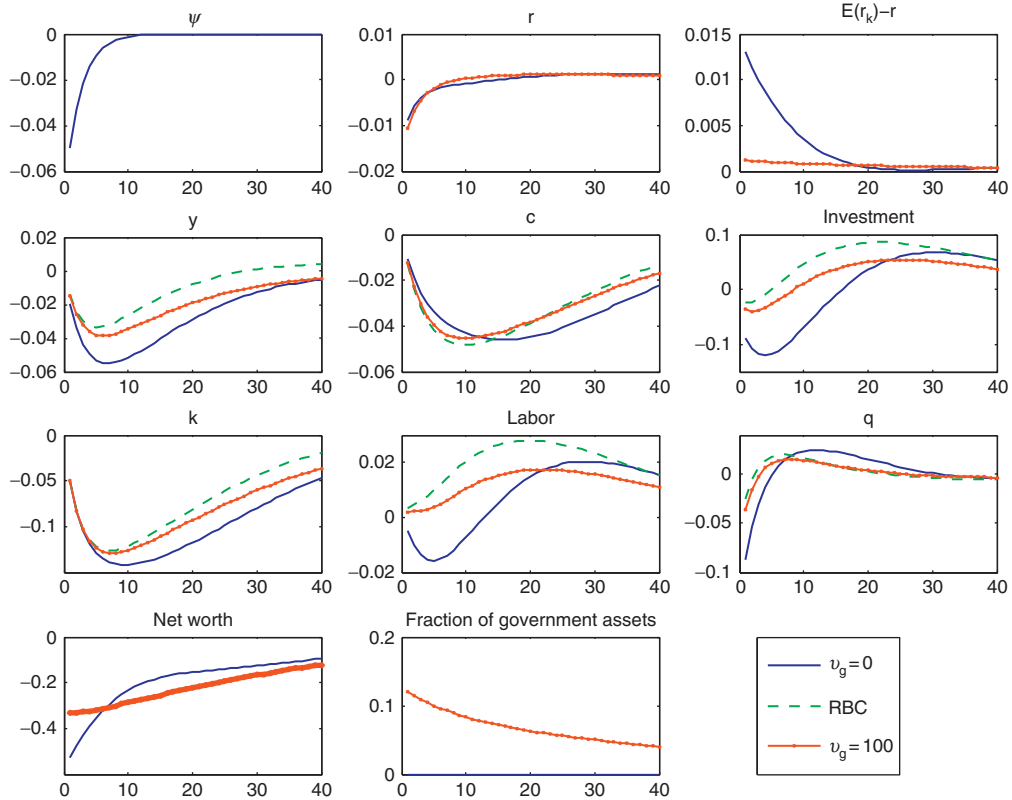


Figure 3 Lending facilities: Perfect interbank market.

allow for features like asset heterogeneity and so forth that would make it clearer how credit market interventions should be allocated between the three approaches.

Finally, although we do not do the exercise here, one can evaluate the net welfare benefits from the credit policy intervention, given different assumptions about the efficiency costs of direct central bank lending, following Gertler and Karadi (2009). As these authors show, however, under reasonable assumptions about these costs, the net benefits to the intervention are large and approximately equal to the gross benefits. They are also increasing in the severity of the crisis.

5. ISSUES AND EXTENSIONS

We now discuss some key issues in the literature that our baseline model does not consider. We also characterize how one might extend our framework to address these issues.

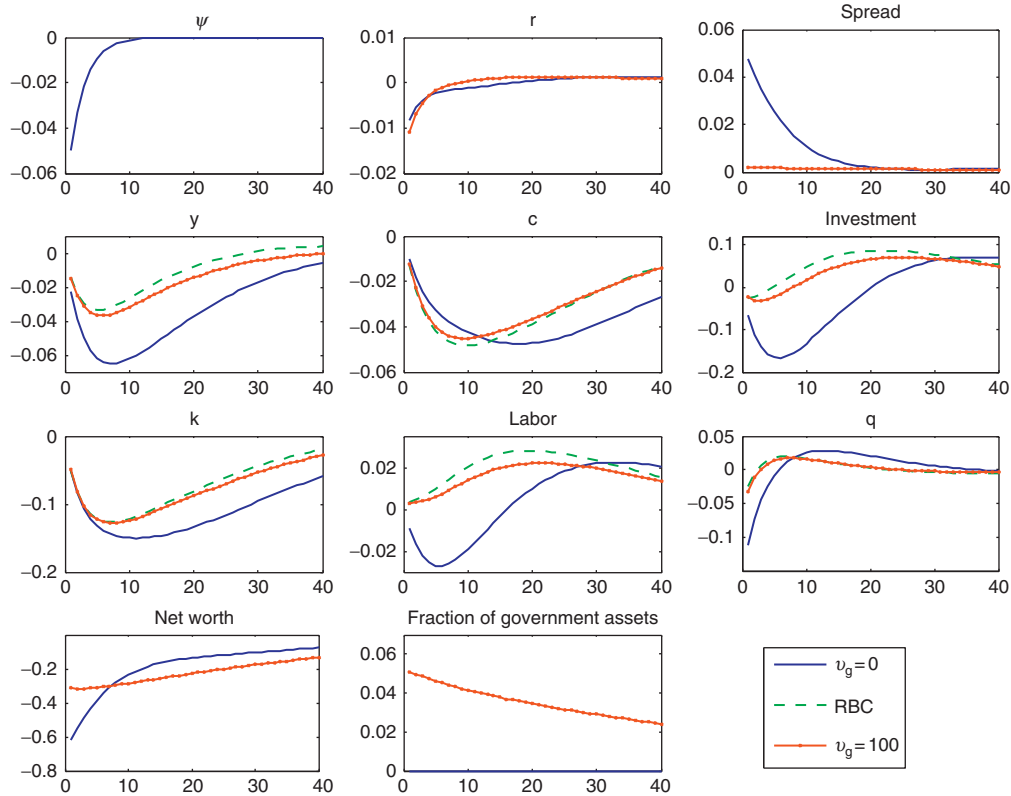


Figure 4 Lending facilities: Imperfect interbank market.

5.1 Tightening margins

Within our baseline model, financial distress is a product of deteriorating intermediary balance sheets: A decline in intermediary net worth forces a decline in the value of assets the intermediary can hold, given the constraint on its leverage ratio induced by the principal-agent problem.

Another complementary way that financial distress can transmit to the real economy is by a tightening of the leverage ratio, as emphasized by Adrian and Shin (2009), Brunnermeier and Pederson (2009), Kiyotaki and Moore (2008), Jermann and Quadrini (2009), Fostel and Geanakoplos (2008), Kurlat (2009), and others.

In the context of our model, any factor that might reduce the fraction of assets that lenders can expect to recover in a default will induce a tightening of margins. Recall that the fraction of assets that depositors can recover is $1 - \theta$, while banks who lend in the interbank can recover the fraction $1 - \theta(1 - \omega)$, with $0 < \omega < 1$. Suppose now that θ and ω might vary. The incentive constraint that determines that maximum leverage ratio becomes

$$V_t(s_t^h, b_t^h, d_t) \geq \theta_t(Q_t^h s_t^h - \omega_t b_t^h), \quad (63)$$

where the t subscripts θ_t and ω_t allow for the possibility of time variation.

An increase in θ_t and/or a reduction in ω_t clearly tightens the incentive constraint. One can then show that this leads to tightening of margins, since lenders will permit less borrowing for any given level of net worth. Kiyotaki and Moore (2008); Del Negro, Eggertsson, Ferrero, and Kiyotaki (2010); and Jermann and Quadrini (2009) used essentially this kind of mechanism to motivate a disruption of financial markets. Intuitively, θ_t is related inversely to the efficiency of the deposit market and the product $\theta_t(1 - \omega_t)$ is related to the efficiency of the interbank market. The less lenders are able to recover from borrowers in either of these markets, everything else equal, the less efficient the financial markets.

In the context of our model, one could imagine forces that lead θ_t and ω_t to move endogenously. For example, a deterioration in overall asset quality might make it more difficult for lenders to recover assets (particularly if the quality decline makes the assets relatively more specific to the borrowers), leading to an increase in θ_t . If the recovery problem is concentrated in the interbank market, then the deterioration in asset quality might induce a reduction in ω_t , causing the interbank market to contract. In either case, an endogenous response of θ_t and ω_t is likely to magnify the crisis.

There is work that attempts to model the tightening of margins explicitly. For example, Eissfeldt (2004) and Kurlat (2009) have frameworks where adverse selection problems are countercyclical. The greater degree of adverse selection in recessions causes a tightening of margins in the secondary financial market in downturns (which is similar to a reduction of ω_t). A much earlier paper by Williamson (1987) motivated something similar to an increase in θ_t in the primary financial market. In this framework, the agency problem that introduces the financial market friction is based on Townsend's (1979) costly state verification (CSV) model. Within the CSV model, the agency costs are expected default costs, which are increasing in the spread of the idiosyncratic shock to the borrower's return distribution. As Williamson showed, if the idiosyncratic risk is countercyclical, agency costs also become countercyclical, which leads a tightening of margins in downturns. Curdia (2007); Christiano, Motto, and Rostagno (2010); and Gilchrist, Yankov, and Zakresjek (2009) incorporated a similar mechanism in contemporary quantitative macroeconomic frameworks. Finally, Fostel and Geanakoplos (2008) also appeal to increases in uncertainty to motivate a tightening of margins, but do so in a setup with heterogeneous beliefs and disagreement.

Another way to allow for tightening of margins is to allow for a precautionary effect on asset holdings. Within our framework, given constant returns at the intermediary level, the leverage ratio is always binding: Banks always hold the maximum level of assets that their respective net worth permits. Aiyagari and Gertler (1999) and Mendoza (2009) relaxed this assumption. As they showed, even if the leverage (or margin) constraint is not

currently binding, an increased likelihood that it could be binding in the future (due possibly to increased uncertainty) can also induce a tightening of margins. Brunnermeier and Sannikov (2009) and He and Krishnamurthy (2008) also presented frameworks where precautionary effects can lead to a tightening of margins. Importantly, within these frameworks, the banks net worth still influences asset holdings.¹⁸ A stronger net worth position, everything else equal, reduces the likelihood the margin constraint will be binding, which encourages the intermediary to expand asset holdings.

5.2 Regulatory arbitrage and securitized lending

Because we are interested in capturing the interaction between banking and the macroeconomic conditions, our representation of the financial intermediary sector is quite parsimonious. We restrict attention to features of financial intermediation that we think are absolutely essential to characterizing this interaction. At the same time, our framework captures three basic aspects of banking that have been emphasized in the literature.¹⁹ First, banks act as delegated monitors. Because evaluating and monitoring borrowers requires specialized expertise, the financial intermediaries within our model operate as conduits that channel funds from households to firms. Second, banks engage in maturity transformation. They issue short-term liabilities and hold long-term assets. Third, they facilitate liquidity provision. Within our framework the interbank market (when it is functioning well) works to ensure that borrowers with idiosyncratic needs for funds receive them.

The banks within our model are best thought of as a consolidated representation of the financial intermediary sector, which includes commercial and investment banks. In this regard our baseline framework does not capture some notable details of the current financial crisis. In particular, a salient feature of the current crisis was the unraveling of the investment banks, which held securitized assets that in many instances were originated and sold off by commercial banks. However, we can extend our model to capture an aspect of this phenomenon.²⁰

In particular, suppose the banker operates a commercial bank that faces binding regulatory capital requirement. In reaction to this regulatory requirement the banker sets up a special purpose vehicle (SPV) that is not subject to the regulatory requirements on capital. The banker places in the SPV assets that the commercial bank originated and securitized. He funds the SPV partly by allocating some of his own net worth to the entity and partly by issuing short-term debt that is a perfect substitute for bank deposits.

¹⁸ These models also have constant returns at the intermediary level. However, they do not restrict attention to log linear approximations of the model and instead they allow for higher order effects of uncertainty on decision making.

¹⁹ See, for example, Diamond, (1984), Diamond and Dybvig (1983); Holmstrom and Tirole (1997); and the survey by Allen, Babus, and Carletti (2009) for discussions of basic aspects of banking.

²⁰ Shleifer and Vishny (2009) also emphasized the role of securitized lending in the crisis.

Think of the overall entity that the banker runs as a universal bank with the commercial bank and the SPV as separate entities. Because it operates off the commercial bank's balance sheet and holds securitized assets, the SPV may be thought of as an investment bank.

The key point is that the universal bank in this case will behave exactly like the financial intermediary in our baseline scenario. In particular, from the standpoint of the universal bank's creditors, what matters is its consolidated balance sheet and not the breakdown of assets and liabilities between the commercial bank and the SPV. Thus, agency problem between the banker and his creditors introduces a maximum permissible leverage ratio for the universal bank as a whole. For simplicity, we abstract from liquidity risks (i.e., $\pi^i = 1$) so that asset prices are equalized across regions. Then it is straightforward to show that the maximum leverage ratio for the universal bank is ϕ , as given by Eq. (20).

Now suppose that the maximum regulatory leverage on the commercial bank ϕ^b is lower than the privately determined value ϕ . In addition, suppose that the SPV is able to operate with a leverage ratio ϕ_t^{spv} that exceeds ϕ :

$$\phi^b < \phi < \phi_t^{spv},$$

where the superscript b denotes commercial bank and the superscript spv denote SPV. Then the universal bank can always find a division of assets and net worth of the commercial bank and the SPV, which satisfies the capital requirement on the commercial bank while at the same time satisfying the privately determined leverage constraint for the universal bank:

$$\begin{aligned} Q_t s_t^b &\leq \phi^b n_t^b \\ Q_t s_t^{spv} &\leq \phi_t^{spv} n_t^{spv} \\ Q_t (s_t^b + s_t^{spv}) &= \phi (n_t^b + n_t^{spv}). \end{aligned} \tag{64}$$

Here, the universal bank uses the SPV and securitization to circumvent the regulation on the commercial bank.²¹ The only binding leverage constraint is the consolidated leverage constraint (Eq. 64), which results from the incentive constraint of the universal bank. Then, while the model now contains securitized lending and assets held off commercial bank balance sheets, the macroeconomic equilibrium is the same as in our baseline framework. Thus, at a first pass, the addition of these features does not alter the predictions of the model about the feedback between the financial and real sectors that magnifies the crisis. Our enriched model will predict that during a crisis, investment banking, securitized lending, and commercial banking will all be disrupted, as happened in practice.

²¹ In practice, a key factor in the growth of investment banks holding securitized assets was the increase in capital requirements on commercial banks, phased in after the banking crises of the 1980s.

Here we have made the strong assumption that the commercial bank and the SPV have a single ownership. It would be interesting to relax this assumption. At the same time, during the crisis, the commercial bank and the SPV did not have a completely arm's length relationship. In many instances as the crisis unfolded commercial banks repurchased securitized assets they had originally sold to other institutions. It would be useful to try to capture this implicit relationship between commercial banks and SPV.

5.3 Outside equity, externalities, and moral hazard

Our baseline presumes that the only type of liability the bank can issue to raise funds is short-term, noncontingent debt. We now explore the possibility that the bank can issue fully state-contingent debt or, equivalently, outside equity. As we show, outside equity issuance is desirable because it provides a hedge to the bank against fluctuations in its net worth. At the same time we consider how an agency problem might limit a bank's use of outside equity financing. We also show that externalities and the anticipation of government credit market intervention can lead a bank to rely too little on outside equity, which introduces a possible role for regulatory capital requirements.

We now allow bankers to issue outside equity. We suppose that a unit of outside equity entitles the holder to the same dividend payout per share as banker's asset. Let q_t be market price of a unit of outside bank equity and e_t the quantity issued. We restrict attention to the case of perfect interbank market (i.e., $\omega = 1$) and refer the reader to the Appendix in this chapter for a more general treatment. Then the bank's balance sheet is given by

$$Q_t s_t = n_t + b_t + d_t + q_t e_t \quad (65)$$

The flow of funds constraint becomes

$$n_t = [Z_t + (1 - \delta)Q_t]\psi_t s_{t-1} - [Z_t + (1 - \delta)q_t]\psi_t e_{t-1} - R_b b_{t-1} - R_t d_{t-1} \quad (66)$$

By issuing outside equity the bank is able to have its creditors share part of the risk in the payoff to its loan portfolio. For example, a negative capital quality shock (fall in ψ_t) is not absorbed entirely by the bank but also by the bank's outside equity holders. Put differently, by issuing outside equity, the bank reduces its leverage ratio and, by doing so, reduces the volatility of its net worth.

Given the hedging value that outside equity affords, everything else equal, the bank would prefer to replace its noncontingent debt with perfectly state-contingent equity. Accordingly, everything else equal, the bank gains by reducing the volatility of its net worth. This then begs the question of why banks do not fund assets with equity or fully state-contingent debt. A classic argument by Calomiris and Kahn (1991) is that short-term debt provides a disciplining device on bank behavior. The need to meet continual

noncontingent payments reduces the degree to which a bank can in any way act against the interest of its creditors to favor its owners.

One way to illustrate the Calomiris and Kahn (1991) argument in the context of our model is as follows: Suppose that it is easier for the banker to divert assets funded by equity than assets funded by deposit. It may take time for outside equity holders to assess whether a suspension or reduction of dividend payments reflects the true condition of bank assets or some malfeasance on the part of the banker. On the other hand, because deposits require immediate payment, it is difficult for the banker to quickly divert funds.

To be concrete, suppose that the bank can divert the fraction $\theta(1 - \omega_e)$ of assets fund by equity where $\omega_e < 0$, but only the fraction θ funded by short-term debt. (The banker cannot divert assets fund by interbank loan since $\omega = 1$ here). We can now express the incentive constraint as:

$$V_t(s_t, b_t, d_t, e_t) \geq \theta(Q_t s_t - \omega_e q_t e_t - b_t) \quad (67)$$

where $V_t(s_t, b_t, d_t, e_t)$ is the bank's continuation value conditional on it raising funds by outside equity as well as by debt. The second term on the right reflects the fact that it is easier for the bank to divert assets funded by equity (as $\omega_e < 0$).

Let R_{et+1} be there turn on bank equity:

$$R_{et+1} = \psi_{t+1} \frac{Z_{t+1} + (1 - \delta)q_{t+1}}{q_t}$$

Then as the Appendix shows, the first-order conditions from the banks portfolio structure problem are given

$$E_t \Lambda_{t,t+1} \Omega_{t+1} (R_{t+1} - R_{et+1}) = (-\omega_e) \cdot E_t \Lambda_{t,t+1} \Omega_{t+1} (R_{kt+1} - R_{t+1}) \quad (68)$$

If the incentive constraint is binding then following the reasoning in Section 2, there are excess returns to bank assets; that is, the expected discounted return to bank assets $E_t \Lambda_{t,t+1} \Omega_{t+1} R_{kt+1}$ exceeds the expected discounted cost of bank deposits, $E_t \Lambda_{t,t+1} \Omega_{t+1} R_{t+1}$. This makes the right side of the equation positive. The left side then implies that for banks to be issuing both deposits and outside equity, the discounted cost of the outside equity, $E_t \Lambda_{t,t+1} \Omega_{t+1} R_{et+1}$, must be less than that of that of deposits.. Intuitively, changing the mix of financing from deposits to outside equity tightens the incentive constraint. For the bank to be indifferent between the financing sources, the cost of outside equity must be less than the cost of deposits.

The household's portfolio decision introduces the following arbitrage relation between the deposit rate and there turn on bank equity.

$$E_t \Lambda_{t,t+1} R_{t+1} = E_t \Lambda_{t,t+1} R_{et+1} \quad (69)$$

Observe that the household discounts the stock return R_{et+1} by the stochastic factor $\Lambda_{t,t+1}$ while the banker uses a discount factor that is augmented by the shadow value

of net worth Ω_{t+1} , which varies countercyclically. The net effect is that the banker's expected discounted cost of issuing equity is less than the household's expected discounted return to holding it. The difference is due to the fact that outside equity provides a hedge for the bank against fluctuations in net worth, something which the bank values directly but the household does not.

To understand the implications for the bank's liability structure, first consider the case where $\omega_e = 0$; that is, shifting from deposit finance to outside equity does not enhance the enforcement problem. It follows from Eq. (68) that for the bank to use both financing options, the cost must be equal to the banker. Otherwise it will exclusively use the lower cost option. Given that the household's arbitrage condition governs the link between the deposit rate and the return on bank equity, it is straightforward to show that, due to its hedging value, outside equity offers the lower cost financing option for the bank. Thus in this instance, the bank would choose to finance exclusively with outside equity (or, equivalently) fully state-contingent debt.

The situation changes, however, if outside equity enhances the incentive problems. If ω_e is sufficiently negative (meaning that outside equity is subject to a significantly greater agency problem than are deposits), the bank may not be able to offer a return on bank equity that is competitive with the return on deposits. In this instance, the bank will resort exclusively to deposit finance.

Thus, one can appeal to an agency problem to motivate why the bank might rely mainly on noncontingent deposits as opposed to outside equity. But here it is important to recognize that there is an externality present in private sector financial structure decisions. In particular, as Section 2 makes clear, the volatility of returns on banks and conversely the volatility of the economy depends on the aggregate balance sheet of the intermediary sector as opposed to the balance sheet of any individual intermediary. That is, it is the leverage ratio of the sector as a whole that makes the financial system vulnerable to disturbances.

Individual banks do not take into account the effects of their own liability structure on the aggregate. At the bank level, this distorts the decision in favor of debt financing and away from the use of outside equity. As a consequence, the aggregate balance sheet features more leverage than a social planner would prefer. This raises the possibility that some form of capital requirements may be optimal. Korinek (2009) and Lorenzoni (2008) have made similar types of arguments.

The introduction of an endogenous choice of equity also raises the issue of moral hazard from the anticipation of policy interventions. The credit policies we described earlier work to stabilize the volatility in banks' shadow value of net worth. Doing so, however, reduces the bank's incentive to resort to outside equity financing. This in turn raises the aggregate leverage in the intermediary sector, increasing the likelihood of another crisis that might require government intervention. Tracing out these moral hazard consequences is an important direction for future research. Some recent work that has explored this issue in a different setting from ours includes Diamond and Rajan (2009), Farhi and Tirole

(2009), and Chari and Kehoe (2010). In our view, capturing the quantitative implications of moral hazard is particularly important for policy evaluation.

6. CONCLUDING REMARKS

If nothing else, we hope that this chapter helps dispel the notion that macroeconomists have not paid attention to the financial sector. As we have seen, over the past twenty years there has been a steady stream of research that incorporates financial frictions into macroeconomic analysis. The recent crisis has precipitated an uptick in the pace of this research and offered many new issues to study.

One difference between research over the past decade as compared to earlier has been an emphasis on developing frameworks suitable for quantitative analysis. We view this as a welcome development since many of the issues involving the role of financial factors in the business cycle and the implications for both credit and regulatory policies ultimately involve quantitative considerations. Our best guess is that at the time the next *Handbook* chapter on this topic is written, the authors will be reviewing macroeconomic models with financial sectors that perform credibly from an empirical standpoint and that provide sharp insights for public policy.

APPENDIX 1

A general model with interbank friction

Here we lay out the general framework with an imperfect interbank market ($\omega < 1$). We abstract from outside equity and government interventions for the exposition. (Appendix 2 will present a framework that includes outside equity and government.) For an equilibrium in which the bank makes loans, issues deposits, and conducts interbank borrowing and lending, the first-order conditions for the bank's choice of (s_t^h, d_t) are Eqs. (14) and (15). The incentive constraint (16) can be rewritten as

$$\begin{aligned} & \{[\theta(1 - \omega) + \mathcal{V}_{bt}]Q_t^h - \mathcal{V}_{st}\}s_t^h \\ & \leq (\mathcal{V}_{bt} - \theta\omega)n_t^h - (\theta\omega + \mathcal{V}_t - \mathcal{V}_{bt})d_t, \end{aligned} \quad (70)$$

where Eq. (70) holds with equality if $\lambda_t^h > 0$, and the strict inequality implies $\lambda_t^h = 0$. For the general case with $\omega < 1$, we have from Eq. (15):

$$\lambda_t^h = \frac{\frac{\mathcal{V}_d}{Q_t^h} - \mathcal{V}_{bt}}{\theta(1 - \omega) - \left(\frac{\mathcal{V}_{st}}{Q_t^h} - \mathcal{V}_{bt}\right)}. \quad (71h)$$

The numerator indicates how much the value of the bank in type h island increases with an additional dollar's worth purchase of a security financed by interbank borrowing ($ds_t^h = 1/Q_t^h$, $db_t^h = 1$). The denominator indicates how much the incentive

constraint is tightened (i.e., RHS minus LHS of Eq. 11 increases) with an additional dollar purchase of the security.

As in the text, we conjecture that the price of security is lower in the investing region than the noninvesting region due to abundant supply: $Q_t^i < Q_t^n$. Then from Eq. (71), we learn

$$\lambda_t^i > \lambda_t^n \geq 0. \quad (72)$$

From Eq. (14), we get

$$\mathcal{V}_{bt} - \mathcal{V}_t = \frac{\theta\omega\bar{\lambda}_t}{1 + \bar{\lambda}_t} > 0. \quad (73)$$

Thus we learn that the marginal cost of interbank borrowing exceeds the marginal cost of deposit, $\mathcal{V}_{bt} > \mathcal{V}_t$. Using these first-order conditions, (70) can be rewritten as

$$Q_t^h s_t^h \leq \frac{1}{\theta(1 - \omega) - \left(\frac{\mathcal{V}_t}{Q_t^h} - \mathcal{V}_{bt}\right)} \left[(\mathcal{V}_{bt} - \theta\omega)n_t^h - \frac{\theta\omega}{1 + \bar{\lambda}_t} d_t \right]. \quad (74)$$

Substituting the first-order conditions and the incentive constraint (74) into the value function (13), we learn

$$V_t(s_t^h, b_t^h, d_t) = [\mathcal{V}_{bt} + \lambda_t^h(\mathcal{V}_{bt} - \theta\omega)]n_t^h + \theta\omega \frac{\bar{\lambda}_t - \lambda_t^h}{1 + \bar{\lambda}_t} d_t.$$

The term $\mathcal{V}_{bt} + \lambda_t^h(\mathcal{V}_{bt} - \theta\omega)$ is the marginal value of net worth to the active banker: With an additional unit of net worth, the banker can reduce the interbank borrowing by one unit (which saves costs by $\mathcal{V}_{bt}$), and relax the incentive constraint by $\mathcal{V}_{bt} - \theta\omega$ (which increases the value of bank by λ_t^h times as much). Substituting this expression for date $t+1$ into the Bellman Eq. (12) yields

$$\begin{aligned} V_t(s_t^h, b_t^h, d_t) &= \mathcal{V}_{st}s_{pt} - \mathcal{V}_{bt}b_t - \mathcal{V}_td_t \\ &= E_t A_{t,t+1} \Omega_{t+1}^{h'} n_{t+1}^{h'}, \end{aligned} \quad (75)$$

where

$$\Omega_t^h = 1 - \sigma + \sigma[\mathcal{V}_{bt} + \lambda_t^h(\mathcal{V}_{bt} - \theta\omega)] \quad (76)$$

is the marginal value of net worth for the banker, who exits with probability $1 - \sigma$ and stays active with probability σ . Applying the method of undetermined coefficient to Eq. (75), we learn

$$\mathcal{V}_{bt} = R_{bt+1} E_t A_{t,t+1} \Omega_{t+1}^{h'}, \quad (77)$$

$$\mathcal{V}_t = R_{t+1} E_t A_{t,t+1} \Omega_{t+1}^{h'} = \frac{R_{t+1}}{R_{bt+1}} \mathcal{V}_{bt}, \quad (78)$$

$$\mathcal{V}_{st} = E_t \Lambda_{t,t+1} \Omega_{t+1}^{h'} [Z_{t+1} + (1 - \delta) Q_{t+1}^{h'}] \psi_{t+1}. \quad (79)$$

Let D_t be aggregate value of deposit of the banks. Then from Eqs. (72) and (74), we have

$$Q_t^i S_t^i = \frac{1}{\theta(1 - \omega) - \left(\frac{\mathcal{V}_{st}}{Q_t^i} - \mathcal{V}_{bt}\right)} \left[(\mathcal{V}_{bt} - \theta\omega) N_t^i - \frac{\theta\omega}{1 + \bar{\lambda}_t} \pi^i D_t \right], \quad (80)$$

$$Q_t^n S_t^n \leq \frac{1}{\theta(1 - \omega) - \left(\frac{\mathcal{V}_{st}}{Q_t^n} - \mathcal{V}_{bt}\right)} \left[(\mathcal{V}_{bt} - \theta\omega) N_t^n - \frac{\theta\omega}{1 + \bar{\lambda}_t} \pi^n D_t \right], \quad (81)$$

where Eq. (81) holds with equality if $\lambda_t^n > 0$, and the strict inequality implies $\lambda_t^n > 0$. The marginal propensity to buy assets with respect to net worth is

$$\phi_t^h = \frac{\mathcal{V}_{bt} - \theta\omega}{\theta(1 - \omega) - \left(\frac{\mathcal{V}_{st}}{Q_t^h} - \mathcal{V}_{bt}\right)}$$

which is the expression for the leverage ratio in the general case of an imperfect inter-bank market. (Observe that this expression becomes Eqs. 27 and 28 if $\omega = 0$).

The rest of the framework is the same as the model in the text. From Eqs. (34)–(36), the aggregate net worth of the bank in investing islands and noninvesting islands satisfies

$$N_t^h = \pi^h \{ [Z_t + (1 - \delta) Q_t^h] \psi_t (\sigma + \xi) S_{t-1} - \sigma R_t D_{t-1} \}. \quad (82h)$$

(A_t, ψ_t) follows an exogenous stochastic process. Then, four prices ($Q_t^i, Q_t^n, R_{t+1}, R_{bt+1}$) and eleven quantities ($Y_t, C_t, L_t, I_t, K_{t+1}, Z_t, D_t, N_t^i, N_t^n, S_t^i, S_t^n$) together with five shadow prices ($\mathcal{V}_t, \mathcal{V}_{bt}, \mathcal{V}_{st}, \lambda_t^i, \lambda_t^n$) are determined as a function of the state variables ($K_t, C_{t-1}, I_{t-1}, A_t, \psi_t, R_t D_{t-1}$) by the sequence of twenty equations: the optimization conditions of households and nonfinancial firms (1, 2, 7, 39, 40), the optimization of banks (71i, 71h, 73, 77 – 81, 82i, 82h), and the market clearing conditions for goods, interbank market funds, securities, and labor (3, 37, 41i, 41n, 42).

Steady state

In the steady state, we have

$$I = \delta K \quad (83)$$

$$C = \left[A \left(\frac{L}{K} \right)^{1-\alpha} - \delta \right] K \quad (84)$$

$$\chi L^\varepsilon = (1 - \alpha) A \left(\frac{K}{L} \right)^\alpha \frac{1 - \beta\gamma}{1 - \gamma} \frac{1}{C} \quad (85)$$

$$Z = \alpha A \left(\frac{L}{K} \right)^{1-\alpha} \quad (86)$$

$$R = \frac{1}{\beta} \quad (87)$$

$$Q^i = 1 \quad (88)$$

We also have

$$N^i = \pi^i \left[(\sigma + \xi)(Z + 1 - \delta)K - \frac{\sigma}{\beta} D \right], \quad (89)$$

$$N^n = \pi^n \left[(\sigma + \xi)[Z + (1 + \delta)Q^n]K - \frac{\sigma}{\beta} D \right], \quad (90)$$

$$N^i + N^n + D = K + \pi^n(Q^n - 1)(1 - \delta)K. \quad (91)$$

The security market equilibrium implies

$$[\delta + \pi^i(1 - \delta)]K = \frac{1}{\theta(1 - \omega) + \mathcal{V}_b - \mathcal{V}_{st}} \left[(\mathcal{V}_b - \theta\omega)N^i + \frac{\pi^i\theta\omega}{1 + \bar{\lambda}} D \right] \quad (92)$$

$$Q^n\pi^n(1 - \delta)K \leq \frac{1}{\theta(1 - \omega) + \mathcal{V}_b - \frac{\mathcal{V}_s}{Q^n}} \left[(\mathcal{V}_b - \theta\omega)N^n + \frac{\pi^n\theta\omega}{1 + \bar{\lambda}} D \right] \quad (93)$$

where equality holds if $\lambda^n > 0$ while the strict inequality implies $\lambda^n = 0$. Concerning the optimization of the bank, we have

$$\lambda^i = \frac{\mathcal{V}_s - \mathcal{V}_b}{\theta(1 - \omega) - (\mathcal{V}_s - \mathcal{V}_b)}, \quad (94)$$

$$\lambda^n = \frac{\frac{\mathcal{V}_s}{Q^n} - \mathcal{V}_b}{\theta(1 - \omega) - \left(\frac{\mathcal{V}_s}{Q^n} - \mathcal{V}_b \right)}, \quad (95)$$

$$\mathcal{V}_b = \beta R_b [1 - \sigma + \sigma \mathcal{V}_b + \sigma \bar{\lambda}(\mathcal{V}_b - \theta\omega)], \quad (96)$$

$$\mathcal{V}_b - \mathcal{V} = \left(1 - \frac{1}{\beta R_b} \right) \mathcal{V}_b = \frac{\theta\omega\bar{\lambda}}{1 + \bar{\lambda}}, \quad (97)$$

$$\mathcal{V}_s = \beta\pi^i (Z + 1 - \delta)[1 - \sigma + \sigma\mathcal{V}_b + \sigma\lambda^i(\mathcal{V}_b - \theta\omega)] + \beta\pi^n [Z + Q^n(1 - \delta)][1 - \sigma + \sigma\mathcal{V}_b + \sigma\lambda^n(\mathcal{V}_b - \theta\omega)] \quad (98)$$

The steady-state equilibrium is recursive: The values of eleven prices and ratio variables $(R_b, Q^n, Z, \lambda^i, \lambda^n, \mathcal{V}_b, \mathcal{V}, \mathcal{V}_s, \frac{N^i}{K}, \frac{N^n}{K}, \frac{D}{K})$ are determined by eleven equations (89)–(98) where (97) has two equations. Then quantity variables (K, I, C, L) are determined by Eqs. (83)–(86).

APPENDIX 2

A general model with outside equity and government intervention

Here we lay out a general framework with an imperfect interbank market ($\omega < 1$) and with outside equity and credit policies. At the beginning of each period (before the arrival of investment opportunity to nonfinancial firms), each bank learns whether to exit or stay active at the end of this period. The active bank raises fund from households by issuing deposit d_t and outside equity e_t at price q_t . The government may buy additional equity $s_{get} - (1 - \delta)\psi_t s_{get}$ from active banks at price Q_{gt} . Outside equity held by households and government pays the same dividend as a security issued by nonfinancial firms. During this period (after the arrival of investment opportunity to nonfinancial firms), the active bank can raise funds by borrowing at interbank market b_t^h and at the discount window m_t^h to partially finance the loan (purchase of security of the nonfinancial firms). The flow of fund constraint of an active bank on type h island is

$$Q_t^h s_{pt}^h = n_t^h + b_t^h + m_t^h + q_t e_t + d_t, \quad (99)$$

where $s_{pt}^h = s_t^h - s_{get}$ is the private holding of the security. The net worth of active bank is defined similarly to Eq. (57) as

$$\begin{aligned} n_t^h = & [Z_t + (1 - \delta)Q_t^h]\psi_t s_{pt-1} - [Z_t + (1 - \delta)q_t]\psi_t e_{t-1} - R_{bt}b_{t-1} \\ & - R_{mt}m_{t-1} - R_t d_{t-1} + (Q_{gt} - Q_t^h)[s_{get} - (1 - \delta)\psi_t s_{get-1}] \end{aligned} \quad (100)$$

The last term is the government “gift” to each banker via an equity injection. Because we assume the government gives the gift to bankers lump sum (including the new entrants), we have $s_{get} = S_{get}/f$. The value of the bank at the end of this period is equal to the expected present value of the future dividend (which is equal to the net worth at the time of exit):

$$V_t = E_t \sum_{i=1}^{\infty} (1 - \sigma)\sigma^{i-1} \Lambda_{t,t+i} \tilde{n}_{t+i}^h$$

where the net worth of the exiting bank does not include the gift:

$$\begin{aligned} \tilde{n}_t^h = & [Z_t + (1 - \delta)Q_t^h]\psi_t s_{pt-1} - [Z_t + (1 - \delta)q_t]\psi_t e_{t-1} - R_{bt}b_{t-1} \\ & - R_{mt}m_{t-1} - R_t d_{t-1}. \end{aligned}$$

The incentive constraint implies the value of the active bank must be at least as large as the value of divertable assets:

$$V_t(s_{pt}^h, b_t^h, m_t^h, e_t, d_t) \geq \theta(Q_t^h s_{pt}^h - \omega b_t^h - \omega_g m_t^h - \omega_e q_t e_t). \quad (101)$$

As in the text, we assume the bank cannot divert assets acquired by government equity injection. On the other hand, the bank can divert the asset financed by outside equity more easily than the deposit; that is, $\omega_e < 0$.

Guessing the value function is linear in the arguments yields:

$$\begin{aligned} V_t^h &= V_t(s_t^h, b_t^h, m_t^h, e_t, d_t) \\ &= \mathcal{V}_{st}s_{pt} - \mathcal{V}_{bt}b_t - \mathcal{V}_{mt}m_t - \mathcal{V}_{et}e_t - \mathcal{V}_td_t + \mathcal{V}_{get}, \end{aligned} \quad (102)$$

and let λ_t^h be the Lagrangian multiplier for the incentive constraint of the bank in h island. Then using Eq. (99), the Lagrangian is

$$\begin{aligned} \mathcal{L} &= V_t^h + \lambda_t^h[V_t^h - \theta(Q_t^h s_{pt}^h - \omega b_t^h - \omega_g m_t^h - \omega_e q_t e_t)] \\ &= (1 + \lambda_t^h)[(\mathcal{V}_{st} - \mathcal{V}_{bt}Q_t^h)s_{pt}^h + (\mathcal{V}_{bt} - \mathcal{V}_{mt})m_t^h + (\mathcal{V}_{bt} - \mathcal{V}_t)d_t \\ &\quad + (\mathcal{V}_{bt}q_t - \mathcal{V}_{et})e_t + \mathcal{V}_{bt}n_t + \mathcal{V}_{get}] \\ &\quad - \lambda_t^h\theta[(1 - \omega)Q_t^h s_{pt}^h + (\omega - \omega_g)m_t^h + (\omega - \omega_e)q_t e_t + \omega(n_t^h + d_t)]. \end{aligned}$$

We focus on the equilibrium in which the bank makes loans, deposits, and conducts interbank borrowing and lending, but may or may not issue outside equity or use the discount window. Then, the first-order conditions for the bank's choice of (s_t^h, m_t^h, e_t, d_t) are given by Eqs. (14) and (15) in the text and

$$(1 + \lambda_t^h)(\mathcal{V}_{bt} - \mathcal{V}_{mt}) \leq \theta(\omega - \omega_g)\lambda_t^h, \quad (= \text{if } n_t^h > 0), \quad (103)$$

$$(1 + \bar{\lambda}_t)(\mathcal{V}_{bt}q_t - \mathcal{V}_{et}) \leq \theta(\omega - \omega_e)\bar{\lambda}_t q_t, \quad (= \text{if } e_t > 0). \quad (104)$$

The incentive constraint (101) can be rewritten as

$$\begin{aligned} \{[\theta(1 - \omega) + \mathcal{V}_{bt}]Q_t^h - \mathcal{V}_{st}\}s_{pt}^h &\leq (\mathcal{V}_{bt} - \theta\omega)n_t^h - (\theta\omega + \mathcal{V}_t - \mathcal{V}_{bt})d_t \\ -[\theta(\omega - \omega_g) + \mathcal{V}_{mt} - \mathcal{V}_{bt}]m_t^h &- [\theta(\omega - \omega_e)q_t + \mathcal{V}_{et} - \mathcal{V}_{bt}q_t]e_t + \mathcal{V}_{get}, \end{aligned} \quad (105)$$

where Eq. (105) holds with equality if $\lambda_t^h > 0$, and the strict inequality implies $\lambda_t^h = 0$.

From Eq. (103), we learn

$$\mathcal{V}_{mt} - \mathcal{V}_{bt} \geq \frac{\theta(\omega_g - \omega)\lambda_t^i}{1 + \lambda_t^i} > \frac{\theta(\omega_g - \omega)\lambda_t^n}{1 + \lambda_t^n}. \quad (106)$$

Thus banks in the noninvesting island do not use the discount window borrowing, while banks in the investing island use it only if the first weak inequality holds with equality. We also learn from Eq. (106) that the marginal cost of the discount window

has to be larger than the marginal cost of interbank borrowing ($\mathcal{V}_{mt} > \mathcal{V}_{bt}$) when both facilities are used. From Eq. (104), we have

$$\mathcal{V}_{bt}q_t - \mathcal{V}_{et} \leq \frac{\theta(\omega - \omega_e)\bar{\lambda}_t}{1 + \bar{\lambda}_t}q_t, (= \text{ if } e_t > 0). \quad (107)$$

Thus, for the bank to issue outside equity to the households, the marginal benefit of saving the cost of interbank borrowing must be larger than the marginal cost of outside equity ($\mathcal{V}_{bt}q_t > \mathcal{V}_{et}$), when the bank can divert the asset more easily when financed by outside equity than interbank borrowing ($\omega > \omega_e$). Using these first-order conditions, Eq. (105) can be rewritten as

$$\begin{aligned} \left[\theta(1 - \omega) - \left(\frac{\mathcal{V}_{st}}{Q_t^h} - \mathcal{V}_{bt} \right) \right] Q_t^h s_t^h &\leq (\mathcal{V}_{bt} - \theta\omega)n_t^h + \frac{\theta(\omega_g - \omega)}{1 + \lambda_t^h} m_t^h \\ - \frac{\theta}{1 + \bar{\lambda}_t} [\omega d_t + (\omega - \omega_e)q_t e_e] + \mathcal{V}_{get} \end{aligned} \quad (108)$$

Substituting the first-order conditions and the incentive constraint (108) into the value function (102), we learn

$$\begin{aligned} V_t^h &= [\mathcal{V}_{bt} + \lambda_t^h(\mathcal{V}_{bt} - \theta\omega)]n_t^h \\ &\quad + \theta \frac{\bar{\lambda}_t - \lambda_t^h}{1 + \bar{\lambda}_t} [\omega d_t + (\omega - \omega_e)q_t e_e] + (1 + \lambda_t^h)\mathcal{V}_{get}. \end{aligned}$$

Substituting this expression for date $t+1$ into the Bellman equation (102), we learn

$$\begin{aligned} V_t &= \mathcal{V}_{st}s_{pt} - \mathcal{V}_{bt}b_t - \mathcal{V}_{mt}m_t - \mathcal{V}_{et}e_t - \mathcal{V}_t d_t + \mathcal{V}_{get} \\ &= E_t \Lambda_{t,t+1} [\Omega_{t+1}^{h'} n_{t+1}^{h'} + \sigma(1 + \lambda_{t+1}^{h'})\mathcal{V}_{get+1}], \end{aligned} \quad (109)$$

where Ω_t^h is given by Eq. (76). Applying the method of undetermined coefficient to Eq. (109), we learn Eqs. (77)–(79) and

$$\mathcal{V}_{mt} = R_{mt+1} E_t \Lambda_{t,t+1} \Omega_{t+1}^{h'} = \frac{R_{mt+1}}{R_{bt+1}} \mathcal{V}_{bt}, \quad (110)$$

$$\mathcal{V}_{et} = E_t \Lambda_{t,t+1} \Omega_{t+1}^{h'} [Z_{t+1} + (1 - \delta)q_{t+1}] \psi_{t+1} \quad (111)$$

$$\begin{aligned} \mathcal{V}_{get} &= E_t \Lambda_{t,t+1} \sigma \{ (1 + \lambda_{t+1}^{h'}) \mathcal{V}_{get+1} + \sigma [\mathcal{V}_{bt+1} + \lambda_{t+1}^{h'} (\mathcal{V}_{bt+1} - \theta\omega)] \\ &\quad \cdot (Q_{get+1} - Q_{t+1}^h) [s_{get+1} - (1 - \delta)\psi_t s_{get}] \}. \end{aligned} \quad (112)$$

Let M_t , $\bar{E}_t$ and D_t be aggregate value of discount window borrowing, outside equity and deposit of the banks. Then from Eq. (108), we have

$$S_{pt}^i = \frac{1}{[\theta(1-\omega) + \mathcal{V}_{bt}]Q_t^i - \mathcal{V}_{st}} \cdot \left\{ (\mathcal{V}_{bt} - \theta\omega)N_t^i + \frac{\theta(\omega_g - \omega)}{1 + \lambda_t^h} M_t - \frac{\pi^i \theta}{1 + \bar{\lambda}_t} [\omega D_t + (\omega - \omega_e)q_t \bar{E}_t] + \pi^i f \mathcal{V}_{get} \right\} \quad (113)$$

$$S_{pt}^n \leq \frac{1}{[\theta(1-\omega) + \mathcal{V}_{bt}]Q_t^n - \mathcal{V}_{st}} \cdot \left\{ (\mathcal{V}_{bt} - \theta\omega)N_t^n - \frac{\pi^n \theta}{1 + \bar{\lambda}_t} [\omega D_t + (\omega - \omega_e)q_t \bar{E}_t] + \pi^n f \mathcal{V}_{get} \right\}, \quad (114)$$

where Eq. (115) holds with equality if $\lambda_t^n > 0$, and the strict inequality implies $\lambda_t^n = 0$. The aggregate net worth of the banks in investing islands and noninvesting islands are similar to Eq. (59) as

$$N_t^h = \pi^h \{ [Z_t + (1-\delta)Q_t^h] \psi_t (\sigma + \xi) S_{pt-1} - \sigma [Z_t + (1-\delta)q_t] \psi_t \bar{E}_{t-1} - \sigma R_{mt} M_{t-1} - \sigma R_t D_{t-1} + \sigma (Q_{gt} - Q_t^h) [S_{get} - (1-\delta) \psi_t S_{get-1}] \}. \quad (115h)$$

The security market equilibrium implies

$$I_t + \pi^i (1-\delta)K_t = S_{pt}^i + S_{gt}^i + \pi^i S_{get} \quad (116)$$

$$\pi^n (1-\delta)K_t = S_{pt}^n + S_{gt}^n + \pi^n S_{get}. \quad (117)$$

The flow of fund constraint of entire banking sector (which implies the interbank market clearing) is

$$Q_t^i S_{pt}^i + Q_t^n S_{pt}^n = N_t^i + N_t^n + M_t + D_t + q_t \bar{E}_t. \quad (118)$$

The rest of the framework is the same as the model in the text, except that the household's budget constraint (5) includes the purchase of the outside equity

$$C_t = W_t L_t + \prod_t - T_t + R_t (D_t + D_{gt}) - (D_{t+1} + D_{gt+1}) + [Z_t + (1-\delta)q_t] \psi_t \bar{E}_{t-1} - q_t \bar{E}_t.$$

Thus the first-order condition for the outside equity purchase is

$$q_t = E_t \{ \Lambda_{t,t+1} [Z_{t+1} + (1-\delta)q_{t+1}] \psi_{t+1} \}. \quad (119)$$

Comparing this expression of household's valuation of equity and the banker's valuation (111), we learn that the household's discount factor is the marginal rate of substitution of consumption $\Lambda_{t,t+1}$, while the banker's discount factor is the marginal rate of substitution *times* the marginal value of net worth $\Lambda_{t,t+1} \Omega_{t+1}'$. And the banker's discount factor is more volatile than the household's over the business cycle.

The government chooses the policy rule to determine $(G_t, T_t, S_{gt}^h, S_{get}, Q_{gt}, D_{gt}, R_{mt+1})$. (A_t, ψ_t) follows an exogenous stochastic process. Then, five prices $(Q_t^i, Q_t^n, q_t, R_{t+1}, R_{bt+1})$ and 13 quantities $(Y_t, C_t, L_t, I_t, K_{t+1}, Z_t, M_t, \bar{E}_t, D_t, N_t^i, N_t^n, S_{pt}^i, S_{pt}^n)$ together with 8 shadow prices $(\mathcal{V}_t, \mathcal{V}_{bt}, \mathcal{V}_{mt}, \mathcal{V}_{st}, \mathcal{V}_{et}, \mathcal{V}_{get}, \lambda_t^i, \lambda_t^n)$ are determined as a function of the state variables $(K_t, C_{t-1}, I_{t-1}, A_t, \psi_t, R_t D_{t-1}, R_t D_{gt-1}, R_{mt} M_{t-1}, \bar{E}_{t-1}, S_{gt-1}, S_{get-1})$. By the sequence of 26 equations: the optimization conditions of households and nonfinancial firms (1, 2, 7, 39, 40, 120), the optimization of banks (71i, 71h, 73, 77–79, 106, 107, 110–114, 116i, 116n), and the market clearing conditions for goods, labor, securities, and interbank market (3, 42, 116, 117, 118).

REFERENCES

- Adrian, T., Shin, H., 2009. Money, liquidity and monetary policy. Federal Reserve Bank of New York and Princeton University, Mimeo.
- Aiyagari, R., Gertler, M., 1999. Overreaction of asset prices in general equilibrium. *Rev. Econ. Dyn.* 2, 3–35.
- Allen, F., Babus, A., Carletti, E., 2009. Financial crises: Theory and evidence. University of Pennsylvania, Mimeo.
- Allen, F., Gale, D., 1994. Limited market participation and volatility of asset prices. *Am. Econ. Rev.* 84, 933–955.
- Allen, F., Gale, D., 2007. Understanding financial crises. Oxford University Press, Oxford, UK.
- Angeloni, I., Faia, E., 2009. A tale of two policies: Prudential regulation and monetary policy with fragile banks. European Central Bank, Mimeo.
- Bagehot, W., 1873. *Lombard Street: A description of the money market*. H. S. King, London, UK.
- Bernanke, B., 2009. The crisis and the policy response. Jan. 13 speech.
- Bernanke, B., Gertler, M., 1989. Agency costs, net worth and business fluctuations. *Am. Econ. Rev.* 79, 14–31.
- Bernanke, B., Gertler, M., Gilchrist, S., 1999. The financial accelerator in a quantitative business cycle framework. In: Taylor, J., Woodford, M. (Eds.), *Handbook of macroeconomics*. Vol. 1. Elsevier, Amsterdam, Netherlands, pp. 1341–1393.
- Brunnermeier, M., 2009. Deciphering the liquidity and credit crunch 2007–2008. *J. Econ. Lit.* 23, 77–100.
- Brunnermeier, M., Pederson, L., 2009. Market liquidity and funding liquidity. *Rev. Financ. Stud.* 22, 2201–2238.
- Brunnermeier, M., Sannikov, Y., 2009. A macroeconomic model with a financial sector. Princeton University, Mimeo.
- Caballero, R., Krishnamurthy, A., 2001. International and domestic collateral constraints in a model of emerging market crises. *J. Monet. Econ.* 48, 513–548.
- Calomiris, C., Kahn, C., 1991. The role of demandable debt in structuring banking arrangements. *Am. Econ. Rev.* 81, 497–513.

- Carlstrom, C., Fuerst, T., 1997. Agency costs, net worth and business fluctuations: A computable general equilibrium analysis. *Am. Econ. Rev.* 97, 893–910.
- Chari, V.V., Kehoe, P., 2010. Bailouts, time consistency and optimal regulation. University of Minnesota, Mimeo.
- Christiano, L., Eichenbaum, M., Evans, C., 2005. Nominal rigidities and the dynamics effects of a shock to monetary policy. *J. Polit. Econ.* 113, 1–45.
- Christiano, L., Motto, R., Rostagno, M., 2005. The Great Depression and the Friedman Schwartz hypothesis. *J. Money Credit Bank* 35, 1119–1198.
- Christiano, L., Motto, R., Rostagno, M., 2010. Financial factors in business fluctuations. Northwestern University, Mimeo.
- Curdia, V., 2007. Monetary policy under sudden stops. Federal Reserve Bank of New York, Mimeo.
- Curdia, V., Woodford, M., 2009a. Credit spreads and monetary policy. Federal Reserve Bank of New York and Columbia University, Mimeo.
- Curdia, V., Woodford, M., 2009b. Conventional and unconventional monetary policy. Federal Reserve Bank of New York and Columbia University, Mimeo.
- Del Negro, M., Eggertsson, G., Ferrero, A., Kiyotaki, N., 2010. The great escape?. Federal Reserve Bank of New York and Princeton University, Mimeo.
- Diamond, D., 1984. Financial intermediation and delegated monitoring. *Rev. Econ. Stud.* 51, 393–414.
- Diamond, D., Dybvig, P., 1983. Bank runs, deposit insurance and liquidity. *J. Polit. Econ.* 91 (3), 401–419.
- Diamond, D., Rajan, R., 2009. Illiquidity and interest rate policy. Mimeo.
- Eisfeldt, A., 2004. Endogenous liquidity in asset markets. *J. Finance* 59 (1), 1–30.
- Faia, E., Monacelli, T., 2007. Optimal interest rate rules, asset prices and credit frictions. *J. Econ. Dyn. Control* 31, 3228–3254.
- Farhi, E., Tirole, J., 2009. Collective moral hazard, systematic risk and bailouts. Harvard University and University of Toulouse, Mimeo.
- Fostel, A., Geanakoplos, J., 2008. Leverage cycles and the anxious economy. *Am. Econ. Rev.* 98, 1211–1244.
- Gertler, M., Gilchrist, S., Natalucci, F., 2007. External constraint on monetary policy and the financial accelerator. *J. Money Credit and Bank* 39, 295–330.
- Gertler, M., Karadi, P., 2009. A model of unconventional. monetary policy. New York University, Mimeo.
- Gilchrist, S., Leahy, J., 2002. Monetary policy and asset prices. *J. Monet. Econ* 49, 75–97.
- Gilchrist, S., Yankov, V., Zakrasjek, E., 2009. Credit market shocks and economic fluctuations: Evidence from corporate bond and stock markets. Boston University, Mimeo.
- Goodfriend, M., McCallum, B., 2007. Banking and interest rates in monetary policy analysis. *J. Monet. Econ.* 54, 1480–1507.
- Gorton, G., 2010. Slapped in the face by the invisible hand: The panic of 2007. Oxford University Press, Oxford, UK.
- Gourio, F., 2009. Disaster risk and business cycles. Boston University, Mimeo.
- He, Z., Krishnamurthy, A., 2009. Intermediary asset pricing. Northwestern University, Mimeo.
- Holmstrom, B., Tirole, J., 1997. Financial intermediation, loanable funds and the real sector. *Q. J. Econ.* 112, 663–691.
- Holmstrom, B., Tirole, J., 1998. Private and public supply of liquidity. *J. Polit. Econ.* 106, 1–40.
- Iacoviello, M., 2005. House prices, borrowing constraints and monetary policy in the business cycle. *Am. Econ. Rev.* 95, 739–764.
- Jermann, U., Quadrini, V., 2009. The macroeconomic effects of financial shocks. University of Pennsylvania and USC, Mimeo.
- Justiniano, A., Primiceri, G., Tambalotti, A., 2009. Investment shocks and business cycles. Northwestern University, Mimeo.
- Kehoe, T., Levine, D., 1993. Debt-constrained asset markets. *Rev. Econ. Stud.* 60, 865–888.
- Kiyotaki, N., Moore, J., 1997. Credit cycles. *J. Polit. Econ.* 105, 211–248.

