Dynamic General Equilibrium Modelling

Burkhard Heer · Alfred Maußner

Dynamic General Equilibrium Modelling

Computational Methods and Applications

With 63 Figures and 24 Tables

Springer

Professor Burkhard Heer
School of Economics and Management
Free University of Bolzano-Bozen
Via Sernesi 1
39100 Bolzano
Italy
burkhard.heer@unibz.it

Professor Alfred Maußner
Department of Economics
University of Augsburg
Universitätsstraße 16
86159 Augsburg
alfred.maussner@wiwi.uni-augsburg.de

Cataloging-in-Publication Data
Library of Congress Control Number: 2004115348

ISBN 3-540-22095-X Springer Berlin Heidelberg New York

Hardcover-Design: Erich Kirchner, Heidelberg

SPIN 11011798 43/3130-5 4 3 2 1 0 – Printed on acid-free paper

Preface

Dynamic General Equilibrium (DGE) models have become the workhorses of modern macroeconomics. Whatever textbook on advanced macroeconomics you consider you will find three kinds of models: the Solow model, the Ramsey model, and the overlapping generations model. The elementary versions of all three models can be studied with paper and pencil methods. But as soon as the researcher starts asking important questions of economic policy, these methods break down.

There are three questions researcher are most interested in. The first concerns transitional dynamics. For example, in growth theory, we are interested in the question of how countries converge to their long-run equilibrium, or, in public finance, we want to understand the behavior of the economy after an enduring tax cut. The second kind of problem concerns economic fluctuations that are caused by supply and demand shocks. Notably stochastic versions of the Ramsey model have been applied successfully to the study of business cycle dynamics. In these models demand and supply shocks trigger intra- and intertemporal substitution between leisure, consumption, and asset holdings and generate patterns in time series that mimic those found in macroeconomic data. The third issue, which has only received limited attention in the recent textbook literature, concerns models with heterogenous agents. Important applications of heterogeneous-agent economies can be found in the theory of income distribution, in the theory of asset pricing or in the field of public finance, to name but a few.

To address any of these issues, the researcher needs to apply computational methods. In recent years, numerical analysis has become one of the standard tools for graduate economics students. The advance of the importance of computational economics is also reflected in the growing number of journals and textbooks on this subject. For the student who is not acquainted with the

solution of numerical problems, the entry into the computer programming of complex dynamic general equilibrium models may not be easy. This book and its accompanying web page is particularly designed for those students with little or no prior computing experience. We start from the scratch and deliberately concentrate on models that are formulated in discrete time so that we are able to bypass the technical complexities that arise when stochastic elements are introduced into continuous time optimizing models. Numerical methods are introduced one after the other and every new method is illustrated with the help of an example. For all examples and applications, the student can download the source code from our homepage `www.wiwi.uni-augsburg.de/vwl/maussner/`. The code is available either in GAUSS or FORTRAN or both. The former computer language is almost identical to MATBLAB and can be translated without any effort. This way, the reader of this book can easily learn advanced programming techniques and, starting from very simple problems, she or he learns to apply them to more complex models, for example, a stochastic growth model with heterogeneous households.

In essence, this book makes the following contributions: 1) it tells the student in a simple way starting from a very basic level how to compute dynamic general equilibrium models. The emphasis is not on formal proofs, but rather on applications with codes and algorithms. Students should be able to start to program their own applications right away. Only some prior knowledge of statistics, linear algebra, and analysis is necessary. The relevant material from numerical analysis is gathered in a sperate chapter for those readers who are unfamiliar with these techniques. 2) The book also emphasizes some problems of the practitioner that have only received limited if any attention at all in the recent textbook literature. For example, we make an extensive effort to discuss the problem of finding a good initial value for the policy function in complex models so that the algorithm converges to the true solution. Likewise, we discuss the problem of modeling the dynamics of the distribution of the individual state variable in heterogeneous-agent economies in detail. Like econom-

rics, for example, numerical analysis is also as much an art as a science, and a young researcher in this field may often wonder why his or her particular computer program does not converge to an equilibrium value or fails to produce a sound solution. In other words, experience is important for the solution of numerical problems and our aim is to share as many as possible of our valuable practical knowledge. 3) Our applications also reflect recent research from the field of business cycle theory. For example, we compute the standard RBC model, monetary business cycle models, or the business cycle dynamics of the asset market. For this reason, the book is also valuable to both the student and the researcher of business cycles. 4) The second part of the book is devoted to the application of numerical methods to the computation of heterogeneous-agent economies. In particular, we consider the heterogeneous-agent extension of the stochastic growth model on the one hand and the overlapping generations model on the other hand. For this reason, the book is also interesting to researchers both in the field of (income and wealth) distribution theory and in the field of public finance.

The book is aimed at graduate students or advanced undergraduates. It may be used for both class-room and self study. It contains a great deal of new research both in the field of computational economics and in the field of macroeconomic theory. The book consists of three parts. Part I studies methods in order to compute representative-agent economies, Part II looks at heterogeneous-agent economies, while we collected numerical and other mathematical tools in part III. We appreciate that this book cannot easily be covered in one semester, but one can conveniently choose parts of it as the basis of a one-semester course. For example, a course on computational methods in business cycle theory may choose the chapters 1 through 3 or 4 where we covered the methods that we judge to be most useful for the computation of representative-agent business cycle and growth models. Chapter 1 introduces the basic techniques to solve the stochastic growth model. Among other, dynamic programming and the numerical techniques of value function iteration and extended deterministic path are presented. A detailed description of numerical tools from

the field of non-linear equations, approximation theory, differential and integration theory or numerical optimization is delegated to chapter 8 that, together with chapter 9 on other mathematical tools, constitutes the part III of the book. Chapter 2 reviews local approximation methods which have been predominantly applied in the analysis of business cycle models. Chapter 3 and 4 cover the methods of parameterized expectations and projection methods, respectively, and the instructor may pick either one or both methods for his course.

Graduate students with prior knowledge of numerical analysis may use chapters 5 through 7 for an introduction to the computation of heterogeneous-agent economies and the theory of income distribution. In chapter 5 and 6, we compute the stationary equilibrium and the dynamics of the distribution function for heterogeneous-agent extensions of the stochastic growth model, respectively. In chapter 7, we look at overlapping generations models. More specifically, we compute the dynamics of the Auerbach-Kotlikoff model and also consider the business cycle dynamics of the stochastic overlapping generations model. Therefore, a one-semester course in computational public finance that is aimed at the computation of Auerbach-Kotlikoff models can be based on chapters 1, 2, and 7.

The field of computational economics is vast, and we do not pretend to survey it. Fortunately, there are several other recent good textbooks that are complementary to ours. Kenneth Judd (1998) is giving a comprehensive survey of computational economics and remains the standard reference, while Miranda and Fackler (2002) have written a book that, like ours, is more directed towards the illustration of examples and algorithms, while their focus, however, is more on continuous time models. Marimon and Scott (1999), finally, have edited a textbook that also illustrates methods in order to compute the stochastic growth model that we have not covered in this book, for example the finite-element method. The textbook by Ljungqvist and Sargent (2000) on recursive macroeconomic theory and the monograph by Stokey and Lucas (1989) on recursive methods may serve as a helpful reference for the economic theory applied in

this book. The presentation, however, is self-contained and the reading of the book is possible without the consultation of other material.

Finally we would like to thank a large number of individuals. This book was written during 2000-2004. We would like to thank students in graduate classes in monetary economics and computational economics that were taught at the universities of Augsburg, Bamberg, Innsbruck and Munich. We received useful comments from Ken Judd, Paul McNelis, José-Victor Ríos-Rull, and Mark Trede. For particular assistance in the preparation of the present text, including critical comments on several drafts, helpful suggestions, and the preparation of the index, we like to thank Jürgen Antony, André de Beisac, Hans-Helmut Bünning, Michael Holsteuer, Nikolai Hristov, Jana Kremer, Dominik Menno, and Sotir Trambev. Burkhard Heer kindly acknowledges support from the German Science Foundation (Deutsche Forschungsgemeinschaft DFG) during his stay at Georgetown University and Stanford University.

October 2004

Burhard Heer, Bolzano, and Alfred Maußner, Augsburg

Table of Contents

Part III. Tools

List of Figures

Part I

Representative Agent Models

Chapter 1

Basic Models and Elementary Algorithms

Overview. This chapter introduces you to the framework of dynamic general equilibrium (DGE) models and to elementary algorithms for solving those models.

The most basic DGE model is the so called Ramsey model, where a single consumer-producer chooses an utility maximizing consumption profile. We begin with the deterministic, finite horizon version of this model, derive the set of first-order conditions (FOCs) that characterizes its solution, and introduce a simple algorithm that numerically solves this set of equations. Then, we consider the infinite horizon version of this model. We characterize its solution along two lines: the Euler equation approach delivers a set of difference equations that determine the optimal time path of consumption; the dynamic programming approach delivers a policy function that relates the agent's choice of current consumption to his stock of capital. Accordingly, we present two different numerical algorithms: the forward iteration approach works on the Euler equations and the value function iterations approach derives an approximation to the policy function. These techniques are readily extended to the stochastic version of the infinite horizon Ramsey model that we introduce in Section 1.3 along with the respective numerical algorithms to solve this model.

After that, you are equipped to study a model that has become the benchmark DGE model in business cycle analysis. We introduce this model in Section 1.4 and use it in Section 1.5 to illustrate the problems of parameter choice and model evaluation.

1.1 The Deterministic Finite Horizon Ramsey Model and Non-linear Programming

1.1.1 The Ramsey Problem

In 1928 Frank Ramsey, a young mathematician, posed the problem "How much of its income should a nation save?"[1] and developed a dynamic model to answer this question. Though greatly praised by Keynes,[2] it took almost forty years and further papers by David Cass (1965), Tjalling Koopmans (1965), and William Brock and Leonard Mirman (1972) before Ramsey's formulation stimulated macroeconomic theory. Today, variants of his dynamic optimization problem are the cornerstones of most models of economic fluctuations and growth.

At the heart of the Ramsey problem there is an economic agent producing output from labor and capital who must decide how to split production between consumption and capital accumulation. In Ramsey's original formulation, this agent was a fictitious planning authority. Yet, we may also think of a yeoman growing corn or of a household, who receives wage income and dividends and buys stocks.

In the following we use the farmer example to develop a few basic concepts. Thus, let time be divided into intervals of unit length indexed by $t = 0, 1, \ldots$. Let K_t and N_t denote the amounts of seed and labor available in period t and assume that they produce the amount Y_t of corn according to

$$Y_t = F(N_t, K_t). \tag{1.1}$$

The properties of the production function F that we require are the usual ones:

[1] Ramsey (1928), p. 543.

[2] Keynes (1930) wrote:

> *... one of the most remarkable contributions to mathematical economics ever made, both in respect of the intrinsic importance and difficulty of its subject, the power and elegance of the technical methods employed, and the clear purity of illumination with which the writer's mind is felt by the reader to play about its subject.*

1) there is no free lunch: $0 = F(0, 0)$,
2) F is strictly increasing in both of its arguments,
3) concave (i.e. we rule out increasing returns to scale),
4) and twice continuously differentiable.

At each period the farmer must decide how much corn to produce, to consume and to put aside for future production. Note, that the amount of next period's seed equals the farmer's future stock of capital K_{t+1}. His choice of consumption C_t and investment is bounded by current production, $C_t + K_{t+1} \leq Y_t$, and aims at maximizing the utility function

$$U(C_0, C_1, \ldots, C_T),$$

where T denotes the farmer's planning horizon. Since leisure does not appear in this function, we shall assume that the farmer works a given number of hours N each period.

In the farmer example seed used for growing corn is not available for future sowing. In other words, capital depreciates fully. When we think of capital in terms of machines, factories, or, even more generally, human knowledge, this is not necessarily so. In these instances the resource constraint is given by $Y_t + (1 - \delta)K_t \geq C_t + K_{t+1}$, where $\delta \in [0, 1]$ is the rate of capital depreciation. In the following, notation will become a bit simpler if we define the production function to include any capital left after depreciation and drop the constant N:

$$f(K_t) := F(N, K_t) + (1 - \delta)K_t. \tag{1.2}$$

Since production without seed is impossible, we assume $f(0) = 0$, while the other properties of F carry over to f.

We are now in the position to state the finite horizon deterministic Ramsey problem formally as follows:

$$\begin{aligned}
&\max_{(C_0,\ldots,C_T)} \quad U(C_0, \ldots, C_T) \\
&\text{s.t.} \\
&\left.\begin{array}{rcl} K_{t+1} + C_t & \leq & f(K_t), \\ 0 & \leq & C_t, \\ 0 & \leq & K_{t+1}, \end{array}\right\} \; t = 0, \ldots, T, \\
&K_0 \quad \text{given.}
\end{aligned} \tag{1.3}$$

In this problem, there is no uncertainty: the farmer knows in advance how much corn he will get when he plans to work N hours and has K_t pounds of seed. Furthermore, he is also sure as to how he will value a given sequence of consumption $\{C_t\}_{t=0}^{T}$. Therefore, we label this problem deterministic. Since we assume $T < \infty$, this is a finite horizon problem.

1.1.2 The Kuhn-Tucker Theorem

Problem (1.3) is a standard non-linear programming problem. The famous Kuhn-Tucker theorem provides a set of necessary and sufficient conditions for a solution to exist. Recall the following (cf. SUNDARAM, 1996, Theorem 7.16, p. 187f.):

Theorem 1.1.1 (Kuhn-Tucker) *Let f be a concave C^1 function mapping U into $\mathbb{R}$, where $U \subset \mathbb{R}^n$ is open and convex. For $i = 1, \ldots, l$, let $h_i : U \to \mathbb{R}$ be concave C^1 functions. Suppose there is some $\bar{\mathbf{x}} \in U$ such that*

$$h_i(\bar{\mathbf{x}}) > 0, \quad i = 1, \ldots, l.$$

Then $\mathbf{x}^*$ maximizes f over $\mathscr{D} = \{\mathbf{x} \in U | h_i(\mathbf{x}) \geq 0, i = 1, \ldots l\}$ if and only if there is $\boldsymbol{\lambda}^* \in \mathbb{R}^l$ such that the Kuhn-Tucker first-order conditions hold:

$$\begin{aligned} &\frac{\partial f(\mathbf{x}^*)}{\partial x_j} + \sum_{i=1}^{l} \lambda_i^* \frac{\partial h_i(\mathbf{x}^*)}{\partial x_j} = 0, && j = 1, \ldots n, \\ &\lambda_i^* \geq 0, && i = 1, \ldots, l, \\ &\lambda_i^* h_i(\mathbf{x}^*) = 0, && i = 1, \ldots, l. \end{aligned}$$

It is easy to see that problem (1.3) fits this theorem if the utility function U and the production function f are strictly concave, strictly increasing, and twice continuously differentiable.[3] Applying Theorem 1.1.1 to problem (1.3) provides the following first-order conditions:[4]

[3] Actually, much weaker assumptions are needed in the present context. Yet for future extensions of the model these stronger conditions will be required.