- Kiyotaki, N., Moore, J., 2008. Liquidity, business cycles and monetary policy. Princeton University and LSE, Mimeo.
- Korinek, A., 2009. Systematic risk-taking amplification effects, externalities and regulatory responses. University of Maryland, Mimeo.
- Krishnamurthy, A., 2003. Collateral constraints and the amplification mechanism. *J. Econ. Theory* 111, 277–292.
- Kurlat, P., 2009. Lemons, market shutdowns and learning. MIT, Mimeo.
- La'O, J., 2010. Collateral constraints and noisy fluctuations. MIT, Mimeo.
- Lorenzoni, G., 2008. Inefficient credit booms. *Rev. Econ. Stud.* 75, 809–833.
- Mendoza, E., 2009. Sudden stops, financial crises and leverage: A Fisherian deflation of Tobin's q . University of Maryland, Mimeo.
- Merton, R., 1973. An intertemporal capital asset pricing model. *Econometrica* 41, 867–887.
- Monacelli, T., 2009. New Keynesian models, durable goods and collateral constraints. *J. Monet. Econ.* 56, 242–254.
- Reinhart, C.M., Rogoff, K., 2009. This time is different: Eight centuries of financial folly. Princeton University Press, Princeton, NJ.
- Reis, R., 2009. Where should liquidity be injected during a financial crisis?. Columbia University, Mimeo.
- Sargent, T.J., Wallace, N., 1983. The real bills doctrine versus the quantity theory of money. *J. Polit. Econ.* 90, 1212–1236.
- Shleifer, A., Vishny, R., 2009. Unstable banking. Harvard University, Mimeo.
- Shleifer, A., Vishny, R., 2010. Asset fire sales and credit easing. Harvard University, Mimeo.
- Smets, F., Wouters, R., 2007. Shocks and frictions in U.S. business cycles: A Bayesian DSGE approach. *Am. Econ. Rev.* 97, 586–606.
- Townsend, R., 1979. Costly state verification. *J. Econ. Theory* 21, 265–293.
- Wallace, N., 1981. A Modigliani–Miller theorem for open market operations. *Am. Econ. Rev.* 71, 267–274.
- Williamson, S., 1987. Financial intermediation, business failures and real business cycles. *J. Polit. Econ.* 95, 1196–1216.

This page intentionally left blank

Financial Intermediaries and Monetary Economics[☆]

Tobias Adrian* and **Hyun Song Shin****

*Federal Reserve Bank of New York

**Princeton University

Contents

1. Introduction	602
2. Financial Intermediaries and the Price of Risk	606
2.1 Model	606
2.2 Pricing of risk	613
2.3 Shadow value of bank capital	614
3. Changing Nature of Financial Intermediation	615
3.1 Shadow banking system and security broker-dealers	615
3.2 Haircuts and VaR	619
3.3 Relative size of the financial sector	621
4. Empirical Relevance of Financial Intermediary Balance Sheets	623
5. Central Bank as Lender of Last Resort	631
6. Role of Short-Term Interest Rates	636
6.1 The risk-taking channel of monetary policy	638
6.2 Two case studies	640
6.3 Related literature	641
7. Concluding Remarks	646
References	648

Abstract

We reconsider the role of financial intermediaries in monetary economics, and explore the hypothesis that the financial intermediary sector is the engine that drives the financial cycle through fluctuations in the price of risk. In this framework, balance sheet quantities emerge as a key indicator of risk appetite and, hence, for the “risk-taking channel” of monetary policy. We document evidence that balance sheets of financial intermediaries provide a window on the transmission of monetary policy through capital market conditions. Short-term interest rates are found to be important in influencing the size of financial intermediary balance sheets. Our findings suggest that the traditional focus on the money stock for the conduct of

[☆] The views expressed in this chapter are those of the authors and do not necessarily represent those of the Federal Reserve Bank of New York or the Federal Reserve System. We are grateful to the editors Benjamin Friedman and Michael Woodford for their advice in guiding the draft through to publication, and to Xavier Freixas for his illuminating comments as discussant.

monetary policy may have more modern counterparts, and suggest the importance of tracking balance sheet quantities.

JEL classification: E44, E52, E53, G01, G18, G2, G24

Keywords

Monetary Economics
Financial Intermediation
Risk Taking Channel
Bank Lending Channel

1. INTRODUCTION

In conventional models of monetary economics commonly used in central banks, the banking sector has not played a prominent role. The primary friction in such models is the price stickiness of goods and services. Financial intermediaries do not play a role, except as a passive player that the central bank uses as a channel to implement monetary policy.

However, financial intermediaries have been at the center of the global financial crisis that erupted in 2007. They have borne a large share of the credit losses from securitized subprime mortgages, even though securitization was intended to parcel out and disperse credit risk to investors who were better able to absorb losses. Credit losses and the associated financial distress have figured prominently in the commentary on the downturn in real economic activity that followed. These recent events suggest that financial intermediaries may be worthy of separate study to ascertain their role in economic fluctuations.

The purpose of this chapter is to reconsider the role of financial intermediaries in monetary economics. In addressing the issue of financial factors in macroeconomics, we join a spate of recent research that has attempted to incorporate a financial sector in a New Keynesian DSGE model. Curdia and Woodford (2009) and Gertler and Karadi (2009) are recent examples. However, rather than phrasing the question of how financial “frictions” affect the real economy, we focus on the financial intermediary sector. We explore the hypothesis that the financial intermediary sector, far from being passive, is instead the engine that drives the boom-bust cycle. To explore this hypothesis, we propose a framework for study to address the following pair of questions. What are the channels through which financial intermediaries exert an influence on the real economy (if at all)? What are the implications for monetary policy?

Banks and other financial intermediaries borrow in order to lend. Since the loans offered by banks tend to be of longer maturity than the liabilities that fund those loans, the term spread is indicative of the marginal profitability of an extra dollar of loans on intermediaries’ balance sheets. The net interest margin (NIM) of the bank is the difference between the total interest income on the asset side of its balance sheet and the interest expense on the liabilities side of its balance sheet. Whereas the term spread indicates the profitability of the marginal loan that is added to the balance sheet, the NIM is an average concept that applies to the stock of all loans and liabilities on the balance sheet.

The NIM determines the profitability of bank lending and increases the present value of bank income, boosting the forward-looking measures of bank capital. Such a boost in bank capital increases the capacity of the bank to increase lending in the sense that the marginal loan that was not made before the boost in bank capital now becomes feasible under the greater risk-bearing capacity of the bank. As banks expand their balance sheets, the market price of risk falls.

In this framework, financial intermediaries drive the financial cycle through their influence on the determination of the price of risk. Quantity variables — particularly the components of financial intermediary balance sheets — emerge as important economic indicators due to their role in reflecting the risk capacity of banking sector and hence on the marginal real project that receives funding. In this way, the banking sector plays a key role in determining the level of real activity. Ironically, our findings have some points of contact with the older theme in monetary economics of keeping track of the money stock at a time when it has fallen out of favor among monetary economists.¹ The common theme between our framework and the older literature is that the money stock is a balance sheet aggregate of the financial sector. Our approach suggests that broader balance sheet aggregates such as total assets and leverage are the relevant financial intermediary variables to incorporate into macroeconomic analysis.

When we examine balance sheet measures that reflect the underlying funding conditions in capital markets, we find that the appropriate balance sheet quantities are of institutions that are marking to market their balance sheets. In this regard, fluctuations in shadow bank and broker-dealer assets are more informative than movements in commercial bank assets. However, as commercial banks begin to mark more items of their balance sheets to market, commercial bank balance sheet variables are likely to become more important variables for studying the transmission mechanism.

Our findings have important implications for the conduct of monetary policy. According to the perspective outlined here, fluctuations in the supply of credit arise from the interactions between bank risk-taking and the market risk premium. The cost of leverage of market-based intermediaries is determined by two main variables — risk and risk-taking capacity. The expected profitability of intermediaries is proxied by spreads such as the term spread and various credit spreads. Variations in the policy target determine short-term interest rates, and have a direct impact on the profitability of intermediaries. For these reasons, short-term interest rates matter directly for monetary policy.

The effect of keeping policy rates low in the aftermath of the financial crisis of 2008 has illustrated again the potency of low policy interest rates in raising the profitability of banks and recapitalizing the banking system from their dangerous low levels. When considering the debates in early 2009 about the necessity (or inevitability) of capital injections into the U.S. banking system, the turnaround in the capital levels of the U.S. banking sector has been worthy of note.

¹ See Friedman (1988) for an overview of the role of monetary aggregates in macroeconomic fluctuations in the United States.

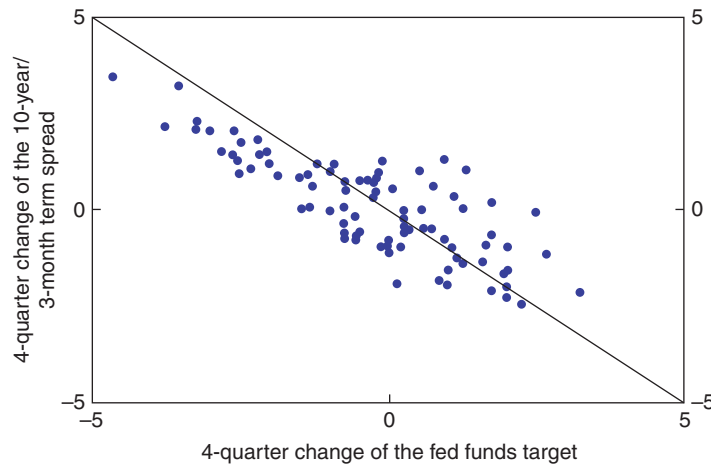


Figure 1 The term spread and the federal funds rate.

Empirically, there is (for the United States) a near perfect negative one-to-one relationship between 4-quarter changes of the federal funds target and 4-quarter changes of the term spread defined as the 10-year/3-month term Treasury spread (Figure 1 uses data from 1987q1 to 2008q3). Thus, shifts in the policy rate translate directly into shifts in the slope of the yield curve. Since the term spread affects the profitability of the marginal loan and the future NIM of the bank, the short rate signals future risk-taking capacity of the banking sector. In this way, variations in the target rate affect real activity because they change the risk-taking capacity of financial intermediaries, thus shifting market risk premiums and the supply of credit. Borio and Zhu (2008) coined the term “risk-taking channel” of monetary policy to describe this set of effects working through the risk appetite of financial intermediaries.

This perspective on the importance of the short rate as a price variable is in contrast to current monetary thinking, where short-term rates matter only to the extent that they determine long-term interest rates, which are seen as being risk-adjusted expectations of future short rates. Current models of monetary economics used at central banks emphasize the importance of managing market expectations. By charting a path for future short rates and communicating this path clearly to the market, the central bank can influence long rates, which then influence mortgage rates, corporate lending rates, and other prices that affect consumption and investment. This “expectations channel,” which is explained in Bernanke (2004), Svensson (2004), and Woodford (2003, 2005), has become an important consideration for monetary policy. In his book on central banking, Alan Blinder (1998, p.70) phrases the claim in a particularly clear way.

central banks generally control only the overnight interest rate, an interest rate that is relevant to virtually no economically interesting transactions. Monetary policy has important macroeconomic effects only to the extent that it moves financial market prices that really matter — like long-term interest rates, stock market values and exchange rates.

In contrast, our results suggest that short-term rates may be important in their own right. Short rates matter because they largely determine the term spread, which in turn determines the NIM and the forward-looking capital of the banking sector. Continued low short rates imply a steep yield curve for some time, higher NIM in the future, and hence higher risk-taking capacity of the banking sector. Conversely, higher short rates imply lower future NIM and a decline in the risk-taking capacity of the banking sector. In particular, an inverted yield curve is a sign of diminished risk-taking capacity, and by extension of lower real activity.

There is empirical support for the risk-taking channel of monetary policy. We find that the growth in shadow bank balance sheets and broker-dealer balance sheets help to explain future real activity. However, we also find that fluctuations in the balance sheet size of shadow banks and security broker-dealers appear to signal shifts in future real activity better than the fluctuations of the larger commercial banking sector. Thus, one lesson from our empirical analysis is that there are important distinctions between different categories of financial intermediaries. In fact, the evolutions of shadow bank and broker-dealer assets have time signatures that are markedly different from those of commercial banks. Our results point to key differences between banking, as traditionally conceived, and the market-based banking system that has become increasingly influential in charting the course of economic events.

Having established the importance of financial intermediary balance sheets in signaling future real activity, we go on to examine the determinants of balance sheet growth. We find that short-term interest rates are important. Indeed, the level of the federal funds target is a key variable: a lowering of short-term rates is conducive to expanding balance sheets. In addition, a steeper yield curve, larger credit spreads, and lower measures of financial market volatility are conducive to expanding balance sheets. In particular, an inverted yield curve is a harbinger of a slowdown in balance sheet growth, shedding light on the empirical feature that an inverted yield curve forecasts recessions. The federal funds target determines other relevant short-term interest rates, such as repo rates and interbank lending rates through arbitrage in the money market. As such, we may expect the federal funds rate to be pivotal in setting short-term interest rates more generally.

These findings reflect the economics of financial intermediation, since the business of banking is to borrow short and lend long. For an off-balance sheet vehicle such as a conduit or structured investment vehicle (SIV) that finances holdings of mortgage assets by issuing commercial paper, a difference of a quarter or half percent in the funding cost may make all the difference between a profitable venture and a loss-making one. This is because the conduit or SIV, like most financial intermediaries, is simultaneously both a creditor and a debtor — it borrows in order to lend.

In this chapter we begin with a simple equilibrium model where financial intermediaries are the main engine for the determination of the price of risk in the economy. We then present empirical results on the real impact of shadow bank and broker-dealer balance sheet

changes, and on the role of short-term interest rates in the determination of balance sheet changes. We also consider the role of the central bank as the lender of last resort in light of our findings. We conclude by drawing some lessons for monetary policy.

2. FINANCIAL INTERMEDIARIES AND THE PRICE OF RISK

To motivate the study of financial intermediaries and how they determine the price of risk, we begin with a stylized model set in a one-period asset market.² The general equilibrium model presented next is deliberately stark. It has two features that deserve emphasis.

First, there is no default in the model. The debt that appears in the model is risk-free. However, as we will see, the amplification of the financial cycle is present. Geanakoplos (2009) highlighted how risk-free debt may still give rise to powerful spillover effects through fluctuations in leverage and the pricing of risk. The model also incorporates insights from Shleifer and Vishny (1997), who demonstrated that financial constraints can lead to fluctuations of risk premia even if arbitrageurs are risk-neutral.³ Adrian and Shin (2007) exhibited empirical evidence that bears on the fluctuations in the pricing of risk from the balance sheets of financial intermediaries.

Second, in the example, there is no lending and borrowing between financial intermediaries. So, any effect we see in the model cannot be attributed to what we may call the “domino model” of systemic risk, where systemic risk propagates through the financial system via a chain of defaults of financial intermediaries.⁴ This is not to deny that interlocking claims matter; however, the benchmark case serves the purpose of showing that chains of default are not necessary for fluctuations in the price of risk.

To anticipate the punch line from the simple model, we show that aggregate capital of the financial intermediary sector stands in a one-to-one relation with the price of risk and the availability of funding that flows to real projects. The larger the aggregate intermediary sector capital is, the lower the price of risk, and the easier the credit.

2.1 Model

Today is date 0. A risky security is traded today in anticipation of its realized payoff in the next period (date 1). The payoff of the risky security is known at date 1. When viewed from date 0, the risky security’s payoff is a random variable $\tilde{w}$, with expected value $q > 0$. The uncertainty surrounding the risky security’s payoff takes a particularly simple form. The random variable $\tilde{w}$ is uniformly distributed over the interval:

$$[q - z, q + z]$$

² A similar model appeared in Shin (2009).

³ Shleifer and Vishny (2009) presented a theory of unstable banking that is closely related to our model.

⁴ See Adrian and Shin (2008b) for an argument for why the “domino model” is inappropriate for understanding the crisis of 2007–2009.

The mean and variance of $\tilde{w}$ is given by

$$E(\tilde{w}) = q$$

$$\sigma^2 = \frac{z^2}{3}$$

There is also a risk-free security, which we call “cash,” that pays an interest rate of i . Let p denote the price of the risky security. For an investor with equity e who holds γ units of the risky security, the payoff of the portfolio is the random variable:

$$W \equiv \tilde{w}\gamma + (1+i)(e - p\gamma) \quad (1)$$

$$= \underbrace{(w - (1-i)p)\gamma}_{\text{risky excess return}} + \underbrace{(1+i)e}_{\text{risk-free ROE}} \quad (2)$$

There are two groups of investors; passive and active. The passive investors can be thought of as nonleveraged investors such as households, pension funds, and mutual funds, while the active investors can be interpreted as leveraged institutions such as banks and securities firms who manage their balance sheets actively. The risky securities can be interpreted as loans granted to ultimate borrowers or securities issued by the borrowers, but there is a risk that the borrowers will not fully repay the loan. Figure 2 depicts these relationships. Under this interpretation, the market value of the risky securities can be thought of as the marked-to-market value of the loans granted to the ultimate borrowers. The passive investors’ holding of the risky security can then be interpreted as the credit that is granted *directly* by the household sector (e.g., through the holding of corporate bonds), while the holding of the risky securities by the active investors can be given the interpretation of *intermediated finance* where the active investors are banks that borrow from the households to lend to the ultimate borrowers.

We assume that the passive investors have mean-variance preferences over the payoff from the portfolio. They aim to maximize

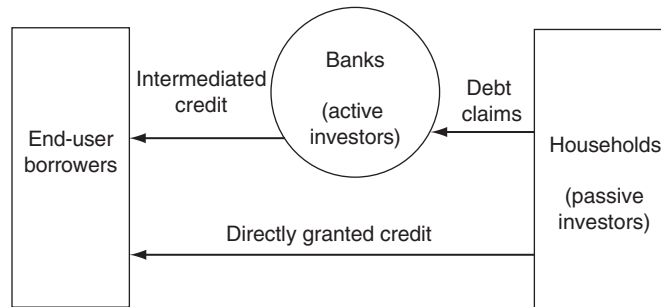


Figure 2 Intermediated and directly granted credit.

$$U = E(W) - \frac{1}{2\tau} \sigma_W^2 \quad (3)$$

where $\tau > 0$ is a constant called the investor's "risk tolerance" and σ_W^2 is the variance of W . In terms of the decision variable y , the passive investor's objective function can be written as

$$U(y) = \underbrace{(q/p - (1+i))}_{\text{Expected Excess Return}} p y + (1+i) e - \frac{1}{6\tau} y^2 z^2 \quad (4)$$

The optimal holding of the risky security satisfies the first order condition:

$$(q - (1+i)p) - \frac{1}{3\tau} y z^2 = 0$$

The price must be below the expected payoff for the risk-averse investor to hold any of the risky security. The optimal risky security holding of the passive investor (denoted by y_p) is given by

$$y_p = \begin{cases} \frac{3\tau}{z^2} (q - (1+i)p) & \text{if } q > p(1+i) \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

These linear demands can be summed to give the aggregate demand. If τ_i is the risk tolerance of the i th investor and $\tau = \sum_i \tau_i$, then Eq. (5) gives the aggregate demand of the passive investor sector as a whole.

Now turn to the portfolio decision of the active (leveraged) investors. These active investors are risk-neutral but face a Value-at-Risk (VaR) constraint, which is common for banks and other leveraged institutions.⁵ The general VaR constraint is that the capital cushion should be large enough so that the default probability is kept below some benchmark level. Consider the special case where that benchmark level is zero. This is an extreme assumption, which we adopt for the purpose of simplifying the model. By setting the VaR constraint to allow no default by the bank, we can treat bank liabilities as a perfect substitute for cash. It would be possible to allow for a less stringent VaR constraint that allows possible default by the bank, but then the modeling has to make allowances for the bank's liabilities of risky debt and that they are priced accordingly. However, the key qualitative features of the model would be unaffected. Thus, in what follows, we will adopt the stringent version of the VaR constraint where the bank holds enough capital to meet the worst case loss, and where the bank's liabilities are risk-free.

⁵ A microfoundation for the VaR constraint is provided by Adrian and Shin (2008a).

Denote by VaR the VaR of the leveraged investor. The constraint is that the investor's capital (equity) e is large enough to cover this VaR . The optimization problem for an active investor is

$$\max_{\gamma} E(W) \quad \text{subject to } \text{VaR} \leq e \quad (6)$$

If the price is too high (i.e., when $p > q/(1+i)$ so that the price exceeds the discounted expected payoff) the investor holds no risky securities. When $p < q/(1+i)$, then $E(W)$ is strictly increasing in γ , so the VaR constraint binds. The optimal holding of the risky security can be obtained by solving $\text{VaR} = e$. To solve this equation, write out the balance sheet of the leveraged investor as

Assets	Liabilities
Securities, $p\gamma$	Equity, e Debt, $p\gamma - e$

For each unit of the risky security, the minimum payoff is $q - z$. Thus, the worst case loss is $(p(1+i) - (q - z))\gamma$. For the bank to have enough equity to cover the worst case loss, we require:

$$(p(1+i) - (q - z))\gamma \leq e \quad (7)$$

This inequality also holds in the aggregate. The left-hand side of Eq. (7) is the VaR (the worst possible loss), which must be met by the equity buffer e . Since the constraint binds, the optimal holding of the risky securities for the leveraged investor is

$$\gamma = \frac{e}{p(1+i) - (q - z)} \quad (8)$$

So the demand from the bank for the risky asset depends positively on the expected excess return to the risky asset $q - (1+i)p$, and positively on the amount of equity that the bank is endowed with e .

Since Eq. (8) is linear in e , the aggregate demand of the leveraged sector has the same form as Eq. (8) when e is the *aggregate capital* of the leveraged sector as a whole.

Replacing the constraint (8) in the amount of debt $p\gamma - e$ allows us to write the new balance sheet as follows:

Assets	Liabilities
Securities, $p\gamma$	equity, e debt, $\frac{q-z}{1+i}\gamma$

(9)

where the debt $\frac{q-z}{1+i}y$ was constructed by substituting $e = -((q/p - (1+i))p - z)y$ into $py - e$. We assume that $q > z$ so as to ensure that the payoff of the risky security is non-negative. The bank's leverage is the ratio of total assets to equity, which can be written as:

$$\text{leverage} = \frac{py}{e} = \frac{p}{p(1+i) - (q-z)} \quad (10)$$

Denoting by y the holding of the risky securities by the active investors and by y_P the holding by the passive investors, the market clearing condition is

$$y + y_P = S \quad (11)$$

where S is the total endowment of the risky securities. Figure 3 illustrates the equilibrium for a fixed value of aggregate capital e . For the passive investors, their demand is linear, with the intercept at $q/(1+i)$. The demand of the leveraged sector can be read off from Eq. (8). The solution is fully determined as a function of e . In a dynamic model, e can be treated as the state variable (see Danielsson, Shin, & Zigrand 2009).

Now consider a possible scenario involving an improvement in the fundamentals of the risky security where the expected payoff of the risky securities rises from q to q' . In our banking interpretation of the model, an improvement in the expected payoff should be seen as an increase in the marked-to-market value of bank assets. For now, we simply treat the increase in q as an exogenous shock. Figure 4 illustrates the scenario. The improvement in the fundamentals of the risky security pushes up the demand curves for both the passive and active investors, as illustrated in Figure 4. However, there is an amplified response from the leveraged investors as a result of marked-to-market gains on their balance sheets.

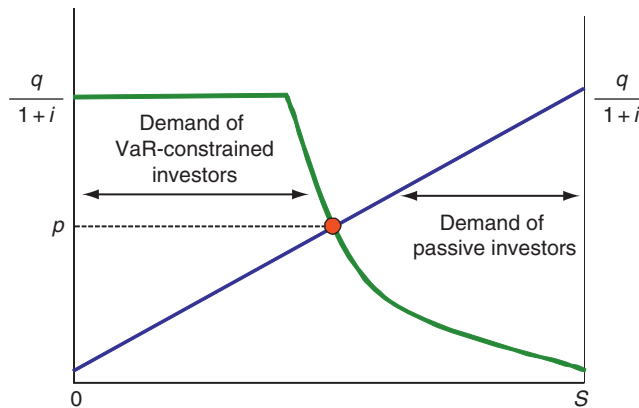


Figure 3 Market clearing price.

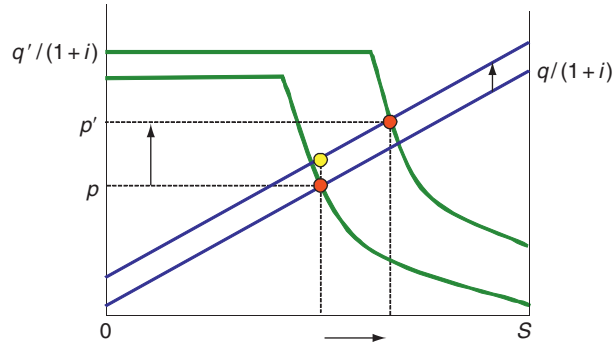


Figure 4 Amplified response to improvement in fundamentals q .

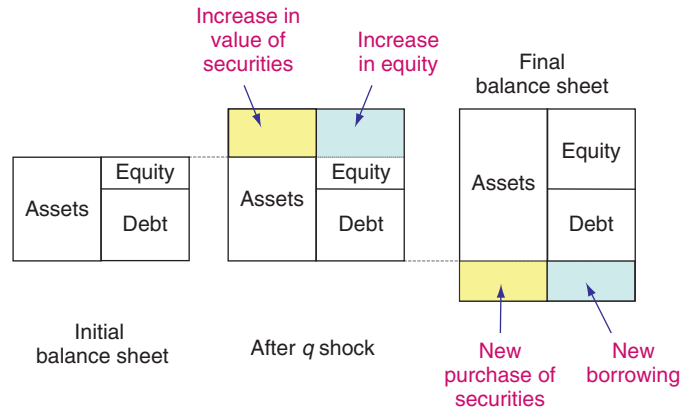


Figure 5 Balance sheet expansion from q shock.

From Eq. (9), denote by e' the new equity level of the leveraged investors that incorporates the capital gain when the price rises to p' . The initial amount of debt was $\frac{q-z}{1+i}\gamma$. Since the new asset value is $p'\gamma$, the new equity level e' is

$$e' = (p'(1+i) - (q-z))\gamma \quad (12)$$

Figure 5 breaks out the steps in the balance sheet expansion. The initial balance sheet is on the left, where the total asset value is $p\gamma$. The middle balance sheet shows the effect of an improvement in fundamentals that comes from an increase in q , but before any adjustment in the risky security holding. There is an increase in the value of the securities without any change in the debt value, since the debt was already risk-free. So, the increase in asset value flows through entirely to an increase in equity. Equation (12) expresses the new value of equity e' in the middle balance sheet in Figure 5.

The increase in equity relaxes the VaR constraint, and the leveraged sector can increase its holding of risky securities. The new holding y' is larger, and is enough to make the VaR constraint bind at the higher equity level, with a higher fundamental value q' . That is,

$$e' = (p'(1+i) - (q' - z))y' \quad (13)$$

After the q shock, the investor's balance sheet has strengthened, and capital has increased without any change in debt value. There has been an erosion of leverage, leading to spare capacity on the balance sheet in the sense that equity is now larger than necessary to meet the VaR. To utilize the slack in balance sheet capacity, the investor takes on additional debt to purchase additional risky securities. The demand response is upward-sloping. The new holding of securities is now y' , and the total asset value is $p' y'$. Equation (13) expresses the new value of equity e' in terms of the new higher holding y' in the right-hand side balance sheet in Figure 5. From Eqs. (12) and (13), we can write the new holding y' of the risky security as

$$y' = y \left(1 + \frac{q' - q}{p'(1+i) - q' + z} \right) \quad (14)$$

From the demand of passive investors Eq. (5) and market clearing,

$$(1+i)p' - q' = \frac{z^2}{3\tau}(y' - S)$$

Substituting into Eq. (14),

$$y' = y \left(1 + \frac{q' - q}{\frac{z^2}{3\tau}(y' - S) + z} \right) \quad (15)$$

This defines a quadratic equation in y' . The solution is where the right-hand side of Eq. (15) cuts the 45 degree line. The leveraged sector amplifies booms and busts if $y' - y$ has the same sign as $q' - q$. Then, any shift in fundamentals gets amplified by the portfolio decisions of the leveraged sector. The condition for amplification is that the denominator in the second term of Eq. (15) is positive. But this condition is guaranteed from Eq. (14) and the fact that $p' > \frac{q' - z}{1+i}$ (i.e., that the price of the risky security is higher than its worst possible realized discounted payoff).

Note also that the size of the amplification is increasing when fundamental risk is small, seen from the fact that $y' - y$ is large when z is small. Recall that z is the fundamental risk. When z is small, the associated VaR is also small, allowing the leveraged sector to maintain high leverage. The higher the leverage is, the greater the marked-to-market capital gains and losses. Amplification is large when the leveraged sector is large relative to the total economy. Finally, note that the amplification is more likely when the passive sector's risk tolerance τ is high.

2.2 Pricing of risk

We now explore the fluctuations in risk pricing in our model. The risk premium in our model is the excess expected return on the risky security, which can be written in terms of the ratio of the discounted expected payoff on the risky security and its price:

$$\text{Risk premium} = \frac{q}{p(i+1)} - 1 \quad (16)$$

Rather than working with the risk premium directly, it turns out to be more convenient to work with a monotonic transformation of the risk premium defined as

$$\pi \equiv 1 - \frac{p(1+i)}{q} \quad (17)$$

The “ π ” stands for “premium.” The variable π is a monotonic transformation of the risk premium that varies from zero (when the risk premium is zero) to 1 (when the risk premium is infinite).

The market-clearing condition for the risky security is $\gamma + \gamma_P = S$, which can be written as

$$\frac{e}{z - q\pi} + \frac{3\tau}{z^2} q\pi = S \quad (18)$$

Our primary interest is in the relationship between total equity e and the risk premium π . Here, e has the interpretation of the total capital of the banking sector, and hence its risk-taking capacity. In our model, the total lending of the banking sector bears a very simple relationship to its total capital e , since the holding of the risky security by the active investors (the banks) is $e/(z - q\pi)$. We impose the restriction that the active investors have a strictly positive total holding of the risky security, or equivalently that the passive sector’s holding is strictly smaller than the total endowment S . From Eq. (5) this restriction can be written as

$$q\pi < \frac{z^2}{3\tau} S \quad (19)$$

By defining $F(e, \pi)$ as below, we can write the market-clearing condition as:

$$F(e, \pi) \equiv e + \frac{3\tau}{z^2} q\pi (z - q\pi) - S(z - q\pi) = 0 \quad (20)$$

We then have

$$\frac{\partial F}{\partial \pi} = q \left(\underbrace{\frac{3\tau}{z^2} (z - q\pi)}_A + \underbrace{S - \frac{3\tau}{z^2} q\pi}_B \right) \quad (21)$$

Both A and B are positive. A is positive since the holding of the risky security by the active sector is $e/(z - q\pi)$, and so $z - q\pi > 0$ in order that the active investors hold positive holdings of the risky security. Another way to view this condition is to note that the market price of the risky security cannot be lower than the lowest possible realization of the payoff of the risky asset, so that

$$p > \frac{q - z}{1 + i} \quad (22)$$

which can be written as $(1 + i)p - (q - z) = z - q\pi > 0$. The second term (term B) inside the big brackets is positive from our condition (19) that the passive investors do not hold the entire supply. Since $\partial F / \partial e = 1$, we have

$$\frac{d\pi}{de} = - \frac{\partial F / \partial e}{\partial F / \partial \pi} < 0 \quad (23)$$

In other words, the market risk premium is decreasing in the total equity e of the banking sector. As stated earlier, we view e as the risk-taking capacity of the banking sector. Any shock that increases the capital buffer of the banking sector will lower the risk premium. We therefore have the following empirical hypothesis.

Empirical Hypothesis 1: Risk premiums fall when the equity of the banking system increases.

This empirical hypothesis is key to our discussion on the role of short-term interest rates on the risk-taking capacity of the banking sector, through the slope of the yield curve, and hence the greater profitability of bank lending. We return to this issue shortly.

2.3 Shadow value of bank capital

Another window on the risk premium in the economy is through the Lagrange multiplier associated with the constrained optimization problem of the banks, which is to maximize the expected payoff from the portfolio $E(W)$ subject to the VaR constraint. The Lagrange multiplier is the rate of increase of the objective function with respect to a relaxation of the constraint, and hence can be interpreted as the shadow value of bank capital. Denoting by γ the Lagrange multiplier, we have

$$\lambda = \frac{dE(W)}{de} = \frac{\partial E(W)}{\partial \gamma} \frac{\partial \gamma}{\partial e} = \frac{q - (1 + i)p}{z - (q - (1 + i)p)} \quad (24)$$

where we have obtained the expression for $dE(W)/d\gamma$ from Eq. (1) and $d\gamma/de$ is obtained from Eq. (8), which gives the optimal portfolio decision of the leveraged investor. Using our π notation, we can rewrite Eq. (24) as

$$\lambda = \frac{q\pi}{z - q\pi} \quad (25)$$

We see from Eq. (25) that as the risk premium π becomes compressed, the Lagrange multiplier γ declines. The implication is that the marginal increase of a dollar's worth of new capital for the leveraged investor is generating less expected payoff. As the risk premium π goes to zero, so does the Lagrange multiplier, implying that the return to a dollar's worth of capital goes to zero.

Furthermore, the shadow value of bank capital can be written as:

$$\lambda = \frac{z(S - \gamma)}{3\tau - z(S - \gamma)} \quad (26)$$

The shadow value of bank capital is decreasing in the size of the leveraged sector, given by γ . Moreover, since there is a one-to-one relationship between γ and the risk premium π , we can also conclude that market risk premiums fall when the size of the intermediary sector increases.

Empirical Hypothesis 2: Risk premiums fall when the size of the banking sector increases.

3. CHANGING NATURE OF FINANCIAL INTERMEDIATION

In preparation for our empirical investigations, we briefly review the structure of financial intermediation in the United States. In particular, we highlight the increasing importance of market-based financial intermediaries and the shadow banking system.

3.1 Shadow banking system and security broker-dealers

As recently as the early 1980s, traditional banks were the dominant financial intermediaries. In subsequent years, however, they were quickly overtaken by market-based financial institutions. Figure 6 plots the size of different types of financial intermediaries for the United States starting in 1985. We see that market-based financial intermediaries, such as security broker dealers and asset-backed securities (ABS) issuers, have become important components of the intermediary sector. The series labeled “shadow banks” aggregates ABS issuers, finance companies, and funding corporations.

In 1985, shadow banks were a tiny fraction of the commercial bank sector, but had caught up with the commercial bank sector by the eve of the crisis. The increased importance of the market-based banking system has been mirrored by the growth of the broker-dealers, who have traditionally played market-making and underwriting roles in securities markets. However, their importance in the supply of credit has increased dramatically in recent years with the growth of securitization and the changing nature of the financial system toward one based on the capital market, rather than one based on the traditional role of the bank as intermediating between depositors and borrowers. Although total assets of the broker-dealer sector are smaller than total assets of the commercial banking sector, our results suggest that broker-dealers provide a

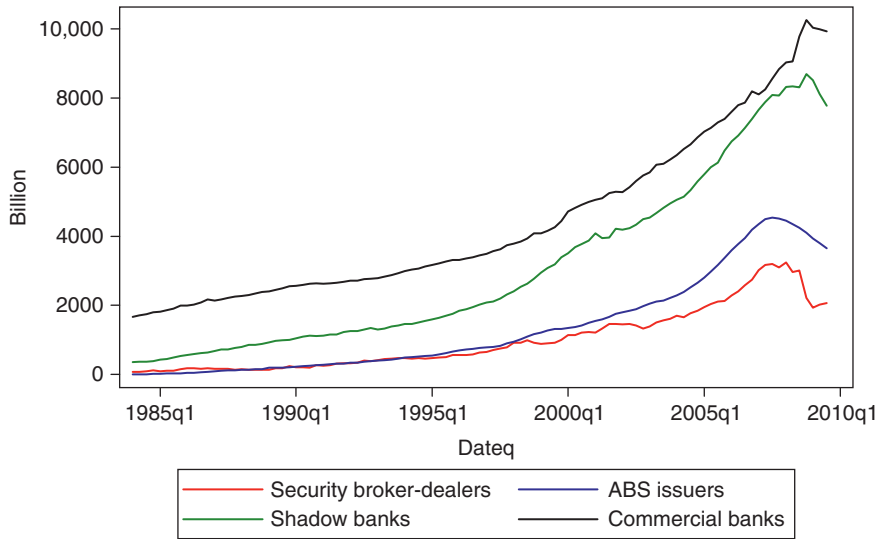


Figure 6 Total assets of commercial banks, shadow banks, and broker-dealers.

better barometer of the funding conditions in the economy, capturing overall capital market conditions. Perhaps the most important development in this regard has been the changing nature of housing finance in the United States. The stock of home mortgages in the United States is now dominated by the holdings of market-based institutions, rather than by traditional bank balance sheets. Broker-dealer balance sheets provide a timely window on this world.

The growth of market-based financial intermediaries is also reflected in the aggregates on the liabilities side of the balance sheet. Figure 7 shows the relative size of the M1 money stock together with the outstanding stock of repos of the primary dealers — the set of banks that bid at U.S. Treasury security auctions, and for whom data are readily available due to their reporting obligations to the Federal Reserve. We also note the rapid growth of financial commercial paper as a funding vehicle for financial intermediaries.

Figure 8 charts the relative size of M2 (bank deposits plus money market fund balances) compared to the sum of primary dealer repos and financial commercial paper outstanding. As recently as the 1990s, the M2 stock was many times larger than the stock of repos and commercial paper. However, by the eve of the crisis, the gap had narrowed considerably, and M2 was only about 25% larger than the stock of repos and financial commercial paper. However, with the eruption of the recent crisis, the gap has opened up again.

Not only have the market-based intermediaries seen the most rapid growth in the run-up to the financial crisis, they were also the institutions that saw the sharpest

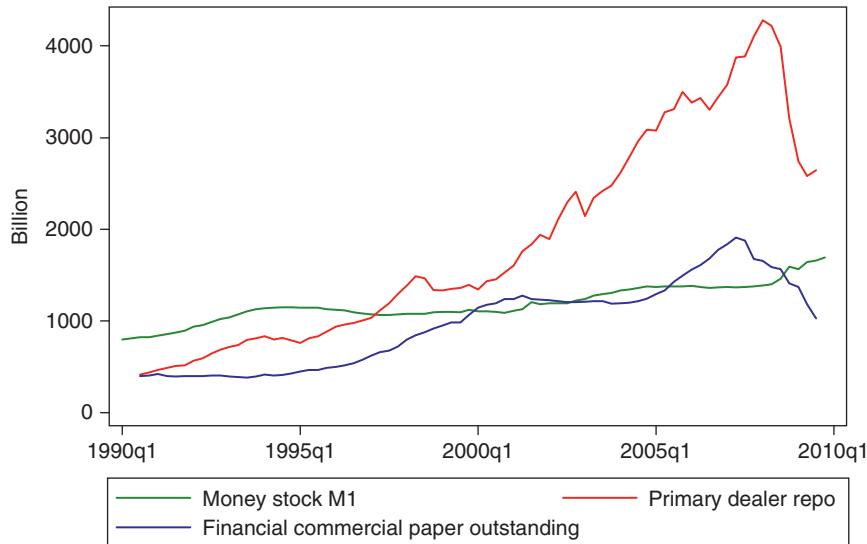


Figure 7 Liquid funding of financial institutions: Money (M1), primary dealer repo, and commercial paper.

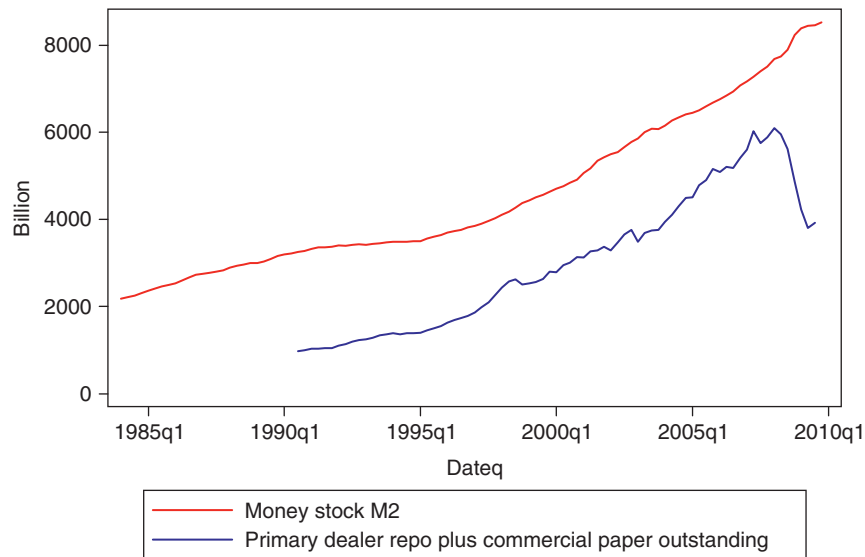


Figure 8 Short-term funding: M2 versus commercial paper + primary dealer repo.

pull-back in the crisis itself. Figure 9 shows the comparative growth rate of the total assets of commercial banks (in red) and the shadow banks (in blue), while Figure 10 shows the growth of commercial paper relative to shadow bank asset growth.

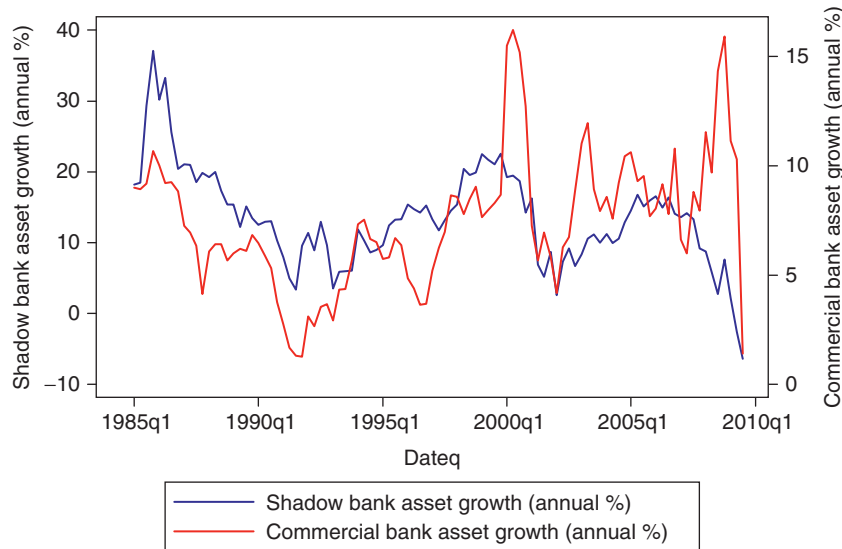


Figure 9 Total asset growth of shadow banks and of commercial banks.

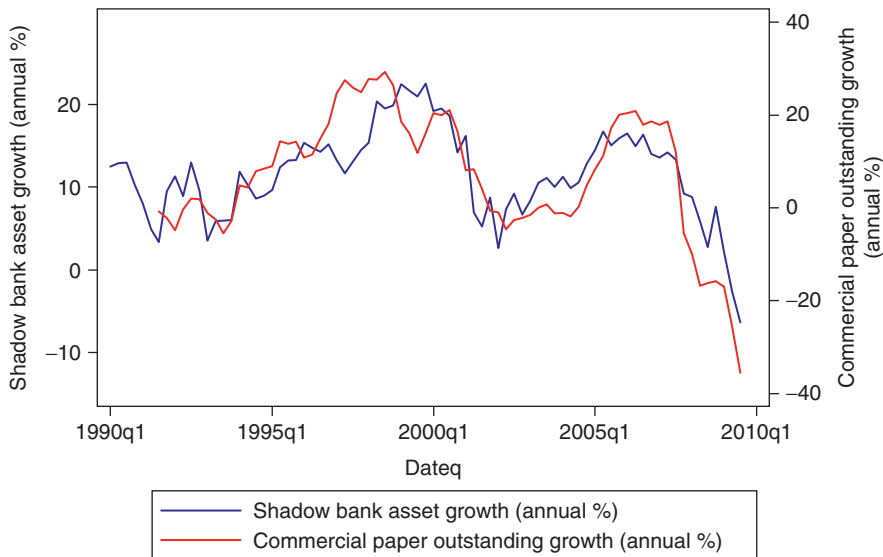


Figure 10 Marginal funding of shadow banks is commercial paper.

We see that, whereas the commercial banks have increased the size of their balance sheet during the crisis, the shadow banks have contracted substantially. Traditionally, banks have played the role of a buffer against fluctuations in capital market conditions,

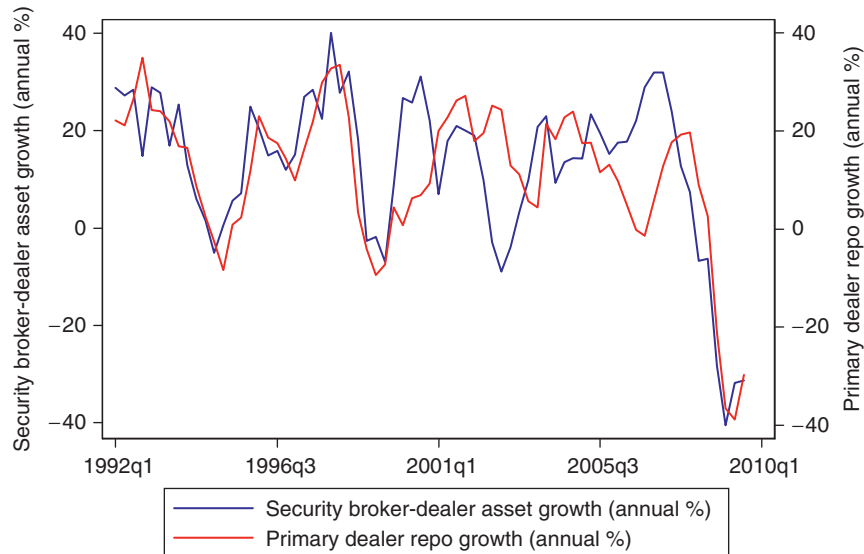


Figure 11 Marginal funding of broker-dealers is repo.

and we see that they have continued their role through the current crisis. As such, looking only at aggregate commercial bank lending may give an overly rosy picture of the state of financial intermediation.

Figure 11 shows that the broker-dealer sector of the economy has contracted in step with the contraction in primary dealer repos, suggesting the sensitivity of the broker-dealer sector to overall capital market conditions. Therefore, in empirical studies of financial intermediary behavior, it is important to bear in mind the distinctions between commercial banks and market-based intermediaries such as broker dealers. Market-based intermediaries who fund themselves through short-term borrowing such as commercial paper or repurchase agreements will be sensitively affected by capital market conditions. But for a commercial bank, its large balance sheet masks the effects operating at the margin. Also, commercial banks provide relationship-based lending through credit lines. Broker-dealers, in contrast, give a much purer signal of marginal funding conditions, as their balance sheets consist almost exclusively of short-term market borrowing and are not as constrained by relationship-based lending.

3.2 Haircuts and VaR

The VaR constraint at the heart of the amplification mechanism in the model in Section 2 characterizes market-based financial intermediaries such as security broker-dealers and shadow banks. The active balance sheet management of financial institutions is documented in Adrian and Shin (2007), who showed that investment banks

exhibit “procyclical leverage”; that is, increases in balance sheet size are associated with increases in leverage. In contrast, the balance sheet behavior of commercial banks is consistent with leverage targeting: for commercial banks, leverage growth is uncorrelated with the growth of balance sheet size.

One useful perspective on the matter is to consider the implicit maximum leverage that is permitted in collateralized borrowing transactions such as repurchase agreements (repos). Repos are the primary source of funding for market-based financial institutions, as well as a marginal source of funding for traditional banks. In a repurchase agreement, the borrower sells a security today for a price below the current market price on the understanding that it will buy it back in the future at a pre-agreed price. The difference between the current market price of the security and the price at which it is sold is called the “haircut” in the repo, and fluctuates together with market conditions. The fluctuations in the haircut largely determine the degree of funding available to a leveraged institution. The reason is that the haircut determines the maximum permissible leverage achieved by the borrower. If the haircut is 2%, the borrower can borrow 98 dollars for 100 dollars worth of securities pledged. Then, to hold 100 dollars worth of securities, the borrower must come up with 2 dollars of equity. Thus, if the repo haircut is 2%, the maximum permissible leverage (ratio of assets to equity) is 50.

Suppose that the borrower leverages up the maximum permitted level. The borrower thus has a highly leveraged balance sheet with leverage of 50. If at this time, a shock to the financial system raises the market haircut, then the borrower faces a predicament. Suppose that the haircut rises to 4%. Then the permitted leverage halves to 25 from 50. The borrower then faces a hard choice. Either it must raise new equity so that its equity doubles from its previous level, or it must sell half its assets, or some combination of both. However, asset disposals have spillover effects that exacerbate the distress for others. The “margin spiral” described by Brunnermeier and Pedersen (2009) models this type of phenomenon.

Considerations of repo haircuts suggest that measured risks will play a pivotal role in the determination of leverage. Adrian and Shin (2008a) presented a contracting model that yields this outcome as a central prediction, and presented empirical evidence consistent with the prediction. Adrian and Shin (2008a) also found that measures of VaR computed from the time series of daily equity returns explains shifts in total assets, leverage, and key components of the liabilities side of the balance sheet, such as the stock of repos.

In the benchmark case where losses are exponentially distributed, the contracting model of Adrian and Shin (2008a) yielded the widely used VaR rule, which stipulates that exposures are adjusted continuously so that equity exactly matches total VaR. Among other things, the VaR rule implies that exposures are adjusted continuously so that the probability of default is kept constant — at the level given by the VaR threshold. Given the ubiquitous use of the VaR rule both by private sector financial institutions and by regulators, this microfoundation of the VaR concept gives a basis for further study.

To be sure, showing that the VaR rule is the outcome of a contracting model says little about the desirability of the widespread adoption of such practices from the viewpoint of economic efficiency. Indeed, there are strong arguments to suggest that risk management tools such as VaR cause spillover effects to other financial institutions that are detrimental to overall efficiency. For instance, the prudent reduction in exposures by the creditors to Lehman Brothers is a run from the viewpoint of Lehman Brothers. The spillover effects are a natural consequence of the contracts being bilateral arrangements. They do not take account of the spillover effects across more than one step in the financial network.

3.3 Relative size of the financial sector

The rapid growth of the market-based intermediaries masks the double-counting involved when adding up balance sheet quantities across individual institutions. Before going further, we therefore note some accounting relationships that can help us to think about the extent of the double-counting.

Let a_i be the total assets of bank i , and x_i the total debt of bank i (i.e., total liabilities minus equity). The total size of the banking sector in gross terms can be written as the sum of all bank assets, given by $\sum_{i=1}^n a_i$. A closely related measure would be the aggregate value of all bank debt, given by $\sum_{i=1}^n x_i$.

Define leverage λ_i as the ratio of total assets to equity of bank i :

$$\lambda_i = \frac{a_i}{a_i - x_i}. \quad (27)$$

Then, solving for x_i and using the notation $\delta_i = 1 - \frac{1}{\lambda_i}$, we have

$$\begin{aligned} x_i &= \delta_i \left(\gamma_i + \sum_j x_j \pi_{ji} \right) \\ &= \delta_i \gamma_i + [x_1 \cdots x_n] \begin{bmatrix} \delta_i \pi_{1i} \\ \vdots \\ \delta_i \pi_{ni} \end{bmatrix} \end{aligned} \quad (28)$$

Let $x = [x_1 \cdots x_n]$, $\gamma = \gamma_1 \cdots \gamma_n$, and define the diagonal matrix Δ as follows.

$$\Delta = \begin{bmatrix} \delta_1 & & \\ & \ddots & \\ & & \delta_n \end{bmatrix} \quad (29)$$

Then we can write Eq. (28) in vector form as:

$$x = \gamma \Delta + x \Pi \Delta$$

Solving for x ,

$$\begin{aligned} x &= \gamma \Delta (I - \Pi \Delta)^{-1} \\ &= \gamma \Delta (I + \Pi \Delta + (\Pi \Delta)^2 + (\Pi \Delta)^3 + \dots) \end{aligned} \quad (30)$$

The matrix $\Pi \Delta$ is given by

$$\Pi \Delta = \begin{bmatrix} 0 & \delta_2 \pi_{12} & \dots & \delta_n \pi_{1n} \\ \delta_1 \pi_{21} & 0 & & \delta_n \pi_{2n} \\ \vdots & & \ddots & \vdots \\ \delta_1 \pi_{n1} & \delta_2 \pi_{n2} & \dots & 0 \end{bmatrix} \quad (31)$$

The infinite series in Eq. (30) converges since the rows of $\Pi \Delta$ sum to a number strictly less than 1, so that the inverse $(I - \Pi \Delta)^{-1}$ is well defined.

Equation (30) suggests what to look for when gauging the extent of double-counting of lending to ultimate borrowers that results from the heavy use of funding raised from other financial intermediaries. The comparison is between γ (the profile of lending to the ultimate borrowers in the economy) and x (the profile of debt values across all banks, which gives a gross measure of balance sheet size). The factor that relates the two is the matrix:

$$\Delta(I + \Pi \Delta + (\Pi \Delta)^2 + (\Pi \Delta)^3 + \dots)$$

This matrix has a finite norm, since the infinite series $I + \Pi \Delta + (\Pi \Delta)^2 + (\Pi \Delta)^3 + \dots$ converges to $(I - \Pi \Delta)^{-1}$. However, for a financial system where leverage is high, and to the extent to which banks are interwoven tightly, the norm can grow without bound. This is because as leverage becomes large, $\delta_i \rightarrow 1$ and, hence, Δ tends to the identity matrix. Moreover, as the degree of interconnectedness between banks becomes large, the norm of the matrix Π converges to 1, since each row of Π will sum to a number that converges to 1. In the limit, as $\Delta \rightarrow I$ and $\|\Pi\| \rightarrow 1$, the norm of the matrix $\Delta(I + \Pi \Delta + (\Pi \Delta)^2 + (\Pi \Delta)^3 + \dots)$ grows without bound.

Consequently, the size of the financial intermediation sector relative to the size of the economy can vary hugely over the financial cycle. We illustrate this phenomenon in Figures 12 and 13, which show the growth of four sectors in the United States from 1954. The four sectors are (i) nonfinancial corporate, (ii) household, (iii) commercial banking, and (iv) security broker-dealer. The data are taken from the Federal Reserve's Flow of Funds accounts. The series are normalized so that the size in 1954Q1 is set equal to 1.