[4] As usual, a prime denotes the first (two primes the second) derivative of a function $f(x)$.

$$0 = \frac{\partial U(C_0, \ldots, C_T)}{\partial C_t} - \lambda_t + \mu_t, \qquad t = 0, \ldots, T, \tag{1.4a}$$

$$0 = -\lambda_t + \lambda_{t+1} f'(K_{t+1}) + \omega_{t+1}, \; t = 0, \ldots T-1, \tag{1.4b}$$

$$0 = -\lambda_T + \omega_{T+1}, \tag{1.4c}$$

$$0 = \lambda_t \left(f(K_t) - C_t - K_{t+1} \right), \quad t = 0, \ldots T, \tag{1.4d}$$

$$0 = \mu_t C_t, \qquad t = 0, \ldots T, \tag{1.4e}$$

$$0 = \omega_{t+1} K_{t+1}, \qquad t = 0, \ldots T, \tag{1.4f}$$

where λ_t is the Lagrangean multiplier attached to the resource constraint of period t,

$$f(K_t) - C_t - K_{t+1} \geq 0,$$

and where μ_t and ω_{t+1} are the multipliers related to the non-negativity constraints on C_t and K_{t+1}, respectively. The multipliers value the severeness of the respective constraint. A constraint that does not bind has a multiplier of zero. For example, if $C_t > 0$ then (1.4e) implies $\mu_t = 0$. If we want to rule out corner solutions, i.e., solutions where one or more of the non-negativity constraints bind, we need to impose an additional assumption. In the present context this assumption has a very intuitive meaning: the farmer hates to starve to death in any period. Formally, this translates into the statement

$$\frac{\partial U(C_0, \ldots, C_T)}{\partial C_t} \to \infty \text{ if } C_t \to 0 \text{ for all } t = 0, \ldots, T.$$

As a consequence, $C_t > 0$ for all $t = 1, \ldots, T$, $\mu_t = 0$ (from (1.4e)), and the Lagrangean multipliers λ_t equal marginal utility of consumption in period t and, thus, are also strictly positive:

$$\frac{\partial U(C_0, \ldots, C_t)}{\partial C_t} = \lambda_t.$$

Condition (1.4d), thus, implies that the resource constraint always binds. Furthermore, since we have assumed $f(0) = 0$, positive consumption also requires positive amounts of seed $K_t > 0$ through period T. However, the farmer will consume his entire crop in the

last period of his life, since any seed left reduces his lifetime utility. More formally, this result is implied by equations (1.4f) and (1.4c), which yield $\lambda_T K_{T+1} = 0$. Taking all pieces together, we arrive at the following characterization of an optimal solution:

$$K_{t+1} = f(K_t) - C_t, \tag{1.5a}$$

$$\frac{\partial U(C_0, \ldots C_T)/\partial C_t}{\partial U(C_0, \ldots C_T)/\partial C_{t+1}} = f'(K_{t+1}). \tag{1.5b}$$

The lhs of equation (1.5b) is the marginal rate of substitution between consumption in two adjacent periods. It gives the rate at which the farmer is willing to forego consumption in t for consumption one period ahead. The rhs provides the compensation for an additional unit of savings: the increase in future output.

1.1.3 Numerical Solutions

In our simple model we gain not many further insights if we numerically solve for the optimal consumption profile. Yet, there are models closely related to this one where simulation is more or less the only analytical tool. For this reason we use use the model to introduce two strategies to obtain numerical solutions.

Direct versus Indirect Methods. There are, in principle, two strategies to numerically solve non-linear programming problems like (1.3). The first approach employs ready-to-use software for maximizing a function subject to constraints. The web site `http://www-fp.mcs.anl.gov/otc/Guide/SoftwareGuide/index.html` and the book by Moré and Wright (1993) are helpful guides to these programs. Yet, sophisticated software is costly and simpler routines shipped with standard programming languages may not work. As an example, consider the specification of preferences and technology given in Example 1.1.1. The GAUSS program `Ramsey1a.g` uses the non-linear programming routine `sqpSolve` to solve this example. The routine requires the user to define the utility function and the set of constraints and to provide starting values for the $2(T+1)$ unknowns $C_0, \ldots, C_T, K_1, \ldots K_{T+1}$. Even with starting values close to the solution, the routine is not able to find an approximate solution.

A second, indirect approach is to solve the first-order conditions of the problem. This, too, may require sophisticated software for solving non-linear equations. But, as we shall show, a simple modification of the well known Newton-Raphson method works amazingly well. We present this method in Section 8.5. Here we apply it to solve Example 1.1.1.

Example 1.1.1

Let U be given by a constant elasticity of substitution function

$$U(C_0, \ldots, C_T) := \left\{ \sum_{t=0}^{T} C_t^{\varrho} \right\}^{1/\varrho}, \qquad \varrho \in (-\infty, 1],$$

and define $f(K_t) := K_t^{\alpha},\ \alpha \in (0,1)$. Given this specification, equations (1.5) become

$$\begin{aligned} K_{t+1} &= K_t^{\alpha} - C_t, && t = 0, 1, \ldots, T, \\ \left[\frac{C_t}{C_{t+1}}\right]^{1-\varrho} \alpha K_{t+1}^{\alpha-1} &= 1, && t = 0, 1, \ldots, T-1. \end{aligned}$$

If we eliminate consumption in the second set of equations using the first $T+1$ equations we arrive at a set of T non-linear equations in the T unknowns $(K_1, K_2, \ldots, K_T)$:

$$\begin{aligned} 0 &= \left(\frac{K_1^{\alpha} - K_2}{K_0^{\alpha} - K_1}\right)^{1-\varrho} - \alpha K_1^{\alpha-1}, \\ 0 &= \left(\frac{K_2^{\alpha} - K_3}{K_1^{\alpha} - K_2}\right)^{1-\varrho} - \alpha K_2^{\alpha-1}, \\ &\vdots \\ 0 &= \left(\frac{K_T^{\alpha}}{K_{T-1}^{\alpha} - K_T}\right)^{1-\varrho} - \alpha K_T^{\alpha-1}. \end{aligned} \tag{1.6}$$

The computer code that solves this problem using the GAUSS programming language is stored in the file `Ramsey1b.g`. The problem is solved for $T = 59$, $\alpha = 0.35$ $\rho = 0.5$, and $K_0 = 0.1$. Our choice of the starting values rests on the following consideration.

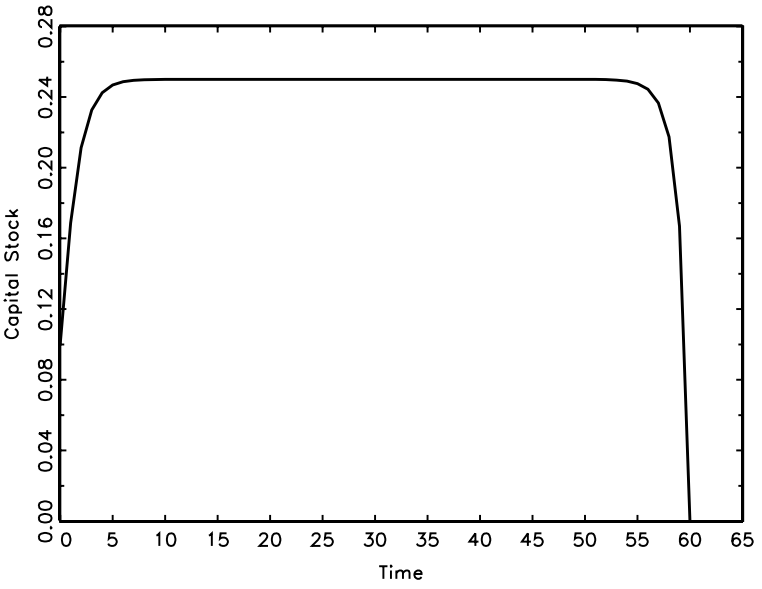

Figure 1.1: Time Path of the Capital Stock in a Finite Horizon Ramsey Model

Since the marginal utility of consumption declines, the farmer prefers a smooth profile. The smoothest profile is the one where the farmer consumes the same amount in each period. In this case equations (1.6) imply a unique capital stock $K^* = \alpha^{1/(1-\alpha)}$ and a related unique level of consumption $C^* = (K^*)^\alpha - K^*$. The farmer cannot choose this profile for three reasons: Firstly, his initial stock of capital K_0 probably differs from K^*, secondly, his desired terminal stock of capital equals zero, and thirdly, the stock of capital changes only with the passage of time but is fixed within any period. Therefore, the choice of the stationary profile $C^*, \ldots, C^*, K^*, \ldots, K^*$ is at best an approximation to the solution. This is confirmed by the actual solution shown in Figure 1.1. On a Pentium III, 796 Mhz Personal Computer it takes only 15 hundredths of a second to compute this solution. The capital stock quickly builds up, stays constant most of the time, and drops to zero within a few periods towards the end of the planning period. It is like driving by car from city A to city B: you minimize travel-

ling time if you drive on the motorway or autobahn that connects the two cities and not on a winding, ordinary road. Therefore, this behavior of the solution path is sometimes referred to as the turnpike property.

1.2 The Deterministic Infinite Horizon Ramsey Model and Dynamic Programming

1.2.1 Recursive Utility

The simple structure of the optimality conditions (1.6) derives from the constant elasticity of substitution (CES) utility function, which implies that the marginal rate of substitution between adjacent dates depends only on consumption at those dates. As a consequence, consumption at the present date depends only on the present capital stock and on future consumption, but is independent of past consumption. Under a more general utility function the marginal rate of substitution depends upon the entire time profile of consumption. In that case it may be difficult to solve the system of necessary conditions. The separability of past decisions from present and future choices is the key element for dynamic programming techniques to work. The time additive separable (TAS) utility function, which may be defined recursively from

$$U_t = u(C_t) + \beta U_{t+1}, \quad \beta \in (0,1), \tag{1.7}$$

shares this property with the CES function. In this definition β is a discount factor and $\beta^{-1} - 1$ is known as the pure rate of time preference. The function $u : [0, \infty) \to \mathbb{R}$ is called the one-period, current-period, or felicity function. Usually we assume that u is strictly increasing, strictly concave and twice continuously differentiable.

The solution to the finite horizon Ramsey model depends upon the chosen terminal date T. Yet, in as far as we want to portray the behavior of the economy with Ramsey type models there is no natural final date T. As a consequence, most models extend the planning horizon into the indefinite future by letting $T \to \infty$.

Iterating on (1.7) we arrive at the following definition of the utility function

$$U_t = \sum_{s=0}^{\infty} \beta^s u(C_{t+s}). \tag{1.8}$$

If we want to rank consumption streams according to this criterion function, we must ensure that the sum on the rhs is bounded from above, i.e., $U_t < \infty$ for every admissible sequence of points $C_0, C_1, \ldots$. This will hold, if the growth factor of one-period utility u is smaller than $1/\beta$. Consider the Ramsey problem (1.3) with infinite time horizon:

$$\begin{aligned} &\max_{C_0, C_1, \ldots} \quad U_0 = \sum_{t=0}^{\infty} \beta^t u(C_t) \\ &\text{s.t.} \\ &\left.\begin{array}{rcl} K_{t+1} + C_t & \leq & f(K_t), \\ 0 & \leq & C_t, \\ 0 & \leq & K_{t+1}, \end{array}\right\} \; t = 0, 1, \ldots, \\ &K_0 \quad \text{given.} \end{aligned} \tag{1.9}$$

In this model we do not need to assume that the one-period utility function u is bounded. It is sufficient to assume that the economy's resources are finite. In a dynamic context this requires that there is an upper bound on capital accumulation, i.e., there is $\bar{K}$ such that for each $K > \bar{K}$ output is smaller than needed to maintain K:

$$K_t > \bar{K} \Rightarrow K_{t+1} < K_t.$$

For instance, using the production function from Example 1.1.1 gives:

$$K \leq K^{\alpha} \Rightarrow \bar{K} = 1^{1/(\alpha-1)} = 1.$$

As a consequence, any admissible sequence of capital stocks is bounded by $K^{\max} := \max\{\bar{K}, K_0\}$. Figure 1.2 makes that obvious: consider any point to the left of $\bar{K}$ such as K_1 and assume that consumption equals zero in all periods. Then, the sequence of capital stocks originating in K_1 approaches $\bar{K}$. Similarly, the sequence starting in K_2 approaches $\bar{K}$ from the right.

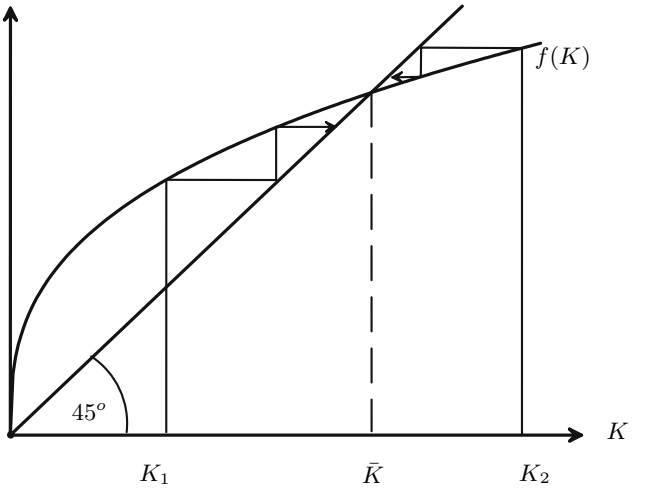

Figure 1.2: Boundedness of the Capital Stock

1.2.2 Euler Equations

There are two approaches to characterize the solution to the Ramsey problem (1.9). The first is an extension of the Kuhn-Tucker method[5] and the second is dynamic programming.[6] According to the first approach necessary conditions may be derived from maximizing the following Lagrangean function with respect to $C_0, C_1, \ldots, K_1, K_2, \ldots$:

$$\mathscr{L} = \beta^t \sum_{t=0}^{\infty} \Big[u(C_t) + \lambda_t \left(f(K_t) - C_t - K_{t+1} \right) + \mu_t C_t + \omega_{t+1} K_{t+1} \Big].$$

Note that in this expression the Lagrangean multipliers λ_t, μ_t, and ω_{t+1} refer to period t values. Period 0 values are given by $\beta^t \lambda_t$, $\beta^t \mu_t$, and $\beta^t \omega_{t+1}$. The first-order conditions for maximizing $\mathscr{L}$ are given by:

$$u'(C_t) = \lambda_t - \mu_t, \tag{1.10a}$$

[5] See, e.g., Chow (1997), Chapter 2 and Romer (1991).
[6] Here, the standard reference is Stokey and Lucas (1989), Chapter 4.

$$\lambda_t = \beta\lambda_{t+1} f'(K_{t+1}) + \omega_{t+1}, \tag{1.10b}$$
$$0 = \lambda_t(f(K_t) - C_t - K_{t+1}), \tag{1.10c}$$
$$0 = \mu_t C_t, \tag{1.10d}$$
$$0 = \omega_{t+1} K_{t+1}. \tag{1.10e}$$

We continue to assume that the farmer hates starving to death, $\lim_{C\to 0} u'(C) = \infty$, so that the non-negativity constraints never bind. Since u is strictly increasing in its argument, the resource constraint always binds. Therefore, we can reduce the first-order conditions to a second order difference equation in the capital stock:

$$\frac{u'(f(K_t) - K_{t+1})}{u'(f(K_{t+1}) - K_{t+2})} - \beta f'(K_{t+1}) = 0. \tag{1.11}$$

This equation is often referred to as the Euler equation, since the mathematician Leonhard Euler (1707-1783) first derived it from a continuous time dynamic optimization problem. To find the unique optimal time path of capital from the solution to this functional equation we need two additional conditions. The period $t = 0$ stock of capital K_0 provides the first condition. The second condition is the so called transversality condition, which is the limit of the terminal condition $\lambda_T K_{T+1} = 0$ from the finite horizon Ramsey problem (1.3). It requires

$$\lim_{t\to\infty} \beta^t \lambda_t K_{t+1} = 0, \tag{1.12}$$

i.e., the present value of the terminal capital stock must approach zero. In the present context, condition (1.12), is a necessary condition,[7] as are conditions (1.11).

1.2.3 Dynamic Programming

We have already seen that the time additive utility function implies that consumption at dates $t = 1, 2, \ldots$ depends upon the capital stock at t but not on consumption before t. All information that is necessary to determine an optimal sequence of capital

[7] See Kamihigashi (2002).

stocks and a related sequence of consumption streams, thus, is incorporated in the stock of capital at date t. Assume that we already know the solution and are able to evaluate U_0. Let $v(K)$ denote life-time utility along the optimal sequence of capital stocks beginning with the given date $t = 0$ capital stock K,[8] and let K^* denote the date $t = 1$ capital stock in the optimal sequence so that $v(K^*)$ is life-time utility at date $t = 1$. Obviously, K^* must solve the following one-stage maximization problem

$$v(K^*) = \max_{0 \leq K' \leq f(K)} \quad u(f(K) - K') + \beta v(K'). \tag{1.13}$$

The first-order condition for this problem is

$$u'(f(K) - K^*) = \beta v'(K^*). \tag{1.14}$$

If we knew the function v, we could compute the solution K^* as a function g of K:

$$K^* = g(K).$$

Since v does not depend on t, the function g must be time invariant, too. In this context g is referred to as the policy function or feed-back rule and v as the value function. Equation (1.13), the so called Bellman equation , is a functional equation in the unknown function v. The mathematical theory of dynamic programming deals with the existence, the properties, and the construction of v and g. Given that both $u(C)$ and $f(K)$ are strictly increasing, strictly concave and twice continuously differentiable functions of their respective arguments C and K, and that there exists a maximum sustainable capital stock $\bar{K}$ as explained above, one can prove the following results:[9]

1) The function v exists, is differentiable, strictly increasing, and strictly concave.
2) The policy function g is increasing and differentiable.