Over this time period, three of the four sectors had grown to roughly 80 times their 1954 size by the end of 2009. These trends, however, have been dwarfed by that of the broker-dealer sector, which had grown to around 800 times its 1954 level at the height of the boom, before collapsing in the recent crisis. Figure 13 is the same chart, but in

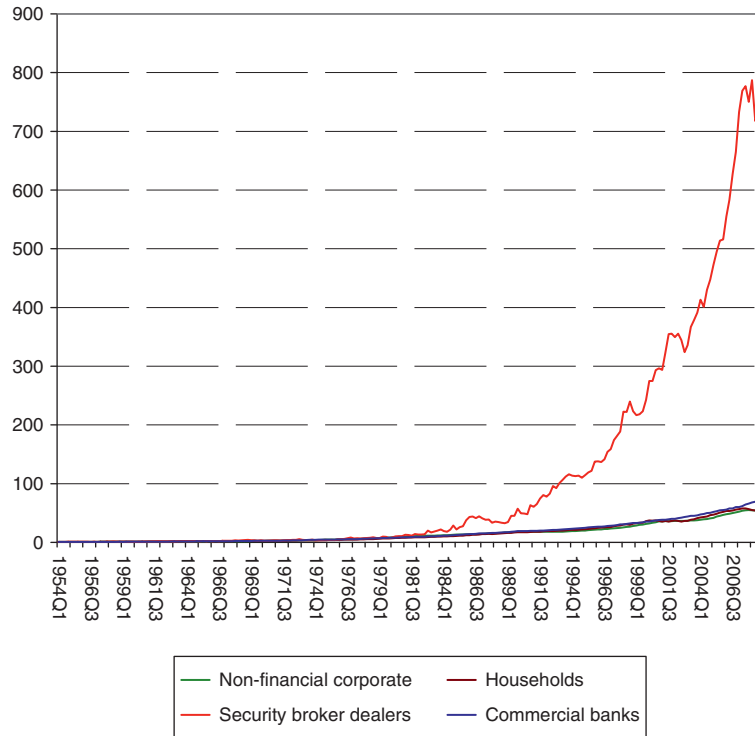


Figure 12 Growth of four U.S. sectors (1954Q1 = 1).

log scale. The greater detail afforded by the chart in log scale reveals that the securities sector kept pace with the rest of the economy up to about 1980, but then started a growth spurt that outstripped the other sectors. On the eve of the crisis, the size of the securities sector was roughly ten times that of the other sectors in the economy.

4. EMPIRICAL RELEVANCE OF FINANCIAL INTERMEDIARY BALANCE SHEETS

Our discussion thus far suggests that financial intermediaries deserve independent study in models of monetary economics due to their impact on financial conditions. Asset prices are influenced by the tightness of balance sheet constraints of financial intermediaries. In this section, we examine empirically whether financial intermediaries' impact on financial conditions can feed through to affect real economic outcomes. The analysis follows Adrian and Shin (2008c) and Adrian, Moench, and Shin (2009).

We label the looseness of balance sheet constraints "risk appetite" following Adrian, Moench, and Shin (2009, 2010) and Adrian, Etula, and Shin (2009). Risk appetite refers to the shadow value of capital of the (leveraged) intermediary sector in the model

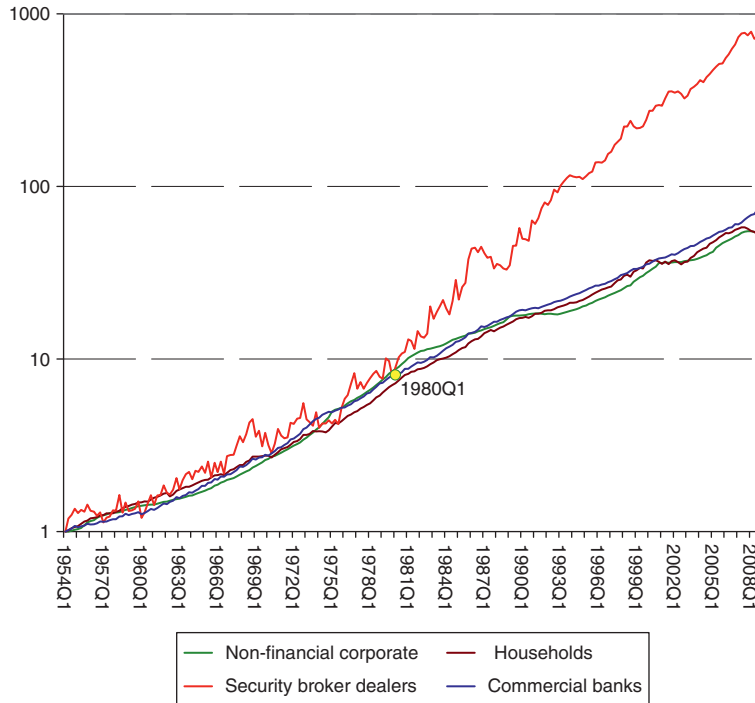


Figure 13 Growth of four U.S. sectors (1954Q1 = 1) (in log scale).

of Section 2. This shadow value of capital indicates the additional profit that the banking sector may earn by having one dollar of extra bank capital. The looser the capital constraint is, the lower the Lagrange multiplier, and hence the higher the risk appetite.

The terminology of “risk appetite” is intended to highlight the apparent change in preferences of the banking sector. We say “apparent” change in preferences, since the fluctuations in risk appetite are due to the constraints faced by the banks rather than their preferences as such. However, to an outside observer, the fluctuations in risk appetite would have the outward signs of fluctuations in risk preferences of the investor.

Adrian, Moench, and Shin (2009) estimated the risk appetite of financial intermediaries for a “macro risk premium.” The risk premium measures the hurdle rate of return for new projects that are financed in the economy, and hence reflects the ease of credit conditions, and corresponds to the risk premium of the model in Section 2. The macro risk premium is estimated from the yield spreads of fixed income securities. In particular, the macro risk premium is estimated as a linear combination of spreads closely tracking GDP growth. In doing so, we allow both term spreads of the Treasury yield curve and credit spreads to enter. Both term spreads and credit spreads are

measures of hurdle rates — the additional yields on longer dated or riskier bonds that induce market investors to fund additional investment or consumption.

To estimate the macro risk premium, Adrian, Moench, and Shin (2009) contemporaneously regress GDP growth on a wide variety of Treasury and credit spreads. They then use the seven constant maturity yields published in the H.15 release of the Federal Reserve Board and compute spreads relative to the federal funds target, and corporate bond spreads for credit ratings AAA, AA, A, BBB, BB, and B from Standard & Poors in excess of the 10-year constant maturity Treasury yield. The analysis of Adrian, Moench, and Shin (2009) starts in the first quarter of 1985, and ends in the fourth quarter of 2009.

An estimate of the macro risk premium is obtained as the fitted value of a linear regression of GDP growth on these corporate and Treasury bond spreads. Hence, empirically, the macro risk premium is a weighted average of spreads where the weights are given by the regression coefficients. The weights can be interpreted as portfolio weights of a portfolio that is tracking GDP growth. Conceptually, the macro risk premium represents the analog of the risk premium term π discussed in Section 2. The estimated macro risk premium, together with GDP growth, is plotted in Figure 14. For ease of interpretation, the macro risk premium is rotated using an affine transformation to match the average level and the volatility of the AA credit spread. The plot shows the strong negative correlation between the macro risk premium and GDP growth.

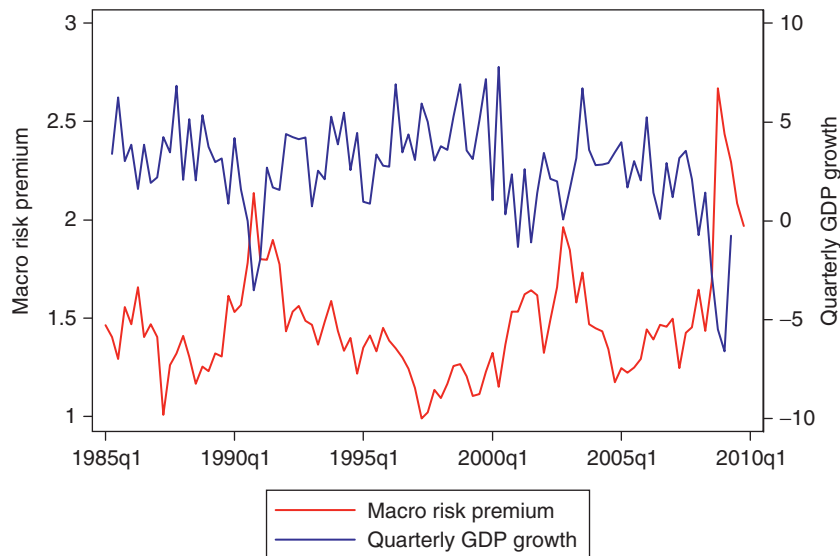


Figure 14 The macro risk premium and GDP growth.

We now turn to our measure of the looseness of financial intermediary capital constraints, which we have called “risk appetite.” As sketched in Section 2, the willingness of banks to lend will be positively associated with the size of intermediary balance sheets. The scenario outlined in Section 2 illustrates how financial intermediaries manage their balance sheets actively by employing a VaR constraint when choosing the size and composition of their portfolio.

Intuitively, the relationship between the macro risk premium, financial intermediary balance sheets, and real activity is as follows. When financial intermediaries have ample balance sheet capacity in terms of higher capital, their balance sheet constraints are loose, risk premia are compressed, and the supply of credit is plentiful, which in turn leads to lower threshold rates of return for real projects, and hence higher GDP growth. Effective risk aversion is low, and real growth is high. Conversely, when financial intermediary funding conditions worsen, their risk appetite declines, leading to lower real growth.

The macro risk premium measure can be interpreted as a portfolio of Treasury and corporate bond yields where the portfolio weights are chosen to maximize the contemporaneous correlation with real GDP growth. In a similar vein, we obtain a measure of intermediary risk appetite by finding the linear combination of one-year lagged balance sheet variables that best predict the (negative) one-year change of the macro risk premium. To a good approximation, the negative change reflects the return on the macro risk premium. A priori, it is not clear which institutions are the most important in determining risk premia for the different asset classes. We therefore build on our related work in Adrian, Moench, and Shin (2010) in which we use subset selection methods to identify the best predictors for excess returns on different asset classes among a large number of potential explanatory balance sheet proxies for various types of financial institutions. Adrian, Moench, and Shin (2010) documented that annual leverage growth of security brokers and dealers is a strong predictor of excess returns on equity and corporate bond portfolios and that quarterly shadow bank asset growth is a strong forecasting variable for excess returns on corporate and Treasury bonds. Adrian, Moench, and Shin (2009) therefore restricted the set of right-hand side variables to these two types of institutions, complemented by commercial banks to highlight the differential impact that balance sheets of market-based intermediaries have with respect to traditional banks. For each of these three types of institutions, one-year lagged asset growth and the growth of net worth are included as potential variables. Adrian, Moench, and Shin (2009) weighted the growth rates of assets and net worth by the relative size of total assets of each intermediary to capture the trends of assets under management across different institutions.

Risk appetite is estimated by regressing the (negative) change of the macro risk premium over one year on the balance sheet measures of the security broker-dealers, the shadow banks, and the commercial banks. These negative changes capture returns to

the risk premia. When the negative change of a risk premium is positive, prices today increase, leading to lower expected returns. The measure of risk appetite is constructed as the fitted value of the regressions of (the negative changes to the macro risk premium) on the intermediary balance sheet variables. The risk appetite measure is displayed together with the macro risk premium in Figure 15. The plot shows that risk appetite is highly negatively correlated with changes to the macro risk premium. Higher risk appetite leads to balance sheet expansions, which are associated with increases in asset prices and hence declines in spreads. Movements in risk appetite are thus strongly negatively correlated with the macro risk premium.

We can also investigate the importance of financial intermediary balance sheets for macroeconomic activity by relating intermediary balance sheets directly to GDP growth. Relative to commercial banks, broker-dealer and shadow bank balance sheets potentially hold more information related to underlying financial conditions, as they are a signal of the marginal availability of credit. At the margin, all financial intermediaries (including commercial banks) have to borrow in capital markets (for instance via commercial paper or repos). For commercial banks, however, their large balance sheets mask the effects operating at the margin. Broker-dealers or shadow banks, in contrast, give a purer signal of marginal funding conditions, as (i) their liabilities are short term and (ii) their balance sheets are closer to being fully marked to market.

In addition, broker-dealers originate and serve as market makers for securitized products, whose availability determines the credit supply for consumers and nonfinancial

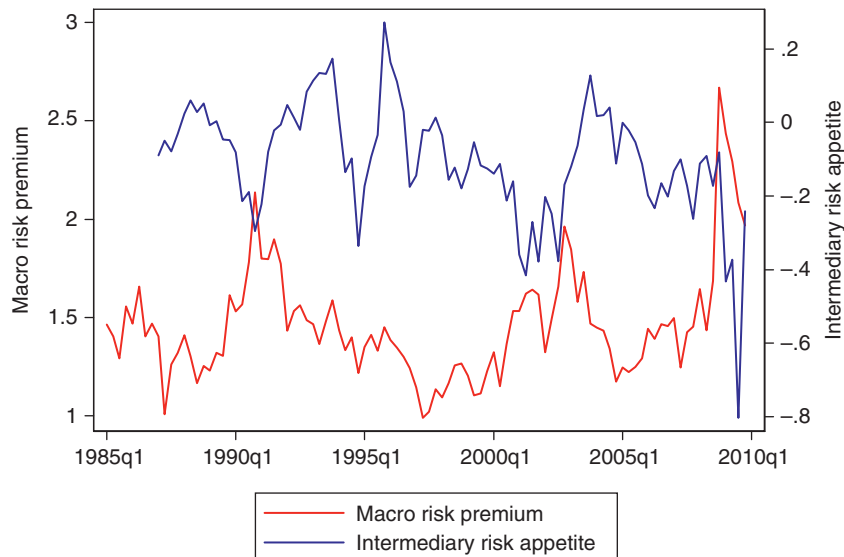


Figure 15 Macro risk premium and intermediary risk appetite.

firms (e.g., for mortgages, car loans, student loans, etc.). Consequently, broker-dealers are important variables not only because they are the marginal suppliers of credit, but also because their balance sheets reflect the financing constraints of the market-based financial system.

To the extent that balance sheet dynamics affect the supply of credit, they have the potential to affect real economic variables. To examine this empirically, we estimate macroeconomic forecasting regressions. In Table 1, we report the results of regressions of the quarterly growth rate of GDP components on lagged macroeconomic and financial variables. In addition, we add the lagged growth rate of total assets and market equity of security broker-dealers. By adding lags of additional financial variables (equity market return, equity market volatility, term spread, credit spread), we offset balance sheet movements that are purely due to a price effect. By adding the lagged macroeconomic variables, we control for balance sheet movements due to past macroeconomic conditions. In Table 1, (and all subsequent tables), * denotes statistical significance at the 10%, ** significance at the 5% level, and *** at the 1% level. All our empirical analysis is done using quarterly data from 1986Q1 to 2009Q2. Variable definitions are given in the data appendix at the end of this chapter. Figure 16 plots the impulse response on GDP growth of a shock to shadow bank asset growth.

Table 1 Impact of Balance Sheets on GDP

	(1) Quarterly GDP growth	(2) Quarterly GDP growth	(3) Quarterly GDP growth
Broker-dealer asset growth (lag)	0.03*		
Broker-dealer equity growth (lag)	0.18		
Shadow banks asset growth (lag)		0.21***	
Shadow banks Equity growth (lag)		0.71***	
Commercial bank asset growth (lag)			0.02
Commercial bank equity growth (lag)			−0.12
GDP growth (lag)	0.03	−0.18	0.09
PCE inflation (lag)	−1.01**	−1.00**	−1.16***
VIX (lag)	0.01	−0.03	−0.02
Credit spread (lag)	−1.37*	−1.81**	−1.01
Term spread (lag)	0.75**	1.18***	0.75*
Fed funds (lag)	0.40	0.19	0.49*
Constant	4.67***	4.94***	4.44**
Observations	93	93	93
R ²	0.288	0.409	0.263

Notes: This table reports regressions of quarterly GDP growth on the total asset and equity growth of broker-dealers, shadow banks, and commercial banks for 1986Q1 to 2009Q2. ***Denotes significance at the 1% level, **denotes significance at the 5% level, and *denotes significance at the 10% level. Significance is computed from robust standard errors.

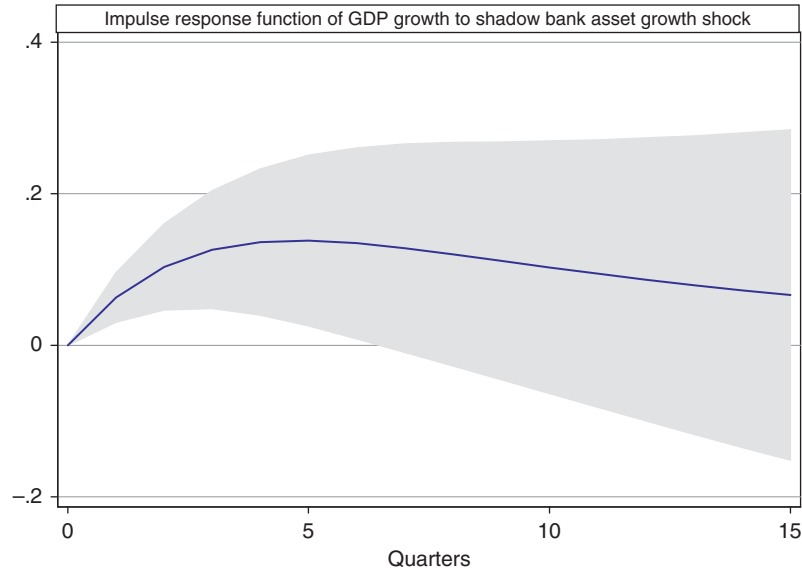


Figure 16 Plots the impulse response GDP growth from a shock to shadow bank total asset growth. The impulse response is estimated from a vector autoregression (VAR) with GDP growth, PCE inflation, shadow bank asset growth, credit spread, VIX, the term spread, and the federal fund target rate as variables, and that ordering is used to produce the impulse response functions using a Cholesky decomposition. The time span is 1986Q1 to 2009Q1.

The growth rate of security broker-dealer total assets has strongest significance for the growth rate of future housing investment and weak significance for total GDP growth (Tables 1 and 2, columns 1). Our interpretation of this finding is that the mechanisms that determine the liquidity and leverage of broker-dealers affect the supply of credit, which, in turn, affects investment and consumption. The total assets and total equity of shadow banks have significant forecasting power for total GDP growth, reflecting their increased role in the total supply of credit for the U.S. economy. Commercial banks, on the other hand, have no forecasting power for GDP, and forecast housing growth with the wrong sign. Adrian, Moench, and Shin (2010) systematically investigated the forecasting power of all financial intermediaries in the U.S. Flow of Funds, and confirm that broker-dealers and shadow banks forecast real activity.

The forecasting power of dealer assets for housing investment is graphically illustrated in Figure 17. The impulse response function is computed from a first order VAR that includes all variables of Table 1, column 1. The plot shows that the response of housing investment to a positive shock in broker-dealer assets growth is positive, large, and persistent.

Table 2 Impact of balance sheets on housing investment

	(1) Quarterly GDP growth	(2) Quarterly GDP growth	(3) Quarterly GDP growth
Broker-dealer asset growth (lag)	0.09***		
Broker-dealer equity growth (lag)	0.10		
Shadow banks asset growth (lag)		−0.00	
Shadow banks equity growth (lag)		0.14	
Commercial bank asset growth (lag)			−0.44*
Commercial bank equity growth (lag)			0.23
Housing growth (lag)	0.89***	0.94***	0.096***
PCE inflation (lag)	−0.30	−0.11	−0.09
VIX (lag)	0.11	0.01	0.03
Credit spread (lag)	−0.92	−0.49	0.01
Term spread (lag)	1.11**	0.60	−0.07
Fed funds (lag)	−0.06	−0.04	−0.27
Constant	−2.53	0.13	3.76
Observations	93	93	93
R ²	0.902	0.881	0.888

Notes: This table reports regressions of quarterly residential investment growth on the total asset and equity growth of broker-dealers, shadow banks, and commercial banks for 1986Q1 to 2009Q2. ***Denotes significance at the 1% level, **denotes significance at the 5% level, and *denotes significance at the 10% level. Significance is computed from robust standard errors.

The differences between the interactions of market based intermediaries and commercial banks with the macroeconomic aggregates are further highlighted in column 3 of Table 1, where we see that commercial bank assets and equity do less well than shadow bank or security-broker-dealer variables as forecasting variables. Our interpretation of these findings is that commercial bank balance sheets are less informative than broker-dealer balance sheets as they (largely) did not mark their balance sheets to market over the time span in our regressions. In addition, in Table 2, we find that growth in commercial bank total assets precedes declines in housing investment. This result is primarily due to the fact that commercial banks offer credit line pre-commitments that tend to be drawn in times of crisis. In fact, in Figure 9, we saw that commercial bank total assets grew at the onset of the recent financial crisis, as structured credit was re-intermediated onto commercial bank balance sheets. The finding that commercial bank assets do not predict future real growth is also consistent with Bernanke and Lown (1991), who used a cross-sectional approach to show that credit losses in the late 1980s and early 1990s did not have a significant impact on real economic growth across states. See Kashyap and Stein (1994) for an overview of the debate on whether there was a “credit crunch” in the recession in the early 1990s.

In the same vein, Ashcraft (2006) found small effects of variations in commercial bank loans on real activity when using accounting based loan data. However, Ashcraft

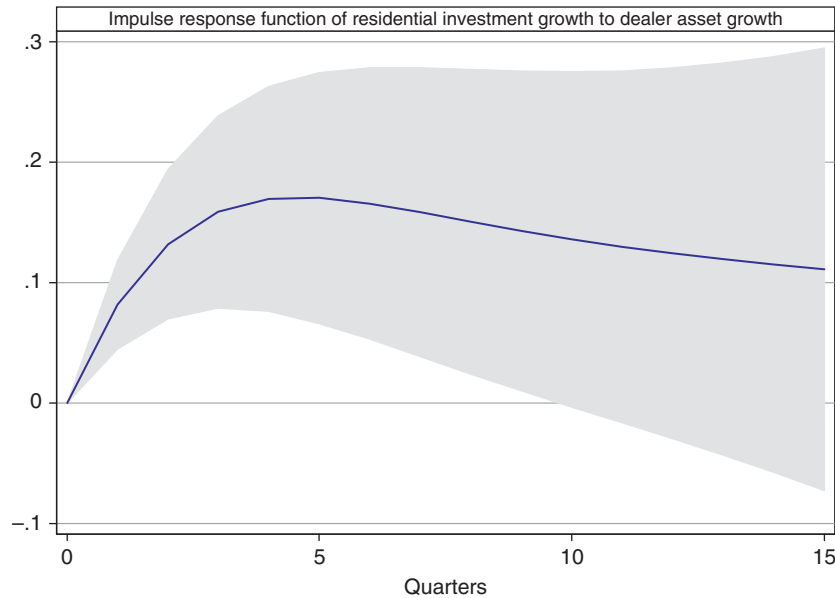


Figure 17 Plots the impulse response of residential investment growth from a shock to broker-dealer total asset growth. The impulse response is estimated from a VAR with residential investment growth, PCE inflation, broker-dealer asset growth, credit spread, VIX, the term spread, and the federal fund target rate as variables, and that ordering is used to produce the impulse response functions using a Cholesky decomposition. The time span is 1986Q1 to 2009Q1.

(2005) found large and persistent effects of commercial bank closures on real output (using FDIC induced failures as instruments). Lown and Morgan (2006) showed that the senior loan officer survey provides significant explanatory power for real activity — again, a variable that (i) is more likely to reflect underlying credit supply conditions and (ii) is not based on accounting data.

The credit supply channel sketched so far differs from the financial amplification mechanisms of Bernanke and Gertler (1989) and Kiyotaki and Moore (1997, 2005). These papers focus on amplification due to financing frictions in the borrowing sector, while we focus on amplification due to financing frictions in the lending sector. We return to a more thorough review of the literature in a later section.

5. CENTRAL BANK AS LENDER OF LAST RESORT

The classical role of the central bank as the lender of last resort (LOLR) is framed in terms of meeting panics that affect solvent, but illiquid, banks. In the simplest case, bank runs arise when depositors fail to achieve coordination in a situation with

multiple equilibria. For example, in Bryant (1980) and Diamond and Dybvig (1983), an individual depositor runs for fear that others will run, leaving no assets in place for those who do not run.

However, in the financial crisis of 2007–2009, the withdrawal of credit was not restricted to one of even a subset of institutions. Instead, entire market sectors were targeted. Figure 18 plots the new issuance of ABS over a three-month interval preceding the measurement date, and clearly illustrates the generalized contraction of credit. If there was a run driven by a coordination failure, it was a simultaneous run from all institutions in the financial system. Albeit, the extent to which each institution suffered from the run depended on its particular vulnerability. In the model outlined in Section 2, it is the interaction between measured risks and the risk-bearing capacity of banks that determines overall lending. Financial institutions that rely on VaR cut back lending when risk constraints become more binding (i.e., when the Lagrange multiplier associate with the constraint increases). From the point of view of an individual bank, this prudent cutting of exposures by creditors closely resembles a “run.” In this sense, the runs on Northern Rock, Bear Stearns, and Lehman Brothers may be better seen as

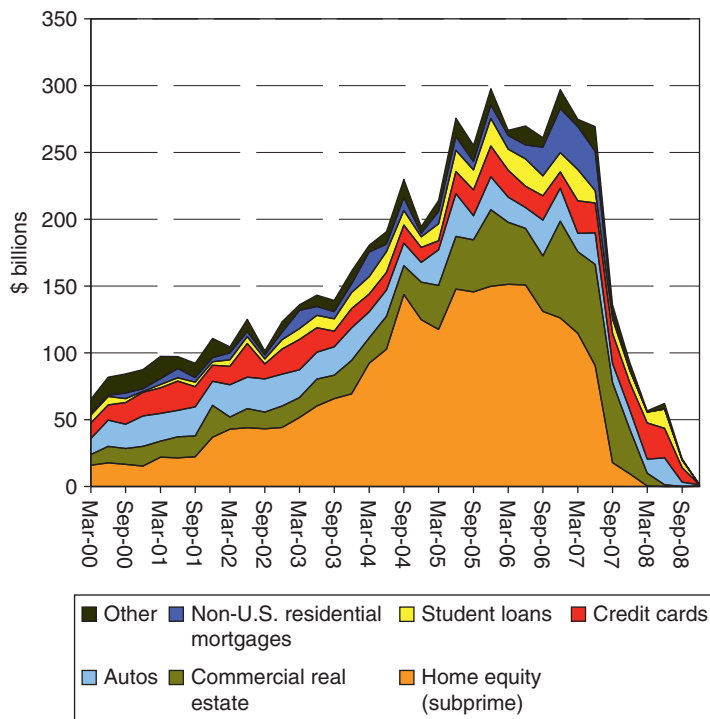


Figure 18 New issuance of ABS in previous three months.

the tightening of constraints on the creditors of these banks, rather than as a coordination failure among them.

We should not draw too fine a distinction between the coordination view of bank runs on the one hand, and the “leverage constraints” view on the other. Coordination (or lack thereof) will clearly exacerbate the severity of any run when a bank has many creditors. The point is, that an explanation of a run on the system needs to appeal to more than just coordination failures. For example, this means that explanations of the runs on Bear Stearns or Lehman Brothers, should make reference to market-wide factors, as well as to the particular characteristics of those firms and their creditors. This is one more instance of the general maxim that, in a modern market-based financial system, banking and capital market conditions cannot be viewed in isolation.

To the extent that the credit crunch can be seen as the consequence of a collapse of balance sheet capacity in the financial intermediary sector, we can interpret the policy response by central banks as an attempt to restore this lost capacity by lending directly into the market. The Federal Reserve has been one of the most aggressive central banks in this context, effectively interposing its own balance sheet between the banking sector and the ultimate borrowers. The Federal Reserve has taken in deposits from the banking sector (through increased reserves) and then lent out the proceeds to ultimate borrowers through the holding of securities (Treasuries, mortgage backed securities, commercial paper, and other private sector liabilities), as well as through currency swap lines to foreign central banks. One indication of the dramatic increases in the Federal Reserve’s balance sheet can be seen in the sharp increase in the holding of cash by U.S. commercial banks, as seen from Figure 19. The increased cash holdings reflect the sharp increase in reserves held at the Federal Reserve — a liability of the Federal Reserve to the commercial banks.

In this way, central bank liquidity facilities have countered the shrinking of intermediary balance sheets and have become a key plank of policy, especially after short-term interest rates were pushed close to their zero bound. The management of the increased Federal Reserve balance sheet has been facilitated by the introduction of interest on reserves on October 1, 2008, which effectively separates the management of balance sheet size from that of the Federal Funds interest rate (see Keister & McAndrews, 2009 for a discussion of the interest on reserve payment on the Federal Reserve’s balance sheet management).

The Federal Reserve has also put in place various other LOLR programs to cushion the strains on balance sheets, and to target the unusually wide spreads in a variety of credit markets. Liquidity facilities have been aimed at the repo market (Term Securities Loan Facility TSLF and Primary Dealer Credit Facility PDCF), the commercial paper market (commercial paper funding facility; CPFF and Asset-Backed Commercial Paper Money Market Fund Liquidity Facility AMLF), and ABS markets (term asset-backed loan facility; TALF). In addition, the Federal Reserve has conducted outright purchases of Treasury

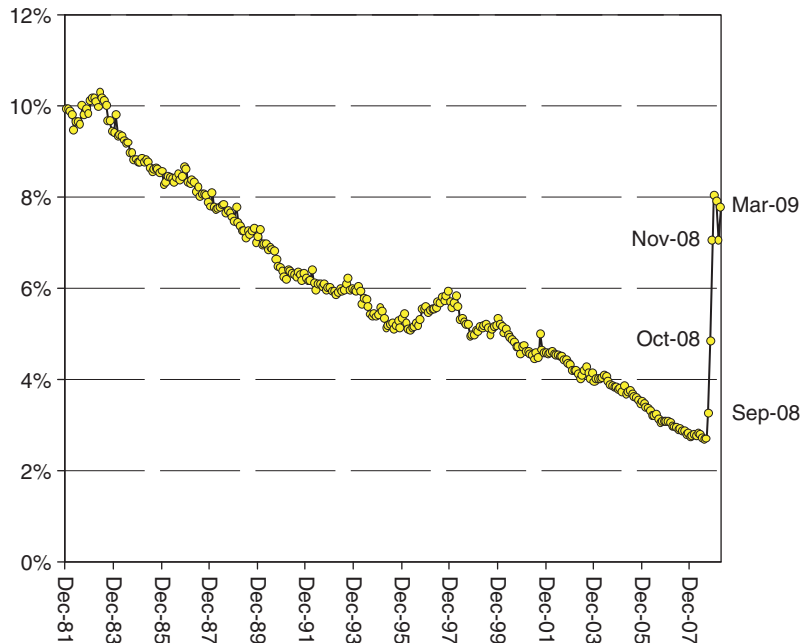


Figure 19 Cash as proportion of U.S. commercial bank assets. (Source: Federal Reserve, H8 database).

and agency securities, and has provided dollar liquidity in the FX futures markets (FX Swap lines). The common motivating element in these policies has been to try and alleviate the strains associated with the shrinking balance sheets of intermediaries by substituting the central bank's own balance sheet. The spirit of these policies differs from that of classic monetary policy in that they are explicitly aimed at replacing the collapse of private sector balance sheet capacity. Since the deleveraging of financial intermediary balance sheets is associated with a widening of risk premia, the effectiveness of balance sheet policies can be judged by the level of risk premia in various financial markets. In practice, the degree to which risk premia are associated with the expansions and contractions of intermediary balance sheets are important indicators for the risk appetite of the financial sector, which, in turn, affect credit supply and real activity. Adrian, Moench, and Shin (2010) used this insight to decompose the risk premia of several asset classes into various components, including those associated with the risk appetite of financial intermediaries.

One instance of the Federal Reserve's liquidity facilities can be seen in Figure 20 (taken from Adrian, Marchioni, and Kimbrough (2010)), which charts total outstanding commercial paper alongside net Federal Reserve commercial paper holdings in the CPFF.⁶ Following the Lehman Brothers bankruptcy in September 2008, the

⁶ See Adrian, Marchioni, and Kimbrough (2010) for a detailed description of the CPFF.

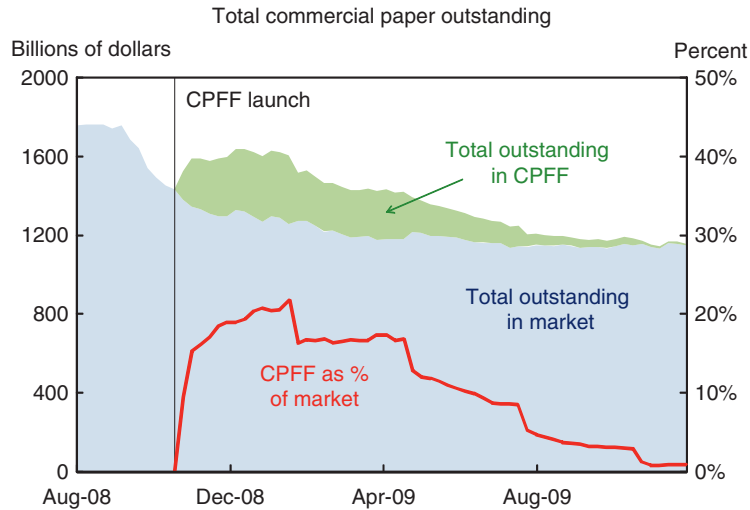


Figure 20 The Federal Reserve's commercial paper funding facility (CPFF).

outstanding amount of commercial paper began to fall precipitously, as can be seen by the sharp downward shift in the shaded blue area. With the creation of the CPFF in October 2008, the Federal Reserve's holdings of commercial paper in the CPFF began to increase rapidly, as shown by the green area in Figure 20. The Federal Reserve's holdings can be seen to replace virtually dollar for dollar the decline in the outstanding amount of commercial paper. In this respect, the Federal Reserve's balance sheet was used to directly replace the decline in balance sheet capacity of the financial intermediary sector. An important feature of the CPFF is that, as the facility was intended to be only a temporary liquidity backstop, it was designed to become more unattractive as market conditions begin to normalize. Accordingly, while the red line in Figure 20 shows that the Federal Reserve held as much as 20% of all outstanding commercial paper at the height of the crisis, CPFF holdings have since fallen steadily as market functioning has improved.

As another example, Figure 18 illustrated how new issuance of ABS had collapsed by the end of 2008. In response, the Federal Reserve instituted TALF, where the central bank provides secured loans to new AAA-rated ABS at a low haircut to private sector investors. TALF was designed specifically to revitalize the ABS market, and its effectiveness can be gauged by Figure 21, which shows the effect on new issuance of ABS before and after the introduction of TALF. The light colored bars on the right show that much of the recent issuance of ABS has been TALF-financed, and that TALF-backed issuance dwarfs the issuance of standard issues.

The balance sheet expansion of the Federal Reserve in response to the financial crisis of 2007–2009 has refocused the monetary policy debate on the role of quantities in

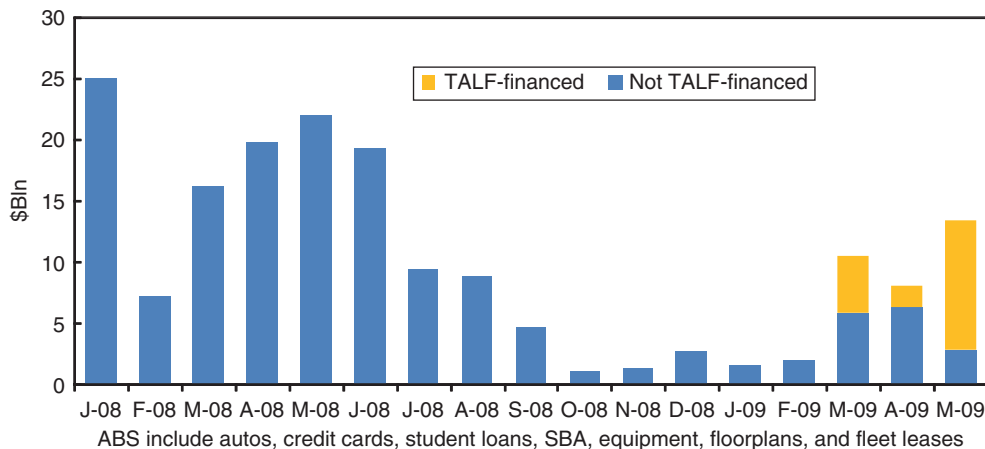


Figure 21 New ABS issuance under TALF.

the monetary policy transmission mechanism. The crisis forcefully demonstrated that the collapse of balance sheet capacity in the financial sector can have powerful adverse affects on the real economy, and, accordingly, the traditional role of the central bank as the LOLR has undergone far-reaching innovations.

6. ROLE OF SHORT-TERM INTEREST RATES

Having established that increases in broker-dealer and shadow bank balance sheets signal increases in real activity, we now investigate the determinants of balance sheet growth. Broker-dealers, shadow banks, and commercial banks fund themselves with short-term debt. Broker-dealers are primarily funded in the repo market (see Figure 11); shadow banks are primarily funded in the commercial paper market (see Figure 10); and the majority of commercial banks' short-term funding is through money (i.e., checking and savings deposits). In the case of broker-dealers, part of the repo funding is directly passed on to other leveraged institutions, such as hedge funds, in the form of reverse repos, while another part is invested in longer term, less liquid securities. Shadow banks, on the other hand, tend to fund holdings of ABS and MBS directly. Commercial banks primarily hold nontradable loans.

Because the majority of the liability side of financial institutions comes from short-term borrowing arrangements, their cost of borrowing is tightly linked to short-term interest rates, such as the federal funds target rate. As broker-dealers and shadow banks hold longer term assets, proxies for their expected returns are spreads — particularly term spreads — which capture the maturity transformation of financial institutions. The leverage of these intermediaries is constrained by risk: in more volatile markets,

leverage is more risky, margins and haircuts are higher, and credit supply tends to be more constrained. We saw in Section 2 how VaR determined balance sheet size, risk premia, and credit supply.

Much of these balance sheet adjustments occur at high frequencies. Although the total assets used in the previous regressions are available only at a quarterly frequency, on the liability side of the balance sheet, there are weekly data available on outstanding repo, outstanding commercial paper, and total money. We use repo data that are collected for the primary dealer universe by the Federal Reserve Bank of New York. Outstanding commercial paper is collected by the depository trust corporation (DTC), and is published at a weekly frequency by the Federal Reserve Board. The broad money measure M2 is also made available by the Federal Reserve Board.

We find that increases in the federal funds target rate are generally associated with a slower growth rate of short-term liabilities. In Table 3, we show regressions of growth rates of repo, repo + commercial paper, and M2 on changes of the federal funds target as well as on other asset prices (and lags of the left hand side variables). The three types of regressions correspond to the funding of the three main financial institutions: broker-dealers, shadow banks, and commercial banks. In each case, increases in the federal funds target are associated with declines in the short-term funding liabilities.

Financial market volatility, as measured by the Chicago Board Options Exchange Volatility Index (VIX), relates negatively to security repo growth and repo + cp

Table 3 Determinants of balance sheet growth

	(1) Repo (Weekly growth)	(2) Repo + CP growth (Weekly growth)	(3) M2 growth (Weekly growth)
Fed funds (1 week change)	−0.630***	−0.355***	−0.054***
Equity return (1 week)	−0.022*	−0.013*	0.001**
VIX (1 week change)	−0.052	−0.027	0.001
Treasury spread (1 week change)	0.703	0.291	0.151**
Credit spread (1 week change)	0.311	0.031	0.337**
Repo growth (1 week lag)	−0.134***	−0.075***	−0.001
CP growth (1 week lag)	0.022	0.028	−0.020
M2 growth (1 week lag)	0.515	0.063	−0.016
Constant	0.136*	0.105**	0.050***
Observations	990	990	989
R ²	0.042	0.032	0.121

Notes: This table reports regressions of repo + commercial paper growth, and M2 growth on their own lags, and asset price variables. The data frequency is weekly from October 3, 1990 to February 3, 2010. Changes refer to one-week changes, and lags to one-week lags. ***Denotes significance at the 1% level, **denotes significance at the 5% level, and *denotes significance at the 10% level. Significance is computed from robust standard errors.

growth. As higher volatility is associated with higher haircuts and tighter capital constraints, both induce tighter constraints on dealer leverage (columns 1 and 2). For M2, we find that higher VIX is associated with larger money growth, which we interpret as flight to quality: in times of crisis, households and nonfinancial corporations tend to reallocate short-term savings to commercial banks (see Gatev, Schuermann, & Strahan, 2009).

Increases in the term spread are associated with higher repo growth. This finding is consistent with the notion that financial intermediaries fund themselves with short-term debt, but lend out longer term, so that a higher term spread increases the carry between assets and liabilities and is associated with larger balance sheets.

6.1 The risk-taking channel of monetary policy

Current models in monetary economics emphasize the importance of managing market expectations. By charting a path for future short rates and communicating this path clearly to the market, the central bank can influence long rates and influence mortgage rates, corporate lending rates, and other prices that affect consumption and investment. In contrast, our findings point to the short-term interest rate as an important price variable in its own right. Empirically, we have seen that the federal funds rate is an important explanatory variable for the growth of balance sheet aggregates. Our model suggests that increasing bank capital increases the risk-taking capacity of the banking system, which in turn leads to a lower equilibrium risk premium, and an increase in the supply of credit by lowering the hurdle rate at which projects are financed.

Banks and other financial intermediaries borrow in order to lend. Since the loans offered by banks tend to be of longer maturity than the liabilities that fund those loans, the term spread is indicative of the marginal profitability of an extra dollar of loans on intermediaries' balance sheets. The NIM of the bank is the difference between the total interest income on the asset side of its balance sheet and the interest expense on the liabilities side of its balance sheet. Whereas the term spread indicates the profitability of the marginal loan that is added to the balance sheet, the NIM is an average concept that applies to the stock of all loans and liabilities on the balance sheet.

The NIM determines the profitability of bank lending and increases the present value of bank income, boosting the forward-looking measures of bank capital. Such a boost in bank capital increases the capacity of the bank to increase lending in the sense that the marginal loan that was not made before the boost in bank capital now becomes feasible under the greater risk-bearing capacity of the bank. As banks expand their balance sheets, the market price of risk falls.

The logic of the argument is illustrated in Figure 22 for the case of monetary easing. A decline in the federal funds target leads to an increase in asset prices and, thus, an increase in the net worth of financial institutions. In response, levered financial

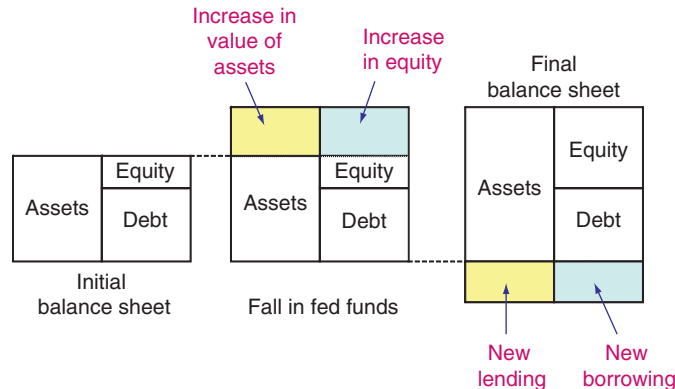


Figure 22 The impact of a decline in the federal funds rate on intermediary balance sheets.

institutions expand their assets by taking on additional leverage. Thus, the impact of changes in short-term interest rates is amplified via the balance sheet management of financial institutions.

For the argument to go through a key assumption the term spread is determined in large part by the short-term interest rate. On this score, the evidence is supportive. Figure 1 highlights the striking fact that there is a near perfect negative one-to-one relationship between 4-quarter changes of the federal funds target and 4-quarter changes of the term spread (the plot uses data from 1987q1 to 2008q3). In this way, variations in the target rate have a one-for-one relationship with the slope of the yield curve. Since the term spread leads the NIM, shifts in the short-rate affect real activity because they change the profitability of financial intermediaries, thus shifting the supply of credit.

The connection between financial intermediary balance sheet management, the slope of the yield curve, and real economic activity in the United States was recently examined by Adrian, Estrella, and Shin (2010), who conducted a VAR study of the relationship between the following variables: quarterly GDP growth as a measure of real activity, the 10-year/3-month term Treasury spread, the NIM of large commercial banks from their Y-9C filings, the quarterly asset growth of shadow banks, the 3-month Treasury yield as a measure of the short-term interest rate, and the quarterly change in the VIX as a measure of risk. The balance sheet aggregate of the intermediary sector is the total assets of shadow banks defined as the sum of total assets of ABS, finance companies, and funding corporations (each component is pulled from the Federal Reserve's Flow of Funds). The VAR includes one lag of each of the variables, as suggested by the Bayesian Information Criterion, and is estimated over the period from 1990Q3 to 2008Q3, where the starting date is determined by availability of the VIX data.

Adrian, Estrella, and Shin (2010) presented empirical results consistent with the following logic. An increase in the term spread tends to increase NIM. This is fairly mechanical as the term spread directly impacts NIM for newly originated loans funded with shorter term liabilities. Higher NIM — a major source of profits for financial intermediaries — leads to an increase in total assets of financial intermediaries: as lending becomes more profitable, the supply of credit is expanded and intermediaries' balance sheets grow. Larger asset growth of intermediaries, in turn, predicts higher GDP growth, which we interpret as a shift in the supply of credit curve. Since the VAR includes only one lag of each variable, the significance levels of the coefficients may also be interpreted as a set of Granger causality tests. These tests are consistent with our hypothesis of a causal chain that runs from the term spread to NIM to lending volume and finally to real growth.

Adrian, Estrella, and Shin (2010) also conducted impulse response studies to verify the main strands in the narrative. In the impulse response studies, a positive shock to the term spread leads to statistically significant increases in NIM over a considerable horizon. The shape of these responses is also consistent with the fact that average NIM tends to trail marginal changes in the term spread, as argued before. Moreover, a positive shock to NIM tends to increase lending by the shadow banking sector. Finally, a shock to asset growth in shadow banking has a quick and significant effect on real economic growth.

The evidence is supportive of the “risk-taking channel” of monetary policy. Variations in short-term interest rates lead real economic outcomes through their impact on the slope of the yield curve. Our interpretation of this evidence is an economic mechanism that operates via the balance sheet management of financial intermediaries who borrow short and lend long. Tighter policy leads to a compression of NIM and causes intermediaries to reduce lending. The flatter the term spread at the end of the tightening cycle, the greater the subsequent reduction in lending activity. This has a direct effect on the supply of credit to the real economy.

6.2 Two case studies

Two recent empirical papers throw more light on the channel of monetary policy that works through changes in the market value of existing loans. Jimenez, Ongena, Peydro, and Saurina (2008) examined a large database of European loans through the detailed information contained in the loan register and showed that a lower short-term interest rate lowers the hazard rate of default on *existing* loans. In addition, they showed that the hazard rate of default for *new loans* increases after the cut in short-term rates.

The fact that the riskiness of existing loans decline may be due to a fall in the interest burden of the borrower. The increased credit quality of the assets will give rise to an increase in the profitability of the lending, and lead to greater lending capacity, as

previously outlined. However, it is the second finding that is more telling. The fact that the riskiness of new lending increases suggests that the new loans are of lower quality, suggesting that the hurdle rate for lending has fallen. Such a combination of (i) greater lending capacity and (ii) erosion of lending standards is consistent with the risk-taking channel of monetary policy.

The same combination of (i) a lowering of a hazard rate of default on *existing* loans and (ii) an increase in the hazard rate of default on *new* loans is also observed in Ioannidou, Ongena, and Peydro (2009). In this study, the authors examined the effect of shifts of the U.S. federal funds rate on the quality of bank loans in Bolivia, which had a banking system close to being dollarized. To the extent that the U.S. federal funds rate is determined independently of the events in Bolivia, the authors regard the effect of short-term interest rate changes as a quasi-natural experiment of the effect of short-term interest rate movements on bank asset quality. As with the paper by Jimenez et al. (2008), the Bolivian study reveals the same combination where a cut in the U.S. federal funds rate leads to an improvement in the quality of existing assets, but new assets are of a lower quality. Paravisini (2008) provided estimates of the impact of bank funding constraints on the supply of bank credit using an instrumental variable approach, and Khwaja and Mian (2008) provided estimates of a bank funding shocks for an emerging market. Freixas (2009) provided an overview of the monetary transmission literature in the context of the 2007–2009 global financial crisis.

This combination of results on existing and new loans suggests that the risk-taking channel is a potentially fruitful avenue for further study. The model in Section 2 provides some of the conceptual background that may be necessary to understand the results.

6.3 Related literature

To highlight what we view as the specific contribution of the “risk-taking channel” of monetary policy, it is important to give an account of the points of contact between our approach and the existing literature in monetary economics and corporate finance. The discussion can be organized along a number of dimensions, but one useful classification is to distinguish those papers that have emphasized the *borrower’s* balance sheet (and the *demand* for credit) from those emphasizing the *lender’s* balance sheet (and the *supply* of credit).

Bernanke and Gertler (1989) is a classic example of an explanation based on the borrower’s balance sheet. Following the earlier work of Bernanke (1983), who argued for the importance of borrower balance sheet distress during the Great Depression, the Bernanke and Gertler model focuses on the agency problems entailed by the asymmetry of information between a nonfinancial corporate borrower and the financial market as a whole.

In the presence of asymmetric information between the borrower and lender, inefficiencies in the optimal contract manifest in the form of deadweight costs and, in

particular, in the spread between the cost of internal funds and that of external funds. The size of the deadweight cost is a function of the net worth of the borrower; that is, the borrower's "skin in the game." Moreover, Bernanke and Gertler (1989) argued that the borrower's net worth is procyclical, and so the funding spread between the internal and external funds should be countercyclical.

An alternative approach that emphasizes the borrower's balance sheet is the work on credit cycles by Kiyotaki and Moore (1997, 2005), who examined the dynamic amplification of credit constraints. In their approach, Kiyotaki and Moore assumed a collateral constraint where the size of the loan that can be obtained by a borrower depends on the current market price of the collateral that can be pledged to the lender. Under such an assumption, the size of the borrower's balance sheet can depend positively on the market price of the asset — in other words, the demand reaction of the borrower can be upward-sloping. When the price of an asset increases, the borrower's funding ability increases, generating larger balance sheets. When the greater demand for the asset pushes up the price of the asset, there is the possibility of amplified responses where asset price increases fuels further investment and aggregate activity, which raises prices further.

The common thread between the work of Bernanke and Gertler (1989) and Kiyotaki and Moore (1997, 2005) is that the focus is on the *borrower's* balance sheet, and the fluctuations in the creditworthiness of the borrower. The supply of lending is determined in the market as a whole, without a separate role for the banking sector.

However, to the extent that the borrower in the Bernanke and Gertler (1989) model can be reinterpreted as a bank, the model can be reoriented in terms of the agency problems in the banking sector. The "double-decker" moral hazard model in Holmström and Tirole (1997) is a good illustration of such a reinterpretation where the banking sector enters the model as a borrower subject to borrowing constraints from its lenders. In the Holmström and Tirole (1997) model, there are two tiers in the agency problem. At the bottom tier, there is a moral hazard problem between a nonfinancial borrower who needs funding to undertake a project and a bank who supplies the funding. The moral hazard problem entails an optimal contract where the incentive constraint stipulates that the borrower has enough of a stake in the project that the good action is taken, rather than the inefficient action that yields private benefit. The importance of borrower net worth, or skin in the game, is a theme that Holmström and Tirole (1997) shared with Bernanke and Gertler (1989). However, the innovation in Holmström and Tirole (1997) is that there is a second tier to the agency problem in which the bank (the lender) is subject to a moral hazard problem, so that there is a constraint on the minimum equity capital that the bank must hold at all times. In this respect, the minimum capital requirement of banks emerges as an endogenous feature of an agency problem where banks must raise funding from depositors and other suppliers of funds.

By reorienting the agency problem so that the focus is on the bank (as borrower) and the financial market (as lenders), the earlier results of Bernanke and Gertler and Kiyotaki and Moore can be transferred to the context of bank distress and bank lending. The recent paper by Gertler and Kiyotaki (2009, Chapter 11 this volume) is a good example of such a reorientation. The agency relationship between the bank and the financial market lender is a moral hazard problem due to the possibility that the bank can steal some portion of the project outcome, so that the bank is required to keep a minimum net worth in place at all times. The bank must then keep a minimum amount of skin in the game, which translates to a minimum capital ratio that the bank must maintain. When credit losses or a fall in the price of assets depletes the capital of the bank, the incentive constraint binds, entailing the withdrawal of lending by the bank.

Adrian and Shin (2008a) take up a similar theme of the binding incentive constraint of the bank, where the agency problem comes in the form of a risk-shifting problem where the bank may take the riskier asset when a lower risk asset may be value-enhancing for the pair as a whole. The bank's market-determined minimum capital requirement arises from the need for the bank (the borrower) to keep sufficient stake in the payoffs from the total balance sheet of the bank. Since the agency problem manifests as a risk-shifting problem, Adrian and Shin (2008a) addressed how second-moment incentives can enter the problem, and how the VaR constraint can emerge as an outcome of the optimal contracting problem.

Brunnermeier and Sannikov (2009) took the moral hazard theme one step further by embedding the moral hazard problem in a dynamic, continuous time contracting environment. In this richer framework, Brunnermeier and Sannikov (2009) examined two separate incentive constraints. One constraint is familiar where the borrower needs to keep sufficient skin in the game, and results in a minimum capital ratio requirement set by the market. The second is a constraint in the spirit of a VaR constraint, which makes the debt instantaneously risk-free. To accommodate a role for both types of constraints, Brunnermeier and Sannikov (2009) incorporated the innovative feature that two types of equity play an essential role. First, there is equity that carries control rights. This first type of equity is the stake of the controlling party. Second, there is equity that is loss-absorbing, but which does not carry control rights. This is the type of equity that is typified by passive investors in hedge funds whose stake can be returned by the controlling investor in the hedge fund. The interaction between these two types of equity is a distinctive feature of the Brunnermeier and Sannikov (2009) model, as well as the dynamic contracting framework, which gives the model considerable richness and complexity. The role of equity as a buffer has an affinity with the work on the effect of regulatory capital on lending as examined by Van den Heuvel (2002).

So far, we have described the set of papers that has their starting point as some type of agency problem between a borrower and a lender. However, there is another strand

of the literature from monetary economics that has emphasized the institutional features surrounding the commercial banking industry, especially for the United States.

A notable example is Bernanke and Blinder (1988), which proposed a simple model of the supply of credit by banks that emphasized the binding nature of the reserve requirement of banks. The constraint is not motivated with further microfoundations; instead, the institution of reserve requirements is taken as given. The reserve requirement stipulates that, for a given amount of deposit funding used by a bank, some minimum amount must be kept on deposit at the Federal Reserve as cash assets of the bank. This is a constraint that links the two sides of the balance sheet, and the assumption is that such a constraint binds all the time.