[8] In the following discussion we need only distinguish the given capital stock K from next period's capital stock K'. Therefore, we drop the time indices and use a prime to denote the latter.

[9] See, e.g., Harris (1987), pp 34-45.

3) The function v is the limit of the following sequence of steps $s = 0, 1, \ldots$:

$$v^{s+1}(K) = \max_{0 \le K' \le f(K)} \quad u(f(K) - K') + \beta v^s(K'),$$

with $v^0 = 0$.

We illustrate these results in Example 1.2.1 and use them to compute the derivative of v so that we can solve condition (1.14).

Example 1.2.1

Let the one-period utility function u and the production function f be given by

$$u(C) := \ln C,$$
$$f(K) := K^\alpha, \quad \alpha \in (0, 1),$$

respectively. In Appendix 1 we use iterations over the value function to demonstrate that the policy function $K^* = g(K)$ that solves the Ramsey problem (1.9) is given by

$$K^* = \alpha\beta K^\alpha.$$

Furthermore, the value function is linear in $\ln K$ and given by

$$v(K) = a + b \ln K,$$
$$a := \frac{1}{1-\beta}\left[\ln(1-\alpha\beta) + \frac{\alpha\beta}{1-\alpha\beta}\ln\alpha\beta\right], \quad b := \frac{\alpha}{1-\alpha\beta}.$$

Thus, let

$$K^* = g(K) := \operatorname*{argmax}_{0 \le K' \le f(K)} \quad u(f(K) - K') + \beta v(K'),$$

and consider the identity

$$v(K) = u(f(K) - g(K)) + \beta v(g(K)).$$

Differentiation with respect to K yields

$$v'(K) = u'(C)\left(f'(K) - g'(K)\right) + \beta v'(K^*)g'(K) = 0.$$

Using the first-order condition (1.14), we find

$$v'(K) = u'(C)f'(K). \tag{1.15}$$

Finally, letting $C' = F(N, K') - K''$ denote next period's consumption, (1.14) and (1.15) may be combined to yield

$$1 = \beta \frac{u'(f(K') - K'')}{u'(f(K) - K')} f'(K'),$$

which is identical to the Euler equation (1.11), except that we used primes instead of time indices.

1.2.4 The Saddle Path

Before we proceed to numerical solutions, we use the phase diagram technique to characterize the solution of the Euler equations (1.11). Substituting the resource constraint $C_t = f(K_t) - K_{t+1}$ into (1.11), the following first-order, non-linear system of difference equations governs the optimal time path of capital accumulation:

$$K_{t+1} = f(K_t) - C_t, \tag{1.16a}$$

$$1 = \beta \frac{u'(C_{t+1})}{u'(C_t)} f'(K_{t+1}). \tag{1.16b}$$

Together with the initial capital stock K_0 and the transversality condition (1.12) these equations determine a unique solution. We use Figure 1.3, to construct it.[10]

Consider first the locus of all pairs (K_t, C_t) along which consumption does not change, i.e., $C_t = C_{t+1}$. According to equation (1.16b) this happens when the capital stock reaches K^*, given by

[10] The time paths shown in this figure are obtained from a numerical simulation. Since they represent the solution of a system of difference equations and not of a system of differential equations they are connected line segments rather than smooth curves.

$$\frac{1}{\beta} = f'(K^*).$$

Since to the right (left) of K^* the marginal product of capital is smaller (larger) than $1/\beta$, consumption decreases (increases) within that region. The vertical arrows in Figure 1.3 designate that behavior.

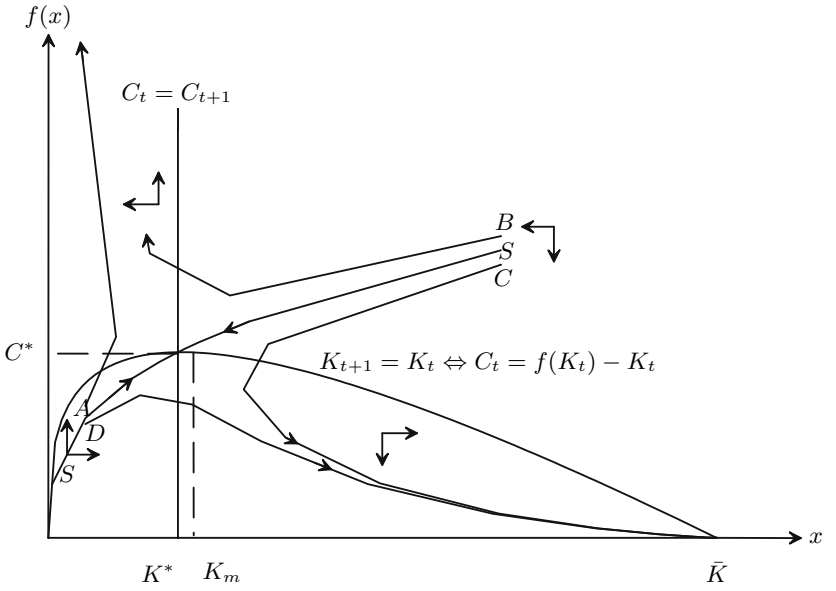

Figure 1.3: Phase Diagram of the Infinite Horizon Ramsey Model

Next, consider the locus of all pairs (K_t, C_t) along which the capital stock does not change, i.e, $K_t = K_{t+1}$. From (1.16a) this yields

$$C_t = f(K_t) - K_t.$$

The graph of this function equals the vertical distance between the function $f(K_t)$ and the 45–degree line. Thus, it starts at the origin, attains a maximum at K_m, defined by $1 = f'(K_m)$, and cuts the K-axis at $\bar{K}$. Points above (below) that locus have a higher (smaller) consumption and, thus, the capital stock declines (increases) in that region, as shown by the horizontal arrows.

The optimal path of capital accumulation is given by the line segment labeled SS, the so called saddle path. Points on that locus converge towards the stationary equilibrium at (K^*, C^*). Along time paths starting at points like A and B the capital stock reaches zero in finite time. To see this, note first that paths originating to the right of K^* but above the locus SS will eventually enter the region to the left of the stationary capital stock. Within this region consumption is increasing and the stock of capital is decreasing. This leaves two possibilities: (1) consumption approaches infinity and the stock of capital approaches zero or (2) the stock of capital becomes zero in finite time. Consider the first case: If consumption approaches an upper limit (possibly ∞), the ratio of the marginal utilities in (1.16b) approaches unity. At the same time the marginal product of capital becomes large and larger. If $f'(K) \to \infty$ for $K \to 0$ condition (1.16b) will eventually be violated.[11] Now consider the case where $K_T = 0$ for some $T < \infty$. Since consumption at $T + 1$ is then necessarily equal to zero the term $u'(C_{T+1})f'(0)$ in (1.16b) jumps to ∞, again violating this condition. Now consider the region to the right of the stationary capital stock. Time paths like D, which originate to the left of K^* and below SS, will approach that region. Within this region they are attracted to the point $\bar{K}$. Yet, along every such path the growth rate of the marginal utility of one-period consumption increases at a rate greater than β, since to the right of K_m the marginal product of capital is less than one:

$$\frac{u'(C_{t+1})}{u'(C_t)} = \frac{1}{\beta f'(K_t)} > \frac{1}{\beta}.$$

Hence, every such path violates the transversality condition

$$\lim_{t\to\infty} \beta^t u'(C_t) K_{t+1} = 0.$$

The line segment SS is the graph of the policy function h that relates the capital stock at each date t to the optimal choice of consumption at this date: $C_t = h(K_t)$. The policy function

[11] A much weaker condition is that the upper bound of $f'(K)$ is larger than β^{-1}.

g is then given by $K_{t+1} = g(K_t) := f(K_t) - h(K_t)$. Given the initial capital stock K_0 the optimal strategy is to choose $C_0 = h(K_0)$ and then to iterate either over the Euler equations (1.16) or, equivalently, over the policy functions h and g. Yet, how do we derive the function h? Unfortunately, the Ramsey model (1.9) admits an analytical solution of the policy function only in two special cases.[12] Both of them require a logarithmic one-period utility function. In the first case output (including depreciated capital, as defined in (1.2)) must be given by the Cobb-Douglas function

$$Y_t + (1-\delta)K_t = AK_t^\alpha, \quad A > 0,\ \alpha \in (0,1).$$

The second case assumes a Cobb-Douglas function for value added,

$$Y_t = AK_t^\alpha, \quad A > 0,\ \alpha \in (0,1),$$

and assumes adjustment-costs of investment that give raise to the following transition equation for capital:

$$K_{t+1} = K_t^{1-\delta} I_t^\delta, \quad I_t = Y_t - C_t.$$

In every other case the policy function or the optimal time path must be approximated numerically.