The idea that reserve requirements bind all the time has been dealt a severe blow by the experience of the aftermath of the recent financial crisis in which commercial banks in the United States held close to one trillion dollars of excess reserves on their balance sheet. Excess reserves have also been a common feature in other parts of the world after the financial crisis. Nevertheless, until recently, the assumption of binding reserve requirements has been an important feature of the academic literature in banking.

Building on the initial short paper by Bernanke and Blinder (1988), their follow-up paper (Bernanke & Blinder, 1992) is an in-depth empirical investigation of the monetary transmission mechanism. In particular, the focus is on how the federal funds rate works through the financial system to influence real activity. The key section (Section 4) of the paper highlights the important empirical role of the federal funds rate in influencing the future loan supply of banks. In particular, Bernanke and Blinder (1992) showed that an increase in the federal funds rate leads to an eventual slowdown of bank lending at a time horizon similar to the impact of the federal funds rate on unemployment. In addition, Bernanke and Blinder (1992) showed that an initial increase in the federal funds rate is met with a rapid adjustment in the bank's portfolio in which the holding of securities first falls, and then is slowly rebuilt.

Bernanke and Blinder (1992) interpreted these findings as showing that the federal funds rate affects the supply of bank lending directly through the portfolio constraints of the bank. In particular, when the federal funds rate is raised, the deposit funding of the bank is squeezed, which puts pressure on the asset side of the bank's balance sheet to contract. Since loans are long-term contractual arrangements, bank lending is initially slow to adjust and all the short-term adjustment is made via the holding of securities. Over time, the holding of securities is built up, but the bank's loan portfolio adjusts slowly to its new (lower) optimum. In this way, a higher federal funds rate is seen to affect bank lending through the squeeze in the deposit funding of the bank, which eventually feeds into the decrease in bank lending.

The bank lending channel examined by Bernanke and Blinder (1992) has close affinities with the risk-taking channel of monetary policy proposed in this chapter.

The common theme is that the federal funds rate has a direct impact on credit supply. However, the differences are also apparent. In Bernanke and Blinder (1992), the mechanism that links the federal funds rate with the supply of bank lending is the binding reserve requirement of the commercial banks. This poses two challenges in the context of the recent crisis. First, reserve requirements have not been binding in the aftermath of the crisis. Second, the Bernanke and Blinder (1992) account focused on the commercial banking sector, since this is the sector for which there is a reserve requirement. However, as we have described in some detail earlier, the credit crunch in the recent financial crisis originated in the shadow banking system and the market-based financial intermediaries that serve it, rather than in the traditional commercial banking sector. Indeed, the commercial banking sector had seen increased balance sheets until the summer of 2009.

Nevertheless, the Bernanke and Blinder (1992) paper stands as a milestone in the literature on the relationship between monetary policy and the banking system. The conjectures that they proposed in their original paper were confirmed in a careful cross-sectional empirical study by Kashyap and Stein (2000), who used a large data set of banks in the United States to examine portfolio changes in response to monetary policy shifts. They investigated the conjecture in Bernanke and Blinder (1992) that banks with less liquid balance sheets (i.e., with fewer securities and more loans) were subject to greater downward impact in their banking lending as a result of monetary tightening. They find strong evidence that less liquid banks are, indeed, subject to greater loan contraction due to monetary tightening. The results are driven in particular by small banks, which form the bulk of their sample.

Relative to the literature surveyed earlier, the risk-taking channel of monetary policy proposed in this chapter has some distinctive features. First, in contrast to Bernanke and Gertler (1989) and other approaches that emphasize the demand for credit and the nonfinancial borrowers' balance sheet, the risk-taking channel emphasizes the role of the *supply* of credit by the financial intermediary sector. In this respect, the risk-taking channel has greater affinity with approaches that emphasize the supply of credit and the constraints that bind on the lenders' side. However, the distinctive feature of the risk-taking channel is the role played by the price of risk and by the market-determined risk premium. The supply of credit is determined by the threshold value of the risk premium charged by the market. In the model sketched in this chapter, the asset choice decision of the banks is determined by an underlying risk-management problem where banks are subject to a VaR constraint. Here we have left this constraint without a further microfoundation. However, such a search would lead naturally to a meeting point with the agency literature that emphasizes the constraint imposed on the borrowers by the market as a whole. Adrian and Shin (2008a) is an example of just such a setting.

7. CONCLUDING REMARKS

We conclude with some implications of our findings for the conduct of monetary policy. Our emphasis on the role of balance sheet aggregates of financial intermediaries leads to policy prescriptions that bear a superficial similarity to an older tradition in monetary economics that emphasized the money stock as a pivotal quantity in monetary policy. The older monetarist tradition emphasized the stock of money because of the supposed direct link between the money stock and real expenditures through the portfolio adjustment of individual consumers who rebalance their portfolios consisting of money and real goods. Monetary aggregates had fallen from favor in the conduct of monetary policy mainly as a backlash against the older monetarist line (see Friedman, 1988).

In this chapter, we have focused on balance sheet aggregates of financial intermediaries, but the rationale is quite different from the older monetarist literature. Our approach has been to emphasize the role of intermediary balance sheets as a determinant for the risk appetite ruling in the economy, and how monetary policy can affect the growth of intermediary balance sheets. Although our rationale for looking at balance sheets differs from the older monetarist literature, our discussion nevertheless suggests that there is a case for rehabilitating some role for balance sheet quantities in the conduct of monetary policy. By influencing the rate of growth of intermediary balance sheets, the monetary authorities can impact real decisions that depend on the price of risk ruling in the economy. Real decisions that are sensitive to financial conditions, such as residential investment, will be particularly susceptible to shifts in the price of risk.

To the extent that monetary policy decisions ripple through to the real economy via the financial system, our discussion also highlights the importance of tracking the institutional underpinnings of the financial system. The instability of money demand functions that undermine the practical use of monetary aggregates in the older monetarist-style analysis is closely related to the emergence of the market-based financial system. As a result of those structural changes, not all balance sheet quantities will be equally useful. The money stock is a measure of the liabilities of deposit-taking banks, and may have been useful before the advent of the market-based financial system. However, the money stock will be of less use in a financial system such as that in the United States. More useful may be measures of collateralized borrowing, such as the weekly series on repos of primary dealers. The model presented in the paper shows that balance sheet quantities of financial intermediaries are closely tied to risk premia and the supply of credit, which, in turn, makes them useful in analyzing the financial conditions that determine the supply of credit. Adrian, Moench, and Shin (2010) presented an empirical analysis that uses balance sheet quantities from a broad range of financial intermediaries in order to gauge financial conditions.

Finally, our results highlight the channel through which monetary policy and policies toward financial stability are linked. When the financial system as a whole holds

long-term, illiquid assets financed by short-term liabilities, any tensions resulting from a sharp pullback in leverage will show up somewhere in the system. Even if some institutions can adjust down their balance sheets flexibly, there will be some who cannot. These pinch points will be those institutions that are highly leveraged, but who hold long-term illiquid assets financed with short-term debt. When the short-term funding runs away, they will face a liquidity crisis. The traditional lender of last resort tools (such as the discount window), as well as the recent liquidity provision innovations, are tools that mitigate the severity of the tightening of balance sheet constraints. However, experience has shown time and again that the most potent tool in relieving aggregate financing constraints is a lower target rate. Past periods of financial stress such as the 1998 crisis were met by reductions in the target rate aimed at insulating the real economy from financial sector shocks. Our findings suggest that, in conducting monetary policy, the potential for financial sector distress should be explicitly taken into account in a forward-looking manner.

APPENDIX

Data sources

Figure 1: The 10-year and 3-month Treasury constant maturity yields as well as the effective federal funds rate are from the Federal Reserve Board's H.15 release.

Figures 6–11: Figures 6, 9, 10, and 11 use total assets of security broker-dealers, ABS issuers, shadow banks (the sum of ABS issuers, finance companies, funding corporations), and nationally chartered commercial banks from the Flow of Funds published by the Federal Reserve Board. In Figures 7 and 8, the money stock measure M1 and M2. Total outstanding and financial commercial paper used in Figures 7, 8, and 10, are from the Federal Reserve Board. Primary dealer repo in Figures 7, 8, and 11 is from the Federal Reserve Bank of New York.

Figures 12 and 13: These figures use total financial assets from the Federal Reserve Board's Flow of Funds.

Figures 14 and 15: These figures are based on computations by Adrian, Moench, and Shin (2009). The macro risk premium as the predicted part of a regression of real GDP growth on constant maturity Treasury yield spreads and corporate bond spreads. The risk appetite variable is obtained by regressing (negative) changes of the macro risk premium on lagged balance sheet variables of security broker-dealers, shadow banks, and commercial banks.

Tables 1 and 2: These tables report regressions of GDP and residential investment growth on the total asset growth of broker-dealers, shadow banks, and commercial banks for 1986Q1 to 2009Q2. Lags are one quarter lags; growth rates are annual. Total assets are from the Federal Reserve Board's Flow of Funds. Shadow banks include ABS issuers, funding corporations, and finance companies. GDP and residential investment

is from the Bureau of Economic Analysis (BEA). PCE inflation is the personal consumption expenditures deflator excluding food and energy as reported by BEA. The equity return is the one quarter return of Standard & Poor's S&P 500 index. The VIX is CBOE's implied volatility index (the VXO from 1986–1989, and the VIX from 1990 onwards). The term spread is the difference between the 10-year constant maturity Treasury yield and the 3-month Treasury bill rate, both are from the Federal Reserve Board. The credit spread is the difference between Moody's Baa spread and the 10-year Treasury rate, both are from the Federal Reserve Board.

Table 3: This table reports regressions of repo growth, repo + commercial paper growth, and M2 growth on their own lags, and asset price variables. The data frequency is weekly from October 3, 1990 to February 3, 2010. Changes refer to one-week changes, and lags to one-week lags. Fed Funds denotes the Federal Funds Target as reported by the Federal Reserve Board. The equity return is the one-week return of Standard & Poor's S&P 500. The VIX is CBOE's implied volatility index for the S&P 500. The term spread is the difference between the 10-year constant maturity Treasury yield and the 3-month Treasury bill rate, both from the Federal Reserve Board. The credit spread is the difference between Moody's Baa spread and the 10-year Treasury rate. Commercial paper growth is the one-week growth rate of total commercial paper outstanding reported by the Federal Reserve Board. Repo growth is the one-week growth rate of primary dealer repo, from the Federal Reserve Bank of New York. M2 growth is the one-week growth of the money measure M2 from the Federal Reserve Board.

REFERENCES

- Adrian, T., Shin, H.S., 2007. Liquidity and leverage. *Journal of Financial Intermediation*. 19 (3), 418–437, 2010. Available as *Federal Reserve Bank of New York Staff Reports*, 328.
- Adrian, T., Shin, H.S., 2008a. Financial intermediary leverage and value at risk. In: *Federal Reserve Bank of New York Staff Reports*, 338.
- Adrian, T., Shin, H.S., 2008b. Liquidity and financial contagion. *Banque de France Financial Stability Review* February.
- Adrian, T., Shin, H.S., 2008c. Financial intermediaries, financial stability, and monetary policy. In: *Federal Reserve Bank of Kansas City 2008 Jackson Hole Economic Symposium Proceedings*.
- Adrian, T., Etula, E., Shin, H.S., 2009. Risk appetite and exchange rates. In: *Federal Reserve Bank of New York Staff Report*, 361.
- Adrian, T., Moench, E., Shin, H.S., 2009. Macro risk premium and intermediary balance sheet quantities. In: *Federal Reserve Bank of New York Staff Report*, 428.
- Adrian, T., Estrella, A., Shin, H.S., 2010. Monetary cycles, financial cycles, and the business cycle. In: *Federal Reserve Bank of New York Staff Report*, 421.
- Adrian, T., Marchioni, D., Kimbrough, K., 2010. The Federal Reserve's Commercial Paper Funding Facility. In: *Federal Reserve Bank of New York Staff Reports*, 423.
- Adrian, T., Moench, E., Shin, H.S., 2010. Financial intermediation, asset prices, and macroeconomic dynamics. In: *Federal Reserve Bank of New York Staff Report*, 422.
- Ashcraft, A., 2005. Are banks really special? New evidence from the FDIC-induced failure of healthy banks. *Am. Econ. Rev.* 95, 1712–1730.
- Ashcraft, A., 2006. New evidence on the lending channel. *J. Money Credit Bank.* 38, 751–776.

- Bernanke, B., 1983. Nonmonetary aspects of the financial crisis in the propagation of the Great Depression. *Am. Econ. Rev.* 73, 257–2576.
- Bernanke, B., 2004. The logic of monetary policy. Remarks before the National Economists Club, December 2, 2004. www.federalreserve.gov/boarddocs/speeches/2004/20041202/default.htm.
- Bernanke, B., Blinder, A., 1988. Credit, money and aggregate demand. *Am. Econ. Rev.* 78, 435–439.
- Bernanke, B., Blinder, A., 1992. The federal funds rate and the channel of monetary transmission. *Am. Econ. Rev.* 82, 901–921.
- Bernanke, B., Gertler, M., 1989. Agency costs, net worth, and business fluctuations. *Am. Econ. Rev.* 79, 14–31.
- Bernanke, B., Lown, C., 1991. The credit crunch. *Brookings Pap. Econ. Act.* 2, 205–247.
- Blinder, A., 1998. Central banking in theory and practice. MIT Press, Cambridge, MA.
- Borio, C., Zhu, H., 2008. Capital regulation, risk-taking and monetary policy: A missing link in the transmission mechanism?. Bank for International Settlements Working Paper 268.
- Brunnermeier, M., Pedersen, L., 2009. Market Liquidity and Funding Liquidity. *Rev. Financ. Stud.* 22, 2201–2238.
- Brunnermeier, M., Sannikov, Y., 2009. A macroeconomic model with a financial sector. Princeton University. Unpublished working paper.
- Bryant, J., 1980. A Model of Reserves Bank Runs and Deposit Insurance. *Journal of Banking and Finance* 4, 335–344.
- Curdia, V., Woodford, M., 2009. Credit spreads and optimal monetary policy. In: Federal Reserve Bank of New York Staff Reports, 385.
- Danielsson, J., Shin, H.S., Zigrand, J.P., 2009. Risk appetite and endogenous risk. London School of Economics and Princeton University. Unpublished working paper.
- Diamond, D., Dybvig, P., 1983. Bank Runs, Deposit Insurance, and Liquidity. *J. Polit. Econ.* 91, 401–419.
- Freixas, X., 2009. Monetary policy in a systemic crisis. University Pompeu Fabra, Spain. Unpublished working paper.
- Friedman, B., 1988. Monetary policy without quantity variables. *Am. Econ. Rev.* 78, 440–445.
- Gatev, E., Schuermann, T., Strahan, P., 2009. Managing bank liquidity risk: How deposit-loan synergies vary with market conditions. *Rev. Financ. Stud.* 22 (3), 995–1020.
- Geanakoplos, J., 2009. The leverage cycle. In: Acemoglu, D., Rogoff, K., Woodford, M. (Eds.), 2009 NBER Macroeconomics Annual 24, University of Chicago Press, Chicago, IL.
- Gertler, M., Karadi, P., 2009. A model of unconventional monetary policy. New York University, Unpublished working paper.
- Gertler, M., Kiyotaki, N., 2009. Financial intermediation and credit policy in business cycle analysis. New York University and Princeton, University Working paper, forthcoming in the Handbook of Monetary Economics.
- Holmström, B., Tirole, J., 1997. Financial intermediation, loanable funds, and the real sector. *Q. J. Econ.* 112, 663–692.
- Ioannidou, V., Ongena, S., Peydro, J.L., 2009. Monetary policy, risk-taking and pricing: Evidence from a quasi-natural experiment. ECB Unpublished working paper.
- Jimenez, G., Ongena, S., Peydro, J.L., Saurina, J., 2008. Hazardous times for monetary policy: What do 23 million bank loans say about the effects of monetary policy on credit risk?. ECB and Bank of Spain Unpublished working paper.
- Kashyap, A., Stein, J., 1994. Monetary policy and bank lending. In: Mankiw, N.G. (Ed.), *Monetary policy*. University of Chicago Press, Chicago, IL.
- Kashyap, A., Stein, J., 2000. What do a million observations on banks say about the transmission of monetary policy? *Am. Econ. Rev.* 90, 407–428.
- Keister, T., McAndrews, J., 2009. Why are banks holding so many excess reserves?. Federal Reserve Bank of New York Staff Report, 380. <http://www.ny.frb.org/research/staffreports/sr380.html>.
- Khwaja, A.I., Mian, A., 2008. Tracing the impact of bank liquidity shocks. *Am. Econ. Rev.* 98, 1413–1442.
- Kiyotaki, N., Moore, J., 1997. Credit cycles. *J. Polit. Econ.* 105, 211–248.

- Kiyotaki, N., Moore, J., 2005. Liquidity and asset prices. *Int. Econ. Rev. (Philadelphia)* 46, 317–349.
- Lown, C., Morgan, D., 2006. The credit cycle and the business cycle: New findings using the loan officer opinion survey. *J. Money Credit Bank.* 38, 1575–1597.
- Paravisini, D., 2008. Local bank financial constraints and firm access to external finance. *J. Finance* 63, 2161–2193.
- Shin, H.S., 2009. Risk and liquidity. Oxford University Press, Oxford, UK 2008 Clarendon Lectures in Finance.
- Shleifer, A., Vishny, R., 1997. The limits of arbitrage. *J. Finance* 52 (1), 35–55.
- Shleifer, A., Vishny, R., 2009. Unstable banking. *J. Financ. Econ.* in press.
- Svensson, L., 2004. Challenges for monetary policy. Paper for the Bellagio Group Meeting at the National Bank of Belgium, January 2004. www.princeton.edu/~svensson.
- Van den Heuvel, S., 2002. The bank capital channel of monetary policy. Wharton School, University of Pennsylvania. Working paper.
- Woodford, M., 2003. Interest and prices: Foundations of a theory of monetary policy. Princeton University Press, Princeton, NJ.
- Woodford, M., 2005. Central bank communication and policy effectiveness. In: *Proceedings of the Federal Reserve Bank of Kansas City Economic Symposium at Jackson Hole, 2005*. <http://www.kc.frb.org/publicat/sympos/2005/sym05prg.htm>.

INDEX-VOLUME 3A

Note: Page numbers followed by f, t and n indicate figures, tables and notes, respectively.

A

- Abbot, W.J., 104n10
- Abel, A., 220
- ABS. *See* Asset-backed securities
- Absence-of-double-coincidence
 - difficulty for, 4
 - pairwise and, 6, 7
- Accelerationist hypothesis. *See* Natural rate hypothesis
- Accelerationist Phillips curve, 425
- Actions, previous/future
 - commitment by, to future action, 5
 - evolution from, 6
 - monitoring of, 6
 - RE for, 172
- Activities, underground
 - regulation of, 5, 5n2
 - taxation of, 5, 5n2
- Adam, K., 219–220, 472
- ADF test, 437
- Adrian, T., 582, 606, 619–620, 623–626, 629, 634, 639–640, 643, 647
- Agents. *See also* Rational inattention; Trade
 - as anonymous, 33
 - assets and, 80
 - in CM, 40n14
 - deviation by, 32–33
 - distribution to, 6, 19, 33, 35, 37–38
 - in DM, 39–41, 43, 48n21, 55
 - economic forecasting by, 174
 - imperfect monitoring by, 5, 158–160, 172–173
 - interaction of, 170
 - monetary behavior by, 36–38, 156, 160, 171–173
 - money and, 156, 171–173
 - production by, 34n6, 36
 - RE by, 172
 - signals for, 206, 206n24
 - specialization of, 31
- Aggregate supply relation. *See* Phillips curve
- Aggregates, monetary. *See also* Federal Reserve, U.S.; Inflation; M1/M2 series; Quantity theory of money
 - analysis of, 146–147
 - cost of, 190
 - CPI and, 190, 191, 396
 - demand for
 - by central banks, 101, 134–135
 - output and, 514–515
 - output shocks for, 200–201, 201f
 - deregulation by, 105–106
 - inflation and, 46, 146, 270–271
 - interest rates and, 98, 136, 141, 144–146, 147
 - measurement for, 147
 - for money, 98
 - price level of, for consumer, 190, 191, 396
 - rational inattention to, 147, 167–168
 - shocks for, 28, 48, 197–200, 207–208, 208f, 267–268
 - supply for, 514–515
 - baseline model of, 186–190, 212
 - demand shock for, 197–200, 207–208, 208f
 - equilibrium model for, 185–196
 - foundations for models of, 191–196
 - model for, 212
 - pricing for, 207
 - strategic complementarities for, 195–196
 - in U.S., 104
- AIG, 551
- Aiyagari, R., 583–584
- Akerlof, G.A., 161, 184, 193
- Akhtar, M.A., 373
- Aliprantis, C., 40n14
- Allen, F., 555n9
- Allocations
 - asset markets for, 88
 - of capital, 88
 - class of, 9–10

Allocations (*cont.*)

- by command, 5, 9
- constraints on
 - as IC, 10–14, 19
 - monitoring for, 11–12
- for counterfeiting, 15
- credit market for, 28
- as efficient, 32
- implementability of, 10, 10n4, 20
- as incentive-feasible, 10–11, 10n4
- insurance market for, 28
- market and, for assets, 88
- Nash bargaining and, 10, 10n4, 35
- perfect ν . imperfect counterfeits as, 8–9, 15
- record keeping for, 28
- of resources, 302
- as sequence in meetings, 10

Altig, D., 288

Altissimo, F., 480, 481

Álvarez, L.J., 238, 256, 266–267

Amador, M., 212

Amato, J.D., 202n22

Anderson, R.G., 104n10, 106–107

Ando, A., 137n31, 376, 379

Andrés, J., 221, 525n40

Andrews, D., 443

Ang, A., 208n29

Angeletos, G.M., 206n24, 212, 218, 220

Angeloni, I., 111, 480

Annual Retail Trade Survey, 47

AR. *See* Autoregressive process

Araujo, L., 7

Arbitrage, regulatory, 384, 588

Area Wide Model, 378, 380

ARIMA modeling, of inflation, 445

Arrow-Debreu model

- costly connections ν ., 7
- equilibrium and, 54
- frictions and, 36
- integration with, 22
- welfare theorem in, 5

Aruoba, B., 30n3, 40n16, 47, 49, 52, 65

Ashcraft, A., 631

Assenmacher-Wesche, K., 112, 129n25

Asset-backed securities (ABS), 615, 632–633, 632f, 635, 636–637, 636f, 640. *See also*
 Intermediaries, financial; Term asset-backed
 loan facility

Assets

- ABS for, 615, 632–633, 632f, 635, 636–637, 636f, 640
 - accumulation of, by banks, 555
 - agents and, 80
 - allocations and market for, 88
 - on balance sheet, 602
 - bank's net worth ν ., 583–584, 584n18
 - bargaining for, 88
 - of broker-dealers, 603, 605, 629
 - as capital, 376–377
 - common stocks as, 379
 - exchange of, 83, 88, 376, 380
 - expansion of, by banks, 555, 584
 - frictions and, 79
 - housing as, 376, 378–379
 - imperfect recognizability of, 4–5, 8
 - as interest-bearing, 101n6
 - intermediation and, 582, 584
 - investment in, 371, 372f, 373, 376–378, 380
 - leverage and, 603
 - into liabilities, 28
 - liquidity of, 79, 88
 - Lucas asset-pricing model for, 46, 80
 - markets for, 26, 27, 29, 30, 79–80, 83–89, 177
 - monetary injections for, 144–145
 - net worth ν ., for banks, 583–584, 584n18
 - pricing of, 34, 79, 80–83, 88, 144–145, 164, 371, 374, 375t, 379, 380
 - quality of, 583
 - rate of return on, 20, 80, 83–88
 - recovery of, 582–583
 - SPV for, 584–585
 - TARP and, 551
 - Taylor principle and, 287, 290
 - trading of, 80–83
 - transformation of, 28
 - yield on, 88, 604, 605
- Atkeson, A., 475
- Automatic transfer system (ATS), 106
- Autoregressive process (AR)
- breakpoints in, 444, 445t
 - CPI and, 442
 - for inflation persistence, 452
 - PCE and, 442
 - root of, 433, 434n17, 437t, 440, 440n24, 441t, 442
 - as univariate, 443–444, 444f
- Azariadis, C., 39

B

- Bacchetta, P., 221
- Bagehot, W., 569
- Bai, J., 443
- Bakhshi, H., 470
- Balance sheets
 - adjustments in, 637
 - assets on, 602
 - of banks, 585, 602–603
 - bank's marketing of, 603
 - for borrowing, 549–550, 549n3, 629–630, 630t, 642
 - of broker-dealers, 629t
 - channel for, 383–385, 384n11, 387, 398–399
 - of Federal Reserve, 633, 636
 - during financial crisis, 549–550, 549n3
 - GDP and, 629t
 - for households, 384–385, 387
 - for intermediaries, 550, 582, 603, 627, 634
 - liabilities on, 602
 - for monetary policy, 646
 - monetary transmission channels for, 383–385, 384n11, 387, 398–399
 - for mortgages, 385, 385n12
 - NIM for, 602–603, 604–605, 604f
 - procyclical leverage and, 603, 619–620
 - risk appetite of, 624–627, 627f
 - for shadow banks, 605
 - term spread for, 602, 603, 604f
- Ball, L., 66, 69, 195, 270, 271–272, 289, 333, 428, 430, 435
- Bank of England, 147
- Banking and Currency, U.S. House Committee on, 352n59
- Bankruptcy
 - laws for, 384n11
 - Lehman Brothers in, 635
- Banks
 - asset accumulation/expansion by, 555, 584
 - assets *v.* net worth for, 583–584, 584n18
 - balance sheet of, 585, 602–603
 - borrowing by, 647
 - business fluctuations for, 555–559, 555n9, 557n10, 558n11
 - capital for, 603, 605
 - capital requirements for, 585n21
 - central bank policy and, 175
 - as commercial/investment entities, 584–586, 585n21
 - cash for, 549, 551, 634, 634f, 636
 - growth of, 624f
 - intermediaries for, 584
 - nontradable loans for, 637
 - procyclical leverage by, 603, 619–620
 - SPV for, 584–586
 - credit supply by, 387
 - creditors *v.* owners of, 587–588
 - as delegated monitors, 584
 - deleveraging by, 383
 - deposits in
 - government insurance on, 75
 - SPV and, 584–585
 - Diamond, D.-Dybvig model for, 27, 28, 30, 71, 75, 77, 588–589, 632
 - disintermediation for, 381
 - equity for, by central banks, 549, 551, 566, 571–573, 586, 593–597, 602, 604
 - financial crisis in, 75, 584, 632
 - hedge funds by, 587–588, 636
 - as illiquid, 632
 - income for, 602–603
 - inflation reports by, 175
 - in interbank markets, 550, 551, 566, 581f, 582f, 583
 - interest rate regulation by, 385–386
 - as intermediaries, 549–551
 - lending channel for, 382, 386–387, 582f, 602
 - leverage ratio by, 586
 - liability structure of, 75, 586, 588, 602, 637
 - liquidity provision by, 584
 - LOLR for, 632–636
 - macroeconomics and, 584
 - maturity transformation by, 584
 - New Monetarism theory for, 75–79
 - NIM for, 602–603, 604–605, 604f, 638–640
 - price of risk for, 549
 - repos for, 620
 - reserve requirements for, 28, 106, 112n17, 135
 - risk by, 602–605
 - risk for, 549
 - as shadow entities, 387, 588, 603, 605, 615–619, 616f, 617f, 618f, 619f, 636–637
 - volatility of net worth for, 586, 588
- Bargaining. *See also* Nash bargaining
 - for assets, 88

- Bargaining. *See also* Nash bargaining (*cont.*)
 by buyer, 35, 88
 monetary theory and, 45
 Mortensen-Pissarides model for, 43
 price taking v ., 43n16
 search model for, 47, 80
 for wages, 497
- Barnes, M., 463
- Barnichon, R., 493
- Barro, R.J., 110, 185, 506
- Barsky, R., 245, 445
- Barter, money v ., 33–34. *See also* Bargaining
- Barth, M.J., III, 289
- Basu, S., 289, 291, 307, 493
- Batini, N., 129n24
- Baumol, W., model of, 40n13
- Bayesian approach. *See also* Econometrics
 channel comparison by, 65
 DSGE model for, 288, 363n66, 416
 to econometrics, 288, 373
 estimation strategy through, 345–351
 for limited information, 315–320, 315nn26–28, 316n29, 316t, 318, 318t, 319, 319t, 320t
 for New Keynesian theory, 362, 362n64, 416, 455n43
 for transparency, 288
 VARs for, 288, 391
- Bear Stearns, 633
- Behavior, monetary
 by agents, 36–38, 156, 160, 171–173
 of central bank, 442
 equation of exchange for, 99, 108
 exchange process for, 98–99, 101, 104–108, 107n15, 110
 GDP and, 137, 491
 as historical, 104–108, 104f, 144
 inertia in, 156, 172, 175, 288, 290–291
 of inflation, 98, 137
 information theory for, 157–160
 Michigan Survey of Consumer Attitudes and Behavior, 207–209, 208n29
 optimizing-agent models of, 37, 156, 171–173
 as policy, 371, 373, 374–375
 policy/theoryfor, 37
 portfolios and, 102, 106–107, 145
 for price level, 137–138
 of private sector, 98, 156, 171–173
 QTM theory and, 121–131
 rational inattention for, 156, 171–173
- Bekaert, G., 208n29
- Bellman equation, 213
- Benabou, R., 30n3
- Benati, L., 113n20, 447, 447n34, 448t, 473
- Berentsen, A., 40n13, 50n22, 53n24
- Bergen, M., 217n43
- Berkelmans, L., 209–210
- Bernanke, B., 27, 383, 401, 403, 411, 415, 548, 549, 549n1, 550, 604, 631–632, 642–643, 644–646
- Beveridge curve, 489–490
- Billion Prices Project, 237
- Bils, M., 255, 263, 269–270, 274, 457n44, 478–481
- Blanchard, O.J., 144, 186n2, 452, 489, 499, 499n17, 516
- Blinder, A.S., 238, 604, 644–645
- BLS. *See* Labor Statistics, U.S. Bureau of
- Bodart, V., 536
- Boivin, J., 255, 269, 391–392, 480
- Bonds
 as corporate, 625
 pricing of, 44
 rating of, 625
 by Treasury, 88, 145, 398, 398n19, 405, 625
 yield on, 398
- Bonomo, M., 217
- Booms
 borrowing during, 550
 as driven by demand, 309
 inflation and, 310
 over future expectations, 309–311
 recessions v ., 305, 305n15, 309–311, 323–324, 323f, 602
- Bordo, M., 144
- Borio, C., 604
- Borrowing
 balance sheets for, 549–550, 549n3, 629–630, 630t, 642
 bank intermediaries for, 584
 booms and, 550
 collateral for, 620
 credit cost for, 549, 550
 evaluation/monitoring of, 584
 lending and, 549, 551
 as short-term, 637

- working capital channel and, 287, 289
- Branch, W.A., 209, 219–220
- Bray, M., 472
- Brayton, F., 379n6
- Broker-dealers
 - assets of, 603, 605, 629
 - balance sheets of, 629t
 - credit for, 615
 - growth of, 624f
 - as intermediaries, 615–619, 616f, 617f, 618f, 619f
 - in repos market, 636
 - in securities, 615–619, 616f, 617f, 618f, 619f
 - in shadow banking, 615–619, 616f, 617f, 618f, 619f
- Brown, A.J., 122
- Brown, C.V., 122, 137n31
- Bruckner, M., 303n12
- Brumberg, R.E., 376, 379
- Brunnermeier, M., 582, 584, 643–644
- Bryant, 632
- Bryant, J., 20
- Buiter, W., 428, 452
- Bullard, J.B., 140, 141n39
- Burdett, K., 67–69, 67n31
- Burns, Arthur, 352n59
- Burstein, A., 242, 249, 268, 269, 274, 452n38, 469–470, 470n59
- Buyer
 - bargaining power of, 88
 - default by, 62
 - in DM, 56
 - seller and, for Walrasian price taking, 75–76
- C**
- Caballero, R.J., 218
- Calibration, aspects of, 30n3
- Calomiris, C., 586–588
- Calvo, G., 67, 67n31, 70, 71, 103, 141n40, 203–205, 217–218, 234, 250, 266, 289–290, 333, 333n41, 334–335, 338, 388, 427, 451, 456, 460, 461, 470, 470n59, 480, 488n1, 507, 512, 521n38. *See also* Phillips curve
- Calza, A., 384–385
- Camera, G., 40n13
- Campbell, J.R., 267
- Canetti, E., 238
- Canova, F., 391–392, 398
- Canzoneri, M.B., 142
- Capital
 - accumulation of, 340–343, 343n51
 - allocations of, 88
 - assets as, 376–377
 - for banks, 585n21, 603, 605
 - channel for, 287, 289, 289n2, 298, 298n5, 302
 - borrowing through, 287, 289
 - VARs and, 289
 - clay aspect of, 18
 - distribution of, 6, 19, 33
 - for DM payments, 51
 - DSGE model for, 341–342
 - Flow of Funds for, 207–210, 207n28, 208f, 209n31, 289, 622, 640
 - insurance for, 18
 - markets in, 551, 615, 634, 634f, 636
 - money and, 30n3
 - motion for, 18
 - regulations for, 584, 585n21
 - user cost of, 376
- Caplin, A., 67n30, 469
- Carlson, J., 219–220
- Carlstrom, C., 384
- Carroll, C.D., 206, 209, 215, 220
- Carvalho, C., 217, 219, 246, 273
- Case, K.E., 378
- Cassese, A., 104n11, 114
- Cavalcanti, R., 8
- Cavallo, A., 237, 266
- CBO. *See* Congressional Budget Office
- Cecchetti, S.G., 242
- Central banks
 - aggregate demand by, 134–135
 - Area Wide Model for, 378, 380
 - behavior of, 442
 - channel system for, 135, 158–160
 - credibility of, 176
 - credit by, 29, 549, 551
 - disinflation by, 428
 - DSGE models for, 373, 375t, 378
 - equity from, 549, 551, 566, 571–573, 586, 593–597
 - expectations channel by, 388, 397–398, 604
 - federal funds rate by, 370–371, 372f, 374, 374n3, 439n22, 443, 455, 463n51, 639, 639f
 - Fedwire for, 28, 74

- Central banks (*cont.*)
 forecasting by, 175–176
 FRB/US model for, 375t, 378, 379n6, 380
 in G7 countries, 209n31
 inflation and, 310–311, 310nn19–20, 430–431, 431n13, 437, 446, 450
 information by, 212
 interaction with, 28–29
 interbank loans by, 135, 566, 583
 interest rate and, 65, 101, 134–137, 138–141, 139n35, 371–372, 372f, 374, 374n3, 375t, 378–380, 385–386
 intermediaries and, 27, 549–551
 intervention by, 65
 liabilities of, 136
 as LOLR, 632–636
 as monetary authority, 28–29
 monetary control by, 29, 58, 134, 141, 142, 144–146
 monetary demand function for, 134–136, 139, 141, 144
 money demand for, 646
 New Area Wide Model for, 378, 380
 nominal spending and, 101, 121–123, 134
 nominal variable for, 99, 100, 100n4, 131
 obsolescence of money from, 134
 policy statement by, 175
 price stickiness of goods/services and, 601
 QTM theory by, 101
 research by, 431n13
 Taylor principle and, 116, 116n21, 119, 140, 142–143, 143n45, 303, 305, 309
 technological improvement by, 134–135
 ToTEM for, 378
 transparency by, 211–212
 volatility and, 310
- Centralized market (CM), 39
 agents in, 40n14
 bond pricing in, 44
 for economy, 39
 as frictionless, 39
 timing for, 40, 40n13, 57, 71–72, 75
- ceteris paribus*, 147. *See also* Quantity theory of money
- Champ, B., 75
- Channels, monetary transmission
 for balance sheet, 383–385, 384n11, 387, 398–399
 for bank capital, 287, 289, 382, 386–387, 582f, 602
 Bayesian approach to, 65
 for borrowing, 287, 289
 capacity for, 158–160
 for central banks, 135, 158–160
 changes in, 385, 385n13
 Coding theorem of information for, 159–160
 for consumption, 374, 375t, 376, 379, 380n7
 for credit, 381
 DSGE model for, 286, 289, 289n2, 298, 298n5, 302
 for exchange rate, 375t, 376, 380, 387t
 for expectations, 388, 397–398, 604
 Friedman channel as, 65
 globalization and, 385n13
 interest rate and, 289, 374, 375t, 376, 379–380
 intermediaries for, 584, 601
 for international trade, 374–375, 376, 380
 for investment, 372f, 374–375, 375t, 376, 377–378, 379–380
 for lending, 382, 386–387, 582f, 602
 for monetary transmission, 158–160, 373–385, 413, 415–416
 New Keynesian channel as, 65
 as pricing, for housing, 386
 for risk-taking, 603–605, 638–646
 Shannon measure and, 135–136, 158–159, 173
 survey of, 385
 Taylor principle and, 306–309, 308f, 388
 for trade, 376
 VARs and, 289
- Chapman, J., 74
- Chari, V.V., 588–589
- Chen, A., 217n43
- Chéron, 489, 503n24
- Chicago Board Options Exchange Volatility Index (VIX), 637, 640
- Chirinko, R.S., 378
- Chiu, J., 40n13
- Cho-Kreps intuitive criterion, 15
- Chowdhury, I., 289, 302n9
- Christiano, L.J., 120, 221, 271–272, 287, 288, 289, 289nn1–2, 305n15, 309, 311, 346, 351, 372, 373n1, 452, 490, 583
- Christoffel, K., 489
- Clarida, R., 29–30, 61, 289, 305n15, 405, 488n1
- Clark, R., 269

- CM. *See* Centralized market
- Cobb-Douglas production function, 552
- Cochrane, J.H., 140, 142
- Coding theorem of information theory, 157–160
- Cogley, T., 437n20, 443, 446, 451, 460n48, 461–463, 462n50, 463n51, 464, 464nn53–54, 468
- Coibion, O., 201–202, 202n21, 209
- Collateral
 for borrowing, 620
 for housing market, 384n11
- Commercial paper funding facility (CPFF), 635, 635f
- Commitment
 discounting and, 6
 to future, 6
 as limited by frictions, 27
- Commodities
 markets for, 54
 pairwise meetings and, 7, 22
 trade in, 7
- Competition
 in markets, 39
 for prices, 38
 for trade, 5
- Computer
 ASCII for, 159
 compression algorithms for, 159
- Confidence tunnels, 321, 321n33, 323, 349
- Congressional Budget Office (CBO), 460
- Construction, residential, 386
- Consumer. *See also* Consumer Price Index
 aggregate price level for, 190, 191, 396
 bargaining by, 35
 costs for, 190
 in DM market, 51
 for durable goods, 385n12
 expectations by, 385n12
 full/imperfect information for, 30, 60f, 186–193, 189nn4–5, 190nn6–8, 196–207, 209–211, 213
 housing prices for, 387
 interest rate and, 208n29, 375t, 376
 Michigan Survey of Consumer Attitudes and Behavior, 207–209, 208n29
 Nash bargaining and, 51
 price *v.* quantity plans by, 193, 193f
 resources of, 379
- Consumer Price Index (CPI)
 aggregate prices and, 190, 191, 396
 AR and, 442
 comeback prices in, 252–253, 253t
 CPI-RDB for, 234–235, 245
 data for, 234, 238, 238n4, 436t
 CPI-X as, 434, 435f, 436t, 439–440, 439f, 440, 440n24, 441–442t
 Expenditure Classes of, 243
 frequency over time for, 260–261, 261f
 for G7, 32t, 111, 129, 130t, 131, 132t, 133, 147, 236t
 for inflation, 114, 114t, 123t, 124f, 125, 126t, 128t, 129, 130t, 132t, 133, 263–266, 264t, 265f
 measures for, 434
 for M1/M2, 114t, 128t, 129, 130t, 131, 132t, 133, 134t, 135t, 147
 mean duration of prices on, 242–243, 243t
 measurement of, 434, 435f, 440, 440n24, 441–442t
 memory in, 251, 254, 277–278
 novel prices in, 252, 253t
 price changes for, 236t, 250, 250t
 QTM theory and, 124f, 125, 128t, 129, 130t, 131, 133
 reference prices for, 251, 252t
 inflation for, 264–266, 264t, 265f
 posted *v.*, 264t, 265–266, 265f
 as stickier/more persistent, 273
 for shelter, 246n8
 sticky prices in, 251
 in U.S., 111–112, 114, 114t, 123t, 124f, 147, 236t, 427f
- Consumption
 channels for, 374, 375t, 376, 379, 380n7
 income shocks and, 387
 interest rates and, 208n29, 375t, 376
 pair-wise meetings for, 10, 18
- Contracting, financial, 28
- Contracts, negotiation of, 427, 427n5
- Cooley, T., 45, 46
- Cooper, R., 196
- Corporate sector
 aggregate demand/output for, 514–515
 bonds of, 625
 growth of, 624f

- Corporate sector (*cont.*)
 inflation in, 497–499, 497n14, 498n16, 501, 501n20
 labor market for, 497, 499–502, 500nn18–19, 501nn20–22, 512, 512n34
 New Keynesian wage inflation equation for, 512, 512n34, 513–514
 as nonfinancial, as capital goods producers, 331–334, 564–565
 price-setting decisions by, 497–499, 497n14, 498n16
- Correia, I., 5
- Costly state verification model (CSV), 583
- Costs
 adjustment in, 344
 aggregate price level for, 190
 Arrow-Debreu model v ., 7
 of capital, 376
 of connections in mechanism-design approach, 5, 7–8
 for consumer, 190
 of credit, 387, 549, 550
 for hiring, 497
 of inflation, 47, 52, 65
 of information as fixed, 172–173
 Keynesian theory and, 156
 monetary shocks for, 255–256
 as opportunity v . production, 32
 of price change, 167–168, 213, 217
 for production of money, 12n5
 productivity and, 189
 technology for saving of, 309
 for transport, 7
 of wages, 190, 256
- Counterfeiting
 allocations for, 15
 Cho-Kreps intuitive criterion for, 15
 equilibrium and, 15
 imperfect recognizability for, 14–16
 of money, 6–7, 8–9, 14–16, 16n6
 perfect recognizability for, 8–9, 15
 pooling with, 15, 16n6
 production for, 16, 16n6
 as threat, 16
 as unprofitable, 15
- Cowles Commission, 373. *See also* Econometrics
- CPFF. *See* Commercial paper funding facility
- CPI. *See* Consumer Price Index
- CPI Research Database (CPI-RDB), 234–235, 245
- CPI-RDB. *See* CPI Research Database
- Craig, B., 30n3, 45
- Credit
 by banks, 387
 for borrowing, 549, 550
 for broker-dealers, 615
 by central banks, 29, 549, 551
 changes for, 411–412
 channels for, 381
 cost of, 387, 549, 550
 democratization of, 387
 credit scores for, 387
 down-payment requirement and, 387
 refinancing costs and, 387
 deregulation of, 385–386
 for economy, 7
 equilibrium in, 34
 equity injection for, 549, 551, 566, 571–573, 586, 593–597, 602, 604
 Federal Reserve policies for, 549, 566–574, 571–573, 580, 586, 593–597, 603
 frictions for, 7
 government intervention in, 381, 574, 586
 for housing, 381
 information technology for, 386–387
 institutional changes in, 385–387
 intermediaries and, 566–574
 M1 controls for, 144–145
 margin tightening for, 582–583
 memory for, 35
 monitoring for, 7
 for mortgages, 381, 386
 obtaining of, 39
 policies for
 direct lending for, 567–569, 567n14, 568n15
 discount window lending as, 569–571, 570n16
 by Federal Reserve, 566–571
 government expenditures/budget constraint for, 574
 during recession, 631, 633
 record keeping for, 62
 risk for, 549, 586
 as substitute for money, 34
 supply of, 603
 in U.S./UK, 144, 146

Crisis, financial

- balance sheets during, 549–550, 549n3
- for banks, 75, 584, 632
- direct lending in credit markets for, 548
- financial intermediaries and, 548–550, 602, 603
- globalization and, 548
- in government, 30, 89, 173, 483, 548–552, 552n6
- interbank services during, 550, 551, 581f, 582f, 583
- intervention by Federal Reserve/central banks/Treasury, 549–551
- investment banks for, 584
- magnification of, 583
- monetary policy after, 374, 647
- monetary transmission during, 416
- in mortgages, 30, 89
- possibility of, 588
- securities demand before, 177
- securitized assets for, 584, 584n20
- shadow banking and, 387, 588, 603, 605
- simulations/policy experiments for, 574
 - calibration for, 575–576, 575t
 - credit policy response to, 579–581, 579f
 - no policy response for, 576–579, 577n17, 578f, 581f

Crucini, M.J., 221

CSV. *See* Costly state verification model

CTW model, of unemployment, 302–303, 303n11, 312–315, 313n25, 320–326, 321f, 321nn33–35

Cumby, R.E., 142

Curdia, V., 383, 384, 583, 602

Currency. *See also* Money

- anonymous transactions in, 134–135
- change in, for shock, 269
- import/export prices in, 241, 268
- M1 currency/demand deposits/OCDs and, 106
- as non-interest bearing, 101n6
- private sector demand for, 134–135, 144
- U.S. House Committee on Banking and Currency and, 352n59

Curtis, E., 35n8

Cyclicality

- for durability of goods, 245–246, 274
- goods for, 233, 260
- price change for, 274
- monetary shock and, 245–246

Cynamon, B.Z., 107

D

Darby, J., 122

Data

- by BLS, 234–235, 240, 244, 247
- for CPI, 234, 238, 238n4, 436t
- CPI-RDB for, 234–235, 245
- disagreement in, 207–210, 207n28, 208f, 209n31
- DSGE for, 286
- equilibrium models for, 172
- as expectation, 210
- Flow of Funds, U.S., for, 207–210, 207n28, 208f, 209n31, 289, 622, 640
- for GDP, 323–324, 323f, 323n36, 324n37
- HAVER database for, 491
- information modeling and, 210
- for Japan, 123t, 125, 126t, 127, 129, 143, 148
- from market scanners, 232, 234, 237, 238t, 241–242, 257, 262, 267, 269, 276
- for output gap, 323–324, 323f
- PPI for, 234–235, 235n2, 237
- for QTM, 111–115, 121
- short *v.* long term for, 172, 173
- sources for, 647–648
- SVAR for, 172
- as unavailable, 185
- variables in, 156

De Grauwe, P., 111

Debt

- clearing/settlement of, 29
- markets for, 551
- of mortgage, 384n11
- repayment of, 72
- as risk-free, 606

Decentralized market (DM)

- agents in, 39, 43, 48n21, 55
- buyers/sellers in, 56
- capital in, 51
- consumer in, 51
- inflation and, 45
- sticky prices in, 62
- timing for, 39–40, 40n13, 57, 71–72, 75
- for trade, 39

Decision making, 38–39

Defection

- IC and, 12
- in pairwise meetings, 11, 18
- payoff for, 12

- Defection (*cont.*)
 in trade, 10–11, 10n4, 18
- Deleveraging, 383
- DelNegro, M., 172, 552n6, 583
- Demand deposits, 105–106, 105n13
- Deposits, bank
 government insurance on, 75
 as interest-bearing, 105n13, 106
 for intermediaries, 550
 M1 for demand in, 105–107, 106n14
 M2 money market accounts as, 104n2, 105,
 107–108, 129n24
 OCDs for, 106–107
 reserves on transaction for, 28, 106, 112n17, 135
 SPV and, 584–585
 UK M1 for, 105n14
- Deregulation, 105–106, 385–386
- Deviatov, A., 16
- Dhyne, E., 249, 254–255, 256, 260–261, 269
- Diamond, D., 27, 28, 30, 71, 75, 77,
 588–589, 632
- Diamond, P., 27, 30n3, 39, 56
- Diba, B., 142
- Disagreement, impulse response of, 208, 208f
- Discounting
 commitment and, 6
 for credit, 569–571, 570n16
 by Federal Reserve, 566, 569–571
 on goods, 232–233
 implementability and, 6
 for trade, 5–6
- Disinflation, 129, 209, 210f, 371, 428, 430–431,
 430f, 447
- Disintermediation, 381. *See also* Banks
- Distribution, of money, 6, 19, 33, 35, 37,
 38, 39, 250, 257–258, 266n15, 316,
 316n29, 317, 320n32
- Divisia M2, 108
- Dixit, 291
- Dixit-Stiglitz framework of
 production, 291, 335
- Dixon, H., 215n35
- DM. *See* Decentralized market
- Domberger, S., 270
- Dominant root, 433, 439–442, 440nn24–26,
 441t, 442t, 460, 461t
- Dong, M., 43n16, 53n24
- Döpke, J., 201n20
- Dorich, J., 108
- Dornbusch, R., 429
- Dotsey, M., 452n38, 469, 470
- Dressler, S., 38
- Druant, M., 256
- DSGE. *See* Dynamic, stochastic general
 equilibrium
- Duffie, D., 80, 83, 220
- Dupor, B., 215n35, 217–218
- Durability, of goods, 232–233, 242–246, 243t,
 244f, 244t, 274, 385, 385n12
- Dutkowsky, D.H., 107
- Dutu, R., 43n16
- Dybvig, P., 27, 28, 30, 71, 75, 77, 632
- Dynamic, stochastic general equilibrium (DSGE)
 activity *v.* policy for, 287, 288
 Bayesian approach to, 288, 363, 363n66,
 416
 for capital, 341–342
 for central banks, 373, 378
 for data, 286
 for economic forecasting, 286
 estimation of, 288, 315, 315n27, 351–361, 356t,
 360f, 361f, 362f, 416, 417t
 Euler equation and, 299, 337, 338, 362,
 379, 469
 evolution of, 490
 expectations channel and, 388, 397–398, 604
 for monetary policy analysis
 capital accumulation, 341–342
 New Keynesian theory for, 286–287, 289,
 298, 298n5, 302, 399–405, 400f, 402t,
 403t, 602, 604
 parameters for, 363, 363f, 406t, 407–11, 412,
 413f, 414f, 416, 417t
 price/wage stickiness/frictions for, 273, 289,
 331, 339, 353, 355–356, 356t, 357t, 359,
 523–526, 524–525f, 525n40
 working capital channel for, 286, 287, 289,
 289n2, 298, 298n5, 302, 306–309, 308f
 as monetary policy/theory, 286–288, 289,
 289n2, 298, 298n5, 302–303, 309,
 311–315, 313n25, 331, 341–342, 363,
 363f, 372, 388, 397–398, 399–405, 400f,
 402t, 403t, 406t, 407–411, 407f, 408f,
 409f, 410f, 412, 413, 413f, 414f, 416,
 417t, 463–469, 465t, 466f, 467f
 for monetary transmission, 373, 375t

- New Keynesian theory for, 286, 399–405, 400f, 402t, 403t
- output gap in, 287–288
- q channel for, 376
- shocks for, 413
- unemployment and, 287, 303, 311–315, 313n25
- Dynan, K., 387
- E**
- Eberly, J., 220
- Econometrics, 288, 373
- Economic Research, National Bureau of (NBER), 323f, 324
- Economy
 - as cashless, 7, 136–138, 140n36, 140n38
 - CM for, 39
 - credit for, 7
 - fluctuations in, 185, 207, 211, 212, 222
 - for goods as traded, 385n13
 - inertia in, 156, 172, 175, 288, 290–291
 - intermediation in, 20–21, 21n8, 21n10, 27–28, 549, 551
 - market exchange for, 31
 - MOE for, 136, 136n30, 140n36, 142
 - monetary neutrality for, 6, 44–45, 59, 67, 70, 70n34, 100–103, 100n2, 147
 - monetary theory and, 4, 4n1
 - money in, 5, 31
 - money stock and, 99–101, 104n10, 118–119, 137, 139, 139n34, 140, 146, 603, 616, 617f, 646–647
 - physical environment for, 552–554, 553nn7–8
 - rational inattention for, 156, 171–173
 - retraction in, 548
 - rigidity in, 195, 195nn12–13
 - search theory and, 39
 - shocks for, 287–288
 - as underground, 5
 - volatility in, 310, 588, 637
- Eden, B., 266–267
- Eggertsson, G., 583
- Ehrmann, M., 480
- Eichenbaum, M., 120, 250–251, 255–256, 262, 264, 268, 269, 271–272, 273, 288, 289, 289n2, 311, 372, 373n1, 452, 490
- Eisfeldt, A., 583
- ELIs. *See* Entry Level Items
- Ellis, C., 481
- Elmendorf, D.W., 387
- Employment
 - Calvo frictions and, 338, 507
 - as cyclical, 488, 491–494, 492t
 - hiring costs for, 497
 - monopoly unions and wages for, 338–340
- Engel, E.M.R.A., 218
- Ennis, H., 40n13, 43n16
- Entry Level Items (ELIs), 240, 242–247
- Environment
 - business fluctuations and, 552–554, 553nn7–8
 - of deregulation, 105–106
 - for economy, 552–554, 553nn7–8
 - money in, 40n14
 - for New Monetarism theory, 39–43
 - triggers for, 40n14
- Equilibrium. *See also* Dynamic, stochastic general equilibrium
 - Arrow-Debreu commodity markets for, 54
 - business fluctuation and, 565–566
 - counterfeiting and, 15
 - in credit system, 34
 - dynamics of
 - demand shocks for, 537
 - derivation of loss function, 542–543
 - efficient steady state for, 529
 - labor market frictions, 345, 363, 363n66, 488–491, 499–502, 520–521f, 520–523, 522–523f, 528
 - linearization of participation condition, 539–540
 - log-linearized equilibrium conditions for, 540–542
 - monetary policy design for, 528–535, 533f, 533n45, 534f
 - monetary policy/technology shocks for, 305n15, 319, 320t, 331–332, 344, 346, 346n54, 355, 359, 359n63, 517–520, 518–519f
 - nominal rigidities for, 528
 - optimal monetary policy for, 529–530
 - price stickiness for, 273, 289, 331, 339, 353, 355–356, 356t, 357t, 359, 523–526, 524–525f, 525n40
 - proof of Lemma for, 537–539
 - real wage rigidities/wage indexation, 535
 - social planner's problem for, 528–529, 588