1.2.5 Numerical Solutions

Strategies. There are two different approaches to approximate solutions of the infinite horizon Ramsey model (1.9) numerically. The dynamic programming characterization of the solution gives raise to methods that compute approximations of the value and of the policy function. Approximate solutions of the Euler equations provide another route of attack. In this chapter we consider value function iteration on a discrete version of the original problem as a member of the former class of methods. From the latter class of methods we consider backward as well as forward iteration.

[12] See, e.g., McCallum (1989), p. 21.

Backward Iteration. In the previous paragraph, we have seen that the optimal path approaches a stationary solution (K^*, C^*). A straightforward procedure is to compute that path by backward iteration of the Euler equations (1.16) from that point onward. However, at (K^*, C^*), per definition, $K_{t+1} = K_t$ and $C_{t+1} = C_t$ for all t. Thus, we must start from a point very close to the optimal path near (K^*, C^*). Yet, if we just pick an arbitrary point near the stationary solution, we have a fair chance to get on a divergent path. To circumvent this problem we make use of a very powerful result from the mathematics of difference equations (see Section 9.1). In the present context it states that the tangent to the saddle path SS at the stationary solution (K^*, C^*) is given by the eigenvector that belongs to the stable eigenvalue of the Jacobian matrix of the system of difference equations (1.16) evaluated at (K^*, C^*).[13] Let $(1, v)'$ denote this vector and let dK denote a small change of the equilibrium capital stock K^*. Then, a point near the saddle path is given by $(K^* + dK, C^* + vdK)$, and we can start the backward iteration at this point.

In the case of Example 1.2.1 the system of difference equations (1.16) is:

$$K_{t+1} = K_t^\alpha - C_t, \tag{1.17a}$$

$$C_{t+1} = \alpha\beta C_t (K_t^\alpha - C_t)^{\alpha-1}, \tag{1.17b}$$

with the stationary solution at

$$K^* = (\alpha\beta)^{1/(1-\alpha)} \quad \text{and} \quad C^* = \left(\frac{1-\alpha\beta}{\alpha\beta}\right)(\alpha\beta)^{1/(1-\alpha)}.$$

At (K^*, C^*) the Jacobian matrix of (1.17) is

$$J = \begin{bmatrix} 1/\beta & -1 \\ \frac{1-\alpha\beta}{\alpha\beta}\frac{\alpha-1}{\beta} & 1 - \frac{1-\alpha\beta}{\alpha\beta}(\alpha-1) \end{bmatrix}.$$

It can be shown analytically that J has one eigenvalue smaller than one and one eigenvalue exceeding one.[14] In the GAUSS program `Ramsey2.g` we use the built-in procedure `gradp` to compute

[13] See Section 8.1 and Section 9.1, if you are unfamiliar with the notions of the Jacobian of a system of non-linear difference equations, of eigenvalues, and of eigenvectors.

[14] See Section 2.3.1 for a prove.

J and the procedure `eigv` to get the two eigenvectors. Since (1.17) has an analytic inverse, it is easy to iterate backwards. Figure 1.4 plots the time path that we get from this procedure when we start at a point very near to the stationary solution ($dK = -K^* 10^{-6}$) and iterate 28 times over the inverse of (1.17). We use the end point of these iterations as initial capital stock K_0 and iterate forward using the exact policy function $K_{t+1} = \alpha\beta K_t^\alpha$. As can be seen from Figure 1.4 both time paths lie indistinguishably close together.

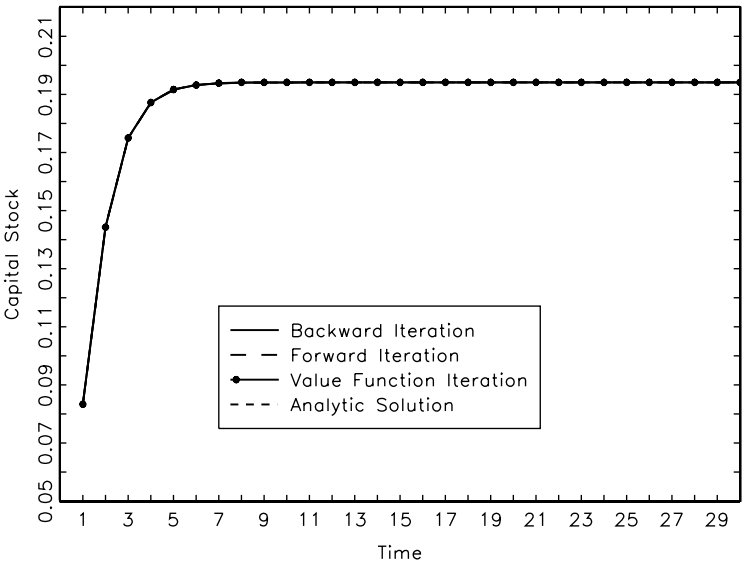

Figure 1.4: Comparison of Approximate Solutions of an Infinite Horizon Ramsey Model

Forward Iteration. A second approach to compute an approximate solution of the Euler equation (1.11) assumes a finite time horizon $T + 1$ at which the solution should be close to K^*. In the case of Example 1.2.1, we must solve the following system of non-linear equations in the T unknowns $K_1, K_2, .., K_T$:

$$0 = \left(\frac{K_1^\alpha - K_2}{K_0^\alpha - K_1}\right) - \alpha\beta K_1^{\alpha-1},$$

$$0 = \left(\frac{K_2^\alpha - K_3}{K_1^\alpha - K_2}\right) - \alpha\beta K_2^{\alpha-1},$$

$$\vdots$$

$$0 = \left(\frac{K_T^\alpha - K^*}{K_{T-1}^\alpha - K_T}\right) - \alpha\beta K_T^{\alpha-1}.$$

One can check whether the given time horizon T is too short for convergence by iterating over T: Increase T to $T' > T$ and solve the enlarged system. If the first T solutions are close to the solutions obtained for the horizon T stop; else increase T' to $T'' > T'$ and continue in this fashion until convergence has been achieved. For our example we used the modified Newton-Raphson algorithm to solve the non-linear system of equations. We set $T = 30$ and find this long enough to get close to the stationary solution. The respective solution is labeled forward iteration in Figure 1.4. It begins at the same capital stock K_0 that we used to compute the exact solution.

Value Function Iteration. The solution method which we consider in this paragraph relies on a discretization of the space to which the sequence of capital stocks is confined. It is simple to program and can be used in all situations where the state of the system is captured by a single variable. Consider, again, Example 1.2.1. We already know that any admissible sequence of capital stocks that starts in $[0, 1]$ must stay in this interval.[15] Thus, we could approximate $[0, 1]$ by an equidistant grid of n points. Since the approximation is the better the more finely meshed the grid is, we must use a large number of points n. To economize on computational time and to keep n reasonably small, the state space should be as small as possible. In our example, $[0, 1]$ is much too large. Since the optimal sequence of capital stocks monotonically approaches the stationary solution K^* from any given initial point K_0, the solution to the continuously valued problem is in $[K_0, K^*]$ or in $[K^*, K_0]$ depending on whether K_0 is smaller or greater than K^*. Note, however, that due to the discretization the stationary

[15] See our discussion on page 12.

solution of the original problem can differ from the stationary solution of the approximate problem. Therefore, we must choose a state space $[\underline{K}, \overline{K}]$ that includes $[K_0, K^*]$ (or $[K^*, K_0]$). Once the policy function has been found, we can confirm our choice. If the optimal policy hits either the upper or the lower bound of the grid, the selected state space is too small.

Having chosen the boundaries of the state space, the next decision concerns the number of grid points. The finer the grid is chosen, the better the approximation is. On the other hand, in each step of the iterations scheme we will describe in a moment, there are n^2 evaluations of the one-period utility function. Thus, computational time places a restriction on n. In the one-dimensional Example 1.2.1 this is no real problem. Yet, as soon as we turn to stochastic or higher-dimensional problems computational time may become a binding restriction.

Given a grid $\mathscr{G} = [K_1, K_2, \ldots, K_n]$, where $K_i < K_j$ for all $i < j \in \{1, 2, \ldots, n\}$, the policy function we look for is a map from points of $\mathscr{G}$ into $\mathscr{G}$ that can be represented by a vector $\mathfrak{g}$ as follows. Let $i \in 1, 2, \ldots, n$ denote the index of $K \in \mathscr{G}$, and let $j \in 1, 2, \ldots, n$ denote the index of $K' = g(K)$, where g is the policy function we wish to approximate. Then, the i–th element of $\mathfrak{g}$ is j, i.e., it is just a pointer to the optimal $K' \in \mathscr{G}$. Analogously, the value function is an n–vector $\mathfrak{v}$, whose i–th element stores the life-time utility associated with an optimal sequence of capital stocks beginning with the i–th element of $\mathscr{G}$. The value function and the policy function can be found in a number of simple steps. First, we initialize the value function. Since we know that the solution to

$$\max_{K'} \quad u(f(K) - K') + \beta \cdot 0$$

is $K' = 0$, we initialize $\mathfrak{v}_i^0$ with $u(f(K_i))\, \forall i = 1, \ldots, n$, and set $\mathfrak{g}_i^0 = i \forall i = 1, \ldots, n$. In the next step we find a new value and policy function as follows: For each $i = 1, \ldots, n$

Step 1: compute

$$\mathfrak{w}_j = u(f(K_i) - K_j) + \beta \mathfrak{v}_j^0,\ j = 1, \ldots, n.$$

Step 2: Find the index j^* such that

$$\mathfrak{w}_{j^*} \geq \mathfrak{w}_j \, \forall j = 1, \ldots, n.$$

Step 3: Set $\mathfrak{g}_i^1 = j^*$ and $\mathfrak{v}_i^1 = \mathfrak{w}_{j^*}$.

In the final step, we check whether the value function is close to its stationary solution. Let $\|\mathfrak{v}^0 - \mathfrak{v}^1\|_\infty$ denote the largest absolute value of the difference between the respective elements of $\mathfrak{v}^0$ and $\mathfrak{v}^1$. Since the value function converges linearly at the rate β, the error from choosing $\mathfrak{v}^1$ as a solution, i.e., $\|\mathfrak{v}^1 - \mathfrak{v}^*\|$, is bounded from above by $\epsilon \in \mathbb{R}_{++}$, if we stop after $\|\mathfrak{v}^0 - \mathfrak{v}^1\|_\infty \leq \epsilon(1-\beta)$ (see Section 8.4).

If one uses a standard programming language (as, e.g., C, Fortran, GAUSS, or Matlab) there is no need to care about finding the maximal element of $\mathfrak{w}$ in Step 2, since there are built-in subroutines (as. e.g., the `maxindc` command in GAUSS or the `MaxLoc` function in Fortran 95).

The algorithm that we have just described is not very smart. We can do much better, if we exploit the structure of our problem. The first thing we can do is to select the initial value function more carefully. We can save on iterations, if the initial value function is closer to its final solution. Using K^* from the continuous valued problem as our guess of the stationary solution, the stationary value function is defined by

$$v^* = u(f(K^*) - K^*) + \beta v^*,$$

and we can use $\mathfrak{v}_i^0 = u(f(K^*) - K^*)/(1-\beta)$ for all $i = 1, 2, \ldots, n$ as our initial guess.

Secondly, we can use the monotonicity of the policy function, i.e.,

$$K_i \geq K_j \Rightarrow K_i' = g(K_i) \geq K_j' = g(K_j).$$

As a consequence, once we find the optimal index j_1^* for K_1, we need no longer consider capital stocks smaller than $K_{j_1^*}$ in the search for j_2^*. More generally, let j_i^* denote the index of the maximization problem in step 2 for i. Then, for $i+1$ we evaluate $u(F(N, K_i) - K_j) + \beta \mathfrak{v}_j^0$ only for indices $j \in \{j_i^*, \ldots n\}$.

Thirdly, we can shorten the number of computations in the maximization Step 2, since the function

$$\phi(K') := u(f(K) - K') + \beta v(K') \tag{1.19}$$

is strictly concave.[16] A strictly concave function F defined over a grid of n points either takes its maximum at one of the two boundary points or in the interior of the grid. In the first case the function is decreasing (increasing) over the whole grid, if the maximum is the first (last) point of the grid. In the second case the function is first increasing and then decreasing. As a consequence we can pick the mid-point of the grid, x_m, and the point next to it, x_{m+1}, and determine whether the maximum is to the left of x_m (if $F(x_m) > F(x_{m+1})$) or to the right of x_m (if $F(x_{m+1}) > F(x_m)$). Thus, in the next step we can reduce search to a grid with about half the size of the original grid. KREMER (2001), pp. 165f, proves that search based on this principle needs at most $\log_2(n)$ steps to reduce the grid to a set of three points that contains the maximum. For instance, instead of 1000 function evaluations, binary search requires no more than 13! We describe this principle in more detail in the following algorithm:

Algorithm 1.2.1 (Binary Search)

Purpose: *Find the maximum of a strictly concave function $f(x)$ defined over a grid of n points $\mathscr{G} = [x_1, ..., x_n]$*

Steps:

Step 1: Initialize: Put $i_{min} = 1$ and $i_{max} = n$.
Step 2: Select two points: $i_l = floor(i_{min} + i_{max})/2$ and $i_u = i_l + 1$, where floor(i) denotes the largest integer less than or equal to $i \in \mathbb{R}$.
Step 3: If $f(x_{i_u}) > f(x_{i_l})$ set $i_{min} = i_l$. Otherwise put $i_{max} = i_u$.
Step 4: If $i_{max} - i_{min} = 2$, stop and choose the largest element among $f(x_{i_{min}})$, $f(x_{i_{min+1}})$, and $f(x_{i_{max}})$. Otherwise return to Step 2.

[16] Since the value function, as well as the utility and the production function, is strictly concave.

Finally, the closer the value function gets to its stationary solution, the less likely it is that the policy function changes with further iterations. So usually one can terminate the algorithm if the policy function has remained unchanged for a number of consecutive iterations. For instance, in the solution of Example 1.2.1, we use $\epsilon = 0.01$. Convergence based on that number requires 223 iterations. Yet, the policy function remains unchanged after 15 iterations.

Putting all pieces together we propose the following algorithm to solve the infinite horizon deterministic Ramsey problem via value function iteration:

Algorithm 1.2.2 (Value Function Iteration 1)

Purpose: *Find an approximate solution of the policy function for the Ramsey model* (1.9)

Steps:

Step 1: Choose a grid of n equally spaced points

$$\mathscr{G} = [K_1, K_2, \ldots, K_n],\ K_i < K_j,\ i < j = 1, 2, \ldots n.$$

Step 2: Initialize the value function: $\forall i = 1, \ldots, n$ set

$$\mathfrak{v}_i^0 = \frac{u(f(K^*) - K^*)}{1 - \beta},$$

where K^* denotes the stationary solution to the continuous-valued Ramsey problem.