- steady state/calibration, 515–517, 516n37
- wage flexibility for, 530, 535–536
- wage stickiness for, 277, 506–512, 508nn31–32, 526–527f, 526–528, 530–535, 531n43
- wealth effect for, 536
- financial accelerator for, 550
- full/imperfect information, 30, 60f, 188–193, 189nn4–5, 190nn6–8, 196–207, 209–211, 213
- inefficiency of, 63
- inflation and, 38, 64
- intermediaries and, 565–566
- models for, 44, 138–139, 172, 344, 606–612, 607f, 610f, 611f, 614–615, 618f, 619–621, 633, 635–636
- for monetary aggregates, 185
- monetary/fiscal authorities and, 344
- in money, 15, 20, 34, 35, 52
- as Nash, 193, 195
- New Monetarism theory for, 44
- nominal rigidities and, 287, 290, 296, 299, 302–309, 310, 316t, 344, 350, 356, 357t, 491
- Ramsey-equilibrium and, 302n10
- rational inattention and, 170–173
- RE for, 140–141, 141n39, 172
- real rate for, 428, 428n8
- resources and, 294–296
- search theory for, 27n2, 31, 39
- as stationary, 48
- sticky prices and, 63
- Taylor principle and, 290, 296–299
- trade in, 72
- Equity
 - from central banks, Federal Reserve for, 549, 551, 566, 571–573, 586, 593–597, 602, 604
 - outside *v.* inside, 586
 - hedging value of, 586, 587–588
- Erceg, C., 144, 335, 513, 530, 531n43
- Ericsson, N.R., 106n14, 148
- Estimation strategy
 - Bayesian approach for, 345–351
 - computation for, 348–349
 - for DSGE, 417t
 - impulse response matching for, 347–348
 - Laplace approximation for, 317, 348, 350–351, 360–362
 - model results for
 - impulse responses for, 358–360
 - parameters for, 355–358, 356n61, 356t, 357t, 358t
 - for output gap, 302
 - VARs for, 342, 345–347, 351–355, 351nn57–58, 395f
- Estrella, A., 429n9, 431, 431n12, 639–640
- Etula, E., 623
- Euler equation, 299, 337, 338, 362, 451, 453–454, 454t
 - DSGE and, 299, 337, 338, 362, 379, 469
 - GMM and, 455, 456t
 - lagged inflation in, 468
- Euro Area
 - inflation for, 480
 - price indexes for, 448
- Evans, C.L., 120, 271–272, 288, 289, 372, 373n1, 452, 490
- Evans, G.W., 140–142, 141n39, 219–220
- Exchange, process of
 - assets for, 83, 88, 376, 380
 - behavior for, 98–99, 101, 104–108, 107n15, 110
 - central banks and, 145
 - channels for rate of, 375t, 376, 380, 387t
 - frictions in, 27
 - globalization for, 385n13
 - in market economy, 31, 34
 - MOE for, 136, 136n30, 140n36, 142
 - money and, 26, 31
 - QTM and, 98–101, 100n2, 108
- Expectations. *See also* Monetary policy/theory; Rational expectations
 - booms and, 309–311
 - changes in, 388
 - channel for, 388, 397–398, 604
 - by consumer, 385n12
 - data for, 210
 - DSGE for, 388, 397–398, 604
 - for inflation, 207, 208n29, 303, 305, 305n15, 372, 373, 374
 - interest rates and, 388
 - management of, 372, 373, 374, 385, 385n13, 388
 - modeling of, 426
 - monetary policy and, 374, 377, 378, 385, 385n13
 - Muth's theory of rational, 426

- New Keynesian theory for, 388, 397–398, 604
 - RE for, 140–141, 141n40, 172, 174–175, 178, 196
 - in reduced-form, 388
 - Taylor principle and, 388
 - Extensions
 - Aruoba extension for, 49
 - for fiscal policy, 49, 535–537
 - for New Keynesian theory, 535–537
- F**
- Fabiani, S., 256
 - Fagan, G., 378
 - Faia, E., 503n24, 530
 - Faig, M., 30n3, 40, 43n16, 47
 - Fair, R.C., 379
 - Farhi, E., 588–589
 - FAVAR model, 372, 390–391, 392–398, 394f, 397f, 399f, 405, 415
 - federal funds, Central banks' rate for, 370–371, 372f, 374, 374n3, 439n22, 443, 455, 463n51, 639, 639f
 - Federal Reserve, U.S.
 - as aggressive, 633
 - balance sheet of, 633, 636
 - credit policies by, 566–574, 580
 - direct lending by, 566–569
 - discount window lending by, 566, 569–571
 - equity injections by, 549, 551, 566, 571–573, 586, 593–597, 602, 604
 - government expenditures/budget constraint for, 381, 566, 574, 586
 - credit risk for, 549
 - crisis intervention by, 483, 548–552, 552n6
 - discounting by, 566, 569–571
 - EDO DSGE model for, 401
 - for equity, 549, 551, 566, 571–573, 586, 593–597, 602, 604
 - Federal Reserve Act for, 566
 - Fedwire for, 28, 74
 - Flow of Funds by, 207–210, 207n28, 208f, 209n31, 289, 622, 640
 - FOMC statements by, 374n3
 - FRB/US model for, 375t, 378, 379n6, 380
 - H.15 release of, 625
 - inflation target by, 309, 430–431, 431n11, 437, 437n20, 446, 447–448, 448t, 473
 - interest-rate instrument by, 105
 - liquidity facilities of, 635, 635f
 - loans by, 566
 - as LOLR, 634
 - M1/M2 series by, 104–105, 104f, 104n10, 107, 107f, 133n27
 - policies for, 566, 571–573
 - Regulation Q and, 105, 381
 - reserve requirement for, 28, 106, 112n17, 135
 - of St. Louis (FRED), 133n27, 147, 148
 - TALF by, 635–636
 - wealth effect and, 375t, 376, 378–380, 379nn5–6
 - Federal Reserve Act, 566
 - Fedwire, 28, 74. *See also* Federal Reserve, U.S.
 - Fernald, J., 332, 493
 - Ferrero, A., 583
 - Ferri, J., 525n40
 - Fiebig, D.G., 270
 - Financial intermediation, theory of, 27, 28
 - Fiscal theory of the price level (FTPL), 141–142, 142n42
 - Fischer, S., 427, 471
 - Fisher, I., 99
 - Fisher, J., 346n54
 - Fisher equation, 44, 102, 102n8
 - Fisher-Konieczny measure, 261
 - Fitzgerald, D., 268
 - Fleming, J.M., 376
 - Float, pairwise meeting for, 7
 - Flow of Funds, U.S., 207–210, 207n28, 208f, 209n31, 289, 622, 640. *See also* Capital
 - Forecasting, economic
 - by agents, 174
 - by central banks, 175–176
 - DSGE for, 286
 - Livingston Survey for, 207–208, 208n29
 - for macroeconomics, 377–378
 - MPS model for, 377–378
 - of other's forecasts, 185
 - by private sector, 174
 - RE effect for, 174
 - revision of, 207
 - Survey of Professional Forecasters, 207–209, 208n29, 398
 - as varying, 207
 - Fostel, A., 582, 583
 - Frain, J.C., 111
 - Francis, N., 493

- FRB/US model, 375t, 378, 379n6, 380, 388. *See also* Central banks; Federal Reserve, U.S.
- FRED. *See* Federal Reserve, U.S.
- Freeman, S., 27, 28–29, 30, 50n22, 71
- Frictions
- Arrow-Debreu model and, 36
 - assets and, 79
 - in business fluctuations, 561–563
 - Calvo model and, 333, 333n41, 334–335, 338, 388, 507
 - commitment as limited by, 27
 - for credit, 7
 - DSGE and, 288
 - for employment, 338, 507
 - in exchange process, 27
 - for households, 334–335
 - inflation and, 501–502
 - for labor market, 345, 363, 363n66, 488–491, 499–502, 520–521f, 520–523, 522–523f, 528
 - modeling of, 27
 - for monetary trade, 5–8, 27, 89
 - New Keynesian theory for, 286, 290–291, 489–490
 - New Monetarism and, 65, 89
 - as nominal, 551–552, 552n6
 - pricing and, 27, 287, 288, 290–291, 333, 333n41
 - in search model, 39, 67n31
 - sticky prices as, 66, 333
 - Taylor principle for, 28, 62, 115, 116, 116n21, 119, 140, 142–143, 143n45, 250, 266, 266n15, 287, 290, 296–299, 302, 303–309, 304nn13–14, 310, 316t, 344, 350, 356, 357t, 491, 502–503, 532–533
 - tractability and, 27
 - for trade, 5, 7–9
 - types of, 27
 - for unemployment, 345, 363, 363n66, 488–491, 499–502, 520–521f, 520–523, 522f, 523f, 528
 - wages and, 338, 507
- Friedman, B.M., 124, 134, 373, 389n14
- Friedman, M., 26, 28, 29, 45, 46, 53, 60–61, 64, 66, 74, 79, 79n36, 99, 100, 100n2, 104, 104n10, 115, 122, 129n24, 143–144, 184, 222, 376
- Friedman channel, 65
- Friedman rule, 45, 46, 61, 64, 66, 74, 79n36
- Frisch labor supply elasticity, 289, 290–291, 299–302, 299n6, 490
- Fritsche, U., 201n20
- FTPL. *See* Fiscal theory of the price level
- Fuerst, T.S., 289n2, 384
- Fuhrer, J.C., 428, 429n9, 430, 431, 431n12, 452, 452n37, 453n40, 456, 457, 457n45, 464n52, 467, 470n59
- FX futures markets, 634
- ## G
- G7 countries
- central banks in, 209n31
 - CPI inflation/monetary growth for, 32t, 111, 129, 130t, 131, 132t, 133, 147, 236t
 - PPI for, 237t
- Gabaix, X., 220
- Gagnon, E., 258–259
- Gale, D., 555n9
- Galenianos, M., 41n15, 43n16
- Gali, J., 29–30, 186n2, 218, 289, 303n11, 305n15, 326, 391, 405, 412, 451n36, 451, 452, 455, 456, 488n1, 489, 489n2, 492, 493, 499, 499n17, 512n34, 516
- Gambetti, L., 391–392, 398, 412
- Gârleanu, N., 80
- Gaspar, V., 111
- Gaussian case. *See also* Shannon measure
- linear-quadratic examples for, 161–168, 172–173, 178–180
 - noise and, 159–163
- GDP. *See* Gross Domestic Product
- Geanakoplos, J., 582, 583
- The General Theory* (Keynes), 222
- Generalized method of moments (GMM), 455, 456t
- Gerlach, S., 112, 129n25
- Gertler, M., 27, 29–30, 218, 250n11, 289, 305n15, 383, 405, 451n36, 455, 456, 488n1, 490, 512, 548, 549, 550, 551, 552n6, 554, 583–584, 602, 631–632, 642–643, 645–646
- Giannoni, M.P., 120, 255, 391–392
- Gilchrist, S., 383, 398, 398n19, 401, 405, 411n20, 548, 583
- Globalization
- exchange channels and, 385n13
 - money transmission and, 385n13
 - recession and, 548

- GMM. *See* Generalized method of moments
- Gold, standard of, inflation under, 444–445
- Goldberg, P.K., 240–241, 247
- Golosov, M., 250n11, 254, 274
- Goodfriend, M., 135, 488n1
- Goodhart, C.A.E., 134–135, 144n45a
- Goods
- central banks and, 601
 - consumer for, as durable, 385n12
 - as cyclical, 232–233, 245–246, 260, 274
 - discounts on, 232–233
 - durability of, 232–233, 242–246, 243t, 244f, 244t, 274, 385, 385n12
 - economy for, as traded, 385n13
 - inflation for, 232–233, 264–266, 264t, 265t
 - labor market for, 233–234
 - in outlets, 262
 - price change for, 232–233
 - price stickiness of, 601
 - production of, 331–334, 564–565
 - reference prices for, 247
 - inflation for., 264–266, 264t, 265f
 - posted v ., 264–266, 264t, 265f
 - as stickier/more persistent, 273
 - sales for, 247, 249, 262–263, 272–273
 - timing of price change for, 233
 - turnover of, 232, 238, 247, 272–273
 - utility from, 32n4
- Gopinath, G., 235, 241, 268, 269, 273
- Gordon, R.J., triangle model by, 425, 425n1, 426t, 428, 460
- Gorodnichenko, Y., 209, 218
- Gorton, G.B., 177
- Gourinchas, P.O., 269
- Government
- bank deposit insurance by, 75
 - financial crisis in, 30, 89, 173, 483, 548–552, 552n6
 - inflation and, 62
 - intervention by, for housing, 381
 - intervention by, in credit markets, 381, 566, 574, 586
 - intervention by, with outside equity, 549, 551, 566, 571–573, 586, 593–597, 602, 604
 - lump-sum transfers of money by, 64
 - money as private v ., 7, 8, 14, 22–23
 - mortgages regulation by, 381
 - private money v ., 7, 8, 14, 22–23
 - taxation by, 20, 22, 100n3
- Gowland, D., 136
- Granger, C., 481
- Gray, J., 427
- Great Depression, 385n12, 548, 642
- Great Moderation, 434, 437, 439, 476
- Greenspan, Alan, 383
- Gross Domestic Product (GDP)
- balance sheets for, 629t
 - broker-dealer's assets growth and, 629
 - CBO and, 460
 - cyclical behavior and, 137, 491
 - data for, 323–324, 323n36, 324n37
 - deflator for, data for, 434, 435f, 436t
 - growth of, 122–123, 122n22
 - as HP filtered, 463
 - impulse response for, 201, 201f, 204f, 208, 208f, 210, 288–289, 345, 346, 347–349, 351, 351n57, 353, 358–362, 362n64, 412, 413f, 414f, 628, 628f, 631f
 - macro risk premium and, 625f
 - mortgage debt ratio to, 384n11
 - output gap data for, 323–324, 323f, 323n36, 324n37
 - in U.S., 122–123, 122n22, 123t, 134t, 135t, 143–144, 143n45, 147–148
- Guerron-Quintana, P., 338
- Guimaraes, B., 262
- Gumbau-Brisa, F., 463
- Gust, C., 305n15
- ## H
- Habit persistence, 337, 337n47, 353–354
- Hafer, R.W., 104n10
- Hagedorn, M., 516
- Hahn, F., 4, 89
- Haircut. *See* Repurchase agreements
- Hall, R., 505, 506
- Haller, S., 268
- Haltiwanger, J.C., 218
- Handbook of Monetary Economics*, 27n2
- Hansen, G.D., 45, 46, 290, 300, 316n29
- Harris, E., 373
- HAVER database, 491
- Hayashi, F., 377
- He, Z., 584
- Head, A., 66, 67n30
- Heckman's sample selection correction, 267

- Hedge funds, 587–588, 636. *See also* Banks
 Hellerstein, R., 240–241, 247
 Hellwig, C., 185n1, 202, 209–210, 242, 268, 274
 Henderson, D.W., 513
 Hendry, D.F., 106n14
 Hernando, I., 256, 267
 Hicks, J.R., 27
 Hirshleifer, J., 191n9
 Hobijn, B., 269
 Hodrick/Prescott filter (HP), 287–288, 303,
 321–323, 321n34, 322f, 322n35, 324, 325,
 326–330, 327f, 328f, 328n39, 329f, 330f,
 443–444, 460, 463n51
 Hoffmann, M., 289, 302n9
 Holmström, B., 642–643
 Hong, H., 288
 Honkapohja, S., 140–141, 141n39, 142
 Hosios, A., 45
 Hosios condition, 529
 House, C.L., 245
 Households
 arbitrage condition of, 588
 balance sheet channels and, 384–385, 387
 budgets for, 338n48
 business fluctuations for, 554–555
 capital accumulation by, 340–343, 343n51
 financial intermediaries and, 554–555, 584
 frictions for, 334–335
 Frisch labor supply elasticity and, 289, 290–291,
 299–302, 299n6, 490
 growth of, 38, 624f
 information for, 209
 intermediaries for, 554–555
 labor market and, 290, 335–338, 337nn45–46,
 495, 495n10
 liquidity effects for, 385n12
 New Keynesian model for, 290, 299, 334–335
 optimization problem for, 343
 productivity of, 496, 496n12
 risk-sharing within, 495
 search model for, 495
 trading by, 39
 utility function of, as search model, 495, 495n9
 Housing
 as asset, 376, 378–379
 collateral for, 384n11
 credit supply for, 381
 demand/construction for, 378
 durables and, 385n12, 647
 government intervention for, 381
 investment in
 balance sheets for, 384–385, 629–630, 630t
 impulse response for, 631f
 monetary transmission for, 416
 mortgages for, 384–385, 384n11, 385n12, 616
 prices for, 378, 385
 pricing channels for, 386
 Howitt, P., 31
 HP. *See* Hodrick/Prescott filter
 Hsieh, C.T., 269
 Hu, T.W., 10n4, 43n16
 Huang, L., 220
 Huggett, M., 38
 Hume, D., 6, 99, 115, 330
 Hybrid models of inflation, 455, 456–459,
 458f, 459t
 Hyperinflation, period of, 112n18, 136, 173.
 See also Crisis, financial
- I**
 Iacoviello, M., 382, 385
 IC. *See* Incentive constraints
 IFS. *See* *International Financial Statistics*
 Ikeda, D., 266n15
 Ilut, C., 309
 IMA. *See* Integrated moving-average
 Impulse response, 201, 201f, 204f, 208, 208f, 210,
 288–289, 345, 346, 347–349, 351, 351n57,
 353, 358–362, 362n64, 412, 413f, 414f,
 628, 628f, 631f
 Inattentiveness, theory of, 213–215, 213n32,
 215n35
 Incentive constraints (IC), 10, 11–14
 Inertia
 definition of, 424
 in economy, 156, 172, 175, 288, 290–291
 in inflation, 172, 175, 288, 290
 in Keynesian theory, 156
 in microeconomics, 172, 175
 in models, 172, 175
 in monetary behavior, 156, 172, 175, 290–291
 in prices, 156
 variables for, 156, 175
 velocity v ., 424
 Inflation
 aggregates and, 46, 145–146, 270–271

- as anticipated, 46
- ARIMA modeling of, 445
- attitudes towards, 45–46, 60
- behavior of, and monetary policy, 445
- booms and, 310
- central banks and, 310–311, 310nn19–20, 430–431
- as core, 439
- in corporate sector, 497–499, 497n14, 498n16, 501, 501n20
- costs of, 47, 52, 65
- CPI for, 114, 114t, 123t, 124f, 125, 126t, 128t, 129, 130t, 132t, 133, 263–266, 264t, 265f
- as cyclical, 491–494, 492t
- with delayed information, 201f
- disinflation and, 430–431, 430f
- distortion by, 45–46, 65, 388
- DM activity and, 45
- effects of, 38
- equilibrium and, 38, 64
- Euler equation for, 299, 337, 338, 362, 379, 451, 453–454, 454t, 455, 456t, 468–469
- expectations for, 207, 208n29, 303, 305, 305n15, 372, 373, 374
- first autocorrelation of, 463t
- frictions and, 501–502
- in G7 countries, 32t, 111, 129, 132t, 133, 147, 236t
- gap in, 446
- under gold standard, 444–445
- for goods, 233, 264–266, 264t, 265f
- Gordon's triangle model of, 425, 425n3, 426t
- government and, 62
- growth and, 108–112, 120t, 121t, 123t
- higher moments and, 270–271
- historical analysis of, 444–446, 448–449
- impulse response for, 201, 201f, 204f, 208, 208f, 210, 288–289, 345, 346, 347–349, 351, 351n57, 353, 358–362, 362n64, 412, 413f, 414f, 628, 628f, 631f
- inertia in, 172, 175, 288, 290
- interest rate and, 101–102, 110, 287, 352n59
- for labor market frictions, 345, 363, 363n66, 488–491, 499–502, 520–521f, 520–523, 522–523f, 528
- lag in, 425, 425n3, 431, 456–459, 468, 470
- Livingston Survey for, 207–208, 208n29
- as long-term concern, 112
- M1/M2 series for, 104f, 109–111, 114t, 124, 124f, 128t, 129, 130t, 131, 133–135, 134t, 135t
- measurement of
 - autocorrelations for, 439–440, 439f
 - CPI for, 434, 435f, 440, 440n24, 441–442t
 - CPI-X for, 434, 435f, 436t, 439–440, 439f, 440, 440n24, 441–442t
 - first-order autocorrelations for, 438–439, 438f
 - GDP deflator for, 434, 435f
 - GDP for, 439–440, 439f, 440, 440n24, 441–442t
 - PCE for, 434, 435f, 440, 440n24, 441–442t
 - PCE-X for, 439–440, 439f
 - persistence for, 425–431, 432f, 432n15, 440, 442, 449
- Michigan Survey of Consumer Attitudes and Behavior, 207–209, 208n29
- monetary policy for, 287–288, 303, 371, 372, 373, 374, 374n3, 405
- money demand and, 101
- money growth and, 98–99, 108–112, 112–134, 112n18, 114t
- New Keynesian theory and, 305, 309–310, 310nn19029, 501–502
- New Monetarism and, 46, 65
- output gap and, 103, 109, 143n45, 287, 302
- payments technology during, 109n16
- peak of, 352–353
- persistence of
 - AR for, 452
 - core CPI for, 444, 444f
 - diving process and, 459–461, 460nn46–48, 461t, 482
 - hybrid model for, 456–459, 458f, 459t
 - as logged, 456–459
 - response to shocks for, 431–432, 432n14, 443
 - SDP for, 469–470
 - shocks for, 433, 433n16, 459
 - trend component for, 437n20, 446, 448, 461–463
 - as unanticipated, 48–49, 57–58
 - in U.S., 111–112, 124f, 236t, 352n59
 - velocity growth and, 110, 112n18, 133
 - volatility of, 219, 415
 - welfare effects of, 29, 45, 46
- Phillips curve and, 184, 425, 428–429, 429f, 443, 457, 468, 476, 489

Inflation (*cont.*)

- for posted price, 264–265
- for PPI, 270
- price changes and, 204, 255f, 259–260, 263–264, 263t, 270–271, 275, 278
- QTM and, 46, 98, 101, 111, 147
- rate of, 44, 62, 101–103
- rational expectations and, 425–431
- shocks for, 117f, 118f, 119f, 341, 359–360, 359n63, 428, 429f
- skewness and, 270–271
- as sluggish, 353
- stabilization of, 119, 414, 415
- Survey of Professional Forecasters, 207–209, 208n29, 398
- target for, 309, 430–431, 431n11, 437, 437n20, 446, 447–448, 448t, 473
- Taylor principle for, 116, 116n21, 119, 140, 142–143, 143n45, 303, 309
- technology shocks and, 359–360, 359n63
- unemployment and, 345, 363, 363n66, 488–491, 499–502, 520–521f, 520–523, 522f, 523f, 528

Inflation Persistence Network (IPN), 235

Information

- Bayesian approach for, as limited, 315–320, 315nn26–28, 316n29, 316t, 318, 318t, 319, 319t, 320t
- by central banks, 212
- Coding theorem of, for channels, 159–160
- for consumer, 30, 60f, 184–193, 189nn4–5, 190nn6–8, 196–207, 209–211, 213
- cost of, 172–173
- as delayed, 185, 196–213, 197n15, 198nn16–17, 200nn18–19, 201f, 208f
- flow rate of, 161, 169
- frequency of, 218–219
- for households, 209
- for inflation, as delayed, 201f
- information theory and, 157–160
- for markets, 169
- modeling of, 210
- for monetary behavior, 160–171
- monetary models for, 172–173, 185, 196–213, 197n15, 198nn16–17, 200–202, 200nn18–19, 201f, 208f, 215, 217, 220
- money's role for, 140, 142–146
- output and, 203, 203n23

- as partial, 184–185, 196–207, 204f
- as perfect, 60f
- Phillips curve and, 184
- on price from firms, 237–238
- as private, 212
- processing of, 168, 173
- for profit, 191, 191n9, 193–194, 194f
- rational inattention to, 170, 186, 215–216, 216nn37–40
- Shannon measure for, 157–158, 161
- as sticky, 168–169, 200–202, 207, 209, 219, 232
- technology for, 386–387
- for unemployment for output gap, 302–303, 311–312, 311nn21–22, 324
- updating of, 169
- value of, 161
- variables in, 207, 207n27

Information theory, 157–160

Inoue, A., 208n29

Insurance

- for allocations, 28
- for capital, 18
- Diamond, D.-Dybvig model for, 75, 77
- by government for deposits, 75

Integrated moving-average (IMA), 445–446, 445n31

Interbank, market for

- asset recovery in, 582–583
- banking in, 550, 551, 566, 581f, 582f, 583
- contraction of, 583
- crisis in, 550, 551, 581f, 582f, 583
- friction for, 589–591
- LIBOR for, 566
- loans and, 135, 566, 583

Interbank lending rate (LIBOR), 566

Interest rate

- adjustment to, 98, 99–100, 110
- aggregate demand and, 98, 136, 141, 144–146, 147
- analysis instruments for, 105, 138, 139n34, 142–143, 143n45, 147n48
- for assets, 101n6
- banking panic and, 75
- central banks and, 65, 101, 134–137, 138–141, 139n35, 371–372, 372f, 374, 374n3, 375t, 378–380, 385–386
- channels and, 289, 374, 375t, 376, 379–380
- consumption and, 208n29, 375t, 376

- contractionary monetary policy and, 384
 - as direct, 376, 380
 - elasticity of, 100n2, 110, 116, 133, 135
 - expectations and, 388
 - Federal Reserve and, 105, 110, 136, 138
 - inflation and, 101–102, 110, 287, 352n59
 - for loans, 550
 - as long-term, 372f, 374n3, 377, 602, 604
 - for M1/M2 series, 104–107, 104f, 105f, 106n14, 107f, 145–146
 - monetary policy effect on, 389–391, 389n14, 603, 604, 605
 - money stock growth rule *v.*, 140, 140n38
 - for mortgage, 386
 - NIM for, 602–603, 604–605, 604f, 638–640
 - as nominal, 101n6, 102, 109–111, 116, 127, 131, 131n26, 138, 144
 - for non-monetary function, 137–138
 - for overnight, 135–136, 141
 - policymakers and, 44, 287
 - real *v.* natural/nominal, 99, 100–101, 137, 138, 138n33, 145–146, 145n46, 147, 374, 374n3, 384
 - regulation of, by banks, 385–386
 - on securities, 102, 105, 106, 135
 - as short-term, 371, 372, 374, 375t, 377, 378–379, 380, 389–391, 389n14, 603, 604, 605, 636–646
 - borrowing and, 637
 - for intermediaries, 602, 603
 - investment spending for, 376, 377, 378
 - variability in, 446, 446n32
 - working capital channel and, 289
 - as zero, 78–79
- Intermediaries, financial. *See also* Banking; Banks
- ABS issuers as, 615, 632–633, 632f, 635, 636–637, 636f, 640
 - assets and, 582, 584
 - balance sheets of, 550, 582, 605, 627, 634
 - bank net worth and, 563–564
 - banks and, 555–559, 555n9, 557n10, 558n11
 - for borrowing, 584
 - broker-dealers as, 615–619, 616f, 617f, 618f, 619f
 - business fluctuations for, 551–566
 - central bank and, 28, 549, 551
 - changing nature of, 615–623
 - as channels, 584, 601
 - commercial/investment banks for, 584
 - credit policies and, 566–574
 - deposits for, 550
 - economic fluctuation and, 602, 603, 605
 - in economy, 20–21, 21n8, 21n10, 27–28, 549, 551
 - equilibrium and, 565–566
 - federal funds rate for, 639, 639f
 - financial crisis for, 548–550, 602, 603
 - as frictionless in business cycle, 549
 - frictionless wholesale financial market and, 559–561
 - growth of, 616–618, 621–23, 623f
 - households and, 554–555, 584
 - leverage by, 603
 - loan rates by, 550
 - as market-based, 616, 617f
 - mortgages and, 602
 - nonfinancial firms and, 564–565
 - physical environment for, 552–554, 553nn7–8
 - price of risk for, 603, 605, 606–615
 - profitability of, 602, 603, 604, 605
 - risk appetite for, 623–627, 627f
 - in shadow banking, 615–619, 616f, 617f, 618f, 619f
 - SIV as, 605
 - structure of, 28
 - symmetric frictions for, 561–563
 - treatment of, 549
 - Walras' Law for, 565
- International Financial Statistics (IFS)*, 125, 147–148
- International-trade, theory for, 7
- Internet, 158, 159
- Intervention, anticipation of, 588–589, 642–643
- Investment
- Aruoba extension for, 49
 - assets and, 371, 372f, 373, 376–378, 380
 - buyer's bargaining power *v.*, 88
 - capital asset and, 376–377
 - channels for, 372f, 374–375, 375t, 376, 377–378, 379–380
 - direct interest rate for, 376, 380
 - long/short-term response for, 373, 378
 - elasticities for, 378
 - in housing, 415, 416, 629–630, 630t, 631f
 - monetary transmission for, 413, 414, 415–416
 - price of, 355, 355n60
 - price taking and, 52

Investment (*cont.*)quantity *v.* price variables for, 378

SIV for, 605

variables in, 378

IPN. *See* Inflation Persistence Network

Ireland, P.N., 107, 110, 139

IS curve, 115–116, 116n21, 118f, 121, 137,
137n31, 139, 288–289, 362, 401, 464

Issing, O., 111

Itskhoki, O., 235, 268

J

Jaimovich, N., 269

Japan, economy of, 123t, 125, 126t, 127, 129, 143,
148

Jerez, B., 43n16, 47

Jermann, U., 582, 583

Jevons, W., 32n4

Jewitt, I., 428, 452, 456

Jiang, J., 43n16

Jimenez, G., 641

Jinnai, R., 219

John, A., 196

Jones, B.E., 107

Jones, R., 27

Jorgenson, D., 376, 376n4

Judd, K., 67–69, 67n31

Julien, B., 35n8

Justiniano, A., 324

K

Kahn, C., 586–588

Kalman filter/gains, 202

Kalman smoother, 311, 312, 312n23, 323

Kara, E., 215n35

Karadi, P., 551, 552n6, 554, 602

Kareken, J., 27

Kasa, K., 202n22

Kashyap, A., 631

Kavajecz, K.A., 104n10

Kehoe, P.J., 262, 267, 268, 272, 588–589

Kennan, J., 10n4, 43n16

Keynes, J.M., 99, 222

Keynesian theory

adjustment costs for, 156

New *v.* Old for, 26, 26n1

NRH and, 103–104

price inertia in, 156

pricing and, 57, 70, 184

rigidity of, 27

sticky prices and, 57, 70, 184

Khan, H., 201n20, 470

Khwaja, A.I., 641

Kiley, M.T., 221

Kilian, L., 208n29

Kim, J.Y., 288, 347

Kimball, M., 245, 271–272, 493

King, R.G., 116, 134n29, 452n38, 469, 488n1

Kiraz, F.B., 208n29

Kircher, P., 41n15, 43n16

Kitamura, T., 217–218

Kiyotaki, N., 22, 27, 31, 32n4, 39, 40, 186n2,
549, 550, 551, 552, 582, 583, 631–632,
642–643Klenow, P., 69n33, 70, 70n34, 207, 217, 240, 243,
244, 247–249, 255, 257, 258–260, 262,
263, 266, 267, 270, 274, 457n44, 478–481

Knotek, E.S., 217n43

Kocherlakota, N., 6, 7, 20, 27–28, 33

Koenig, E.F., 220

Konieczny, J.D., 270

Korenok, O., 221

Korinek, A., 588

Krause, M., 502

Krishnamurthy, A., 584

Krusell, P., 38

Kryvtsov, O., 70, 74n34, 240, 247–249, 257,
258–260, 262, 266, 274

Kumar, A., 67n30

Kurlat, P., 582, 583

Kwan, Y.K., 288

LLabor market. *See also* Wages, labor

adjustment of, 277

BLS for, 234–235, 240, 244, 247

for corporate sector, 497, 499–501, 499–502,
500nn18–19, 501nn20–22, 512, 512n34

cost of, 190, 256

as cyclical, 488, 491–494, 492t

determination of

as flexible, 503–506, 504nn26–27

as sticky, 277, 506–512, 508nn31–32,

526–527f, 526–528, 530–535, 531n43

as flexible, 490n5

fluctuations in, 488–489, 491

- frictions in, 345, 363, 363n66, 488–491, 499–502, 520–521f, 520–523, 522f, 523f, 528
- Frisch labor supply elasticity and, 289, 290–291, 299–302, 299n6, 490
- for goods, 233–234
- households and, 290, 335–338, 337nn45–46, 495, 495n10
- inflation for, 345, 363, 363n66, 488–491, 499–502, 520–521f, 520–523, 522–523f, 528
- output and, 189, 189n4
- prices and, 137, 233–234, 276–277
- productivity of, 190, 355, 492, 492n6
- as real, 492, 492n6
- rigidity of, 489
- Rogerson model for, 53, 53n24
- Taylor principle for, 491
- in UK, 122
- unemployment in, 45, 53, 53n24, 287, 302–303, 303n11, 312–315, 313n25, 320–326, 321f, 321nn33–35, 345, 363, 363n66, 488–491, 489n4, 499–502, 520–521f, 520–523, 522f, 523f, 528
- variability of, 256
- wages for, 489, 491
- Labor Statistics, U.S. Bureau of (BLS), 234–235, 240, 244, 247
- Lach, S., 262, 270
- Laforte, J.P., 221
- Lagos, R., 21–22, 29, 39, 50n22, 79n36, 80n38
- Lagos–Wright model, 21, 39, 47
- Laibson, D., 220
- Langot, 489, 503n24
- La'O, J., 218, 220
- Lapham, B., 67n30
- Laplace approximation, 317, 348, 350–351, 360–362
- Leahy, J., 250n11, 469
- Lebow, D., 238
- Leeper, E.M., 142
- Lehmann Brothers, 551, 566–567
- Lender of last resort (LOLR), 632–636. *See also* Central banks; Federal Reserve, U.S.
- Lending
 - borrowing and, 549, 551
 - channels for, 382, 386–387, 582f, 602
 - credit policies for, 566–569, 567n14, 568n15
 - by Federal Reserve, 566–571
 - during financial crisis, 548
 - monetary transmission for, 415
 - by securities market, 386–387
 - as securitized, 584
- Lester, B., 50n22
- Lettau, M., 380n7
- Leverage, 582, 586, 588, 603, 619–620
- Levin, A.T., 144, 447, 513
- Levy, D., 217n43, 250n12
- Lewis, K.F., 216n39, 220
- Li, N., 269
- Li, Y., 15
- Li, Z., 30n3
- LIBOR. *See* Interbank, market for
- Lie, D., 463
- Lindé, J., 288
- Linzert, T., 489
- Liu, H., 220
- Liu, L., 53n24, 66
- Livingston Survey, 207–208, 208n29
- Loans
 - by Federal Reserve, 566
 - as interbank, 135, 566, 583
 - Interest rates for, 550
 - as marginal, 602, 603, 604
 - as nontradable loans, 637
 - rates for, 550
 - as subprime, 602
 - TALF for, 635–636, 636f
- LOLR. *See* Lender of last resort
- Lopez-Salido, D., 221, 392, 502
- Lorenzoni, G., 202, 220, 588
- Lothian, J.R., 104n11, 114, 148
- Lown, C.S., 382, 631
- Lubik, T.A., 502
- Lucas, R., 27, 29, 43, 45, 57–58, 60–61, 84
- Lucas, R.E., Jr., 99, 103–104, 104n11, 105n13, 108, 109, 110, 112, 113, 114, 118, 121, 184–185, 206n25, 250n11, 254, 274, 341, 373, 426
- Lucas asset-pricing model, 46, 80
- Lucca, D.O., 343
- Ludvigson, S.C., 380n7
- Lump-sum transfers, 5, 38, 64, 76
- Lünnemann, P., 237
- Luo, Y., 176, 220

M

M1/M2 series. *See also* Aggregates, monetary;

Federal Reserve, U.S.; Money

definitions for, 104–105

Divisia M2 and, 108

by Federal Reserve, 103–105, 103n10, 104f

growth in, 104–105, 104f, 106f, 111

inflation for, 109–111, 114t, 124, 124f, 128t,
129, 130t, 131, 133

interest rates for, 104–107, 104f, 105f, 106n14,
107f, 145–146

as international, 105, 105n14, 125, 126t

label for, 104n10

M1 in

CPI for, 128t

credit controls and, 144–146

currency/demand deposits/OCDs within, 106
disinflation for, 129

growth/inflation for, 109–112, 114t, 124, 124f

interest payments in, 136

interest sensitivity of, 129

money demand for, 110, 131, 133, 134t, 646

money market deposit accounts in, 107, 129n24

money stock of, 616, 617f

OCD interest paid on, 105, 145–146

QTM for, 109

reserve requirement for, 28, 106, 112n17, 135

stability of, 107

sweeps programs for, 106–107

velocity in, 98, 105–107, 107f

velocity of, 107, 107f, 135t

M2 in

CPI inflation on, 114, 129, 130t, 131, 132t,
133, 134t, 135t, 147

growth of, 129n25

interest sensitivity of, 106

market accounts under, 104n2, 105, 107–108,
129n24

money demand for, 108

money stock of, 616, 617f

size of, 616, 617f

money demand by, 131, 133–134, 133n28, 135t,
646

movement of, 104–106

Regulation Q for, 105, 381

velocity for, 107, 108f

Mackowiak, B., 171, 176, 177, 206, 206n26, 207,
234n1, 235, 254

Macroeconomics

banking and, 584

cyclicality of price change for, 274

FAVAR and, 396, 397f

forecasting for, 377–378

frequency of price changes in, 271–272, 278

heterogeneity in prices changes for, 273–274

lack of synchronized prices for, 258–262,
275–276, 278

modeling for, 171–173

monetary policy and, 39, 52, 136, 371, 374, 380

Phillips curve for, 29, 30n3, 53–55, 57–61,
115–116, 121, 137, 139, 184, 289, 297,
298, 425, 428–429, 429f, 443, 455, 457,
468, 470, 476, 489

price age and, 276

price interpretation by, 235, 235n2

price setting for, *v.* microeconomics, 277, 277t

price/wage changes for, 267–268, 275–278

product turnover for, 232, 238, 247, 272–273

research in, 589, 602, 603, 604

shocks in, 412–414, 413f, 414f

size of change in prices for, 274–275

sticky reference prices for, 26, 27, 29, 30, 30n3,
34, 57, 203–205, 217, 217n43, 235, 240,
271–273, 289, 331, 339, 353, 355–356,
356t, 357t, 359, 523–526, 524–525f,
525n40

unemployment in, 45, 53, 53n24, 287, 302–303,
303n11, 312–315, 313n25, 320–326, 321f,
321nn33–35, 345, 363, 363n66, 488–491,
489n4, 499–502, 520–521f, 520–523,
522f, 523f, 528

variables in, 156

wealth effect for, 375t, 376, 378–379, 379nn5–6

Malin, B., 69n33, 263

Mankiw, G., 66–67, 69, 193, 200, 200n19, 206,
207, 209, 220, 221, 221n46, 249, 270, 330,
351n58, 353, 470, 472f

Mankiw, N.G., 168

Manovskii, I., 516

Marcet, A., 472

Marginal cost, 449, 451, 454, 455, 460,
460nn47–48462, 463n51, 464n53, 480, 483

Margins, tightening of, 582–583

Markets

for assets, 29, 30, 79–80, 83–89, 177

balance sheets for, 603

- broker-dealers in, for repos, 636
- as CM, 39, 40, 40nn13–14, 44, 57, 71–72, 75
- in commercial paper, 551, 634, 634f, 636
- for commodities, 54
- competition in, 39, 56, 177
- as competitive, 39, 56
- for credit, 548
- for debt, 551
- as DM, 39–40, 40n13, 43, 45, 48n21, 51, 55, 56, 57, 62, 71–72, 75
- as emerging, 232
- equilibrium in, 54, 490–491, 499–502, 520–521f, 520–523, 522–523f, 528
- fluctuation in, 174
- frictions in, for labor, 345, 363, 363n66, 488–491, 499–502, 520–521f, 520–523, 522f, 523f, 528
- FX futures markets as, 634
- government intervention in, 381, 566, 574, 586
- for housing, 384n11
- imperfections in, 374, 380–381, 415
- information and variation in, 169
- in interbank, 550, 551, 566, 581f, 582f, 583
- labor for, 497, 499–502, 500nn18–19, 501nn20–22, 512, 512n34
- M2 accounts for, 107–108
- for mortgage-backed securities, 386–387, 551
- for mortgages, 30, 89, 384–385, 384n11, 385n12, 386–387, 386t, 416, 602, 604, 605, 616
- as OTC, 80
- power in, 334–335
- rational inattention for, 177
- risk premium for, 603, 604
- scanners for, 232, 234, 237, 238t, 241–242, 257, 262, 267, 269, 276
- search model for, 56
- segmentation of, 269
- shares in, 84–88, 84n40
- trading volume in, 79
- variables in, 169, 176
- volatility of, 310, 588, 637
- Markov-Chain Markov Chain, 348, 440n24, 455n43, 476
- Markup shock, 464, 468, 468n56
- Matching model. *See also* Search model
 - as random, 31
 - search and, 489, 489n3, 496
- Mateika, F., 169–171, 176, 177, 216n37, 216n39, 217n43
- Matsuyama, K., 343
- Mauskopf, E., 373, 379n6
- Maximum likelihood, 455
- McCallum, B.T., 102, 109, 112n17, 122, 139, 140, 142–143, 142n41, 145
- McCandless, G.T., 114
- McCarthy, J., 411
- McGough, B., 219–220
- Mechanism-design approach, 4–9, 22
- Medium of exchange (MOE), 136, 136n30, 140n36, 142
- Meetings
 - allocation in sequence of, 9
 - as anonymous, 33
 - as pairwise, 7, 9–10, 11, 16, 17, 18, 22
 - production in, 10
 - as single *v.* double coincidence, 32–33, 32n4, 33n5, 37
 - for trade, 7–10
- Meiselman, D., 104n10
- Meltzer, A.H., 105n12
- Memory
 - in CPI, 251, 254, 277–278
 - record keeping and, 62
- Mendoza, E., 583–584
- Menzio, G., 38n12, 56, 66
- Merz, M., 495n9
- Metrick, A., 177
- Metropolis algorithm, 317, 348, 350, 357, 361, 363f
- Meyer-Gohde, A., 200n19, 221n47
- Mian, A., 641
- Michaels, R., 144
- Michigan Survey of Consumer Attitudes and Behavior, 207–209, 208n29
- Microeconomics
 - modeling for, 171–173
 - price-setting for, 277, 277t
 - sources of inertia and, 172, 175
- Midrigan, V., 242, 257–258, 262, 267, 268, 273, 274, 275
- Mihov, I., 255
- Minetti, R., 382, 385
- Miron, J.A., 387
- Mitra, K., 140, 141n39
- Modeling, for microeconomics, 171–173

Models, monetary

- for aggregate supply, 184–196, 212
- Aruoba-Waller-Wright as, 30n3, 40n16, 47
- for asset markets, 30, 79–80, 83–89, 177
- Ball-Mankiw as, 66, 69
- as benchmark
 - environment for, 39–43
 - monetary equilibrium in, 44
 - quantifying for, 47
 - sticky prices into, 66
 - tractability for, 56
- Berentsen-Rocheteau as, 50n22
- Burdett-Judd as, 67–69, 67n31
- Calvo model as, 67, 67n31, 70, 71, 103, 141n40, 203–205, 217–218, 234, 250, 266, 289–290, 333, 333n41, 334–335, 338, 388, 427, 451, 456, 480, 488n1, 507, 512, 521n38
- Caplin-Spulber as, 67
- cash-in-advance as, 45
- Chapman model for, 74
- Cooley-Hansen as, 45, 46
- CSV as, 583
- CTW as, 302–303, 303n11, 312–315, 313n25, 320–326, 321f, 321nn33–35
- with delayed information, 185, 196–213, 197n15, 198nn16–17, 200nn18–19, 201f, 208f
- Diamond, D.-Dybvig for, 27, 28, 30, 71, 75, 77, 588–589, 632
- as domino effect, 606
- Dong-Jiang as, 43n16
- Dooley-Hansen as, 46
- Dornbusch's overshooting as, 429, 429n9
- Dutu as, 43n16
- for equilibrium, 44, 138–139, 172, 344, 606–612, 607f, 610f, 611f, 614–615, 618f, 619–621, 633, 635–636
- Faig-Huangfu as, 43n16
- Faig-Jerez as, 43n16, 47
- family *v.* market for, 39
- FAVAR model as, 372, 390–391, 392–396, 394f, 415
- with fixed cost of information, 172–173
- frictions in model for search, 39, 67n31
- Galeanos-Kircher as, 43n16
- general equilibrium as, 606–612, 607f, 610f, 611f
- haircut in, 620, 635–636
- repos in, 620
- risk-free debt in, 606
- shadow value of bank capital for, 614–615, 618f
- VAR constraint in, 619–621, 633
- Hu-Kennan-Wallace, W., as, 43n16
- with imperfect information, 196–197, 217, 220
- inertia/noisiness in, 172, 175
- with interbank friction, 589–591
- Lagos-Wright as, 21, 39, 41, 47
- Lucas asset-pricing as, 46, 80
- Lucas-Prescott as, 43
- Mankiw-Reis model for, 470, 472f
- by Molico, 40
- money-in-the-utility-function as, 45
- Mortensen-Pissarides as, 43, 53n24
- Nosal-Rocheteau as, 15, 28–29, 74
- as optimizing-agent, for behavior, 37, 156, 171–173
- with outside equity/government intervention, 549, 551, 566, 571–573, 586, 593–597, 602, 604
- overlapping-generations as, 45, 75
- for payments, 71–74
- of private sector behavior, 98, 156, 171–173
- for rational inattention, 147, 156–173
- with reduced form, 63n29, 388, 431–449, 482
- Rocheteau-Wright as, 43n16, 56
- Rogerson as, 53–54
- Sanches-Williamson, S., as, 43n16
- Shi-Trejos-Wright as, 40
- Sidrauski-Brock for, 102
- as state-dependent/independent, 176, 205, 213–214, 214n33, 218, 233, 469–470, 591–593
- sticky-information for, 200–202, 215, 217
- Taylor principle as, 28, 62, 115, 116, 116n21, 119, 140, 142–143, 143n45, 250, 266, 266n15, 287, 290, 296–299, 302, 303–309, 304nn13–14, 316t, 344, 350, 356, 357t, 491, 502–503, 532–533
- theory *v.* empirical evidence for, 140n37
- for trend inflation, 437n20, 446, 448, 461–463
- Walrasian pricing as, 43, 44, 48n21, 51, 53, 65–66, 71–76, 497, 522
- Williamson-Wright as, 50n22
- of worker-shopper pair, 39

- Models of Monetary Economics* (Kareken; Wallace), 27
- Modigliani, F., 376, 379
- Moench, E., 623–626, 629, 634, 647
- Mojon, B., 481
- Molico, M., 37, 38, 40, 40n13
- Monacelli, T., 385
- Mondria, J., 176, 177, 210
- Monetary policy shocks, 352–354, 352f, 358–362, 359n63, 360f, 361f, 362f, 372, 373n1, 413–416, 428, 429f, 430
- Monetary policy/theory
- agents for, 37, 156, 171–173
 - aggregates for, 98
 - balance sheets for, 646
 - for bargaining, 45
 - barter *v.* money and, 33
 - as basic, 31
 - behavior in, 37, 98–99, 101, 106, 137, 138, 145, 156, 160–173, 175, 207–209, 208n29, 371, 373–375
 - booms and, 309–311
 - capital and, 30n3
 - as cashless, 7
 - changes for, 370
 - channels of, 373–385
 - as contractionary, 384
 - control of, by central bank, 28, 58, 134
 - creation of, 4
 - after crisis, 373, 647
 - demand function for, 100–101, 100n2, 109–110, 115–116, 131, 133, 133n28, 139–140, 143
 - DSGE as, 286–288, 289, 289n2, 298, 298n5, 302–303, 309, 311–315, 313n25, 331, 341–342, 363, 363f, 372, 388, 397–398, 399–405, 400f, 402t, 403t, 406t, 407–411, 407f, 408f, 409f, 410f, 412, 413, 413f, 414f, 416, 417t, 463–469, 465t, 466f, 467f, 602, 604
 - economics and, 4, 4n1, 370
 - environment for, 40n14
 - equilibrium in, 15, 20, 34, 35, 52
 - equilibrium models for, 44, 140, 172, 344, 606–612, 607f, 610f, 611f, 614–615, 618f, 619–621, 633, 635–636
 - as essential, 4–7, 27–28
 - imperfect monitoring for, 5
 - nonessentiality *v.*, 33
 - numerical methods for, 40n13
 - search-type frictions for, 39, 67n31
 - evolution of, 396–397, 397f
 - for exchange, 31, 34
 - as expansionary, 378–379, 385n12
 - expectations management by, 374, 378, 385, 385n13, 388
 - experimental models for, 27
 - fiscal *v.*, 5
 - frictions and, 4–8, 27, 89
 - future conduct of, 415–416
 - inflation behavior and, 445
 - macroeconomics and, 39, 52, 373
 - mechanism-design approach to, 4, 5, 22
 - microfoundations for, 27n2
 - money as commodity for, 15, 20, 37
 - monitoring of, 7
 - movement in, 185
 - MPS model for, 377, 379n6
 - output and, 189
 - prices/expenditure categories to, 396
 - as private *v.* government, 6–7, 8, 14, 22–23
- QTM theory
- business cycle frequency for, 116, 121–123
 - by central banks, 99
 - CPI and, 124f, 125, 128t, 129, 130t, 131, 133
 - data for, 112–115, 121–131
 - deregulation and, 105
 - Gaussian noise on, 159–163
 - GDP growth and, 122–123, 122n22
 - historical behavior/data of, 103–108, 104f, 121–131, 144
 - inflation and, 46, 98, 101, 110–111, 147
 - money neutrality and, 100
 - nominal income growth and, 114–116, 122–124, 123t
 - quantitative models for, 115–121, 117f, 118f, 119f, 120t, 121t
 - as unconditioned, 113n19
- research/progress in, 27, 370
- reserves on transaction deposits of, 28, 106, 112n17, 135
- risk-taking channel for, 603–605, 638–646
- role for, 134–136
- shocks by, 352–354, 352f, 358–362, 359n63, 360f, 361f, 362f, 372, 373n1, 413–416
- short term interest rates for, 371, 372, 374, 375t, 377, 378–379, 380, 389–391, 389n14, 636–646

Monetary policy/theory (*cont.*)

stability of

M2 for, 108

money demand and, 109–111, 646

money growth/inflation link to, 108–111

Taylor principle and, 28, 62, 115, 116, 116n21,
119, 140, 142–143, 143n45, 250, 266,
266n15, 287, 290, 296–299, 302, 303–309,
304nn13–14, 310, 316t, 344, 350, 356,
357t, 491, 502–503, 532–533

timing of, 370

tracking devices for, 38

transparency in, 175–176, 211–213

in underground economy, 5

unemployment in, 45, 53, 53n24, 287, 302–303,
303n11, 312–315, 313n25, 320–326, 321f,
321nn33–35, 345, 363, 363n66, 488–491,
489n4, 499–502, 520–521f, 520–523, 522f,
523f, 528

volatility and, 287, 309

Monetary vector autoregressions, 430

Money

counterfeiting of, 6–7

allocations for, 15

Cho-Kreps intuitive criterion for, 15

equilibrium and, 15

imperfect recognizability for, 14–16

perfect recognizability for, 8–9, 15

pooling with, 15, 16n6

production of, 16, 16n6

as threat, 16

as unprofitable, 15

distribution of, to agents, 6, 19, 33, 35, 37

as divisible, 36n9, 38

as fiat, 6, 33, 37, 99, 136, 445

growth rate of, 44–45, 79, 101

haircuts and, 637

holdings of, 6, 17

imperfect divisibility of, 7

inflation and, 98–99, 101–104, 108–112, 287,
303, 370, 405

informational role for, 145–146

as inside *v.* outside, 8, 14, 16–18, 20, 28

issuance of, 8, 10

as metallic, 99

neutrality of, 6, 29, 44–45, 67, 70, 70n34,
100–103, 100n2, 125, 147

non-neutrality of, 125, 196, 245

portability of, 33

production costs of, 12n5

quantity of, 16, 65

rate of return on, 5, 101n6

real *v.* nominal quantities of, 99

recognizability of, 14, 33

record keeping for, 28

shortage of, 75

storability of, 33

as substitute for credit, 34

superneutrality of, 102–103

transactions using, 5, 75, 135

transfer of

IC allocations for, 10–14, 19

imperfect monitoring for, 6

as lump sum, 38

uniformity of, 5, 8, 14–18

as utility, 142, 142n43

value of, 15

as wealth, 18

Money demand stability, 108–111

Money market, 107–108, 129n24

Monitoring

of actions, 6

by agents, 6, 158–160, 172–173

for borrowing, 584

of cashless economy, 7

for credit system, 7

as endogenous, 19–20

for IC, 11–12

as imperfect *v.* perfect, 6, 22, 33, 172–173

mechanism-design approach to, 4–5, 22

of money, 6

of pairwise meetings, 7–8

Monopoly, power of

Ramsey-equilibrium and, 302n10

Taylor principle and, 307n16

Monte Carlo exercise. *See* Markov-Chain Markov
Chain

Moore, G., 428, 429n9, 430, 452

Moore, J., 39, 549, 550, 551, 552, 582, 583,
631–632, 642–643

Moral hazard, 588–589, 642–643. *See also*

Intervention, anticipation of

Morgan, D.P., 382, 631

Morris, S., 211, 212, 218

Mortensen, D., 43

Mortgages, markets for

- crisis in, 30, 89
- differences in, 384, 384n11
- funding for, 386, 386t
- GSEs for, 386
- for housing, 384–385, 384n11, 385n12, 616
- interest rates for, 386
- as marketable securities, 386–387, 551
- as residential, 381, 384–386, 384n11, 385n12, 416, 616
- securitization in, 386–387
- Moscarini, G., 216
- Motion, Newton's second law of, 424
- Motto, R., 309, 583
- Moving-average representation of inflation, 433
- MPS model, 377, 379n6. *See also* FRB/US model
- Mumtaz, H., 481
- Mundell, R.A., 374
- Muth, J., 426
- N**
- NAIRU. *See* Nonaccelerating inflation rate of unemployment
- Nakamura, E., 69n33, 234, 235, 236t, 237t, 238t, 240, 242, 246, 247, 249, 255–256, 257, 260–261, 262, 266, 267, 269, 273, 457n44, 479
- Nash bargaining
 - allocation and, 10, 10n4, 35
 - Aruoba alternatives to, 43n16
 - consumer/producer for, 51
 - equilibrium for, 193, 195
 - flexible wage economy for, 505–506
 - wages and, 491, 507, 510, 516
 - Walrasian pricing *v.*, 43, 48n21, 51, 65–66
- Natural rate hypothesis (NRH), 103–104
- NBER. *See* Economic Research, National Bureau of
- Nelson, C.R., 122
- Nelson, E., 122, 129n24, 142, 142n41, 221
- Neoclassical growth theory, 39
- Net interest margin (NIM), 602–603, 604–605, 604f, 638–640. *See also* Banks
- New Area Wide Model, of ECB, 377, 378
- New Classicists, 66
- New Keynesian channel, 65
- New Keynesian theory
 - analysis by, 488, 488n1, 489n2
 - applications by, 29, 30n3
 - Bayesian impulse response for, 362, 362n64, 416
 - Calvo sticky-price model for, 289, 333, 333n41, 507
 - CTW for, 302–303, 303n11, 312–315, 313n25, 320–326, 321f, 321nn33–35
 - Dixit-Stiglitz for, 291, 335
 - DSGE model and, 286–288, 289, 289n2, 399–405, 400f, 402t, 403t, 602, 604
 - financial sector in, 602, 604
 - with price-setting frictions for, 286
 - Euler equation and, 299, 337, 338, 362, 379, 451, 453–454, 454t, 455, 456t, 468–469
 - expectations channel for, 388, 397–398, 604
 - extensions for, 535–537
 - Frisch labor supply elasticity in, 289, 290–291, 299–302, 299n6, 490
 - household production/labor/capital for, 290, 299, 334–335
 - inflation and, 305, 309–310, 310nn19–20, 501–502
 - IS curve for, 115–116, 116n21, 118f, 121, 137, 137n31, 139, 288–289, 362, 401, 464
 - labor market frictions for, 345, 363, 363n66, 489–490, 499–502, 520–521f, 520–523, 522f, 523f, 528
 - non-neutrality of money and, 125, 196, 245
 - output gap in, 287–288
 - HP filter for, 287–288, 303, 321–323, 321n34, 322f, 324, 325, 326–330, 327f, 328f, 328n39, 329f, 330f, 460
 - output *v.* input for, 289
 - Phillips curve for, 29, 30n3, 53–55, 57–61, 115–116, 121, 137, 139, 184, 289, 297, 298, 425, 428–429, 429f, 443, 455, 457, 468, 470, 476, 489
 - price-setting frictions in, 287, 290–291
 - QTM relations between inflation/money
 - growth for, 113, 119–121, 119f, 120t, 121t
 - RE equilibrium in, 140–141, 141n39, 172
 - rigidity of, 65
 - shocks and, 117f, 119f, 287–288
 - sticky price model for, 57, 61–66, 70, 184, 289, 333, 333n41, 507
 - Taylor principle and, 28, 62, 115, 116, 116n21, 119, 140, 142–143, 143n45, 250, 266, 266n15, 287, 290, 296–299, 302, 303–309, 304nn13–14, 310, 316t, 344, 350, 356, 357t, 491, 502–503, 532–533

New Keynesian theory (*cont.*)
 unemployment under, 287, 489–490, 489n4
 versatility of, 303n11
 wage inflation equation for, 513–514

New Monetarism theory
 frictions and, 65, 89
 inflation and, 46, 65
 model as benchmark for
 banking for, 75–79
 environment for, 39–43
 monetary equilibrium in, 44
 quantifying for, 47
 tractability for, 56
 sticky prices for, 66–71

New *v.* Old Monetarism and, 26, 26n1, 28

New York Times, 155

Newton, Isaac, second law of motion by, 424

Nicolini, J., 5

Nie, J., 176

Niehans, J., 99, 136

NIM. *See* Net interest margin

Nimark, K., 207, 218

Nishioka, S., 266n15

Nominal interest rate, 101n6, 102, 109–111, 116,
 127, 131, 131n26, 138, 144

Nonaccelerating inflation rate of unemployment
 (NAIRU), 425

Normal distribution, 316, 316n29, 317, 320n32

Northern Rock, runs on, 633

Nosal, E., 15, 28–29

Nowak, L., 104n11, 114, 148

NRH. *See* Natural rate hypothesis

O

OCD. *See* Other checkable deposit

“Of Money” (Hume), 330

Office for National Statistics, United Kingdom,
 147–148

Ohanian, L., 475

Okun, Arthur, 425n1

Olivei, G., 463

Ongena, S., 641

Optimizing-agent, models for, 37, 156, 171–173

O’Reilly, G., 448

Orphanides, A., 143n44, 472

Ortiz, A., 398, 398n19, 411n20

Ostroy, J., 6, 27n2

OTC. *See* Over-the-counter

Other checkable deposit (OCD), 106

Output gap
 aggregate demand shocks for, 200–201, 201f
 attitudes towards, 60
 CTW model for gap in, 320–326, 321f,
 321nn33–35
 data for, 323–324, 323f
 delayed information and, 203, 203n23
 DSGE and gap in, 287–288
 gap in
 data for, 323–324, 323f
 definition of, 302
 estimation of, 302
 HP filter for, 287–288, 303, 321–323,
 321n34, 322f, 324, 325, 326–330, 327f,
 328f, 328n39, 329f, 330f, 460
 inflation and, 103, 109, 116, 143n45, 287,
 302, 430–431, 430f
 as latent variable, 302
 U.S. data for, 323–324, 323f
 GDP and, 323–324
 impulse response of, 201, 201f, 204f
 input *v.*, 289
 monetary policy and, 189, 405
 persistence and, 428–430, 429f
 potential for, 302, 325–330, 325f, 328f, 330f
 price and, 191, 201
 productivity and, 190–191, 201
 quantity of, 191, 198
 shock and, 326, 416, 428
 stability in, 415–416
 Taylor principle for, 143n45, 287, 302
 unemployment and, 302–303, 311–312,
 311nn21–22, 324
 volatility of, 212
 wages and, 189, 189n4

Overshooting model, of R. Dornbusch, 429

Over-the-counter (OTC), 80

P

Panageas, 220

Paravisini, D., 641

Pareto optimum, 190

Parkin, M., 111

Patinkin, D., 99, 100, 136

Patman, Wright, 352n59

Paustian, M., 384

Pavan, A., 212, 220

- Payments, systems/technology for, 27, 28–29, 30, 71–74, 73f, 108, 109n16, 111
- PCE. *See* Personal consumption expenditures
- Peach, R.W., 411
- Pederson, L., 80, 582
- Peneva, E., 256
- Peng, L., 176
- Perron, P., 443
- Persistence
 of inflation
 in reduced form, 63n29, 388, 431–449, 482
 of shocks on, 433, 433n16
 measurement of, 433–443, 434n17, 435f, 436t, 437t, 438f, 439f, 441t, 442t
 microeconomic evidence on, 478–482
 output and, 428–429
 structural sources of, 449–473
 analytics of, 452
 anchored expectations for, 473–478, 475f, 477f, 478t
 Calvo/Rotemberg model for, 451
 disinflation/supply shocks for, 450–451
 DSGE model for, 463–469, 465t, 466f, 467f
 inherited/intrinsic characteristics for, 449, 452–453, 456–459, 482
 learning models as, 471–473
 unit root test for, 435, 435n19, 437, 437n20, 437t
 Personal consumption expenditures (PCE), 434
 data for, PCE-X for, 434, 435f, 436t, 444, 444t
 Peydro, J.L., 641
 Phelps, E.S., 184
 Phillips, A.W., 184–185, 460. *See also* Phillips curve
 Phillips curve. *See also* New Keynesian theory
 conditions for, 55
 dis-inflationary boom and, 333
 Euler equation and, 454–456, 454n41
 Gordon's style of, 425, 425n3, 426t
 in hybrid form, 218
 imperfect information and, 185
 importance of, 184
 inflation and, 184, 425, 428–429, 429f, 443, 457, 468, 476, 489
 information and, 184
 IS shock and, 115, 121, 137, 139
 with lagged inflation, 470
 as long run, 29, 30n3, 53–55, 184
 for macroeconomics, 29, 30n3, 53–55, 57–61, 115, 121, 137, 139, 184
 money exclusion from, 137
 in New Keynesian model, 29, 30n3, 53–55, 57–61, 115–116, 121, 137, 139, 184, 289, 297, 298, 425, 428–429, 429f, 443, 455, 457, 468, 470, 476, 489
 as Old Monetarist, 57–61
 price *v.* wage for, 339–340, 340n50
 shocks to, 115, 121, 137, 139, 302n10, 429f
 as short run, 29, 30n3, 57, 115, 184, 222
 unemployment and, 489
 variables for, 184
 Phillips-Perron test, 437
 Piger, J., 447
 Pissarides, C., 43, 56
 Pivetta, F., 438
 Poisson process, 215, 250n11, 276
 Polan, M., 111
 Porter, R.D., 129n24
 Portfolios, monetary
 adjustment of, 156
 behavior on behalf of, 102, 106–107, 145
 macro risk premium measure for, 625f, 626, 627f
 money demand and, 145, 646
 Rational inattention to, 210
 risk aversion by, 177
 Power
 of buyer for bargaining, 88
 labor union and, 334–335
 in markets, 334–335
 of monopoly, 302n10, 307n16
 PPI. *See* Producer Price Index
 Prescott, E., 43, 295n3, 332, 341
 Price
 adjustment in, 70, 70n34, 203–204, 213–214, 217, 217n42, 233, 234, 238–242
 staggering of, 258
 types of, 247
 age of, 266
 Heckman's sample selection correction and, 267
 size *v.*, 266–267
 for aggregate supply, 207
 for assets, 34, 79, 80–83, 88, 145, 164, 374, 426
 of bonds, 44
 Calvo model for, 67, 67n31, 70, 71, 103, 141n40, 203–205, 217–218, 234, 250, 266, 289–290, 333, 333n41, 334–335, 338, 388, 427, 451, 456, 461, 470, 480, 488n1, 507, 512, 521n38

Price (*cont.*)

change in

average magnitude for, 257
 costs for, 167–168, 213, 217
 CPI for, 236t, 250, 250t
 frequency of, 238–242, 239t, 241f, 246f,
 254–256, 255f, 271–272, 278
 heterogeneity in, 273–274
 increases *v.* decreases for, 257
 inflation and, 204, 255f, 259–260, 263–264,
 263t, 270–271, 275, 278
 inventory build-up for, 262–263
 lack of synchronization for, 258–262,
 275–276, 278
 regular *v.* sale for, 262–263
 seasonality for, 260–261, 274
 size of, 274–275
 as transitory, 267–268, 275, 278
 in CM, 40n14, 44
 as comeback prices, 252–253, 253t
 as competitive, 38
 as constant, 156–157, 170
 contraction of, 427
 discrimination in, 262
 dispersion of, 67n31
 distribution for, 38
 kurtosis and, 257–258
 timing and, 250
 duration of, 247–248, 248t
 as endogenized, 35, 38
 firms' information on, 232, 234, 237, 238t,
 241–242, 257, 262, 267, 269, 276
 fiscal theory for level of, 142
 as fixed, 35, 65
 flexibility in, 65, 240, 245, 427
 goods for variation in, 232–233
 hazard rate of, 23, 266, 276
 for housing, 378, 385
 increase in, from more money, 99
 indexing and, 205
 inertia in, 156
 inflation and, 204, 255f, 259–260, 263–264,
 263t, 270–271, 275, 278
 interpretation of, 235, 235n2
 of investment, 355, 355n60
 Keynesian theory and, 57, 70, 184
 Lucas-Prescott model as, 43
 mean duration of, 242–243, 243t

monetary behavior for, 137–138
 monetary theory and, 4, 26, 28, 29–30
 as novel, 252, 253t
 output and, 189, 191, 201
 quantity plans *v.*, 193, 193f
 rational inattention and, 170
 in relation to money, 101
 response of, to shocks, 427
 rigidity in, 521
 setting of, 268
 size of change in, 257–258
 as sluggish, 157
 stability in, 415
 staggered setting of, 258
 as state-dependent, 233
 as sticky, 26, 27, 29, 30, 30n3, 34, 57, 203–205,
 217, 217n43, 233, 235, 240, 271–273, 289,
 331, 339, 353, 355–356, 356t, 359,
 523–526, 524–525f, 525n40
 timeframe for, 147
 transaction's UOA and, 136
 wages and, 137, 234, 276–277
 from Web sites, 237

Price taking, 43n16, 52

Primiceri, G, 324, 391–392, 460n48

Producer Price Index (PPI), 234–235, 235n2,
 237t, 270

Productivity

by agents, 34n6, 35
 into cash, 36
 costs/prices and, 189
 by households, 496, 496n12
 by labor, 190, 355, 492, 492n6
 monetary shocks and, 205–207, 212, 355
 output and, 190–191, 201
 shocks to, 205–207, 212, 355

Products. *See* Goods

Professional Forecasters, Survey of, 207–209,
 208n29, 398

Profitability, 191, 191n9, 193–194, 194f

Punishment, for trader, 32

Q

QTM. *See* Quantity theory of money

Quadrini, V., 582, 583

Quah, D., 493

Quantity theory of money (QTM)
ceteris paribus for, 147

- as cross-country, 110–111, 122–123, 125, 126t, 127–128
 - definition of, 147
 - equation of exchange and, 98–99, 108
 - exchange process under, 98–99, 108
 - financial sector's technological improvement and, 134
 - Friedman's conception of, 100n2
 - meanings for, 99
 - money growth and
 - data averaging for, 112–115
 - demand for, 131–134
 - historical data for, 121–131
 - New Keynesian model for, 113, 115–121, 117f
 - quantitative models for, 115–121
 - as unconditioned, 113n19
 - money *v.* inflation in, 46, 98, 101, 110–111, 147
 - neutrality of money and, 6, 29, 44–45, 67, 70, 70n34, 99, 147
 - New Keynesian theory for, 113, 115–121, 117f, 120t, 121t
 - supply/demand for, 100, 100n5
 - for U.S., 112, 122, 123t, 127
- R**
- Rabanal, P., 493
 - Rafael, 525n40
 - Rajan, R., 588–589
 - Ramey, V.A., 289, 493
 - Ramsey, 297
 - Ramsey-efficient equilibrium, 297, 297n4, 302, 302n10
 - Randomization, 10, 16, 28
 - Rational expectations (RE)
 - agents' actions for, 172
 - equilibrium model for, 140–141, 141n39, 172
 - for forecasting, 174
 - imperfect information for, 196
 - inflation persistence and, 426–431, 443, 443n27
 - as misleading, 178
 - for policy evaluation, 174–175
 - shocks for, 174n8
 - Rational inattention
 - competitive markets for, 177
 - entropy for, 215, 215n36
 - Gaussian-linear-quadratic examples for, 161–168, 172–173, 178–180
 - market fluctuations/variables for, 174, 176
 - mutual information for, 213, 215–216, 216nn37–40
 - partial information and, 185
 - rational expectation and, 164–166
 - responses to, 163–164
 - Shannon measure and, 156–157
 - as slow response, 147, 167–168
 - uncertainty and, 166–167, 176
 - Ravenna, F., 269, 289, 302n9, 307, 447–448
 - Ray, S., 213n43
 - RE. *See* Rational expectations
 - Recession
 - adverse selection in, 583
 - booms *v.*, 305, 305n15, 309–311, 323–324, 323f, 602
 - credit crunch in, 631, 633
 - globalization and, 548
 - Reduced-form persistence, 63n29, 388, 431–449, 482
 - Regulation Q, 105, 381. *See also* Federal Reserve, U.S.; M1/M2 series
 - Reifschneider, D., 379n6, 388
 - Reis, R., 168, 191, 200, 200n19, 205, 206, 207, 209, 212, 213–215, 214n34, 215n35, 219, 220, 221, 221n46, 221n48, 249, 438, 470, 472f
 - repos. *See* Repurchase agreements
 - Repurchase agreements (repos), 620, 634, 636, 637, 638t
 - Reserve Bank of New Zealand, 138
 - Reserves, requirement for, 28, 106, 112n17, 135
 - Resources, economic
 - allocation of, 302
 - of consumer, 378
 - equilibrium and, 294–296
 - Retailers, markups by, 47
 - Reynard, S., 108
 - Rigidity, strategic complementarity and, 195–196
 - Rigobon, R., 235, 237, 241, 268
 - Risk
 - appetite for, 623–627, 627f
 - for banks, 549, 602–605
 - channels for, 603–605, 638–646
 - for credit, 549, 586
 - premium for, 603, 604, 606–615, 625f
 - sharing of, 14, 18, 21–22, 495

Risk (*cont.*)

as systemic, 606

Roberts, J., 428n6, 457

Roca, M., 212

Rocheteau, G., 15, 28–29, 30n3, 43n16, 45, 50n22, 53n24, 56, 74, 80n38

Rogerson, R., 53–55, 290, 300

Romer, C.D., 387

Romer, D., 195, 271–272

Rondina, G., 202n22

Rosen, S., 343

Rostagno, M., 309, 583

Rotemberg, J.J., 115, 116n21, 427, 451, 460, 471, 479, 488n1

Rudd, J., 238, 455

Rudolf, B., 470

Rupert, P., 53n24

S

Sacrifice ratio, for unemployment, 425, 425n1, 450

Sala, L., 490

Samuelson, P., 27, 99, 101, 184

Sanches, D., 33, 43n16

Sannikov, Y., 584, 643–644

Sargent, T.J., 5, 103, 113, 113n20, 312n23, 426, 460n48, 472

Saurina, J., 641

Savings, rate of return on, 20

Sbordone, A., 437n20, 443, 451, 460n48, 461–463, 462n50, 463n51, 464, 464nn53–54, 468

Schabert, A., 289, 302n9, 303n12

Schorfheide, F., 30n3, 65, 172

Schwartz, A.J., 104, 122

Schwartzman, F., 219

SDP. *See* State-dependent pricing models

Search model

bargaining and, 47, 80

competitive markets *v.*, 56

economy and, 39

for equilibrium, 27n2, 31, 39

frictions in, 39, 67n31

households for, 495, 495n9

matching model and, 489, 489n3, 496

monetary policy shock and, 522n39

Securities

ABS as, 615, 632–633, 632f, 635, 636–637, 636f, 640

broker-dealers in, 615–619, 616f, 617f, 618f, 619f

before financial crisis, 177

interest rate on, 102, 105, 106, 135

lending by, 386–387

as mortgage-backed, 386–387, 551

Shadow banks. *See also* Banks

ABS/MBS for, 636–637

asset growth of, 618f

balance sheets for, 605

broker-dealers and, 615–619, 616f, 617f, 618f, 619f

in commercial paper market, 636

financial crisis and, 387, 588, 603, 605

intermediaries for, 615–619, 616f, 617f, 618f, 619f

monetary model for, 614–615, 618f

Shannon measure

channel capacity for, 135–136, 158–159, 173

in communications engineering, 158

definition of mutual information by, 157–158

with fixed capacity, 173

Gaussian case for, 161–168, 172–173

information processing by, 157–158, 161

processing rate variations for, 173

rational inattention and, 156–157

research on, 176

Shares, in markets, 84–88, 84n40

Sheedy, K.D., 262

Shell, K., 54

Shi, S., 8, 21–22, 21n9, 35, 36, 38, 39, 40, 56

Shiller, R.J., 378

Shimer, R., 503n25, 505, 516

Shin, H. S., 202n22, 211, 212, 218, 582, 606, 619–620, 623–626, 629, 634, 639–640, 643, 647

Shintani, M., 221

Shleifer, A., 584n20, 606, 606n3

Shocks, monetary

for aggregate, 28, 48, 197–200, 207–208, 208f, 267–268

consumption and, 387

for cost, 255–256

currency change for, 269

as cyclical, 245–246

by DSGE model, 413

for economy, 288, 291

to Euler equation, 453–454, 454t

- identification of, 345–346, 346n54, 392
- for income, 387
- for inflation, 341, 428
 - dynamics of, 431–432, 432n14
- as IS curve, 115–116, 116n21, 118f, 121, 137, 137n31, 139, 288–289, 362, 401, 464
- as liquidity, 81, 550
- as long-term, 271
- in macroeconomics, 412–414, 413f, 414f
- by monetary policy, 351–352, 352–354, 352f, 358–362, 359n63, 360f, 361f, 362f, 391–392
- by money demand, 144, 646
- in New Keynesian model, 117f, 119f, 287
- nominal demand for, 204f, 214n33, 232–235, 271
- as nonpolicy, 101, 116
- output and, 326, 416
- as permanent, 266
- Poisson process for, 215, 250n11, 276
- as preference/technology, 31, 43, 51, 57, 81n39
- price response to, 427
- to productivity, 205–207, 212, 355
- RE *v.* rational inattention for, 174n8
- as real, 57–58
- response to, 117f, 268–270, 413
- tax rates for, 269
- in technology, 305n15, 319, 320t, 327–332, 344–346, 346n54, 353–355, 353f, 359–360, 359n63, 492, 494f, 517–520, 518–519f
 - inflation and, 359–360, 359n63
- as transitory, 267–268
- volatility of, and price change, 255
- Sichel, D.E., 387
- Signal extraction, 29–30, 30n3, 202n22, 207n28, 212
- Silva, J., 516
- Sims, C.A., 142, 169–170, 176, 213, 215–216, 346, 351
- SIV. *See* Structured Investment Vehicle
- Skrzypacz, A., 270
- Slacaleck, J., 201n20, 206, 220
- Slobodyan, S., 472
- Small, D.H., 129n24
- Smets, F., 113, 144, 172, 480
- Smith, A., 31
- Smith, B., 5
- Smith, T., 38
- Solow, R., 184, 295n3
- Sommer, M., 220
- Special purpose vehicle (SPV), 584–585
- Spulber, D., 67n30
- SPV. *See* Special purpose vehicle
- Standard & Poors, 625
- Starr, R., 27n2
- State-dependent pricing models (SDP), 258, 275, 276, 469–470
- Stein, J., 631
- Steinsson, J., 69n33, 234, 235, 236t, 237t, 240, 242, 246, 247, 249, 255–256, 257, 260–261, 266, 267, 273, 457n44, 479
- Sticky prices. *See also* Frictions
 - for benchmark model, 66
 - in CPI, 251
 - in DM, 62
 - duration of, 276
 - equilibrium and, 63
 - as friction, 66, 333
 - macroeconomics and, 273
 - New Keynesian theory and, 57, 61–66, 70, 184, 289, 333, 333n41, 507
 - New Monetarism theory for, 66–71
 - for shelter, 246n8
 - theory for, 26, 27, 29, 30, 30n3, 34, 57, 203–205, 217, 217n43, 233, 235, 240, 271–273, 289, 331, 339, 353, 355–356, 356t, 359, 523–526, 524–525f, 525n40
- Sticky wages, 506, 530
- Stiglitz, 291
- Stock, J.H., 434, 437n20, 445, 446
- Stracca, L., 385
- Strategic complementarity, rigidity and, 195–196
- Structural vector autoregression (SVAR), 172
- Structured Investment Vehicle (SIV), 605.
 - See also* Intermediaries, financial
- Stuart, A., 143
- Sun, H., 56, 220
- Superneutrality, of money, 102–103
- Surico, P., 113, 113n20
- Surveys
 - Annual Retail Trade Survey, 47
 - Livingston Survey, 207–208, 208n29
 - Michigan Survey of Consumer Attitudes and Behavior, 207–209, 208n29
 - Survey of Professional Forecasters, 207–209, 208n29, 398

SVAR. *See* structural vector autoregression
 Svensson, L.E.O., 103, 112, 115, 212, 604
 Swanson, N.R., 221
 sweep program and, 106–107

T

TALF. *See* Term asset-backed loan facility
 Tambalotti, A., 269
 TARP. *See* Troubled Assets Relief Program
 Taxation
 budget constraints and, 20, 22, 100n3
 restriction on, 20, 22
 of underground activities, 5, 5n2
 Taylor, J.B., 116n21, 143, 234, 258, 266, 266n15,
 427, 428, 464, 468n56, 474, 479, 506
 Taylor principle
 assets and, 287, 290
 central banks and, 115, 474
 contracting models for, 428
 equilibrium for, 290
 expansion of, 538
 expectation channel and, 388
 historical comparisons using, 143
 inflation and, 118, 140, 303
 output gap and, 143n45, 287, 302
 targeting of, 309
 for labor market, 491
 log-linearized equilibrium with, 296–299
 monetary models and, 28, 62, 115, 116, 116n21,
 119, 140, 142–143, 143n45, 250, 266,
 266n15, 287, 290, 296–299, 302, 303–309,
 304nn13–14, 310, 316t, 344, 350, 356,
 357t, 491, 502–503, 532–533
 monopoly power and, 307n16
 New Keynesian theory and, 28, 62, 119,
 143n45, 250, 266n15, 290, 296–299, 302,
 303–309, 304nn13–14, 502–503
 output gap and, 143n45, 287, 302
 price adjustment for, 266, 266n15
 price rigidities/labor market frictions and, 287,
 290, 296, 299, 302–309, 310, 316t, 344,
 350, 356, 357t, 491
 problems for, 287
 stability under, 305
 time-dependent pricing by, 234, 258
 working capital channel and, 306–309, 308f
 TDP. *See* Time-dependent pricing models
 Technology

 central banks and, 134
 as cost saving, 309
 for credit information, 386–387
 shocks for, 305n15, 319, 320t, 327–332,
 344–346, 346n54, 353–355, 353f, 359–360,
 359n63, 492, 494f, 517–520, 518–519f
 inflation and, 359–360, 359n63
 VAR analysis for, 332
 Teles, P., 5
 Telyukova, I., 40n13
 Term asset-backed loan facility (TALF), 635–636,
 636f
 Term spread, 603–605, 604f, 637–640
 Tetlow, R., 379n6
 Thomas, C., 489, 501n22, 508n31, 511n33, 512
 Time-dependent pricing models (TDP), 258–259,
 275, 276
 Tirole, J., 588–589, 642–643
 Tobin, J., 40n13, 374, 377
 Toledo, M., 516
 Tootell, G.M.B., 467
 Topel, R., 343
 ToTEM, at Bank of Canada, 377
 Townsend, R.M., 6, 27–28, 185, 583
 Trabandt, M., 221n46, 287
 Trade
 channels for, 376
 in commodities, 7
 as competitive, 5
 defection in, 10–11, 10n4, 18
 discounting for, 5
 DM for, 39
 equilibrium in, 72
 by households/groups, 39
 mechanism for, 5
 meetings for, 7, 9–10
 as monetary
 command allocation for, 5
 as essential, 4, 5–7
 fiat *v.* commodity as, 6
 frictions for, 5–9
 settings for, 4
 pairwise meetings for, 7, 9–10, 22
 partners for, 39
 price signals for, 173
 punishment for, 32
 risk-sharing in, 14, 18, 21–22
 as specialists, 31

- Transactions. *See also* Channels, monetary transmission
 as anonymous, 135
 bank liabilities for, 75
 money in, 5, 75, 135
 prices for, 156–157
 reserves for, 28, 106, 112n17, 135
 as short *v.* long term, 172, 173
 UOA for, 136
- Transmission, monetary. *See also* Channels, monetary transmission
 changes in, 370, 405, 406t, 407–411, 407f, 408f, 409f, 410f
 channels for, 135–136, 158–160, 373–385
 as neoclassical, 374, 374n2, 375t, 376, 415
 as non-neoclassical, 373, 374, 375t, 380–381, 415
 spending in, 374
 DSGE model for, 372
 globalization and, 385n13
 lending crisis and, 416
 over Internet, 158, 159
 for residential investment, 415, 416
 VAR/FAVAR approach to, 372, 390–391, 392–396, 394f, 416
- Transparency
 Bayesian approach for, 288
 by central banks, 211–212
 as harmful, 212
 in monetary policy, 175–176, 211–213
- Treasury, U.S. Department of, 88, 145, 398, 398n19, 405, 549, 566, 604, 604f, 625
- Trejos, A., 8, 22, 35, 36, 40
- Triangle model, of inflation by R. Gordon, 425, 425n1, 426t, 428, 460
- Trigari, A., 489, 490, 503n24, 512, 522n39
- Tristani, O., 111
- Troubled Assets Relief Program (TARP), 551
- Tsiddon, D., 262, 270
- Tsuruga, T., 215n35, 217–218, 221
- Tutino, A., 177, 216n39, 220
- U**
- UK. *See* United Kingdom
- Unemployment
 CTW model for, 302–303, 303n11, 312–315, 313n25, 320–326, 321f, 321nn33–35
 as cyclical, 488, 491–494, 492t
 DSGE and, 287, 303, 311–315, 313n25
 fluctuations in, 488–489, 491
 frictions for, 345, 363, 363n66, 488–491, 499–502, 520–521f, 520–523, 522f, 523f, 528
 Gali model for, 303n11
 inflation and, 29
 information content of, 311–312
 as involuntary, 488, 489n4, 497
 in macroeconomics, 45, 53, 53n24, 287, 302–303, 303n11, 312–315, 313n25, 320–326, 321f, 321nn33–35, 345, 363, 363n66, 488–491, 489n4, 499–502, 520–521f, 520–523, 522f, 523f, 528
 in monetary theory/policy, 45, 488–489, 489n4
 natural rate of, 53, 222
 New Keynesian model for, 287, 489–490, 489n4
 output gap and, 302–303, 311–312, 311nn21–22, 324
 Phillips curve and, 489
 Rogerson labor model for, 53, 53n24
 sacrifice ratio for, 425
 stabilization of, 491
 as variable, 324, 325n38
 volatility of, 506
- Union, monopoly
 Calvo frictions for, 338, 507
 employment and, 338–340, 507
 market power and, 334–335
 optimization by, 338–339
 wages and, 338–340
- Unit of account (UOA), 136
- Unit root tests, 433, 435
- United Kingdom (UK), 105n14, 122, 123t, 143, 146–148
- United States (U.S.)
 CPI in, 111–112, 114, 114t, 123t, 124f, 147, 236t, 427f
 credit/money growth in, 146
 economic openness of, 385n13
 GDP growth for, 122–123, 122n22, 144
 gold standard in, 444–445
 Great Depression in, 385n12, 548, 642
 inflation in, 52n59, 111, 114t, 124f, 236t, 313
 monetary aggregates in, 104
 money growth rates in, 146
 QTM for, 112, 122, 123t, 127
 unknown breakpoint test, 434n17, 443–444
 Utility, from goods, 32n4

V

- Valles, J., 392
- van Nieuwerburgh, S., 176, 210
- van Rens, T., 512n34
- van Wincoop, E., 221
- Variables
 - in data, 156
 - for inertia, 156, 175, 424
 - in information, 207, 207n27
 - for investment, 378
 - in markets, 169, 176
 - Phillips curve and, 184
 - rational inattention and, 174, 176
 - scale of disturbance for, 172
 - for VARs, 640
- Vector autoregressions (VARs)
 - assessment of, 360–362, 360f, 361f, 362f
 - for Bayesian approach, 288, 391
 - as constraint, 619–621, 633
 - estimation strategy through, 342, 345–347, 351–355, 351nn57–58, 395f
 - FAVAR approach to, 393–396, 394f, 395f, 416
 - for impulse response functions, 288
 - lag length for, 346, 346n55
 - Laplace approximation and, 317, 348, 350–351, 360–362
 - for monetary transmission, 372, 390–391, 392–396, 394f, 416
 - technology shocks and, 332
 - variables for, 640
 - for working capital channel, 289
- Veldkamp, L., 176, 185n1, 210
- Velocity
 - inertia v ., 424
 - as monetary, 99, 106, 107–111
- Venkateswaran, V., 209–210
- Vermeulen, P., 241, 247, 256
- Vincent, N., 269
- Vishny, R., 584, 606, 606n3
- VIX. *See* Chicago Board Options Exchange Volatility Index
- Volatility
 - in banks' net worth, 586, 588
 - central banks and, 310
 - haircuts and, 637
 - of inflation, 219, 415
 - in markets, 310, 588, 637
 - in monetary theory/policy, 287, 309

- in output gap, 212
- of shocks/price change, 255
- of unemployment, 506
- Volcker, Paul, 209, 210f, 371, 391, 447

W

- Wages, labor, 301–302, 489, 491, 507, 510, 516.
 - See also* Labor market; Nash bargaining
- Waldman, M., 218
- Walentin, Karl, 287
- Wallace, N., 5, 8, 10n4, 15, 16, 20, 27, 58, 426
- Wallace, W., 43n16
- Waller, C., 40n13
- Wallich, Henry, 138
- Walras' Law, 565
- Walrasian price taking, 43, 44, 48n21, 51, 53, 65–66, 71–76, 497, 522
- Walsh, C.E., 289, 302n9, 307, 488n1, 489, 489n2, 497, 503n24, 521–522
- Watson, M.W., 116, 144, 434, 437n20, 445, 446
- Weber, W.E., 114
- Wei, M., 208n29
- Weibull distribution, 266n15
- Weil, D.N., 387
- Weill, P.O., 212
- Werning, I., 206n24
- Whelan, K., 448, 455
- Wicksell, K., 11, 32n4, 99, 136–137, 136n30, 138–140, 140n36
- Wiederholt, M., 171, 176, 177, 206, 206n26, 207
- Wiener process, 164
- Wiener-Kolmogorov formulae, 202n22
- Williams, J., 379n6, 472, 473, 475, 475f, 475n67
- Williamson, S., 27, 28, 33, 40n13, 43n16, 50n22, 583
- Willis, J.I., 207, 217, 262, 267, 270, 274
- Wintr, L., 237
- Wolfers, J., 207, 209
- Wolman, A., 234n1, 452n38, 469, 488n1
- Woodford, M., 29, 61–62, 65, 113, 115, 116n21, 120, 122, 135, 136, 137, 139–141, 140n36, 140n38, 142, 176, 202, 212, 218, 249, 275, 289, 384, 457, 488n1, 489n2, 602, 604
- Worker-shopper pair. *See* Models, monetary
- Working capital channel, 286, 287, 289, 289n2, 298, 298n5, 302, 306–309, 308f
- Wouters, R., 113, 144, 172, 332, 373, 400, 472, 490

Wright, R., 8, 21–22, 27, 29, 31, 32n4, 35, 35n8,
36, 39, 40, 40n13, 44, 47, 50n22, 53n24,
54, 56, 66

Wulfsberg, F., 249, 257, 260

X

Xiong, W., 176

Y

Yankov, V., 583

Yellen, J.L., 161, 193

Yield, on assets, 88, 398, 604, 605

Young, E.R., 176

Yun, T., 103, 294–296, 295n3, 488n1

Z

Zabczyk, P., 481

Zaffaroni, P., 481

Zakrajsek, E., 398, 398n19, 411n20, 583

Zbaracki, M.J., 191n10

Zerom, D., 269

Zha, T., 346

Zhu, H., 604

Zhu, T., 10n4

Zhu, Z., 201n20

This page intentionally left blank

INDEX-VOLUME 3B

Note: Page numbers followed by f, t and n indicate figures, tables and notes, respectively.

A

- Acar, M., 1450
- Account reversals, 1481
- Accountability measures, 854–855
- Acemoglu, D., 1019
- Active rules, 960
- Adam, K., 703
- Adao, B., 947n12
- Adaptive control, 1274n42
 - model, 1124
 - robust control v., 1132
- Adaptive learning, 824
 - baseline model calibration under, 1075–1076
 - inflation persistence in, 1073–1089
 - IT and, 1071–1073
 - macroeconomic outcomes and, 1076t
 - for monetary policies, 1057
 - monetary policy rules/stability under, 1065–1071
 - optimal monetary policy under, 1071–1089
 - structural change transition dynamics in, 1090
 - structural relations and, 1062n7
- Adaptive models, 1123–1125
- Adaptive optimal policy (AOP), 1274
- Adolfson, M., 1252n18, 1255n25, 1269
- Adrian, T., 1294n68, 1430
- Advanced economies, 1244
- After-tax revenues, 762, 763
- Aggregates, monetary. *See also* Federal Reserve, US; Inflation
 - colinearity exceptions, 1326–1328
 - colinearity of, 1326
 - demand, 1450n23
 - private consumption and, 787n73
 - supply relations, 726–727
 - disturbances, 805
 - fluctuations, 702
 - growth rates of, 1353
 - liquidity effect using, 1369–1370
 - monetary regimes and, 1325–1328
 - resource constraint, 665, 676, 687
 - supply
 - curve, 757
 - demand relation, 726–727
 - inflation rate satisfying, 794
 - log-linear, 779–780, 786
 - Phillips curve and, 799
 - short-run, 764–765
 - two pillars, 1325–1326
 - underground economy activity levels of, 674
- Aghion, P., 1464n76, 1480
- Aguiar, M., 1465
- Ahmed, S., 1450
- AIM algorithm, 1255
- Aizenman, J., 1458n54, 1477n119, 1479n134, 1496n200, 1497n203
- Akhmedov, A., 1033
- Albenesi, S., 994
- Alesina, A., 1017, 1021–1022, 1027, 1029, 1031–1032, 1035, 1039–1041, 1498n210
- Algorithms, in Ramsey problem, 679–680
- Alternative monetary policy instruments, 842n5
- Altig, D., 697, 702
- Alvarez, F., 993
- Amano, R., 1287
- Amato, J., 1459n55
- Ambler, S., 1287
- An, S., 1231
- Anchor country, 1036n41
- Anderson, 1430
- Anderson, E., 1112n17, 1114, 1145–1146, 1148, 1148n40
- Anderson, G. S., 1255n25
- Angeloni, I., 1394, 1399–1400
- Angeriz, A., 1247
- Announcement effect, 1387, 1397f
- Anticipated utility model, 1124
- Anticipation effect
 - interest rates announcement with, 1387–1388
 - reserve management
 - influenced by, 1392–1399

- Aoki, K., 742, 803, 811, 812n96, 1272, 1272n40
 AOP. *See* Adaptive optimal policy
 Ardagna, S., 1041
 Arestis, P., 1247
 Arifovic, J., 1090
 Ascari, G., 1198–1199
 Ashcraft, A., 1425
 Assenmacher-Wesche, K., 1290
 Assets
 of central banks, 1422f
 demand/returns of, 1386
 endowment economy and, 1122
 frictionless markets in, 864–865, 869–915
 international markets in, 877–879
 markets, 915–928, 1187
 nominal government, 943
 -price bubble, 1294n69
 price stabilization in, 844n6
 prices, 843, 1496–1498
 stabilization of, 844n6
 toxic, 1442–1443
 trade, 878–879
 Asso, F., 854
 Asymmetric disturbances
 price stickiness and, 807n92
 sectoral heterogeneity and, 803–815
 three types of, 804
 Asymmetry
 financial crisis with, 1011–1012
 pricing patterns with, 877n9
 Asymptotic fluctuations, 772
 Atkeson, A., 947, 951
 Australia
 economic variables checklist in, 1180
 Interest rates/inflation/output of, 1162–1163f
 Autocorrelation, 971t
 Average inflation, 1187t
Aversion to uncertainty, 1102–1103
B
 Baba, N., 1213
 Bacchetta, P., 908, 1464n76, 1480n135
 Backus, D., 1266
 Backward-looking model, 1115, 1253n21
 constraints in, 768
 distortions in, 1121
 Bade, R., 1017
 Bahmani-Oskooee, M., 1450
 Balance of payments, 1488–1493
 Balance sheet effect, 1451–1453, 1493f
 Balance sheet recession, 1215–1216
 Balassa, B., 1447–1448
 Balassa-Samuelson relationship, 1448
 Baldwin, R. E., 1319
 Ball, L., 798, 801, 848, 1115, 1133, 1142, 1142f,
 1247, 1249, 1307–1309, 1315–1316,
 1334–1337
 Ball's model, 1141–1143
 Ball-Sheridan methodology, 1307
 Band Basket Crawl (BBC), 1464
 Banerjee, R., 1480n135
 Bank credit, 1497
 Bank of Canada, 1278n49
 Bank of England
 constant interest rate of, 1263n31
 interest rates and, 1278n49
 monetary policy committee of, 1023
 Bank of Japan (BoJ), 1209
 criticism of, 1210–1211
 daylight overdrafts and, 1383n44
 policy measures of, 1420
 QE policy of, 1212
 reserve demand shocks to, 1372
 zero interest rate policy pursued by, 1211–1212
 Bank of Korea, 1455
 Bank reserves, 1172–1173
 BoJ demand shocks of, 1372
 central bank's quantity of, 1364–1366
 central bank's supply of, 1411n77, 1423–1424,
 1426–1427
 demand for, 1362–1363
 demand within maintenance period of,
 1404–1409
 Euro Area's excess demand of, 1382t
 Eurosysteem demand on, 1373
 financial flows and, 1372
 interest rate relationship with, 1349, 1434
 interest rates and, 1375, 1384f, 1388–1392
 Japan's excess demand of, 1383t
 liquidity effect and, 1348–1349, 1370–1371
 monetary policy demand of, 1365f
 overnight interest rates and, 1376f
 policy interest rates relationship with,
 1374–1383, 1385–1386
 quantity of, 1358
 short-term foreign debts and, 1452n34

- short-term interest rates and, 1384f
- supply changes of, 1352–1353, 1365–1366
- supply of, 1348
- supply within maintenance period of, 1409–1413
- supply-demand equilibrium for, 1360
- target interest rates and, 1378f
- United States' demand for, 1379–1380, 1380t, 1424n88
- United States requirements of, 1360n11, 1365n18
- Bank runs theory, 1483n157
- Bankruptcy law, 1494
- Banks. *See also* Central bank(s); European Central Bank; Federal Reserve
 - Japan's lending reluctance of, 1214–1215
 - Norges, 1269
 - quantity of reserves of, 1358
 - reserves chain of causation of, 1172
 - United States currency holdings of, 1361n13
- Bansal, R., 1122
- Barillas, F., 1132n33
- Barnett, S., 1468n95
- Barro, R., 1004, 1006, 1035, 1040, 1479n127
- Barsky, R. B., 1177
- Barsky, R. T., 1145
- Bartolini, L., 1394n55, 1477n121
- Basar, T., 1106
- Basel Accord, 1219
- Basel Committee on banking supervision (BCBS), 1218
- Baseline closed-economy models, 864
- Baseline model calibration, 1075–1076
- Baseline monetary model, 870–886
- Bassetto, M., 949, 951
- Batini, N., 837, 1249, 1285, 1459n55
- Bayesian approach, 845, 1101, 1124
- Bayesian decision theory, 1109
- Bayesian estimation, 1231
- Bayesian model, 1113–1117
 - detection probabilities in, 1114–1116
 - learning and, 1117–1119
 - reservations/extensions in, 1116–1117
- Bayesian optimal policy (BOP), 847–848, 1274, 1274n43
- Bayesian optimal simple three-parameter rule, 846, 848
- Bayesian probability, 1123f
- Bayesian-Kalman filtering, 1130
- BBC. *See* Band Basket Crawl
- BCBS. *See* Basel Committee on banking supervision
- Bean, C. R., 1289n58
- Bebczuk, 1452
- Beck, G. W., 1274n43
- Bekaert, G., 1475, 1479n131, 1479n134
- Belief changes, 1122–1123
- Bellman equation, 1118–1119
 - adaptive model misusing, 1125
 - robust decision rules induced by, 1118, 1121
- Bellman-Issacs condition, 1108
- Benati, L., 1207, 1228, 1231
- Benchmark parameter values, 905t, 984t, 1076t
- Benhabib, J., 1070
- Benigno, P., 690n7, 760, 764n48, 765, 775, 786, 788, 811n95, 816n102, 820, 927, 976, 977n37, 986–987, 992
- Berg, A., 1462
- Bergin, P., 954
- Bergo, J., 1269, 1284
- Bernanke, B. S., 1056, 1191, 1210–1211, 1213, 1221, 1290, 1359n10, 1368
- Bernhard, P., 1106
- Beyer, A., 1178
- Biased estimators, 1334
- Bienen, H., 1449n21
- Bilateral debt, 1494n187
- Billi, R. M., 703
- Bils, M., 697
- Biscarri, J. G., 1480n135
- Blackwell, D. A., 1103n7
- Blake, A. P., 746–748, 1273
- Blanchard, O., 865, 955, 1025, 1321
- Blenck, D., 1388n48
- Blinder, A. S., 1022, 1061, 1359n10
- Block, S., 1033
- Bohn, H., 964, 966, 967–968
- Boivin, J., 1059n6, 1196
- BoJ. *See* Bank of Japan
- Bonds
 - government, 943n7
 - liquidity services of, 961–963
 - price support of, 947–948
- BOP. *See* Bayesian optimal policy

Bordo, M., 1161, 1487n178
 Borensztein, E., 1462, 1485n166
 Borio, C., 1294, 1388, 1397n61, 1425
 Borrowed reserve target, 1179, 1379n39
 Boskin Commission, 712, 714
 Bounded processes, 774
 Brainard, W., 1101, 1366
 Branch, W., 1090
 Brayton, F., 844
 Brazier, A., 1090
 Breakdown point, 1138–1139, 1151
 Brender, A., 1033, 1466
 Brock, W. A., 847, 1127
 Brock-Sidrauski model, 976
 Brookings Institute, 834
 Brown, G., 1161n3, 1185
 Brumm, H. J., 1018
 Brunner, A. D., 1357
 Brunner, K., 1174n17
 Bruno, M., 1453
 Bryant, R., 831, 833–834, 1174n17
 Budget constraints
 flow, 660, 944
 of government, 939
 of households, 685, 938–939
 individual flow, 874, 878–879
 in New Keynesian analysis, 874
 present value, 939
 sequential government, 661, 665, 676
 Bugamelli, M., 1042n48
 Buiter, W., 949, 950n14, 951, 977n38, 1044n51
 Bulgaria, 1330f
 Bullard, J., 1057, 1065, 1068–1069, 1090
 Bulow, J., 1493
 Bundersbank, of Germany, 1179
 Bundesbank, of Germany, 1168
 Bureaucrats, 1021–1022
 Burns, A., 1177
 Burnside, C., 1483n157, 1484n162
 Burstein, A., 907n24, 985n49, 1445n12, 1446n14, 1451n27
 Business, opportunistic cycles, 1047–1050

C

Caballero, R., 1458n54
 CACs. *See* Collective action clauses
 Cagan, Phillip, 1177

Cagetti, M., 1131, 1145
 Calibration
 baseline model, 1075–1076
 for robustness, 1109–1117
 Calvo, G., 684, 686, 761, 976, 985, 993, 1450, 1452n29, 1454, 1461, 1482, 1484n162, 1485n167
 Calvo price setting, 815, 938, 979–980, 985n49, 1061
 Calvo-Phillips, 1357–1358
 Calvo-Yun model, 688, 792
 Campbell, J. Y., 703
 Campillo, M., 1017–1018
 Canada
 central bank of, 1180
 interest rates/inflation/output of, 1162–1163f
 transfer of balances of, 1361n12
 United States interest rates influencing, 1180n38
 Canova, F., 1195
 Canzoneri, M., 725, 942, 948, 949, 954, 959, 961, 965, 1007
 Capacity utilization gap, 839
 Capital. *See also* Bank reserves
 accumulation/sticky prices with, 684–689
 controls, 1477, 1477n119
 financial crisis requirements of, 1218–1220
 flight, 1331, 1333t
 inflows, 1479–1481
 markets, 1320–1321, 1366–1367
 Capital flows, 1472–1481
 emerging markets, 1472–1477
 capital controls in, 1477
 financial integration and, 1472
 integration legal barriers in, 1472–1473
 market prices and, 1473–1475
 sterilization/offset in, 1475–1476
 financial openness
 capital inflows and, 1479–1481
 welfare improvement from, 1477–1480
 procyclicality and, 1465–1466
 Carlstrom, C., 941, 944
 Carpenter, S. B., 1370, 1370n32, 1381n41, 1401–1402, 1402n68, 1406–1408, 1412
 Carter, C. K., 1229
 Carter, Jimmy, 1172
 Carter, T., 1287
 Carvalho, A., 1249
 Cash/credit goods model, 973–974, 994

- benchmark parameter values in, 984t
- nominal interest rate implied in, 975
- nonzero interest rates in, 985–986
- optimal inflation and, 985, 989f
- optimal inflation/interest rates in, 988f, 990f
- optimal monetary/fiscal policy and, 980–984
- policy variables in, 987t
- price stability and, 977–980
- wage stickiness and, 987n52
- Cash-in-advance model, 938–939
- Castelnuovo, E., 1061
- Catao, 1454
- Causality, 1019–1020
- Cavallo, E., 1452n29
- Cavallo, M., 1451
- CBI. *See* Central bank independence
- CCDL, 962f, 963f
- Cecchetti, S. G., 853–854, 1315
- Central bank(s)
 - asset-price bubble of, 1294n69
 - bank reserve supply changes used by, 1352–1353, 1365–1366
 - of Canada, 1180
 - commitment value of, 733–737
 - equilibrium predictions of, 823–824
 - Euro Area's assets of, 1422f
 - Euro Area's liabilities of, 1421f
 - financial crisis shut down of, 1220–1221
 - fixed reserves quantity of, 1364–1366
 - flexible inflation targeting and, 740–741
 - forecast targeting of, 738
 - forward path projections of, 814–815
 - future policy implications for, 1221–1223
 - future policy rate of, 1387n47
 - gap-adjusted price level failure of, 759
 - government policy coordination
 - problem with, 948–949, 955–963
 - inflation increase/real interest rate
 - increase of, 945–946
 - inflation objectives of, 654–655
 - inflation suppressed by, 940
 - inflation targets of, 940, 975n33
 - inflation/output expectations of, 725–726
 - information set of, 758n40
 - interest rate below normal of, 1358
 - interest rate expectations of, 1394n57
 - interest rates of, 835–836, 853–854, 893–894, 1427
 - interest rates set by, 1347–1351, 1366–1367, 1383–1399
 - intertemporal trade-off facing, 1084–1087
 - Japan's assets of, 1422f
 - Japan's liabilities of, 1421f
 - loss function of, 835n1, 1006
 - money supply control of, 726–727
 - operating procedures of, 1389t
 - optimal inflation target of, 713
 - optimal policy theory of, 757–758, 851–852
 - output gap stabilization and, 1088–1089
 - policy rate changes by, 1351–1352
 - price stabilization of, 944, 1167
 - private-sector expectations of, 825, 825n114
 - recipe for success of, 1189
 - reputation loss of, 1009
 - reserves supply varying of, 1411n77, 1423–1424, 1426–1427
 - reserves/policy interest rates and, 1385–1386
 - response constraint of, 1013
 - rules deviation of, 1007, 1009–1010
 - stabilization policies of, 657, 701–702
 - standing facilities of, 1394n55
 - target criterion of, 791
 - target variables of, 1250
 - targeted asset purchases of, 1428
 - Taylor principles obeying, 954–955
 - United States' assets of, 1422f
 - United States' liabilities of, 1421f
 - zero lower bound constraining, 749n29
- Central bank independence (CBI), 1013–1027, 1031n37
 - causality and, 1019–1020
 - contracting approach in, 1016–1017
 - democratic deficit from, 1020–1022
 - during financial crisis, 1014–1015, 1023–1025, 1047
 - inflation and, 1455–1456
 - inflation's negative
 - relationship with, 1020n23
 - instrument v. goal independence in, 1016
 - loss function minimized by, 1046–1047
 - macroeconomic performance and, 1017–1019
 - measuring degree of, 1017–1018
 - political cycles and, 1031–1032
 - rules in, 1015–1017
 - rules/discretion in, 1013–1014
- Certainty-equivalence theorem, 1259n28

- Chang, R., 1483n157
 Chao, C. C., 1450
 Chari, A., 1479n131
 Chari, V., 664, 947, 973, 975, 976, 983, 984, 986, 987n51, 994
 Chen, Z., 1146
 Chernoff, H., 1111, 1113
 Chernoff entropy, 1113–1114
 China, 1186t, 1189
 Chinn, M., 1479n134, 1484n162, 1485n166
 Choi, H., 1463n75
 Chou, W. I., 1450
 Chow, G. C., 1272
 Chrétien, Jean, 1183
 Christiano, L., 664, 697, 702, 944, 973, 986, 987n51, 1135, 1191n48, 1197, 1368, 1370, 1372
 Chugh, S., 987n52
 Claessens, S., 1394n55, 1479n131
 Clarida, R., 726, 728, 765, 853, 947, 1058, 1063n8, 1068, 1176, 1196, 1198, 1201
 Clarida-Galí-Gertler New Keynesian model, 1347
 Clarke, G., 1463n75
 Classical model detection, 1112–1113
 Clearing balances, 1388n49
 Closed-economy model, 885
 Clouse, J. A., 1391n53, 1394n55
 Cobb-Douglas aggregator, 910, 917, 924
 Cobb-Douglas function, 1447
 Cochrane, J., 703, 942, 943, 947, 950, 964–965, 969–971, 1145
 Coefficients
 feedback rules changing, 966, 967n28
 on lagged interest rates/interest rates/output gap, 849f
 optimal/inflation rates/unemployment gap, 850f
 pass-through, 1445–1446, 1445n11
 reaction, 853
 regression, 960
 simple policy rules, 845t
 Coenen, G., 911n33
 Cogley, T., 695, 979, 1118–1119, 1121, 1189, 1191, 1228–1230, 1274n42
 Cohen, G. D., 1027
 Colacito, R., 1118, 1274n42
 Cole, H. L., 869, 916
 Coletti, D., 1287
 Colinearity, 1326–1328
 Collard, F., 976n36
 Collective action clauses (CACs), 1494
 Commercial Paper Funding Facility (CPFF), 1419
 Commitment
 optimal, 743–756, 767–776, 791, 796–797, 808, 811, 818–825
 value of, 733–737
 Commodities, 1467–1468
 Competitive devaluation, 867–868, 870, 909–915
 Competitive equilibrium, 676, 711
 from Friedman rule, 662
 in optimal inflation rate, 677
 primal form of, 665–666, 671
 sticky prices/money demand and, 695–696
 Complete-market model, 893
 Composition of capital inflows, 1497–1498
 Congress, 1024
 Connolly, M., 1450
 Constant long-run level, 812
 Constrained-optimal policy, 744, 768
 Consumer price index (CPI), 1239, 1460
 inflation rate overstated by, 658
 inflation rates of, 1164f
 IT and, 1468–1469
 measurement error in, 706
 monetary policy/theory and, 1460
 quality improvements in, 712–713
 United States inflation expectations of, 1194f
 Consumption
 Cobb-Douglas aggregator of, 910, 917, 924
 demand, 868
 Dixit-Stiglitz aggregator of, 670, 760–761
 growth rates of, 715
 inflation tax on, 674
 private, 787n73
 real exchange rates and, 902
 steady-state, 913n35
 tax/optimal monetary policy and, 984–990
 transaction costs of, 659–660
 United States growth in, 1123f
 Consumption Euler equation, 946
 Contagion, 1485
 Contingent rules, 1007–1008
 Contracting approach, 1016–1017
 Contractionary devaluation, 1492–1493
 Controls on capital outflows, 1477

- Controls on inflows, 1477
 - Cooper, R., 1449
 - Cooperative welfare-maximizing policies, 888–894, 910
 - Corana, A., 1231
 - Corners hypothesis, 1465
 - Correia, I., 664, 937–938, 947n12, 973–974, 977, 977n38, 980–981, 983, 984, 994
 - Corridor system, 1390n51
 - of New Zealand, 1426n91
 - quantitative easing in, 1426f
 - Corsetti, G., 865, 869, 908, 910, 911n31, 924, 1484n158, 1484n162
 - Cosimano, T., 1174n17
 - Cosine shocks, 1137
 - Costa, O. L. V., 1274n42
 - Cost-minimization, 707
 - Cost-push effects, 809
 - of disturbances, 788n74, 789, 807n92
 - markup shock and, 892n17
 - Cost-push shocks, 1081–1083, 1083–1084f
 - impulse responses to, 733t, 797f
 - in optimal equilibrium dynamics, 729–733
 - price level raised by, 796n82
 - Countries
 - anchor, 1036n41
 - currency unions joined by, 1039
 - Euro inflation rates, 1323–1324f
 - Euro output growth of, 1322f
 - hard currency pegs adopted by, 1329t
 - Inflation by, 1162–1163f
 - by Interest rates, 1162–1163f
 - IT adopted by, 1245t, 1314t
 - IT appraisal by, 1242n7
 - IT not adopted by, 1242
 - OECD/inflation/ IT of, 1246f, 1247n11, 1249n14
 - OECD/inflation of, 1249n14
 - OECD/long-term inflation
 - expectations of, 1060f
 - output by, 1162–1163f
 - policy regimes by, 1310t
 - standard deviations by, 1203–1204t
 - Covered interest differentials, 1474
 - CPFF. *See* Commercial Paper Funding Facility
 - CPI. *See* Consumer price index
 - Credit frictions, 822n109
 - Crisis, in emerging markets, 1481–1498
 - Crisis management, 1485–1488
 - early warning indicators in, 1495–1498
 - international financial institutions in, 1486–1488
 - private sector involvement in, 1486
 - Cross-border supply spillovers, 910n30
 - Cross-country demand imbalances, 915–928
 - Cross-country output gap stabilization, 866, 902
 - Cross-country output spillover, 892, 895
 - Cross-country regressions, 1479n132
 - Cross-country terms, 890
 - Crow, John, 1183
 - Crowe, C., 1018–1019, 1020n23, 1456
 - Cukierman, A., 1018, 1454, 1455–1456
 - Cultural rigidities, 1210
 - Cumby, R., 725, 949, 954, 965
 - Curdia, V., 821–822, 1293n67, 1425
 - Currency
 - crisis, 1482n153
 - mismatch, 1451–1453, 1452n33
 - Currency misalignments
 - cross-country demand imbalances and, 915–928
 - international demand imbalances and, 868–869
 - Currency unions, 1034–1041
 - countries joining, 1039
 - monetary policies from, 1038n42
 - multilateral, 1037–1039
 - optimal monetary policy for, 1037–1038
 - trade benefits of, 1040–1041
 - unilateral adoptions and, 1035–1036
 - unilateral/financial crisis and, 1036–1037
 - Currie, D., 1266, 1271
- ## D
- Dages, G., 1483n157
 - Dale, E., 815
 - Daniel, B., 950
 - Daniel, J., 1468n95
 - Dasgupta, A., 1484n158
 - Davig, T., 943, 949, 959–961, 967, 1273
 - Davis, J., 1468n95
 - Daylight overdrafts, 1383n44, 1388n50
 - de Carvalho Filho, I. E., 1242n7
 - de Gracia, F. P., 1480n135
 - de Grauwe, P., 1090
 - de Gregorio, J., 1459
 - de Haan, J., 1018
 - de Mello, L., 1285

- De Paoli, B., 913n34
- Dealing rate, 1398
- Debt intolerance, 1444n7
- Debt overhang, 1494n189
- Decision rules
 - frequency decompositions under, 1142f
 - in linear-quadratic problem, 1154–1155
 - robustness, 1118, 1121
 - worst-case model and, 1133f
- Decision theory, 1100–1101
- Decisionmakers
 - commitment of, 734n12
 - decision rules robustness for, 1118, 1121
 - model detection problem of, 1114
- Default swap, 1474
- Defaults, 1493–1495
- Deflation
 - avoiding, 811n94
 - discretionary policy resulting in, 752–753
 - Friedman rule association to, 663n2
 - Friedman rule with, 975–976
 - monetary policy rules and, 1070–1071
 - nominal interest rates and, 986n50
- Deflationary liquidity trap, 1289
- Del Negro, M., 697
- Dell’Ariccia, G., 1025
- Dellas, H., 976n36
- Demand, 1450n23
 - for credit/Japan’s decline in, 1215–1216
 - gap, 923
 - imbalances, 868–869
 - for money functions, 1178
 - in optimal monetary policy, 918–925
 - private consumption and, 787n73
 - procyclicality and, 1466–1467
 - for real balances, 659
 - supply relations, 726–727
- Demand-for-money functions, 1171
- Demiralp, S., 1370, 1370n32, 1381n41, 1401–1402, 1402n68, 1402n71, 1406, 1408, 1412
- Democratic deficit, 1020–1022
- Denes, M., 752n33
- Deposit facility rate, 1425n90
- Destination market, 896
- Detection error probability, 1116f
- Detection probabilities, 1114–1116
- Determinacy, 1066–1067, 1200n59, 1200t
- Deterministic path, 819
- Deterministic sequence, 966
- Devaluation
 - balance sheet effect in, 1451–1453
 - competitive, 867–868, 870, 909–915
 - contractionary, 1492–1493
 - contractionary influence of, 1449–1453
 - currency mismatch in, 1451–1453, 1452n33
 - expansionary, 1489–1492
 - of goods, 1445–1453
 - political costs of, 1449–1450
 - price pass-through and, 1450–1451
- Devereux, M., 865, 908, 924, 927, 1460n59
- Dewald, W. G., 832
- Diamond, D., 1483n157
- Diaz-Alejandro, C., 1450n23, 1484
- Diba, B., 725, 942, 948, 949, 954, 959, 961, 964–965
- Differences-in-differences, 1315
- Discount loss function, 728
- Discount window, 1391, 1391n53
- Discretion, 734n11
 - benefits, 1004–1005
 - in CBI, 1013–1014
 - equilibrium, 1005, 1008, 1266–1270
 - optimal monetary policy under, 1062–1063
 - optimization under, 1266–1270
 - simple rules v. loss of, 1010
- Discretionary optimization, 734n13, 737
- Discretionary policy, 733–738
 - deflation/negative output gap result of, 752–753
 - inflation path under, 736f
 - optimal, 831
 - optimal policy commitment compared with, 754f
 - of policy makers, 1015
- Discretionary policymakers, 911n31
- Distortions
 - in backward-looking model, 1121
 - financial factors causing, 1293n67
 - forward-looking, 1121
 - international prices influenced by, 911n31
 - large steady-state, 816n102
 - monopoly, 911n31, 976
 - real/nominal, 870–871
 - relative price, 806–807
 - steady state, 778n62
 - taxation and, 666–667, 699

Disturbances

- aggregate, 805
 - alternative parameterization of, 756n37
 - asymmetric, 807n92
 - asymmetric/sectoral heterogeneity and, 803–815
 - cost-push influence of, 788n74, 789, 807n92
 - economy influenced by, 810–811
 - monetary stabilization policy influenced by, 724–725
 - natural real wages shifted by, 817–818
 - nonzero cost-push influence of, 788–789
 - in output, 729
 - price level, 809
 - with targeting rules, 992–993
 - types of, 804
- Disyatat, P., 1397n61, 1425
- Divine coincidence, 890–891
- Dixit-Stiglitz aggregator, 670, 760–761
- Dixit-Stiglitz price index, 793–794
- do Val, J. B. R., 1274n42
- Dollar pricing, 877n9
- Domestic currency
- foreign demand for, 675–684, 680t, 717–720
 - primal form, 716–720
- Domestic/foreign goods, 917
- Dooley, M., 1473n99, 1484, 1484n162
- Dornbusch, R., 1448, 1453, 1482
- Dotsey, M., 792
- Dow, J., 1394n55
- Downward nominal rigidities, 657–658, 704–706
- Drazen, A., 1007, 1020, 1027, 1031n37, 1033, 1466, 1477n121
- Driffill, J., 1266
- DSGE models, 1196, 1198–1199, 1252n18
- Dupor, W., 787
- Dupuis, P., 1110
- Dutch disease, 1467–1468
- Dvorak, T., 1321
- Dynamic stochastic simulations, 833–835

E

- Early warning indicators, 1495–1498
- Easterly, W., 1453, 1454
- Eaton, J., 1494
- ECB. *See* European Central Bank
- Econometric defense for filtering, 1139–1140
- Economic Monetary Union (EMU), 1059

Economy

- advanced, 1244
 - disturbances influencing, 810–811
 - endowment, 1122
 - hard currency pegs in, 1332
 - infinite-lived households in, 760
 - integration of, 1319–1321, 1331–1332
 - Italian, 1042n49
 - market, 878
 - in Rational Partisan Theory, 1027–1028
 - multiple shocks in, 1009–1010
 - stagflation in, 1004
 - state-contingent evolution of, 774–775
 - structural transformations, 1204f
 - Swedish, 1043
 - uncertainty of, 1270–1274
 - variables in, 1134–1135, 1180
- Edison, H., 1458n54, 1478, 1479n131, 1479n132, 1480
- Edwards, S., 1449n21, 1450, 1454, 1480n135, 1480n144
- Efficient allocations, 880–884
- Efficient exchange rate, 864n2
- Efficient steady state, 777–782
- Efficient/inefficient shocks, 891
- Eggertsson, G. B., 742, 750, 752n33, 753–754, 842
- Ehrmann, M., 1206, 1315
- Eichenbaum, M., 697, 702, 907n24, 1197, 1368, 1445n12, 1446n14, 1483n157, 1484n162
- Eichengreen, B., 1042, 1285, 1444n8, 1465, 1481n146, 1484n159, 1495
- Eijffinger, S. C. W., 1274n43
- Ejerskov, S., 1374, 1399–1400, 1401
- Elastic price standard, 741
- Elasticity of demand, 1379–1383
- Ellison, M., 1274n43
- Ellsberg, D., 1102, 1103–1104
- Ellsberg paradox, 1099, 1102
- Ellsberg urn, 1116f
- Emerging economies
- inflation of, 1248t
 - IT adopted by, 1314t
 - IT/inflation and, 1246f
- Emerging markets, 1441–1442
- account reversals in, 1481
 - asset prices in, 1496–1498
 - balance of payments shock in, 1488–1493
 - bank credit of, 1497

- Emerging markets (*cont.*)
- capital flows of, 1472–1477
 - composition of capital inflows in, 1497–1498
 - contagion in, 1485
 - crisis in, 1481–1498
 - crisis management, 1485–1488
 - early warning indicators in, 1495–1498
 - international financial institutions in, 1486–1488
 - private sector involvement in, 1486
 - defaults in, 1493–1495
 - foreign exchange reserve holding in, 1496–1497
 - IT and, 1459n55
 - IT preconditions for, 1285–1286
 - models for, 1443–1445
 - speculative attacks models and, 1482–1484
 - sudden stops in, 1482, 1482n152
- EMS. *See* European Monetary System
- EMU. *See* Economic Monetary Union; European Monetary Union
- Endogenous capital accumulation, 820n106
- Endogenous variables
- asymptotic fluctuations in, 772
 - equilibrium dynamics of, 775
 - quadratic function of, 784–785
 - state-contingent evolution of, 758, 807
- Endowment economy, 1122
- Endowment process, 1127f, 1128f
- Engel, C., 865, 901, 908
- EONIA. *See* European overnight interest average
- Epstein, L. G., 1099, 1100, 1102, 1104, 1146
- Equalization of expected returns, 1474
- Equilibrium
- determination, 1264–1265
 - efficient exchange rate v., 864n2
 - of endogenous variables, 775
 - FTPL proposing, 949–951
 - predictions, 823–824
 - in Ramsey problem, 767–768
 - real interest rate, 839, 847
 - real wages, 816
 - resource loss in, 661
- Erceg, C., 815, 976n36
- Erhard, Ludwig, 1168
- ERM. *See* Exchange rate mechanism
- Error correction, 819
- E-stability, 1067–1068
- determinacy and, 1066–1067
 - extensions and, 1068–1070
 - in New Keynesian model, 1066–1070
- Estimated impulse response, 1381n41
- Euler equations, 874, 939
- Euro, 1041–1046, 1461n62
- capital markets and, 1320–1321
 - countries adopting, 1306
 - economic integration and, 1319–1321
 - Europe's transition to, 1204–1208
 - during financial crisis, 1043–1044
 - IT adoption and, 1306–1311, 1311t
 - monetary regimes, 1318–1325
 - output fluctuations, 1321–1322
 - political/monetary union and, 1045–1046
 - pre-financial crisis of, 1041–1043
 - price levels and, 1322–1325
 - robustness of, 1338t
 - trade determinants and, 1319–1321
- Euro Area
- central bank assets of, 1422f
 - central bank liabilities of, 1421f
 - excess reserve demand for, 1382t
 - excess reserves/short-term interest rates of, 1384f
 - inflation in, 1061f
 - inflation rates of, 1323–1324f
 - interest rates/inflation/output of, 1162–1163f
 - macroeconomic performance of, 1208, 1208f
 - output growth of, 1322f
 - policy interest rates of, 1418f
 - reserves/overnight interest rates of, 1376f
 - reserves/target interest rates in, 1378f
 - target/market interest rates of, 1393
- Europe
- elasticity of demand in, 1379–1383
 - EMU convergence process in, 1205–1206
 - Euro transition of, 1204–1208
 - European Monetary Union (EMU)
 - inflation persistence disappearance of, 1207
 - long-term inflation expectations of, 1206
 - structural changes under, 1206–1207
 - long-term inflation expectations of, 1206
 - political unification in, 1045–1046
 - price-level dispersion of, 1324–1325
- European Central Bank (ECB), 1003, 1059
- monetary policies decided by, 1167
 - policy tensions related to, 1042
 - price stability goal of, 1325–1326
- European Monetary System (EMS), 1179, 1306

- European Monetary Union (EMU), 1034, 1167
 - convergence process toward, 1205–1206
 - inflation persistence disappearance of, 1207
 - long-term inflation expectations of, 1206
 - structural changes under, 1206–1207
 - European overnight interest average (EONIA), 1373
 - European Parliament, 1045
 - Eurosystem
 - deposit facility rate of, 1425n90
 - high-frequency reserve demand in, 1373
 - liquidity effect for, 1372–1374
 - reserves/policy interest rates of, 1375–1379
 - Evans, C., 697, 702, 1197, 1368
 - Evans, G. W., 824, 825n114, 1057, 1065, 1066, 1067, 1069–1070, 1089–1090
 - Exchange rate
 - fixed, 913n34, 1461
 - flexible, 1465n86
 - floating, 1461–1462
 - import prices moving with, 907n24
 - international adjustment mechanism and, 869
 - in international monetary transmission, 886–887
 - New Keynesian analysis determination of, 877–879
 - nominal effective, 1166f
 - pass-through, 895, 908–909
 - relative, 1180
 - targeting, 1457
 - volatility in, 1443n5
 - Exchange rate mechanism (ERM)
 - collapse of, 1204–1205
 - crisis of, 1180–1181
 - failure of, 1167
 - Exchange rate regimes, 1461–1465
 - categorizing, 1463
 - corners hypothesis in, 1465
 - evaluating choices in, 1462–1464
 - fixed exchange rates, 1461
 - floating exchange rates, 1461–1462
 - Exogenous decline, 893f
 - Exogenous disturbance process, 761, 763, 764n48
 - Expansionary devaluation, 1489–1492
 - Expectations, of inflation, 1341
 - Expected utility theory, 1103
 - Explosive solutions, 947–948
 - Export price shocks, 1469
 - Extensions, 1068–1070
 - External balance curve, 1491–1492f
- F**
- Fackler, P., 1075
 - Fair, R. C., 832
 - Fang, W. S., 1249n14
 - Farmer, R. E. A., 1273
 - Farr, H., 1174n17
 - FDI. *See* Federal Deposit Insurance
 - FDIC. *See* Federal Deposit Insurance Corporation
 - Fear of Floating, 1461
 - Feasibility constraint, 665
 - Federal Deposit Insurance (FDI), 1025n31, 1495
 - Federal Deposit Insurance Corporation (FDIC), 1419
 - Federal funds rate, 1349f, 1350
 - Federal Open Market Committee (FOMC), 854, 1179, 1406n72
 - Federal Reserve, US. *See also* Central bank(s)
 - balances of, 1370n32
 - bond price support of, 947–948
 - Congress limiting, 1024
 - discount window of, 1391
 - monetary policy implemented by, 1370
 - new facilities created by, 1429n93
 - overstepping mandate by, 1023–1024
 - reserves requirements of, 1172–1173
 - target funds rate announced by, 1350
 - Taylor rule abandoned by, 1012
 - Feedback coefficients, 966
 - Feedback rules
 - coefficients changing with, 967n28
 - of fiscal/monetary policy, 955
 - in monetary policy, 947n12
 - Feinman, J., 1402, 1402n70, 1412n78
 - Feldstein, M., 1025, 1027, 1042
 - Feldstein-Horioka regression, 1473n99
 - Ferguson, T. S., 1103n7
 - Fernandez-Arias, E., 1481n146
 - Fernández-Villaverde, J., 1197–1198
 - Ferrero, G., 1072, 1090–1091
 - Filtering, 1130–1132
 - Finance
 - derivatives, 1415
 - domestic/global implications of, 916–918
 - flows of, 1372
 - integration of, 1472
 - markets for, 1011, 1366–1367
 - monetary policies and, 1025–1027

- Finance (*cont.*)
 openness in, 1477–1481, 1479n134
 repression in, 1443
- Financial autarky, 878
 flexible-price allocation under, 917
 Home output differential in, 924
 international price misalignments in, 925n40
 monetary policy trade-offs in, 919–922
 natural allocation under, 915–916
- Financial crisis, 1289n58
 asymmetry during, 1011–1012
 capital requirements during, 1218–1220
 CBI during, 1014–1015, 1023–1025, 1047
 central banking shut down of, 1220–1221
 central bank's price stability function in, 1167
 the euro during, 1043–1044
 financial stability during, 1216–1221
 framework for, 1292n65
 liquidity during, 1218
 lower inflation forecasts during, 1011
 monetary policy during, 1216–1221, 1288–1291
 monetary regimes and, 1312–1313
 rules v. discretion during, 1010–1013
 shock variance during, 1011
 2007–2009, 1414–1431
 unilateral currency unions, 1036–1037
- Financial frictions/imperfections, 864, 915–926
 imports currency price stability and, 866–867, 894–909
- Financial stability, 1026
 achieving, 1167–1168, 1291
 during financial crisis, 1216–1221
 with IT, 1287–1295
 monetary policies and, 1291–1293
- Financial stability committee (FSC), 1222
- Firms, 709–710
- First-order conditions (FOCs), 750
 for inflation, 898–899
 optimal equilibrium dynamic solutions and, 787
 of Ramsey problem, 691–693
- Fiscal policy, 698f
 active rules for, 960
 cash/credit goods model and, 980–984
 CCDL stable set in, 962f
 constraints on, 989–990
 inflation rate/money demand and, 664–667
 modeling frictions in, 983–984
 monetary policy v., 938–941, 1022
 money demand and, 716–717
 nominal anchor provided by, 944
 nominal variable indeterminacy of, 950–951
 non-Ricardian, 944–945
 as passive, 956, 959, 991
 price determination in, 937
 price stability in, 936–937
 procyclicality of, 1466
 Ramsey optimal, 984–985
 Ricardian, 948, 952
 as Ricardian, 765n51, 945
 Ricardian/non-Ricardian, 963–972
 specific feedback rules of, 955
 taxes available in, 994–995
- Fiscal theory of the price level (FTPL), 937
 equilibrium proposed by, 949–951
 monetarist arithmetic contrasted with, 942–943
 monetarist doctrine compatibility with, 952–953
 money supply rules and, 951–952
 multiple fiscal authorities in, 953–955
 nominal government assets in, 943
 non-Ricardian regimes and, 949–955
 PIR focus of, 943–944
- Fiscal variables, 1033
- Fischer, B., 1326–1328
- Fischer, S., 1183, 1453–1454, 1477
- Fisher, I., 830
- Fisher equation, 660–661
- Fitzgerald, T., 944, 1135, 1191n48
- Fixed exchange rates, 913n34, 1461, 1463
- Fixed rate mortgage, 1430f
- Fixed-horizon commitment, 741
- Fleming, W., 1108
- Flexible exchange rates, 1463, 1465n86
- Flexible inflation targeting criterion, 740–741, 822, 1009
- Flexible prices, 699, 917, 980–983, 986–987
- Flexible rules, 1003, 1011
- Flex-price allocation, 891
- Floating exchange rates, 1461–1462
- Flood, R., 1483
- Flow budget constraints, 660, 944
- Flow loss function, 897
- FOCs. *See* First-order conditions
- FOMC. *See* Federal Open Market Committee
- Forbes, K., 1477n123, 1485n166
- Forecast targeting, 1240n3

- of central banks, 738
 - in monetary policies, 1239–1240
 - in optimal monetary policy, 737–742
 - Forecast Taylor curve, 1260–1262, 1261f
 - Forecasting and Policy System (FPS), 1276
 - Foreclosures, 1416
 - Foreign exchange reserve holding, 1496–1497
 - Forward-looking distortions, 1121
 - Forward-looking variables, 1252, 1271
 - Fourier transforms, 1136–1137
 - Four-period market price, 1147f
 - FPS. *See* Forecasting and Policy System
 - Fractionalized systems, 1019–1020
 - Fraga, A., 1459
 - Fragoso, M. D., 1274n42
 - France, 1182f
 - Frankel, J., 1039–1040, 1043, 1319–1320, 1445n12, 1449, 1452n29, 1463n75, 1473n99, 1482n153
 - Fratzscher, M., 1206
 - FRB/US large-scale rational expectations model, 837, 838f, 848, 1275
 - Freedman, C., 1286
 - Frequency decompositions, 1142f
 - Frequency domain details, 1136–1140
 - Friberg, R., 908
 - Friction model, 1402
 - Frictionless asset markets, 864–865, 869–915
 - Friedman, M., 830, 864, 888, 937, 974, 976, 985, 994, 1101, 1103–1104, 1140–1144, 1174n17, 1177, 1304, 1325
 - Friedman rule
 - competitive equilibrium outcome from, 662
 - deflation association of, 663n2
 - with distortion taxation, 666–667, 669
 - failure of, 677–679
 - flexible prices in, 980, 986–987
 - inflation rate set in, 699–700
 - low deflation rate from, 975–976
 - lump-sum taxation with, 662–664
 - optimal deviation from, 679–681
 - optimal inflation and, 989n55
 - optimal rate of inflation of, 655–656
 - optimality of, 658–659
 - price-stability trade-off v., 695–701
 - as Ramsey optimal, 680–681
 - Ramsey policy satisfying, 975–976, 975n34
 - Ramsey problem solution of, 674
 - return to scale/imperfect competition/tax evasion from, 670t
 - sticky prices and, 897
 - untaxed income causing failure of, 667–675
 - Fries, G., 1174n17
 - FSC. *See* Financial stability committee
 - FTPL. *See* Fiscal theory of the price level
 - Fuerst, T., 941, 944
 - Fuhrer, J. C., 836, 846
 - Fuhrer model, 846
 - Fujiwara, I., 1213
 - Full regression results, 1337t
 - Furman, J., 1486, 1487n173
 - Future policy rate, 1387n47
- ## G
- Gagnon, J., 1431
 - Galasso, V., 1041
 - Gali, J., 726, 853, 864, 947, 1058, 1068, 1176, 1204
 - Galindo, 1452
 - Gambetti, L., 1195, 1204
 - Game of chicken, 942–943
 - Game theory, 1484n158
 - Gap-adjusted price level, 759
 - Garber, P., 1483
 - Garleanu, N., 1425
 - Garrett, B., 1174n17
 - Gascon, 1430
 - Gaspar, V., 824, 1071, 1074–1075, 1078, 1081–1082, 1091, 1178
 - Gatti, R., 1031–1032
 - Gaussian distributions, 1150
 - GDP-deflator inflation, 891
 - Geirberding, C., 1178
 - Gelos, R. G., 1485n166
 - Generalized method of moments (GMM), 1099
 - Geraats, P., 1307
 - Gerlach, S., 1290, 1459n55
 - Gerlach-Kristen, P., 1023
 - Germany, 1168, 1175–1176, 1182f
 - Geromel, J. C., 1274n42
 - Gersovitz, M., 1449n21, 1494
 - Gertler, M., 726, 853, 947, 1058, 1068, 1176, 1221, 1247, 1292n65, 1307, 1425
 - Ghosh, A., 1463

- Giannoni, M., 739–740, 775, 820, 823, 848, 851–852, 1196, 1257, 1260, 1263
- Giavazzi, F., 1035, 1473n99
- Gilboa, I., 1099, 1102–1103, 1103n7, 1104, 1106, 1109, 1110
- Gilboa–Scmeidler axioms, 1103
- Gillum, G., 1174n17
- Giovannini, A., 1035, 1043
- Girshick, M. A., 1103n7
- Glick, R., 1485n166
- Global equilibrium, 882
- Global imbalances, 1289
- Global inflation, 655t, 1174
- Global output gap, 899–900
- Global trade-offs, 890–891
- Global trading system, 1189
- GMM. *See* Generalized method of moments
- Goldberg, L., 1483n157
- Goldfajn, I., 1459, 1482
- Gonçalves, C. E. S., 1247, 1247n12, 1249, 1313, 1315–1316, 1339
- Gonzalez, M., 1033
- Goodfriend, M., 684, 976n36, 1196, 1424n88, 1425
- Goodhart, C. A. E., 1174n17, 1243
- Goodhart's Law, 1178
- Goods
- cash/credit for, 958f, 970f
 - devaluation contractionary influence on, 1449–1453
 - law of one price and, 1445–1446
 - NKPC and, 896
 - nontraded, 1446–1449
 - pass-through coefficients and, 1445–1446
 - price inflation of, 817
 - pricing/devaluation of, 1445–1453
 - services prices and, 1187
 - sticky export prices of, 1446
- Gopinath, G., 1465
- Gordon, D. B., 1004, 1006, 1367–1368
- Gorodnichenko, Y., 741
- Gourinchas, P. O., 1479n130
- Government
- bonds, 943n7, 1321n6
 - budget constraints of, 939
 - central banks policy coordination problem with, 948–949, 955–963
 - debt increase of, 1012
 - inflation/liabilities and, 955
 - interest rates/securities of, 1364
 - liabilities of, 944–945
 - lump-sum taxation of, 681–684, 690, 696
 - nominal debt of, 950
 - nominal value of liabilities of, 942–943
 - political business cycle and, 1466–1467
 - purchases, 789
 - Ricardian policies reactions of, 967
 - sequential budget constraint of, 661, 665, 676
- Gravity equations, 1319
- Great Depression, 749n29, 830
- Great Inflation, 1018, 1174, 1175–1177, 1191
- Great Moderation, 1313
- Great Moderation period, 1033, 1161, 1185
- accountability measures during, 854–855
 - global imbalances in, 1289
 - learning from, 852–855
 - macroeconomic uncertainty of, 1191
 - NICE years and, 1189–1204
 - in United States, 853
- Greece, 1044, 1318n5, 1321n6
- Greenspan, A., 1042, 1185–1186, 1191, 1195, 1197, 1210, 1290
- Grilli, V., 1017, 1039
- Gropp, R., 1043
- Growth rate
- of monetary aggregates, 1353
 - shift, 1127n28
- Guerrón-Quintana, P., 1197
- Guidotti, P., 664, 976, 1482n152, 1498n210
- Guidotti rule, 1496
- Gulde, A. M., 1463
- Gürkaynak, R. S., 1206, 1206n66, 1247, 1318
- Guse, E., 1070
- Gust, C., 1450
- Guthrie, G., 1390n51
- Gutiérrez, E., 1456
- ## H
- Haan, J., 1456
- Haldane, A., 837
- Hall, R. E., 741
- Hamilton, J. D., 1273, 1370–1371, 1372n34, 1401
- Handbook in Economics* (McCallum), 832
- Handbook in Macroeconomics* (Taylor, J. B., Woodford), 832

- Handbook of Macroeconomics* (Bordo, Schwartz), 1161
- Hanes, C., 1369n29, 1391n53
- Hansen, L. P., 1100, 1101, 1104, 1108–1110, 1114, 1119, 1122–1123, 1125–1126, 1126f, 1127f, 1128f, 1131, 1132n33, 1136, 1138–1139, 1144, 1145–1146, 1147, 1147f, 1148
- Hansen, S., 1023
- Hard currency pegs, 1329n7, 1333t
 countries adopting, 1329t
 economic integration, 1331–1332
 economic performance of, 1332
 inflation control with, 1330–1331
 monetary regimes with, 1328–1332
- Harrison, R., 1090
- Harvey, C., 1475, 1479n131, 1479n134
- Hasako, H., 1388n48
- Hausmann, R., 1443n5, 1452n30, 1483n157
- Hayashi, F., 1373, 1401
- Hellwig, C., 985n49
- Helpman, E., 869, 916
- Henderson, D., 834, 976n36
- Henry, P., 1478, 1479n131
- Hibbs, D. A., 1027
- High pass-through, 1445n11
- Higher target inflation rate, 843
- High-frequency noise filter, 1191n48
- Hilton, S., 1388n48, 1401
- Ho, C., 1388
- Holmsen, A., 1284
- Homes
 appreciation, 905
 consumption demand, 868
 depreciation, 897
 markups, 893f, 903f, 914f
 monetary policy, 895
 output differential, 924
 preference shock, 926f
 prices of, 1220, 1416
 productivity, 904f
 productivity shock, 912f
 terms of trade, 887, 892
- Honkapohja, S., 824, 825n114, 1057, 1065, 1066, 1067, 1069–1070, 1089–1090
- Hooper, P., 831
- Households
 budget constraints of, 685, 938–939
 flow budget constraint of, 660, 944
 labor income of, 684–685
 labor supply of, 708
 monetary stabilization policy and, 759–760
 New Keynesian analysis/decisions/preferences of, 871–873
 optimal policy commitment and, 821
 quantity of goods of, 707–708
- Hoxha, I., 1479n130
- Hrung, W. B., 1401
- Hu, Y., 1315
- Hungary, 1044
- Huntley, J., 1450
- Husain, A., 1464n76
- Hyperinflation, 947, 1453–1454
 model, 1070
 monetary policy rules and, 1070–1071
- ## I
- IMF. *See* International Monetary Fund
- Imperfect competition, 670–672, 670t
- Imperfect information, 756–759
- Implementability constraint, 666, 671, 673, 982n43
- Imports
 financial frictions/imperfections and, 907–908
 local currency price stability of, 866–867, 894–909
 price shocks, 1469–1470
 prices, 907n24
- Impulse responses
 to cost-push shocks, 733t, 797f
 to economic variables, 1134–1135
 Fourier transforms of, 1136
 of output gap, 808f
 to shocks, 1143f
 to target criterion variables, 813f
 of transitory component, 1126f
- Income tax, 665
- Indexation scheme, 693–695, 697–698
- India, 1189
- Individual flow budget constraint, 874, 878–879
- Inefficient shocks, 891
- Infinite-horizon control problem, 1105
- Infinite-lived households, 760, 821
- Inflation. *See also* Zero inflation
 average, 1187t
 bias, 736–737, 1004
 in Bulgaria, 1330f

Inflation. *See also* Zero inflation (*cont.*)

CBI and, 1455–1456
 CBI's negative relationship with, 1020n23
 central bank suppressing, 940
 central bank/output expectations and, 725–726
 central bank's objectives of, 654–655
 central bank's target path of, 940
 consumption tax and, 674
 by country, 1162–1163f
 direct costs of, 994
 under discretionary policy/Ramsey policy/
 timeless perspective, 736f
 of emerging economies, 1248t
 in Euro Area, 1061f
 Europe's expectations of, 1206
 expectations, 1247–1248, 1341
 financial crisis forecasts of, 1011
 first-order conditions for, 898–899
 forecast targeting, 1240n3
 Friedman rule's optimal rate of, 655–656
 global, 655t, 1174
 government liabilities and, 955
 hard currency pegs controlling, 1330–1331
 hyperinflation and, 1453–1454
 increase/real interest rate increase, 945–946
 IT macroeconomics and, 1246–1247
 lagged, 1086f
 lagged output gap influencing, 1073n23
 long-range target of, 746n25
 loss function and, 897–898
 monetary policy influence on, 771–772
 monetary policy response of, 1357–1358
 in monetary policy rules, 837
 monetary policy rules stabilizing, 834–835
 monetary policy targeting, 802
 monetary targetry explanations of, 1174–1177
 monopoly profits taxed by, 985n47
 negative, 657, 664, 667
 NICE years expectations of, 1194–1195
 nonzero trend, 1198–1199
 nutter, 1239n1, 1354n5
 in OECD countries, 1060f
 OECD countries and, 1246f
 optimal long-run average rate of, 732
 optimal response to, 840f
 optimal state-contingent path of, 732
 output gap/optimal monetary policy and,
 821–822

output gap/optimal simple rules expectations of,
 837–838

positive, 657
 pragmatic monetarism reducing, 1168
 private-sector expectations of, 1059–1061
 quasi-difference of, 1080f
 Ramsey policy's volatile, 995
 under rational expectations, 1061–1065
 reaction coefficient to, 853
 sectoral, 885
 shock impulse responses of, 1143f
 short-run aggregate supply and, 764–765
 stabilization program for, 1454–1455
 standard deviations of, 1187t
 steady growth and, 1167
 sticky prices/variability of, 995
 structural, 705
 tradeoffs, 918–925
 United States positive, 656

Inflation persistence, 1340

structural, 792–798
 adaptive learning with, 1073–1089
 EMU disappearance of, 1207
 in IT, 1316–1317
 output gap and, 1079f, 1081f, 1083–1084f
 output gap stabilization and, 1088f
 policy function output gap/lagged inflation and,
 1086f
 price stickiness and, 1087–1088

Inflation rate. *See also* Optimal inflation rate

aggregate-supply relations and, 794
 central banks stabilization policies and,
 701–702
 consumer price index overstating, 658
 of CPI, 1164f
 cross-sectional standard deviations of, 1205f
 of Euro countries, 1323–1324f
 Friedman rule with, 699–700
 higher target, 843
 impulse responses in, 733
 money demand/fiscal policy and, 664–667
 in neo-Keynesian model, 698
 optimal, 656–657
 optimal coefficients on, 850f
 positive, 796–797
 price index and, 811n95
 Ramsey optimality and, 665, 703
 two sectoral, 809–810

- Inflation Targeting (IT), 1026, 1061, 1183–1184, 1240n3
- adaptive learning and, 1071–1073
 - advanced economies adopting, 1244
 - of central bank, 975n33
 - countries adopting, 1245t
 - countries appraisal of, 1242n7
 - countries not adopting, 1242
 - CPI and, 1468–1469
 - emerging countries adopting, 1314t
 - emerging markets and, 1459n55
 - Euro adoption and, 1306–1313, 1311t
 - full regression results in, 1337t
 - future
 - financial stability in, 1287–1295
 - flexible IT, 1293–1295
 - price-level targeting, 1286–1287
 - history of, 1243–1244
 - import price shocks and, 1469–1470
 - inflation persistence in, 1316–1317
 - influence summary of, 1249–1250
 - international monetary system and, 1242n6
 - low inflation/steady growth of, 1167
 - macroeconomics, 1244–1250
 - history, 1242–1250
 - inflation and, 1246–1247
 - inflation expectations in, 1247–1248
 - output in, 1248–1249
 - monetary policy with, 1003, 1457–1459
 - money-growth targeting alternative to, 1242–1243
 - New Zealand adopting, 1183–1184, 1238–1239, 1243–1244, 1276–1277
 - numerical, 1239
 - OECD countries inflation and, 1246f, 1247n11, 1249n14
 - output performance from, 1249f
 - policy-rate path in, 1251n16
 - practice, 1275–1286
 - developments in, 1276–1278
 - emerging-market economies preconditions, 1285–1286
 - interest-rate path, 1279
 - Norges Bank, 1281–1284
 - Riksbank, 1280–1281
 - private sector knowing, 1073n22
 - research on, 1338
 - robustness of, 1338t
 - theory, 1250–1275
 - commitment to, 1269–1270
 - discretion equilibrium in, 1266–1270
 - equilibrium determination in, 1264–1265
 - forecast Taylor curve in, 1260–1262
 - judgment in, 1275
 - linear-quadratic model as, 1252–1257
 - optimal policy choice in, 1258–1260
 - optimal policy projections in, 1262–1263
 - optimization under discretion in, 1266–1270
 - projection model in, 1257–1258, 1268–1269
 - targeting rules in, 1263–1264
 - transparency/accountability of, 1240–1241
 - uncertainty, 1270–1274
 - state of economy and, 1270–1274
 - transmission mechanism and, 1272–1274
- Information set, 758n40
- Initial conditions control, 1315
- Instrument rules, 1058n5, 1064–1065
- Instrument v. goal independence, 1016
- Instrumental variables, 1315–1316
- Interbank market rates, 1375n38
- Interest and Prices* (Woodford), 833, 1347
- Interest rates. *See also* Lagged interest rates;
- Nominal interest rates; Policy interest rates;
 - Real interest rates; Target interest rates
- anticipation effect from, 1387–1388
 - of Bank of England, 1278n49
 - of Bank of England/Riksbank, 1263n31
 - bank reserve supply and, 1375
 - bank reserves and, 1384f, 1388–1392
 - bank reserves relationship with, 1349, 1434
 - in cash/credit goods model, 988f, 990f
 - central bank varying, 1427
 - of central banks, 835–836, 853–854, 893–894
 - central banks below normal, 1358
 - central banks expectations on, 1394n57
 - central banks setting, 1347–1351, 1366–1367, 1383–1399
 - coefficients on, 849f
 - by country, 1162–1163f
 - deviation from, 1412n78
 - government securities and, 1364
 - market, 1360–1367
 - natural, 729, 839
 - near-zero, 1382n43
 - negative nominal, 841n4
 - nominal, 668n3, 727, 742

- Interest rates. *See also* Lagged interest rates;
 Nominal interest rates; Policy interest rates;
 Real interest rates; Target interest rates
 (*cont.*)
 overnight, 1376f
 pegging of, 943–944, 1154, 1355
 short movements of, 1294n68
 short-term, 834
 smoothing, 837–838
 supply-induced, 1385f
 Taylor principle violated by, 947–948, 1068
 ZLB on, 841–843
- Interest-rate lower bound, 748–756
- Interest-rate path, 1279
- Interest-rate reaction function, 822
- Intermediate goods, 670–671
- Internal balance curve, 1490f, 1493f
- Internal ratings based approaches (IRB), 1219
- International adjustment mechanism, 869
- International asset markets, 877–879
- International borrowing/lending, 925–927
- International demand imbalances, 868–869
- International financial institutions, 1486–1488
- International Monetary Fund (IMF), 1044, 1248, 1291
- International monetary system, 1242n6
- International monetary transmission, 886–887
- International prices
 distortions influencing, 911n31
 manipulation of, 910–911
 misalignments in, 925n40
- International relative prices, 866–867, 869–915
- International transmissions
 of home markups exogenous decline, 893f
 optimal monetary policy, 894
 home markup exogenous decline in, 903f
 home productivity/preference shocks in, 904f
 of shocks, 882
- Intertemporal effects, 1091–1092
- Intertemporal implementability conditions, 984n46
- Intertemporal trade-offs, 1084–1087
- Intra-maintenance period reserve demand/supply, 1403–1404t
- Intratemporal effects, 1091–1092
- Intra-temporal trade-offs, 1082–1084
- Inverted Wishart distribution, 1229
- Investment prices, 702–703
- IRB. *See* Internal ratings based approaches
- IS curve, 1064, 1359n10
- IS-LM model, 1347, 1355
- Issing, O., 1175, 1178
- IT. *See* Inflation Targeting
- Italian economy, 1042n49
- Italy, 1182f
- Ito, H., 1479n134
- Izquierdo, A., 1452n29, 1482
- ## J
- Jacobson, D. H., 1103n7, 1145, 1154
- Jácome, L., 1456
- Jacquier, E., 1230
- Jagannathan, R., 1145
- James, M. R., 1110
- Jansen, D., 1174n17
- Japan, 1209–1216
 balance sheet recession in, 1215–1216
 bank lending reluctance in, 1214–1215
 central bank assets of, 1422f
 central bank liabilities of, 1421f
 demand for credit decline in, 1215–1216
 elasticity of demand in, 1379–1383
 excess reserve demand for, 1383t
 excess reserves/short-term interest rates of, 1384f
 interest rates/inflation/output of, 1162–1163f
 liquidity effect for, 1372–1374
 macroeconomic data of, 1186t, 1209f
 monetary policies role in, 1210–1214
 near-zero interest rates of, 1382n43
 policy interest rates of, 1418f
 reserves/overnight interest rates of, 1376f
 reserves/policy interest rates of, 1375–1379
 reserves/target interest rates in, 1378f
 structural/cultural rigidities of, 1210
 target/market interest rates of, 1393
- Jappelli, T., 1473n99
- Jeanne, O., 1479n130, 1496n200, 1497n201
- JEC. *See* Joint Economic Committee
- Jensen, C., 746–748
- Jinjarak, Y., 1458n54
- Jinushi, T., 1372
- Johnson, D. R., 1317
- Johnson, H. G., 832, 1485
- Johnson, S., 1019, 1477n124
- Joint Economic Committee (JEC), 1177
- Jordá, O., 1402n71
- Judd, J., 853

Judd, K. L., 1075

Judgment

in monetary policies, 1275

by Riksbank, 1282f

Judson, R. A., 675

Juster, T., 1145

Justiniano, A., 1197

K

Kacperczyk, M., 1429

Kahn, M. S., 1287

Kalemli-Ozcan, S., 1479n130

Kalman filter, 1128–1130, 1230, 1258

Kamin, S., 1450

Kaminsky, G., 1480n137, 1483n157, 1485n166

Kaplan, E., 1477n124

Kapur, 1463n75

Karadi, P., 1425

Karantounias, A. G., 1148

Kasa, K., 1135, 1148

Kashyap, A., 1043

Kehoe, P., 664, 947, 973, 986, 987n51, 993

Keister, T., 1424n88

Keynes, J. M., 1100

Khan, A., 695, 789, 789n75, 976n36

Khan, C., 955

Khemani, S., 1033

Kiley, M., 1059n6

Kilian, L., 1177

Kim, C. J., 1230, 1273

Kim, J., 705–706

Kim, S., 957

Kimball, M., 1144, 1145

Kimbrough, K., 664, 1430

Kimura, T., 1213

King, M., 940, 1090, 1216, 1239n1, 1240n3, 1354n5

King, R., 684, 690n7, 695, 789, 976n36

Kinney, D., 1483n157

Kisselev, K., 1451

Kitamura, T., 801

Kiyotaki, N., 1292n65

Klein, M., 1478, 1480

Klein, P., 1255

Klenow, P., 697

Kletzer, K., 1484n162

Klomp, J. G., 1018

Kneebone, R. D., 1033

Knight, F. H., 1100–1102, 1102n6

Knightian uncertainty, 1147f

Kobayashi, H., 1213

Kocherlakota, N., 949, 952–953

Kohn, D., 855, 1290, 1294n69

Kohn, R. P., 1229

Kollmann, R., 927n46

Koo, R. C., 1216

Kooi, W. J., 1456

Kopecky, K., 1174n17

Kose, M. A., 1478, 1479n134

Kostyshyna, O., 1090

Kreps, D. M., 1124

Krishnamurthy, A., 1458n54

Krugman, P., 1211, 1451, 1465n86, 1483, 1484, 1484n159, 1494n189

Kryvtsov, O., 697

Kuester, K., 847

Kydland, F., 1004

L

Labor contracts, 1042

Labor income

of households, 684–685

price stickiness influencing, 987–988

taxes/Ramsey optimality and, 669–670

Labor markets, 820, 979

Labor supply, 682

of households, 708

income tax rate and, 665

Lagged inflation, 1086f

Lagged interest rates

coefficients on, 849f

optimal response to, 840f

Lagged output gap, 1073n23

Lagrange multiplier, 730

backward-looking constraint associated with, 768

bounded processes constructed for, 774

implementability constraint in, 982n43

in Ramsey problem, 668

second-order conditions and, 770n55

Taylor expansions of, 786–787

unique bounded evolution for, 739n17

Lags, long variable, 1140–1143

Landström, M., 1456

Lane, P., 1320, 1321n6, 1322, 1460n59

Lane, T., 1174n17

- Large steady-state distortions, 784–786, 816n102
- Laséen, S., 1252n18, 1255n25, 1260, 1269
- Law of one price (LOOP), 877, 894–897, 1445–1446
- Lawson, N., 1181
- Laxton, D., 1249, 1285, 1459n55
- LCP. *See* Local currency pricing
- Leaning against the wind, 1293
- Learning, 1117–1132
 - adaptive control v. robust control and, 1132
 - adaptive models in, 1123–1125
 - Bayesian models and, 1117–1119
 - belief changes and, 1122–1123
 - Bellman equation and, 1121
 - Kalman filter and, 1128–1130
 - ordinary filtering/control and, 1130
 - robust filtering/control and, 1130–1132
 - specification doubts and, 1119
 - state prediction and, 1125–1128
 - two risk-sensitivity operators and, 1119–1121
- Learning algorithms, 1067n16
- Lee, C. S., 1249n14
- Lee, J., 1496n200
- Leeper, E., 942, 943, 949, 955–961, 961n26, 967, 1273, 1367–1368
- Lehman Brothers, 1328, 1416
- Leland, H., 1144
- Levchenko, A., 1473n100
- Levin, A., 695, 697, 815, 831, 837, 844, 846–847, 851, 855, 976, 979, 984n46, 1121, 1176, 1316–1317, 1318, 1340–1341, 1340n9
- Levine, P., 1266, 1271
- Levine, R., 1480
- Levy-Yeyati, E., 1463
- Liabilities, government, 944–945
- Lifetime utility function, 708
- Limited Information Maximum Likelihood strategies, 1020n23
- Lin, S., 1247, 1316, 1339
- Lindé, J., 697, 702, 1252n18, 1255n25, 1269
- Lindsey, D., 1174n17
- Linear filters, 1274n42
- Linear quadratic model, 780, 817, 1272n40
- Linear rational expectations model, 836
- Linear target criterion, 747–748
- Linear-quadratic model, 1252–1257
- Linear-quadratic problem, 1105–1106
 - decision rules in, 1154–1155
 - in optimal monetary policy, 726–729
- Linear-quadratic-Gaussian problem, 1109
- Linear-quadratic approximation, 776, 782, 784, 787, 794, 806, 826
- Linearization
 - of optimal dynamics, 772–774
 - of structural equations, 760, 764–772, 779, 793, 805
- Liquidity
 - of bonds, 961–963
 - during financial crisis, 1218
 - preference function, 667
- Liquidity effect, 1352, 1369n31
 - absence of, 1401
 - aggregate time series used by, 1369–1370
 - bank reserves and, 1348–1349, 1370–1371
 - disappearance of, 1370
 - for Eurosystem, 1372–1374
 - for Japan, 1372–1374
 - nonborrowed reserves generating, 1368, 1368n26
 - in United States, 1367–1372
- Loayza, 1445n12
- Loayza, N., 1497
- Local currency pricing (LCP), 866
 - Home appreciation and, 905
 - imports/stability of, 894–909
 - monetary policy/endogeneity of, 908–909
 - optimal monetary policy targeting rules in, 866–867
 - optimal monetary policy under, 903
 - price setting under, 877
- Local-currency price stability, 866–867, 894–909
- Log-linear aggregate supply, 779–780, 786
- Log-linear equations, 772
- Lohmann, S., 1015
- Loisel, O., 947n12
- London School of Economics (LSE), 1220
- Long Term Capital Management (LTCM), 1185
- Long-range inflation target, 746n25
- Long-run expectations, 1317–1318
- Long-run expected values, 809–810
- Long-run inflation expectations, 1176f
- Long-term forward rates, 1206n66
- LOOP. *See* Law of one price
- Lopez-Salido, D., 949, 959, 984n46
- Lora, 1449n21

- Loss function
 CBI minimizing, 1046–1047
 of central banks, 835n1, 1006
 flow, 897
 inflation and, 897–898
 quadratic, 740n19, 786, 792, 816n102
 welfare-based, 776–786
- Lower bound, 742, 748–756
- Loyo, E., 949
- LSE. *See* London School of Economics
- LTCM. *See* Long Term Capital Management
- Luangaram, P., 1458n54
- Lubik, T., 947, 1196, 1199–1200, 1200n59, 1231
- Lucas, R. E., 975, 975n34, 993, 994, 1099, 1148, 1195
- Lucas, R. E., Jr., 831
- Lucas-supply curve, 1274n43
- Lump-sum taxation, 977
 cutting, 945
 in Fiscal policy, 991
 of government, 662–664, 681–684, 690, 696
 seignorage losses covered by, 663–664
- Lundblad, C., 1479n134
- M**
- Macro volatility, 905–907
- Macroeconomic interdependence
 under asset market imperfections, 915–928
 baseline monetary model, 870–886
 budget constraints in, 874
 exchange rate determination in, 877–879
 household decisions in, 871–873
 international asset markets in, 877–879
 open-economy Phillips curve in, 884–886
 price-setting decisions if, 874–877
 in global equilibrium, 882
- Macroeconomic model
 new type of, 831
 robustness in, 1133–1134
- Macroeconomics
 CBI and, 1017–1019
 of China, 1186t
 of Euro Area, 1208, 1208f
 of France, 1182f
 of Germany, 1175–1176, 1182f
 global imbalances in, 1289
 inflation expectations in, 1247–1248
 IT influence on, 1242–1250
 of Italy, 1182f
 of Japan, 1186t, 1209f
 output in, 1248–1249
 persistence, 1076–1081
 rational expectations/adaptive learning and, 1076t
 standard deviations by country, 1203–1204t
 standard deviations/uncertainty of, 1192–1193f
 uncertainty, 1191
 of United Kingdom, 1171f, 1182f
 of United States, 1170f, 1175–1176, 1175t
 of West Germany, 1169f
- MA/FP. *See* Monetary active/fiscal passive
- Magud, N., 1477
- Main refinancing operation (MRO), 1374
- Maintenance period
 bank reserves demand within, 1404–1409
 bank reserves supply within, 1409–1413
 reserve demand/supply within, 1399–1413
- Maisel, S. J., 854
- Mankiw, G., 798, 801
- Mann, C., 831
- Marcet, A., 1070, 1273
- Marchesi, M., 1231
- Marchioni, D., 1430
- Marder, A. N., 1318
- Marimon, R., 1273
- Marion, N., 1496n200
- Market economy, 878–879
- Market interest rates, 1360–1367, 1381n41
- Market prices, 1147–1148, 1473–1475
- Markov jump-linear-quadratic (MJLQ), 1272–1274
- Markov-Chain Monte Carlo algorithm, 1227–1228, 1231
- Markov-perfect equilibrium, 733f, 750–751, 753
- Markup
 factor, 762n46
 level of, 672
 shock/cost-push effects and, 892n17
- Marques, R. P., 1274n42
- Marshall-Lerner conditions, 1446
- Martin, P., 1480n137
- Martin, W. M., 1177
- Martinez, P., 1483n157
- Martini, C., 1231

- Marumo, K., 1213
- Masaki, K., 1388n48
- Mascaro, A., 1174n17
- Masciandaro, D., 1017, 1456
- Masson, P., 1007, 1285, 1459n55, 1485
- Mathematical foundations, 1098–1099
- Mathieson, D., 1473n99
- Mauro, 1025, 1454, 1462
- McAndrews, J., 1424n88
- McCallum, B., 746–748, 831–832, 839–840, 941, 947, 949, 952–953, 1174n17, 1210, 1356–1358, 1358, 1359n10
- McConnell, M., 1204
- MCI. *See* Monetary Conditions Index
- Mckenzie, K. J., 1033
- McKibbin, W., 834, 1460
- McKinnon, R. I., 1177, 1210
- McMahon, M. F., 1023
- McNees, S. K., 832
- Meade, E., 1018–1019, 1020n23, 1456
- Mean squared gaps, 1281n52
- Measurement issues, 706, 838–841
- Medium-term refinancing operations (MRO), 1401
- Mejia, L. F., 1452n29, 1482
- Meltzer, A., 1174n17, 1210, 1357
- Mendoza, E., 1458n54, 1484n162, 1485n167
- Meulendyke, A. M., 1369n29, 1388n48, 1391m53
- Meyer, L., 832, 854
- Microfoundations, 760–765
- Mihov, I., 1368
- Milani, F., 1061
- Milesi-Ferretti, G. M., 1320
- Miller, B. L., 1144
- Miller, G., 1456
- Miller, M., 1458n54
- Miller, S. M., 1249n14
- Minella, A., 1459
- Minimum state variable (MSV), 1065
- Minsky, H., 1216, 1217
- Miranda, M. J., 1075
- Mirman, L., 1127
- Miron, J. A., 1017–1018
- Mishkin, F., 831, 1059n6, 1247, 1315–1316
- Miteza, I., 1450
- Mitra, K., 1057, 1065, 1068–1069
- Mitton, T., 1477n124
- Miyanoya, A., 1388
- MJLQ. *See* Markov jump-linear-quadratic
- Model detection problem, 1114
- Model misspecification, 1106–1107
- decision theory and, 1100–1101
- with filtering, 1135–1136
- types of, 1107–1108
- Models. *See also* Bayesian model; Cash/credit goods model; Macroeconomic model; New Keynesian model; Open-economy model
- adaptive, 1123–1125
- anticipated utility, 1124
- backward-looking, 768, 1115, 1121, 1253n21
- Ball's, 1141–1143
- baseline, 1075–1076
- baseline calibration of, 1075–1076
- baseline closed-economy, 864
- baseline monetary, 870–886
- Bayesian, 1117–1119
- Brock-Sidrauski, 976
- Calvo, 815
- Calvo-Woodford-Yun, 688
- Calvo-Yun, 792
- cash/credit goods, 973–974, 994
- cash-in-advance, 938–939
- Clarida-Gali-Gertler New Keynesian, 1347
- closed-economy, 885
- complete-market, 893
- detection, 1112–1113
- detection problem and, 1114
- DSGE, 1196, 1198–1199
- for emerging markets, 1443–1445
- FRB/US, 837, 838f, 848, 1275
- frictions in, 983–984, 1402
- Fuhrer, 846
- Hicks-Keynes IS-LM, 1347
- hyperinflation, 1070
- IS-LM, 1355
- linear quadratic, 1272n40
- linear rational expectations, 836
- linear-quadratic, 1252–1257
- Mundell-Fleming, 1475
- neo-Keynesian, 698
- nonlinear DSGE, 1252n18
- nonlinear structural, 744–745
- partisan, 1032
- perfect knowledge, 847
- policy rules evaluated by, 833–844
- projection, 1257–1258, 1268–1269

- Quarterly Projection, 1276
- robust control, 1117
- Rudebusch-Svensson, 846
- Salter-Swan, 1448
- simple, 707–709
- speculative attacks, 1482–1484
- sticky information, 798–802
- stochastic growth, 1127
- structural relations in, 729
- three-asset demand and supply, 1392–1393
- two-sector, 803, 803n88
- two-state, 752n33
- vector autoregression, 1367, 1368–1370
- worst-case, 1133f
- Modified optimal policy, 852
- Mody, A., 1444n8, 1464n76
- Molnar, K., 1071, 1073
- Monetarist arithmetic
 - FTPL compatibility with, 952–953
 - FTPL contrasted with, 942–943
 - as noncooperative game, 938
 - price stability/instability through, 939–941
- Monetary active/fiscal passive (MA/FP), 959–961
- Monetary Conditions Index (MCI), 1277n48
- Monetary non-neutrality, 656, 715
- Monetary policy rules
 - accountability measures in, 854–855
 - under adaptive learning, 1065–1071
 - dynamic stochastic simulations of, 833–835
 - hyperinflation/deflation and, 1070–1071
 - inflation measures in, 837
 - interest rates ZLB in, 841–843
 - optimal simple rules of, 835–838
 - output gap measurement issues for, 838–841
 - robustness of, 824, 844–848
 - stabilizing inflation/output gap in, 834–835
 - variables in, 843–844
 - ZLB's implications on, 842
- Monetary policy/theory. *See also* Optimal monetary policy
 - as active, 956, 959
 - adaptive learning for, 1057
 - of anchor country, 1036n41
 - Bank of England committee for, 1023
 - bank reserve demand and, 1365f
 - central bank's commitment to, 733–737
 - by committee, 1022–1023
 - complete price stability in, 806–807
 - CPI and, 1460
 - from currency union, 1038n42
 - ECB deciding, 1167
 - exchange rate targeting in, 1457
 - Federal Reserve implementing, 1370
 - feedback rules in, 947n12
 - during financial crisis, 1216–1221, 1288–1291
 - financial regulation and, 1025–1027
 - financial stability and, 1291–1293
 - fiscal policy v., 938–941, 1022
 - flexible rules needed in, 1003, 1011
 - forecast targeting in, 1239–1240
 - future conduct of, 1423–1426
 - Great Moderation rules/shocks to, 1201
 - historical background of, 830–832
 - implementation of, 1360f
 - important limitations of, 838
 - inflation responding to, 1357–1358
 - inflation targeted in, 802, 1003
 - inflation/output restrictions from, 771–772
 - intra/intertemporal effects in, 1091–1092
 - IS curve and, 1359n10
 - IT in, 1457–1459
 - Japan's, 1210–1214
 - judgment in, 1275
 - LCP endogeneity and, 908–909
 - models evaluating rules of, 832–844
 - money growth control in, 1353–1354
 - nominal anchor from, 956
 - nominal targets for, 1456–1460
 - observability of, 1007
 - open-economy model analysis of, 862–863
 - optimal inflation rate and, 663–664
 - optimal taxation framework in, 994
 - outcomes achievable through, 777–778
 - paid reserve rate in, 1363
 - policy frontier in, 837
 - price-level rule for, 742, 789
 - procyclicality of, 1467
 - Ramsey optimality in, 661–662, 984–985
 - recurring shocks to, 1185
 - reputation building in, 1004–1007
 - rules v. discretion in, 1010–1013
 - shocks in, 1290n62
 - specific feedback rules of, 955
 - stabilization role of, 1008
 - sticky prices/wages and, 815–818
 - Taylor rules posterior distributions in, 1202f

- Monetary policy/theory. *See also* Optimal monetary policy (*cont.*)
- time inconsistency in, 715–716
 - trade-offs, 919–922
 - traditional view of, 1361
 - uncertainty of, 1270–1274
- Monetary regimes
- background of, 1306–1307
 - data of, 1309–1311
 - Euro, 1318–1325
 - capital markets and, 1320–1321
 - economic integration and, 1319–1321
 - output fluctuations, 1321–1322
 - price levels and, 1322–1325
 - financial crisis and, 1312–1313
 - hard currency pegs, 1328–1332
 - countries adopting, 1329t
 - economic integration, 1331–1332
 - economic performance of, 1332
 - inflation control with, 1330–1331
- Inflation Targeting (IT), 1313–1318
- inflation persistence in, 1316–1317
 - initial conditions control in, 1315
 - instrumental variables in, 1315–1316
 - means/variances in, 1314–1316
 - propensity score matching, 1316
- long-run expectations of, 1317–1318
- methodology, 1307–1309
- three periods/three regimes, 1308–1309
 - two periods/two regimes, 1307–1308
- monetary aggregates, 1325–1328
- colinearity exceptions, 1326–1328
 - colinearity of, 1326
 - two pillars, 1325–1326
- regime shifts influence, 1334–1337
- biased estimator in, 1334
 - three regimes, 1336
 - three time periods in, 1335–1336
 - unbiased estimator in, 1335
- results, 1311–1312
- Euro influence, 1312
 - IT influence, 1311–1312
- robustness of, 1312–1313
- short-run expectations of, 1317
- traditional, 1306
- Monetary stabilization policy, 1008, 1009
- disturbances influencing, 724–725
 - household utility in, 759–760
- Monetary targetry, 1168–1183, 1353
- demise of, 1178
 - great inflation explanations with, 1174–1177
 - pragmatic monetarism and, 1177–1183
 - Volcker regime change and, 1168–1174
- Monetary transmission mechanism, 656, 818–820, 894–897, 1059n6
- Monetary union, 1045–1046
- Money. *See also* Currency; Domestic currency
- cash, 841n4
 - for goods, 958f, 970f
 - holding benefits of, 1359n10
 - holding/opportunity cost of, 658–659
 - monetary policy growth control of, 1353–1354
 - zero nominal return of, 727n5
- Money, Interest, and Prices* (Patinkin), 833, 1347
- Money demand
- elasticity, 700–701
 - fiscal policy and, 716–717
 - friction motivations in, 658–659
 - inflation rate/fiscal policy and, 664–667
 - optimal inflation rate and, 658–664
 - primal form with, 716–720
 - sticky prices and, 695–696
 - transactions cost function and, 669
- Money supply
- central banks control of, 726–727
 - constant growth of, 976n35
 - rules/ FTPL and, 951–952
- Money velocity, 662, 671–672
- Money-growth targeting, 1242–1243
- Mongelli, F. P., 1322
- Monopoly
- competition, 803n88
 - distortions, 911n31, 976
 - production power of, 870–871
 - profits, 985n47
- Montiel, P., 1481n146
- Moore, G., 1255n24
- Morck, R., 1485n167
- Morgan, J., 1022
- Morgenstern, 1098, 1100, 1102–1103
- Morris, S., 1484, 1484n158
- Moss, C. M., 1374
- MRO. *See* Main refinancing operation; Medium-term refinancing operations
- MSV. *See* Minimum state variable
- Muldoon, Robert, 1183, 1243

- Multilateral currency unions, 1037–1039
- Multiple fiscal authorities, 953–955
- Multiple steady states, 841–842
- Multiplier problem, 1149
- Mundell, R., 1034, 1041
- Mundell-Fleming model, 910, 1475
- Muranaga, J., 1213
- Muth, J., 1056, 1098
- N**
- Nakajima, T., 952n17
- Nakamura, E., 697
- Nakayama, T., 1213
- Nash equilibrium, 867, 909, 911–915
- Nash gaps, 912f, 914f
- Natalucci, F. M., 1316
- Natural allocation, 915–916
- Natural interest rates, 729, 839
- Natural outputs, 888
- Natural rate allocations, 880–884
- Natural rate of unemployment, 847
- Natural real wages, 815, 817–818, 818n104
- Natural relative price, 807
- Near-zero interest rates, 1382n43
- Negative correlations, 971
- Negative inflation, 664, 667
- Negative interest elasticity, 1382
- Negative nominal interest rates, 841n4
- Negative output gap, 755
 - discretionary policy resulting in, 752–753
 - target criterion and, 756n36
- Negative rates of inflation, 657
- Nelson, C., 1230, 1273
- Nelson, E., 949, 953, 1243
- Neo-Keynesian model, 698
- Neuman, M. J. M., 1315
- Neumeyer, P. A., 993
- Neves, J., 1446n14
- New Keynesian model
 - adaptive learning
 - baseline model calibration under, 1075–1076
 - inflation persistence in, 1073–1089
 - IT in, 1073–1075
 - optimal monetary policy under, 1071–1089
 - Bayesian estimation in, 1231
 - Calvo-Phillips relation of, 1357–1358
 - Clarida-Gali-Gertler, 1347
 - E-stability, 1066–1070
 - determinacy and, 1066–1067
 - extension and, 1068–1070
 - framework used, 1057–1058
 - generalizations about, 790–818
 - inflation dynamics, under rational
 - expectations, 1061–1065
 - microfoundations of, 760–765
 - optimal monetary policy
 - under adaptive learning, 1071–1089
 - baseline model calibration under, 1075–1076
 - under commitment, 1063–1064
 - under discretion, 1062–1063
 - intertemporal trade-offs in, 1084–1087
 - intra-temporal trade-offs in, 1082–1084
 - macro-economic performance/persistence
 - under, 1076–1081
 - optimal instrument rules in, 1064–1065
 - sensitivity analysis of, 1087–1089
 - solution method for, 1074–1075
 - optimal monetary policy in, 703, 726–759
 - optimal policy commitment in, 755–756
 - predetermined variables in, 1265n32
 - price adjustment models and, 790–802
 - price index stabilizing, 802–818
 - sectoral heterogeneity/asymmetric
 - disturbances in, 803–815
 - sticky wages/prices and, 815–818
 - stabilization theory in, 863
 - structural parameters of, 1199t
 - Taylor rule with, 1068
- New Keynesian open-economy analysis
 - budget constraints/Euler equations in, 874
 - exchange rate determination in, 877–879
 - household decisions/preferences in, 871–873
 - international asset markets in, 877–879
 - local currency pricing (LCP), price setting
 - under, 877
 - natural/efficient allocations in, 880–884
 - price-setting decisions in, 874–877
 - producer currency pricing (PCP), price setting
 - under, 875–876
 - real/nominal distortions in, 870–871
- New Keynesian Phillips curve (NKPC), 732, 884–886, 1074
 - goods/destination market and, 896
 - open-economy, 863–864
 - price stability and, 684

New Zealand

- corridor system of, 1390n51, 1426n91
- IT adopted by, 1183–1184, 1238–1239, 1243–1244, 1276–1277

Neyapti, B., 1018, 1455–1456

NICE years, 1160–1161, 1185–1204

- Great Moderation, 1189–1204
 - DSGE models nonzero trend inflation and, 1198–1199
 - DSGE models results and, 1196
 - indeterminacy of, 1200–1201
 - inflation expectation re-anchored during, 1194–1195
 - literature review of, 1196–1198
 - monetary policy rules/shocks of, 1201
 - structural change during, 1202–1204
 - structural VARs and, 1195–1196
 - volatility/uncertainty during, 1189–1194

Nicoletti-Altimari, S., 1061

Nicolini, J., 672, 674, 937, 1070

Niepelt, D., 949, 950, 950n16

Nikolov, K., 742

Nishioka, N., 1213

Nishioka, S., 1213

Nixon, Richard, 1030

NKPC. *See* New Keynesian Phillips curve

No tax without representation, 1024

Nominal anchor, 1354–1355

- fiscal policy providing, 944
- from monetary policy, 956
- price stability, policy coordination providing, 941–962

Nominal debt, 950

Nominal effective exchange rate, 1166f

Nominal government assets, 943

Nominal income targeting, 1460

Nominal interest rates, 668n3, 727

- cash/credit goods model implied in, 975
- deflation and, 986n50
- lower bound on, 742
- means/standard deviations of, 1188t
- in Taylor principles, 1068
- zero lower bound on, 765

Nominal targets, 1456–1460

Nominal value of liabilities, 942–943

Nominal variable indeterminacy, 950–951

Nominal wages, 657, 704–706

Non-Bayesian approach, 1124

Nonborrowed reserves

- federal funds rate changes and, 1349f, 1350
- liquidity effect generated by, 1368, 1368n26
- target funds rate and, 1350–1351f

Noncooperative game, 938

Noninflationary, consistently expansionary
macroeconomic performance (NICE).
See NICE years

Nonlinear DSGE models, 1252n18

Nonlinear structural model, 744–745

Nonperforming loans (NPLs), 1214

Non-predetermined variables, 1253, 1253n19

Nonquality-adjusted prices, 709–713

Non-Ricardian regimes

- fiscal policy, 944–945, 963–972
- FTPL and, 949–955
- plausibility of, 964–965
- surplus negative correlation in, 969–970

Nontradable goods (NTG), 1446–1449

Nonzero cost-push effect, 788–789

Nonzero interest rates, 985–986

Nonzero trend inflation, 1198–1199, 1231

Nordhaus, W. D., 1029–1030

Norges Bank, 1269, 1278n49, 1281–1284, 1283f

Normative theory, 973–995

NPLs. *See* Nonperforming loans

NTG. *See* Nontradable goods

Nunnenkamp, P., 1450

O

Oatley, T., 1018

Obstfeld, M., 869, 875, 910, 911n32, 914, 916, 952n17, 993, 1041, 1477, 1483, 1497n203

OCR. *See* Official Cash Rate

Oda, M., 1213

Oda, N., 1431n97

OECD. *See* Organization for Economic
Cooperation and Development

OECD countries

- IT in, 1249n14
- IT/inflation and, 1246f, 1247n11
- long-term inflation expectations in, 1060f

Official Cash Rate (OCR), 1277n48

Offset coefficient, 1476

Oil prices, 1177

OIS spread, 1429f

Okina, K., 1213, 1375n38

Olivera, 704–705, 1449n21

- Onatski, A., 695, 697, 979, 1121, 1272
- OPEC oil price increases, 1177
- Open-economies
 - classical view of, 886–894
 - divine coincidence of, 890–891
 - efficient international relative price adjustments in, 886–888
 - optimal policy in, 888–894
- Open-economy model
 - monetary policy analysis in, 862–863
 - optimal monetary policy in, 927
 - production goods in, 863–864
- Open-economy Phillips curve, 884–886
- OPP. *See* Optimal Policy Projections
- Opportunistic cycles
 - business/political, 1047–1050
 - in political business cycles, 1029–1031
 - pooling equilibrium in, 1049–1050
 - separating equilibrium in, 1049
- Opportunity cost, 658–659
- Optimal Bayesian policy rules, 847–848
- Optimal deviation, Friedman rule, 679–681
- Optimal discretionary policy, 831
- Optimal dynamics, 769–776, 782n66
- Optimal equilibrium dynamics, 729–733
 - FOC solutions and, 787
 - policy authority's actions in, 739
- Optimal exchange rate, 901
- Optimal fiscal policy, 990–993
- Optimal inflation dynamics, 773
- Optimal inflation rate, 656–657, 698f, 974
 - in cash/credit goods model, 988f, 990f
 - cash/credit goods model determining, 985, 989f
 - central banks, 713
 - competitive equilibrium in, 677
 - domestic currency, foreign demand for, 675–684
 - Friedman rule and, 989n55
 - high markup in, 672
 - monetary policy and, 663–664
 - money demand and, 658–664
 - money demand elasticity and, 700–701
 - price stickiness and, 697–700, 989
 - quality bias and, 706–714, 712t
 - relevant considerations in, 715
 - sticky prices and, 684–695, 700f
- Optimal long-run average rate of inflation, 732
- Optimal monetary policy, 990n58
 - under adaptive learning, 1071–1089
 - advance commitment to, 734
 - baseline closed-economy models and, 864
 - cash/credit goods model and, 980–984
 - central banks theory of, 757–758
 - under commitment, 1063–1064
 - commitment in, 733–737
 - for currency unions, 1037–1038
 - discount loss function in, 728
 - under discretion, 1062–1063
 - discretionary optimization differs from, 737
 - efficient/inefficient shocks in, 891
 - error correction in, 742n21, 819
 - flexible inflation targeting rule of, 1009
 - forecast targeting in, 737–742
 - global trade-offs in, 890–891
 - home markup exogenous decline in, 903f
 - home preference shock and, 926f
 - home productivity/preference shocks in, 904f
 - under imperfect information, 756–759
 - implementing, 990–993
 - inflation evolution/output gap in, 821–822
 - inflation tradeoffs/demand imbalances in, 918–925
 - in international transmissions, 894
 - intertemporal trade-offs in, 1084–1087
 - intra-temporal trade-offs in, 1082–1084
 - under LCP, 903
 - linear-quadratic model of, 1252–1257
 - linear-quadratic problem in, 726–729
 - macro-economic performance/persistence under, 1076–1081
 - in New Keynesian model, 703, 726–759
 - no consumption tax in, 984–990
 - nonlinear structural model and, 744–745
 - in open-economies, 888–894
 - in open-economy models, 927
 - optimal equilibrium dynamics in, 729–733
 - output-gap adjusted-price level in, 796–797, 801–802
 - price levels/disturbances in, 809
 - Ramsey solution to, 990–991
 - rational expectations in, 824
 - relative price misalignment in, 897–905
 - sensitivity analysis of, 1087–1089
 - shock response of, 1057–1058
 - simple policy rules v., 848–852, 848f
 - small steady-state distortions in, 806, 816–817
 - special parameterization of, 924

- Optimal monetary policy (*cont.*)
 target criterion in, 739–740, 775–776, 814–815, 818–819
 targeting rules in, 866–867
 timeless perspective in, 743–748
 volatilities under, 906–907, 906t
- Optimal policy
 choice, 1258–1260
 under discretion, 1062–1063
 functions, 1265n34
 prescription, 891
 problem, 765–769
 projections, 1262–1263
 rules, 1069
 theory, 851–852
- Optimal policy commitment
 discretionary policy compared with, 754f
 household utility and, 821
 in New Keynesian model, 755–756
 target criterion and, 754–755
- Optimal Policy Projections (OPP), 1275
- Optimal price stability, 788–790
- Optimal Quantity of Money*, 994
- Optimal response, 840f
- Optimal simple rules, 835–838, 927n46
- Optimal stabilization policy, 745
 with frictionless asset markets, 869–915
 in global Nash equilibrium, 911–915
 macro volatility and, 905–907
- Optimal state-contingent path of inflation, 732
- Optimal target criterion, 780–781, 791, 900–901, 900n20
- Optimal taxation, 994
- Optimal wage tax rate volatility, 988n54
- Optimality, of price stability, 974
- Optimization conditions
 of Euler equations, 939
 as Fisher equation, 660–661
 second-order conditions for, 786–787
- Optimization under discretion, 1266–1270
- The Optimum Quantity of Money* (Friedman), 974
- Organization for Economic Cooperation and Development (OECD), 936, 1003, 1060f, 1061, 1309
- Orphanides, A., 837, 839–840, 847, 851–852, 948n13, 1058, 1072, 1076–1077, 1089, 1091, 1271
- Ossowski, R., 1468n95
- Ötöker-Röbe, I., 1286
- Oudiz, G., 1266
- Output
 by country, 1162–1163f
 disturbances in, 729
 Euro countries growth of, 1322f
 fluctuations of, 1321–1322
 IT performance in, 1249f
 in macroeconomics, 1248–1249
 monetary policy influence on, 771–772
 short-run aggregate supply and, 764–765
 target level of, 773
- Output gap
 adjusted price-level target, 781
 coefficients on, 849f
 impulse responses of, 808f
 inflation persistence and, 1079f, 1081f, 1083–1084f
 inflation/optimal monetary policy and, 821–822
 inflation/optimal simple rules expectations of, 837–838
 measurement issues of, 838–841
 monetary policy rules stabilizing, 834–835
 relative price gap and, 810
 response, 816
 stabilization, 732, 779, 786, 1088–1089, 1088f
 two-state model predicting, 752n33
 zero-inflation steady-state and, 783, 785n70
- Output gap adjusted-price level
 deterministic path for, 819
 in optimal monetary policy, 796–797, 801–802
- Output gap-adjusted price level targets, 891n16
- Overnight funds, 1394n56
- Overnight interest rates, 1376f, 1395–1397
- Overnight market, 1386–1387
- P**
- Pagan, A., 1369–1370, 1369n31
- Pagano, M., 1035, 1473n99
- Palenzuela, D. R., 1061
- Panizza, 1443n5, 1452
- Papaioannou, E., 1321
- Pappa, E., 1195
- Parameterization, 756n37
- Parkin, M., 1017, 1356–1357
- Parsley, D., 1445n12
- Partisan cycles, 1027–1029
- Partisan models, 1032

- Party system, 1019–1020
- Pass-through coefficients, 1445–1446, 1445n11
- Patinkin, D., 832, 1347
- Paulson Report, 1161
- PCP. *See* Producer currency pricing
- Pearlman, J., 1271
- Pedersen, L. H., 1425
- Peg the export price (PEP), 1470
- Pegged interest rate solution (PIR), 943–944, 1154, 1355
- PEP. *See* Peg the export price
- Perceived law of motion (PLM), 1067
- Perez-Quiros, G., 1204
- Perfect capital mobility, 1465n86
- Perfect knowledge model, 847
- Permanent relative-price shock, 813–814
- Perotti, R., 964
- Perri, F., 1451
- Persson, M., 993, 994
- Persson, T., 993, 1016, 1030, 1033, 1040
- Perturbations, 1145n39
- Pesenti, P., 865, 869, 908, 910, 911n31, 924, 1459n55, 1484n158, 1484n162
- Peterson, I. R., 1110
- Phelan, C., 949, 952–953
- Phelps, E., 664, 699, 937, 974–975, 994
- Phillips curve, 799, 805
 - sectoral inflation in, 885
 - targeting rules combined with, 891–892
- Phillips-curve trade-off, 790
- Piger, J. M., 1379
- PIR. *See* Pegged interest rate solution
- PLM. *See* Perceived law of motion
- Polemarchakis, H., 952n17
- Policy
 - cash/credit goods model variables of, 987t
 - central banks changing rates of, 1350–1352
 - cooperation deviations, 909–915
 - determinations/target criterion in, 738–739
 - discretionary policy of, 1015
 - dynamic stochastic simulations of, 833–835
 - frontier, 836f, 837
 - function output gap, 1086f
 - inertia, 837–838
 - instruments, 1006n9, 1488–1493
 - model evaluating, 832–844
 - model structural relations objectives of, 729
 - Norges Bank options of, 1283f
 - optimal equilibrium dynamics of, 739
 - rate path, 1251n16, 1265n34
 - regimes by country, 1310t
 - of Riksbank, 1280f
 - Target Agreement, 1278n49
 - tensions, 1042
 - 2007–2009 financial crisis response of, 1414–1422
 - welfare-based analysis of, 790
- Policy interest rates
 - bank reserve changes with, 1374–1376, 1385–1386
 - reserves relationship with, 1374–1383, 1385–1386
 - of United States/Euro Area/Japan, 1418f
- Policymakers
 - Bayesian approach of, 845, 1101–1102
 - GDP-deflator inflation from, 891
 - global output gap/price inflation changes of, 899–900
 - international price manipulation of, 910–911
 - International relative price misalignments and, 866–867
 - terms of trade manipulated by, 911n32
- Political budget cycles, 1033
- Political business cycles, 1027–1034, 1466–1467
 - fiscal variables in, 1033
 - opportunistic cycles in, 1029–1031
 - partisan cycles in, 1027–1029
- Political cycles, 1031–1032
- Political opportunistic cycles, 1047–1050
- Political unification, 1045–1046
- Politicians, elected, 1021–1022
- Pollard, P., 1022
- Polson, N. G., 1230
- Poole, W., 1056, 1174n17
- Pooling equilibrium, 1049–1050
- Porter, R., 675, 1174n17
- Portes, R., 1321
- Posen, A. S., 1019
- Positive correlation, 964
- Positive shocks, 963f
- Positive theory, 937–973
- PPP. *See* Purchasing power parity
- PPT. *See* Producer price targeting
- Practical monetarism, 1178
- Pragmatic monetarism, 1163, 1168, 1177–1183
- Prasad, E., 1473n100, 1478, 1479n132, 1479n134
- Prati, A., 1394n55

- Precautions, 1143–1144
- Pre-commitments, 768
- Predetermined variables, 1252–1253, 1265n32
- Preference shocks, 904f
- Preference specifications, 1148n40
- Prescott, E., 1004
- Present value budget constraint (PVBC), 939
- Present-value constraint, 666
- Preston, B., 825n114
- Price adjustments
 - alternative models of, 790–802
 - Calvo model of, 815
 - Calvo-Yun model of, 792
 - frequency of, 807n93
- Price index, 909
 - complete stability of, 812n96
 - constant long-run level of, 812
 - inflation rate and, 811n95
 - stabilizing which, 802–818
 - in sticky-price sector, 811–812
 - target criterion and, 781n65
 - two log sectoral, 808f
- Price levels
 - adjustments to, 1451n27
 - cost-push shocks raising, 796n82
 - disturbances influencing, 809
 - Euro and, 1322–1325
 - gap, 814
- Price stability, 684
 - cash-in-advance model in, 938–939
 - with distortionary taxation, 699
 - ECB's goal of, 1325–1326
 - financial crisis and, 1167
 - in fiscal policy, 936–937
 - through monetarist arithmetic, 939–941
 - in monetary policy, 806–807
 - nominal anchor
 - basic FTPL in, 942–943
 - coordination problem and, 955–962
 - FTPL criticisms and, 949–955
 - non-Ricardian fiscal policies in, 944–945
 - pegged interest rate solution in, 943–944, 1154, 1355
 - policy coordination in, 941–962
 - price determinacy in, 945–948
 - normative theory of, 973–995
 - cash/credit goods model and, 977–980
 - no consumption tax and, 984–990
 - optimal fiscal/monetary policy in, 990–993
 - Ramsey optimal policy in, 993–994
- optimality of, 788–790, 974
- positive theory of, 937–973
- Ricardian/non-Ricardian fiscal policy in, 963–972
- Price stickiness, 698f
 - asymmetric disturbances and, 807n92
 - inflation persistence and, 1087–1088
 - labor income influenced by, 987–988
 - on optimal inflation, 989
 - optimal inflation rate and, 697–700
 - sensitivity analysis and, 1087f
- Price-levels
 - dispersion, 1324–1325
 - for monetary policy, 742, 789
 - target criterion, 756
 - target of, 741, 755, 810–811, 842
 - targeting/IT and, 1286–1287
- Price-quantity representations, 1425
- Prices
 - asset, 843, 1496–1498
 - central banks stabilization of, 944
 - discrimination in, 885–886
 - dispersion of, 985n49
 - fiscal policy determination of, 936–937
 - flexibility of, 814
 - goods devaluation and, 1445–1453
 - indexation scheme and, 693–695
 - inflation, 899–900
 - inflation of goods, 817
 - model of, 799
 - pass-through, 1450–1451
 - rigidity of, 833
 - two-state model predicting, 752n33
- Price-setting
 - Calvo, 938
 - decisions, 874–877
 - problem, 709–710
- Price-stability trade-off, 695–701
- Price-taking assumption, 1445n10
- Primal form, 716–720
- Primiceri, G. E., 1189, 1195, 1197, 1227–1229, 1230
- Priors, 1228–1229, 1231
- Prisoners dilemma, 1483
- Private consumption, 787n73
- Private sector
 - central banks expectations of, 825, 825n114

crisis management involvement of, 1486
 forward-looking behavior of, 726
 inflation expectations of, 1059–1061
 inflation target known by, 1073n22
 Procyclicality, 1465–1472
 capital flows and, 1465–1466
 commodities and, 1467–1468
 demand policy and, 1466–1467
 export price shocks and, 1469
 of monetary policy, 1467
 PEP/PPT and, 1470–1471
 political business cycle and, 1466–1467
 product price index and, 1471–1472
 product-oriented choices and, 1468–1472
 Producer currency pricing (PCP), 865, 875–876
 Producer price targeting (PPT), 1470
 Product market, 1042
 Product price index, 1471–1472
 Production
 monopoly power in, 870–871
 open-economy model goods of, 863–864
 zero inflation subsidies in, 689–690
 zero inflation without subsidies in, 690–693
 Productive expenditure, 821
 Productivity growth, 1042n48
 Product-oriented choices, 1468
 Profits, 762, 763
 Projection model, 1257–1258, 1268–1269
 Propensity score matching, 1316
 Public
 rules deviation of, 1007, 1009–1010
 sector liabilities, 950
 zero inflation expectations of, 1005
 Purchasing power parity (PPP), 866
 PVBC. *See* Present value budget constraint

Q

QE. *See* Quantitative Easing
 QPM. *See* Quarterly Projection Model
 Quadratic flow loss, 888–889
 Quadratic function, 784–785
 Quadratic loss function, 740n19, 786, 792, 816n102, 1255, 1259
 Quadratic objective, 776–786
 Quality bias
 optimal inflation rate and, 706–714, 712t
 simple model of, 707–709
 Quality of goods, 712

Quality-adjusted prices, 713–714
 Quantitative Easing (QE), 1212
 Quantity of goods, 707–708
 Quarterly Projection Model (QPM), 1276
 Quasi-difference of inflation, 1080f, 1081f
 Querubin, P., 1019
 Quinn, D., 1472
 Quintyn, M., 1456
 Qvigstad, J. F., 1284

R

Radecki, L., 1174n17
 Radelet, S., 1483n157, 1487n173
 Rajan, R., 1450, 1479n132, 1480n157
 Ramsey allocation, 983
 Ramsey policy, 747
 Friedman rule as, 680–681
 implementing, 993–994
 inflation rate and, 665, 703
 labor-income tax rate and, 669–670
 monetary policy, 661–662, 748–750, 767, 773, 801, 818
 monetary/fiscal policies and, 984–985
 domestic currency/foreign demand in, 680t
 Friedman rule satisfied by, 975–976, 975n34
 inflation path under, 736f
 volatile inflation in, 995
 Ramsey problem
 equilibrium relations in, 767–768
 first-order condition of, 691–693
 flexible price competitive economy and, 981–982
 Friedman rule solution to, 674
 intertemporal implementability conditions of, 984n46
 Lagrange multiplier in, 668
 money velocity in, 671–672
 numerical algorithm in, 679–680
 steady state in, 683, 696–697
 underground economy and, 673–674
 Ramsey solution, 990–991
 Ranciere, R., 1464n76, 1480n137, 1496n200, 1497
 Random-Walk Metropolis, 1231
 Rasche, R. H., 1174n17
 Raskin, M., 1431
 Rational expectations, 1066, 1101–1102
 discretionary equilibrium imposing, 1008
 Inflation dynamics and, 1061–1065

- Rational expectations (*cont.*)
 - macroeconomic outcomes and, 1076t
 - in optimal monetary policy, 824
 - recursive learning algorithms and, 1065
 - revolution, 1030
- Rational Partisan Theory, 1027–1028, 1032
- Rational-expectations equilibrium, 735, 823
- Ravn, M., 702
- Rawls, J., 748
- Razin, A., 869, 916
- RBNZ. *See* Reserve Bank of New Zealand
- RE equilibrium (REE), 1066
- Reaction coefficient, 853
- Reaction functions, 1063n8
- Reagan, Ronald, 948, 1172
- Real exchange rates
 - consumption preferences and, 902
 - with targeting rules, 922
 - terms of trade and, 895
- Real GDP growth, 1165f, 1187t
- Real interest rates
 - equalization of, 1475
 - increase/inflation increase, 945–946
 - means/standard deviations of, 1188t
- Real wages, 815
- Rebelo, S., 907, 1445n12, 1446n14, 1483n157, 1484n162
- Recursive learning algorithms, 1065
- Reduced-form VAR innovations, 1190f
- REE. *See* RE equilibrium
- Regression coefficient, 960
- Regulatory capture, 1026–1027
- Reifschneider, D., 704, 842, 1275
- Reinhart, C., 1444n7, 1445n11, 1450, 1461, 1463, 1477, 1477n123, 1481n146, 1483n157, 1485n166
- Reinhart, V. R., 1213
- Reis, R., 798, 801
- Relative demand gap, 868
- Relative entropy, 1149, 1150
- Relative exchange rate, 1180
- Relative prices, 791n77
 - adjustments, 868, 1354–1356
 - distortions, 806–807
 - gap, 810
 - misalignment, 897–905
- Remache, J., 1431
- Reputation, 1004–1007, 1009
- Reserve Bank Act of 1989, 1243
- Reserve Bank of Australia, 1278n49
- Reserve Bank of New Zealand (RBNZ), 1183, 1276, 1278n49
- Reserve demand/supply
 - within maintenance period, 1399–1413
 - market interest rates and, 1360–1367
- Reserve management, 1392–1399
- Reserve rate, 1363
- Reserve remuneration, 1427f
- Reserves market, 1386, 1424f, 1425
- Residential mortgage lending, 1415
- Resources loss, 661
- Responsiveness, 1134–1136
- Return to scale, 668–670, 670t
- Revenues, 762
- Rey, H., 1480n137
- Rhee, M. W., 1479n131
- Riboni, A., 1023
- Ricardian policies, 765n51, 945, 948, 952, 963–972, 1070–1071
 - government reactions to, 967
 - surplus/debt response of, 967–969
- Ricardian regime, 969n30
- Ricardo, D., 830
- Ricci, L., 1478, 1480
- Ridella, S., 1231
- Rigidities of rules, 1004–1005
- Rigobon, R., 1443n5, 1444n8
- Riksbank, 1275, 1278n49, 1280–1281
 - constant interest rate of, 1263n31
 - judgment by, 1282f
 - policy options of, 1280f
- Risk
 - aversion, 1144–1148
 - management, 1218
 - market price of, 1147–1148
 - sensitive joint filtering, 1132n32
 - sensitivity interpretations, 1148n40
- Risk, Uncertainty and Profit* (Knight), 1100
- Risk sharing
 - under complete markets, 877–878
 - under incomplete markets, 878–879
 - mechanism of, 869
- Robertson, J., 1019, 1369–1370, 1369n31
- Robust control, 1132
- Robust control model, 1117
- Robust control techniques, 852

Robustness, 1104–1109

Ball's model with, 1141–1143
 Bayesian model detection and, 1113–1117
 calibrating for, 1109–1117
 classical model detection and, 1112–1113
 control with, 1132
 decision rules with, 1118, 1121
 econometric defense for filtering and,
 1139–1140
 of Euro, 1338t
 filtering/control and, 1130–1132
 frequency domain details and, 1136–1140
 of IT, 1338t
 to learning, 851f
 limiting version of, 1138–1139
 long variable lags and, 1140–1143
 market price of risk and, 1147–1148
 of monetary policy rules, 824, 844–848
 of monetary regimes, 1312–1313
 precautions and, 1143–1144
 reasonable preference for, 1111
 responsiveness and, 1134–1136
 risk aversion and, 1144–1148
 in simple macroeconomic model, 1133–1134
 standard control theory and, 1104–1106
 standard errors of, 1311n2
 state evolution in, 1111–1112
 Rodrik, D., 1477n124, 1478
 Rogoff, K., 875, 910, 911n32, 914, 952n17, 1013,
 1015, 1031–1032, 1444n7, 1445n11, 1463,
 1464n76, 1478, 1493
 Røisland, Ø., 1284
 Rojas-Suárez, L., 1483n157
 Ropele, T., 1198–1199
 Rose, A., 1039, 1040, 1043, 1242n6, 1319, 1461,
 1481n146, 1482n153, 1485n166,
 1494n187, 1496n196, 1497n203
 Rosenthal, H., 1029
 Rossi, P., 1230
 Rotemberg, J., 684, 705, 778n63, 831, 844,
 976n36, 991
 Roubini, N., 1027, 1451, 1484n158, 1484n162
 Rubio-Ramírez, 1197
 Rudebusch, G., 835n1, 837, 839–840, 846, 853,
 1253n21
 Rudebusch-Svensson model, 846
 Ruge-Murcia, 705–706, 1023
 Rules v. discretion, 1004–1013

in CBI, 1013–1014
 during financial crisis, 1010–1012
 in monetary policies, 1012–1013
 reputation in, 1004–1007

S

Sachs, J., 1266, 1483n157, 1487n173, 1494n189
 Sack, B., 1206n66, 1213, 1431
 Sahay, R., 1453
 Sales revenues, 762
 Salles, J. M., 1247, 1247n12, 1313, 1315–1316, 1339
 Salter-Swan model, 1448
 Samuelson, P., 1447–1448
 Santaella, J., 1449n21
 Santoro, S., 1071, 1073
 Sargent, T., 937, 939–943, 943, 945, 952, 952n17,
 1099, 1101, 1104, 1108–1109, 1110, 1114,
 1118–1119, 1122–1123, 1131, 1132n33,
 1136, 1138–1139, 1145–1146, 1148, 1189,
 1191, 1228–1230, 1274n42, 1355, 1357
 Sasson, D., 1479n131
 Savage, L. J., 1098, 1100, 1101–1102
 Savastano, 1285, 1444n7, 1445n11, 1459n55
 Sbordone, A. M., 695, 979
 Schaling, E., 1274n43
 Schaumburg, E., 1269
 Schivardi, F., 1042n48
 Schlesinger, Helmut, 1168
 Schmeidler, D., 1099, 1102–1103, 1103n7, 1104,
 1106, 1109, 1110
 Schmidt-Hebbel, K., 1247, 1315–1316, 1445n12,
 1454
 Schmitt-Grohé, S., 664, 669, 670, 675, 693,
 695–698, 702–703, 706, 725, 961, 975n33,
 976, 977, 984n46, 985n47, 986–987,
 986n50, 987n51, 989n55, 991–992, 1070
 Schmukler, S., 1480n137, 1483n157, 1485n166
 Schnabl, P., 1429
 Schorfheide, F., 697, 947, 1196, 1199–1200,
 1200n59, 1231
 Schwartz, A., 937, 1161, 1487n178
 Schweickert, R., 1450
 Second-order conditions
 Lagrange multiplier and, 770n55
 for optimality conditions, 786–787
 Sectoral heterogeneity, 803–815
 Sectoral inflation, 885
 Sectoral price level, 806–807

- Security and Exchange Commission, 1025n31
- Seignorage income, 675
- Seignorage losses, 663–664
- Sensitivity analysis, 1087–1089, 1087f
- Separating equilibrium, 1049
- Separation principle, 1271
- Sequential budget constraint, 661, 665, 676
- Shapiro, M. D., 741, 1145
- Sharma, S., 1285, 1459n55
- Sheedy, K. D., 792, 797
- Shen, C. H., 1450
- Sheridan, N., 1247, 1249, 1307–1309, 1315–1316, 1334–1337
- Shi, M., 1033
- Shin, H. S., 1294n68, 1484, 1484n158
- Shioji, E., 1372
- Shirakawa, K., 1213
- Shirakawa, M., 1375n38
- Shiratsuka, S., 1213, 1375n38
- Shock therapy, 1451n26
- Shock vector, 1109
- Shocks, 1253n20
 - covariance between, 1039
 - in different time periods, 1134–1135
 - in economy, 1009–1010
 - financial crisis with, 1011
 - inflation impulse responses to, 1143f
 - in monetary policies, 1290n62
 - to monetary policies, 1185
 - optimal monetary policy response to, 1057–1058
- Short-run aggregate supply, 764–765
- Short-run expectations, 1317
- Short-term discount rate, 1113
- Short-term foreign debts, 1452n34
- Short-term interest rates, 834
- Shrestha, S., 1484n162
- Sibert, A., 1031, 1044
- Simple policy rules
 - alternative specifications of, 842
 - coefficients, 845t
 - optimal monetary policy v., 848–852, 848f
 - other variable responses in, 843–844
- Simple rules
 - abandonment caveats of, 1011–1012
 - contingent rules and, 1007–1008
 - discretion loss and, 1010
- Sims, C., 911n33, 942, 943, 950, 964–965, 1139–1140, 1195, 1255
- Singh, K., 1460
- Sloek, T., 1478, 1480
- Small, D. H., 1213
- Smets, F., 697, 702, 824, 839, 1071, 1074–1075, 1078, 1197
- Smith, A., 830
- Smith, R. T., 1477n123
- Social planner's problem, 663, 1008
- Söderlind, P., 1266
- Soderstrom, U., 1043
- Solberg-Johansen, K., 1284
- Soledad, M., 1483n157
- Solow residual, 1129f
- Souganidis, P., 1108
- Sovereign spread, 1474
- Specification doubts, 1119
- Spectral analysis, 1139
- Speculative attacks, 1482–1484
- Spiegel, M., 1494n187, 1495n196, 1497n203
- Stability of velocity, 1178
- Stabilization
 - of asset prices, 844n6
 - of central banks, 657
 - cross-country output gap, 866
 - inflation programs for, 1454–1455
 - inflation rates in, 701–702
 - monetary policies role of, 1008
 - in New Keynesian model, 863
 - of optimal exchange rate, 901
 - output gap, 779, 786
 - of price index, 802–818
 - welfare and, 759–790
- Stable solution, 947–948
- Stagflation, 1004
- Staggered pricing, 761
- Standard control theory, 1104–1106
- Standard portfolio theory, 1362, 1362n15
- Standing facilities, 1394n55, 1400
- State evolution, 1111–1112, 1116
- State prediction, 1125–1128
- State-contingent evolution, 743, 749, 758, 774–775, 807
- Static valuation problem, 1150–1152
- Statistical detection theory, 1112n17
- Steady-state, 777–782
 - consumption, 913n35
 - distortions, 778n63, 782–784
 - large, 784–786, 816n102

- in optimal monetary policy, 806, 816–817
 - Kalman filter, 1130
 - in Ramsey problem, 696–697
 - Steinsson, J., 697
 - Sticky information model, 798–802
 - Sticky prices. *See also* Friction model
 - with capital accumulation, 684–689
 - friction in, 656–657
 - Friedman rule and, 897
 - inflation variability of, 995
 - monetary policy/sticky wages and, 815–818
 - money demand and, 695–696
 - monopoly distortions and, 976
 - nonquality-adjusted prices and, 709–713
 - optimal inflation rate and, 684–695, 700f
 - quality-adjusted prices with, 713–714
 - Ramsey allocation with, 983
 - sector, 811–812
 - Sticky wages, 815–818
 - Stigler, G., 1026
 - Stiglitz, J., 1486, 1487n173
 - Stochastic difference equation, 730
 - Stochastic discount factor, 762
 - Stochastic growth model, 1127
 - Stochastic volatility, 1226–1228
 - Stochastically switching policy regimes, 959–961
 - Stock, J., 853–854, 1185, 1195
 - Stockton, D. J., 1275
 - Stokey, N. L., 975, 975n34, 993, 994
 - Storgaard, P. E., 908
 - Stracca, L., 1374
 - Strategic interactions, 867–868
 - Strategic manipulations, 870, 909–911
 - Strategic monetary interactions, 911n33
 - Strongin, S. H., 1368–1369, 1368n26
 - Structural inflation, 705
 - Structural inflation inertia, 792–798, 797f
 - Structural parameters, 1199t
 - Structural reforms, 1042
 - Structural rigidities, 1210
 - Structural VARs, 1195–1196
 - Sturzenegger, F., 1463
 - Sudden stops, 1482, 1482n152, 1492f
 - Summers, L., 701, 1017
 - Supply. *See also* Demand; Labor supply;
 - Money supply
 - demand equilibrium, 1360
 - induced interest rates, 1385f
 - log-linear aggregate, 779–780, 786
 - short-run aggregate, 764–765
 - spillovers, 910n30
 - Surplus
 - debt dynamics, 964
 - Ricardian policies response to, 967–969
 - in Ricardian regime, 969n30
 - in United States, 968f
 - GDP and, 971t
 - negative correlation of, 969–970
 - regressions, 964
 - Sutherland, A., 927
 - Svensson, J., 1033, 1058n5, 1210
 - Svensson, L., 757, 837, 846, 993, 1239n1, 1240n3,
 - 1252n18, 1252n21, 1255n25, 1258–1260,
 - 1263, 1265, 1265n33, 1269, 1271–1272,
 - 1272n40, 1273–1274, 1275, 1281n52,
 - 1354n5
 - Swanson, E. T., 1206, 1206n66, 1318
 - Swedish economy, 1043
- T**
- Tabellini, G., 1016, 1017, 1021–1022,
 - 1033, 1454
 - TAF. *See* Term Auction Facility
 - Takeda, T., 1372
 - TALF. *See* Term Asset-Backed Securities Loan Facility
 - Tallarini, T., 1109, 1145
 - Tambalotti, A., 1269
 - Target criterion
 - of central banks, 791
 - linear, 747–748
 - negative output gap and, 756n36
 - optimal, 780–781, 791
 - in optimal monetary policy, 814–815,
 - 818–919
 - optimal monetary policy and, 739–740,
 - 775–776
 - optimal policy commitment and, 754–755
 - output-gap-adjusted price level targets in,
 - 891n16
 - in policy determinations, 738–739
 - price indices and, 781n65
 - price-level, 756
 - time-invariant, 812–813
 - variable impulse responses, 813f
 - Target fund rates, 1350–1351f

- Target interest rates
 - bank reserves and, 1378f
 - of Japan/Euro Area/United States, 1393
 - target rate change and, 1413f
- Target levels, 773, 1254
- Target rates, 1398, 1406n72, 1413f
- Target shortfall, 755
- Target variables, 1250
- Targeted asset purchases, 1428
- Targeting rules, 1058n5, 1264
 - in cross-country terms, 890
 - disturbances with, 992–993
 - in IT, 1263–1264
 - in optimal monetary policy, 866–867
 - Phillips curve combined with, 891–892
 - real exchange rate with, 922
- Taxes
 - evasion of, 670t, 672–675
 - fiscal policies with, 994–995
 - profits from, 989
 - rate volatility of, 990n57
 - system incomplete in, 656
- Taxpayers, 1024
- Taylor, L., 1451
- Taylor, A., 1479n134
- Taylor, J. B., 703, 831, 832, 833–836, 844–845, 854–855, 908, 1012, 1176, 1251, 1260, 1357
- Taylor curve, 1251
- Taylor expansions
 - of Lagrange multiplier, 786–787
 - quadratic terms in, 784–785
 - zero-inflation steady-state in, 782n67
- Taylor principles, 946–947
 - central bank policy obeying, 954–955
 - interest rate violation of, 947–948, 1068
 - nominal interest rates in, 1068
- Taylor rule, 836, 1090, 1357
 - deviating from, 853–854
 - Federal Reserve abandoning, 1012
 - interest-rate reaction function of, 822
 - modern-day policy rules of, 831
 - New Keynesian model with, 1068
 - short-term interest rates in, 834
- Taylor rules posterior distributions, 1202f
- Teles, P., 664, 937, 947n12
- Tenreiro, S., 1040
- Tequila crisis, 1332, 1333t
- Term Asset-Backed Securities Loan Facility (TALF), 1419
- Term Auction Facility (TAF), 1419
- Term Securities Lending Facility (TSLF), 1419
- Terms of trade
 - policymakers manipulating, 911n32
 - real exchange rates and, 895
 - strategic manipulations of, 870, 909–911
 - transmission channel, 883
- Terrones, 1454, 1458n54
- Tesfaselassie, M. F., 1274n43
- Tetlow, R. J., 1275
- TG. *See* Tradable goods
- Thornton, Henry, 830
- Three-asset demand and supply
 - model, 1392–1393
- Tille, C., 911n31
- Time inconsistency, 1006n9
- Time periods, with shocks, 1134–1135
- Time-invariant policy, 746
- Time-invariant solutions, 768
- Time-invariant target criterion, 812–813
- Timeless perspective, 736f, 743–748
- Time-varying parameters VAR
 - stochastic volatility
 - estimation procedure of, 1228–1231
 - posterior distribution simulation of, 1229–1231
 - stochastic volatility with, 1226–1228
- Tinbergen, J., 1354
- Tinbergen principle, 1184, 1217
- Tinsley, P. A., 1174n17
- TIP. *See* Treasury Investment Program
- Tobin, J., 705, 1366
- Tornell, A., 1480n137
- Toxic assets, 1442–1443
- Tracking problems, 1124
- Tradable goods (TG), 1447
- Trade
 - benefits, 1040–1041
 - determinants, 1319–1320
 - flows, 1320f
- Transaction costs, 659–660, 669
- Transfer of balances, 1361n12
- Transfer payments, 976
- Transitory component, 1126f
- Transmission channel, 883
- Transmission mechanism, 1272–1274

- Treasury Investment Program (TIP), 1398
 - Treasury-Federal Reserve Accord, 947–948
 - Trichet, J. C., 1056, 1326–1327
 - Tryon, R., 844
 - TSLF. *See* Term Securities Lending Facility
 - Turmuhambetova, G. A., 1104
 - Two log sectoral price index, 808f
 - Two pillars, 1325–1326
 - Two risk-sensitivity operators, 1119–1121
 - T¹ operator in, 1119–1120
 - T² operator in, 1120–1121
 - Two sectoral inflation rates, 809–810
 - 2007–2009 financial crisis, 1414–1431
 - Two-country open-loop Nash equilibrium, 912–913
 - Two-period valuation problem, 1153–1155
 - Two-person, dynamic game, 1106
 - Two-person, zero-sum game, 1106, 1109–1110
 - Two-player, zero-sum game
 - breakdown suffered in, 1138–1139
 - shock vector distribution in, 1109
 - worst-case evolution equation in, 1111–1112
 - Two-player game, 1107
 - Two-sector model, 803, 803n88
 - Two-state model, 752n33
- U**
- Ueda, K., 1211–1212, 1213, 1431n97
 - Uesugi, I., 1373, 1401
 - Ugai, H., 1212–1213, 1431
 - UIP. *See* Uncovered interest parity
 - Unbiased estimators, 1335
 - Uncertainty, 1283f
 - Uncovered interest parity (UIP), 927n46
 - Underground economy, 672–673
 - aggregate activity levels in, 674
 - Ramsey problem and, 673–674
 - Unemployment gap, 839
 - optimal coefficients on, 850f
 - optimal response to, 840f
 - Unemployment rate, 840f
 - Unilateral adoptions, 1035–1036
 - Unique bounded evolution, 739n17
 - United Kingdom
 - interest rates/inflation/output of, 1162–1163f
 - macroeconomic data of, 1171f, 1182f
 - United States
 - bank reserves demand in, 1379–1380, 1380t, 1424n88
 - bank reserves requirements of, 1360n11, 1365n18
 - bank reserves/interest rates of, 1388–1392
 - bank's currency holdings of, 1361n13
 - Canada/interest rates of, 1180n38
 - central bank assets of, 1422f
 - central bank liabilities of, 1421f
 - consumption growth in, 1123f
 - CPI inflation expectations of, 1194f
 - economic structural transformations in, 1203–1204f
 - elasticity of demand in, 1379–1383
 - excess reserve demand for, 1380t, 1424n88
 - excess reserves/short-term interest rates of, 1384f
 - Great Inflation of, 1176–1177
 - Great Moderation period in, 853
 - house prices in, 1416
 - Interest rates/inflation/output of, 1162–1163f
 - investment prices in, 702–703
 - liquidity effect in, 1367–1372
 - long-run inflation expectations of, 1176f
 - long-term forward rates of, 1206n66
 - macroeconomic data of, 1170f, 1175–1176, 1175t
 - partisan models supported in, 1032
 - policy interest rates of, 1418f
 - positive inflation target of, 656
 - recent monetary history of, 1198
 - reserves demand within maintenance period of, 1404–1409
 - reserves/overnight interest rates of, 1376f
 - reserves/policy interest rates of, 1375–1379
 - reserves/target interest rates in, 1378f
 - residential mortgage lending in, 1415
 - seignorage income of, 675
 - surplus/debt dynamics in, 968f
 - target/market interest rates of, 1393
 - toxic assets originating in, 1442–1443
 - Volcker disinflation of, 1173
 - Untaxed income, 667–675
 - Upadhyaya, K., 1450
 - Uribe, M., 664, 669, 670, 675, 693, 695–698, 702–703, 706, 725, 961, 975n33, 976, 977, 984n46, 985n47, 986–987, 986n50, 987n51, 989n55, 991–992, 1070
 - Utility function, 659
 - Utility loss, 1014

V

Valdes, R., 1482
 Valla, N., 1274n43
 Value-at-Risk (VaR), 1218
 van Wijnbergen, S., 1451n26
 van Wincoop, E., 908, 1480n135
 VAR. *See* Vector autoregression models
 VaR. *See* Value-at-Risk
 Vázquez, F., 1456
 Vector autoregression models (VAR), 1367, 1368–1370
 Vega, M., 1316, 1339
 Vegh, C., 664, 1453, 1454
 Velasco, A., 1483n157
 Vestin, D., 824, 1071, 1075, 1078
 Volatilities, 906–907, 906t
 Volcker, P., 948, 1056, 1068, 1160–1161, 1168–1174, 1172, 1191, 1197
 monetary targetry and, 1168–1174
 United States disinflation and, 1173
 Vollrath, D., 1479n130
 von Hagen, J., 1315
 von Neumann, 1099, 1100, 1102–1103
 von Neumann-Morgenstern-Savage foundation, 1099
 VonZurMuehlen, P., 1174

W

Wage rigidity, 833, 979
 Wage stickiness, 987n52
 Waggoner, D. F., 1273
 Wagner, A., 1498n210
 Wallace, N., 937, 939–940, 942–943, 945, 952, 952n17, 1099, 1355, 1357
 Wallis, 1139
 Wallstein, S., 1463n75
 Walsh, C., 831, 843, 1016, 1061, 1244, 1293
 Walters, Alan, 1181
 Wang, N. E., 1131
 Wang, T., 1099–1100, 1102, 1104
 Warnock, F., 1479n131
 Watson, M., 853–854, 1185, 1195
Wealth of Nations (Smith), 830
 Webb, S., 1018, 1455
 Wei, S. J., 1445n12, 1478
 Welfare
 -based analysis, 790
 improvement, 1477–1480

 optimal policy problem and, 765–769
 quadratic objective based on, 776–786
 -relevant gaps, 888
 stabilization policies and, 759–790
 West Germany, 1169f
 Westermann, F., 1480n137
 Whiteman, 1134–1135
 Whittle, 1132n32, 1145
 Wicksell, K., 830, 1348, 1352, 1355–1358, 1357–1358, 1427–1428
 Wieland, 831, 833, 847, 1118, 1274n43
 Wilcox, D. W., 1275
 Williams, J., 695, 697, 704, 831, 836, 837, 839, 842–843, 844, 846–847, 848, 851–852, 979, 1058, 1072, 1076–1077, 1089, 1091, 1121
 Williams, N., 697, 979, 1104, 1121, 1131, 1272
 Williamson, J., 1451n26, 1465n86
 Winkelreid, D., 1316, 1339
 Wolf, H., 1463
 Wolman, A., 690n7, 695, 789, 792, 976n36
 Woodford, M., 684, 686, 690n7, 739–740, 742, 750, 753–754, 757, 760, 764n48, 765, 775, 778n63, 782, 786, 788, 807n94, 816n102, 820–823, 824, 831–833, 838, 842, 844, 846, 848, 851–853, 863, 940–942, 944–945, 947, 948–949, 950, 950n14, 951, 953–955, 960–961, 967, 975n33, 976, 976n36, 977n37, 986–987, 991–992, 1057–1058, 1061, 1063, 1090–1092, 1198, 1257, 1258–1260, 1263–1265, 1265n33, 1271–1272, 1272n40, 1273–1274, 1293n67, 1347, 1390n51, 1425, 8663
 Worst-case evolution equation, 1111–1112
 Worst-case model, 1133f
 Wouters, R., 697, 702, 1197
 Wright, J., 1390n51
 Würtz, F. R., 1373–1374, 1399–1400
 Wyplosz, C., 1322

X

Xu, J., 1460n59

Y

Yang, D., 1463n75
 Yaron, A., 1122
 Yasuhide, 1372
 Yates, T., 1090
 Ye, H., 1247, 1316, 1339

Yeung, B., 1485n167
 Yoshida, T., 1213
 Yu, W., 1485n167
 Yun, T., 684, 686, 775

Z

Zampolli, F., 1273
 Zero inflation
 with production subsidies, 689–690
 without production subsidies, 690–693
 public's expectations of, 1005
 Zero interest rate policy (ZIRP), 1211–1212, 1382
 Zero lower bound (ZLB), 701–704, 750–751, 832, 1070
 central banks constrained by, 749n29
 history-perspective and, 750–751
 interest rates with, 841–843
 monetary policy rules implications of, 842
 multiple steady states implied in, 841–842
 negative rates of inflation and, 657
 on nominal interest rates, 765, 841–843
 in reserves market, 1424f
 Zero nominal interest rate, 937
 Zero nominal return, 727n5
 Zero steady-state inflation, 690n7
 Zero-inflation steady-state, 773, 793
 output gap and, 783, 785n70
 in Taylor expansions, 782n67
 Zero-sum games, 1103n7
 Zha, T., 1195, 1273
 Zhu, H., 1294n68
 Zhuravskaya, E., 1033
 Zingales, L., 1026
 ZIRP. *See* Zero interest rate policy
 Zizza, R., 1042n48
 ZLB. *See* Zero lower bound