Step 3: Compute a new value function and the associated policy function, $\mathfrak{v}^1$ and $\mathfrak{g}^1$, respectively: Put $j_0^* \equiv 1$. For $i = 1, 2, \ldots, n$, and j_{i-1}^* use Algorithm 1.2.1 to find the index j_i^* that maximizes

$$u(f(K_i) - K_j) + \beta \mathfrak{v}_j^0$$

in the set of indices $\{j_{i-1}^*, j_{i-1}^* + 1, \ldots, n\}$. Set $\mathfrak{g}_i^1 = j_i^*$ and $\mathfrak{v}_i^1 = u(f(K_i) - K_{j_i^*}) + \beta v_{j_i^*}^0$.

Step 4: Check for convergence: If $\|\mathfrak{v}^0 - \mathfrak{v}^1\|_\infty < \epsilon(1-\beta)$, $\epsilon \in \mathbb{R}_{++}$ *(or if the policy function has remained unchanged for a number of consecutive iterations) stop, else set* $\mathfrak{v}^0 = \mathfrak{v}^1$ *and* $\mathfrak{g}^0 = \mathfrak{g}^1$ *and return to step 3.*

The GAUSS program `Ramsey2.g` solves Example 1.2.1 using a grid of 500 equally spaced points that starts with K_0 and extends slightly over K^*. When we neglect the structure of our problem and iterate over all grid points all the time until the policy function has been unchanged in more than 30 consecutive iterations the program takes about 84 seconds (as compared to 5 hundredths of a second for the forward iteration). Algorithm 1.2.2, that we implement in the GAUSS procedure `SolveVIDc` in the file `DGE.src`, takes about 3 seconds to converge. Figure 1.4 shows that the approximate solution is close to the analytic solution, too. Yet, due to the given grid the maximal percentage deviation from the analytic solution of 0.076 is by far greater than those for both the backward and the forward iteration solution.

Linear Interpolation Between Grid Points. Our Algorithm 1.2.2 evaluates the value function at a number of discrete grid points. We have learned several methods to speed up the computation, e.g., by exploiting the concavity of the rhs of the Bellman equation or the monotonicity of the policy function. You got to know another useful programming device for faster computation in the form of the binary search algorithm 1.2.1. Yet, even with a fine grid of 500 points, the percentage deviation from the true solution can be still significant as we found in the solution of Example 1.2.1. For a multidimensional state space, Algorithm 1.2.2 may often be too inaccurate and may only serve as an initial guess. Furthermore, in Example 1.2.1, we were able to restrict the state space to a small interval containing the initial capital stock K_1 and the steady state capital stock K^*. In the second part of the book, you will encounter heterogenous-agent economies. Agents will be different with regard to their productivity and income and, therefore, accumulate different levels of savings. As a consequence, wealth is distributed unequally among agents and the state space (wealth K) cannot be confined to a small interval.

In summary, the method of value function iteration necessitates the discretization of the state space. The finer the grid, the closer we get to the true solution of the problem. As soon as the problem becomes more complex, however, the computational time becomes a binding constraint. Of course, this will be the case as soon as the number of the continuous state variables increases. Therefore, the value function iteration approach introduced above may not be very satisfactory for many purposes. However, we may be able to get a much more accurate solution, if we allow the next-period state variable to be points off the grid.

How do we accomplish this? Consider Step 3 of Algorithm 1.2.2, where we maximize the rhs of the Bellman equation (1.19) with respect to K'. Assume that K_j is this solution. Since the value function is increasing and concave, the true maximizer must lie in the interval $[K_{j-1}, K_{j+1}]$. If we were able to evaluate the rhs of the Bellman equation at all $K' \in [K_{j-1}, K_{j+1}]$, we could pick the maximizer of $\phi(K')$ in this interval. Two things are necessary to achieve this goal: an approximation of the value function over the interval $[K_{j-1}, K_{j+1}]$ and a method to locate the maximum of a continuous function. We consider function approximation in Section 8.2 and refer the reader to Section 8.6.1 that describes a simple numerical maximization tool.

Note, that we can evaluate $u(f(K) - K')$ for every $K' \in [K_{j-1}, K_{j+1}]$. Furthermore, we have estimates of the value function at the points $v(K_{j-1})$, $v(K_j)$, and $v(K_{j+1})$. The simplest method to approximate the value function over the entire interval $[K_{j-1}, K_{j+1}]$, thus, is to draw straight lines between these three points, i.e., to use linear interpolation.[17]

We are now able to modify Step 3 of Algorithm 1.2.2 in the following way: first, assume that j_i^* is the index neither of the first

[17] In the present example, linear interpolation is not a problem. For many problems, however, linear interpolation is not warranted. For example, in many problems in the theory of finance, e.g., the equity premium puzzle, we need to model the risk aversion of the agents. Risk aversion, however, depends on the second derivative of the utility function $u(\cdot)$ and, hence, on the curvature of the value function. In such applications, we need to model the concavity of the value function by high-order polynomials with shape-preserving methods.

nor of the last grid point so that the optimum of

$$\phi(K_j) := u(f(K_i) - K_j) + \beta v(K_j)$$

is bracketed by $[K_{j_i^*-1}, K_{j_i^*+1}]$. With linear interpolation, $\phi(K_j)$ is approximated by the continuous and concave function

$$\begin{aligned}\phi(K_j) \simeq \hat{\phi}(K_j) &:= u(f(K_i) - K_j) \\ &+ \beta \left[\mathfrak{v}^0_{j_i^*-1} + \frac{\mathfrak{v}^0_{j_i^*+1} - \mathfrak{v}^0_{j_i^*-1}}{K_{j_i^*+1} - K_{j_i^*-1}} (K_j - K_{j_i^*-1}) \right].\end{aligned} \tag{1.20}$$

We use the golden section search algorithm 8.6.1 to find the maximizer of this function. Differently from Algorithm 1.2.2 we must now store K_j and not its index. If the index of the optimal point is $j_i^* = 1$ ($j_i^* = n$), we consider a point close to the right of K_1 (left of K_n). If we find that $\hat{\phi}(K_j)$ returns a smaller value at this point than at K_1 (K_n), we accept K_1 (K_n). Otherwise we use golden section search to find K_j in $[K_1, K_2]$ ($[K_{n-1}, K_n]$). This gives raise to the following modification of Algorithm 1.2.2

Algorithm 1.2.3 (Value Function Iteration with Linear Interpolation)

Purpose: *Find an approximate solution of the policy function for the Ramsey model* (1.9)

Steps:

Step 1: Choose a grid of n equally spaced points

$$\mathscr{G} = [K_1, K_2, \ldots, K_n],\ K_i < K_j\ \forall i < j = 1, 2, \ldots n.$$

Step 2: Initialize the value function: $\forall i = 1, \ldots, n$ set

$$\mathfrak{v}^0_i = \frac{u(f(K^*) - K^*)}{1 - \beta},$$

where K^* denotes the stationary solution to the continuous valued Ramsey problem.

Step 3: Compute a new value function and the associated policy function, $\mathfrak{v}^1$ and $\mathfrak{g}^1$, respectively.

Step 3.1: *Put* $j_0^* \equiv 1$. *For* $i = 1, \ldots, n$, *and* j_{i-1}^* *find the index* j_i^* *that maximizes*

$$u\left(f(K_i) - K_j\right) + \beta \mathfrak{v}_j^0$$

in the set of indices $\{j_{i-1}^*, j_{i-1}^* + 1, \ldots, n\}$. *Algorithm 1.2.1 can be used to accomplish this.*

Step 3.2: *Find the solution* $\tilde{K}$ *off the grid: determine whether* $j_i^* = 1$ *or* $j_i^* = n$. *If so, add (subtract) a small fraction of the distance* Δ *between* K_1 *and* K_2 (K_{n-1} *and* K_n) *to* K_1 (K_n). *If* $\hat{\phi}(K_1 + \Delta) < \hat{\phi}(K_1)$ ($\hat{\phi}(K_n - \Delta) < \hat{\phi}(K_n)$ *put* $\tilde{K} = K_1$ ($\tilde{K} = K_n$), *otherwise use Algorithm 8.6.1 to locate the maximizer of* (1.20), *say* $\tilde{K}$, *in* $[K_1, K_2]$ ($[K_{n-1}, K_n]$). *If* $1 < j_i^* < n$ *use Algorithm 8.6.1 to find the maximizer of* (1.20) *in the interval* $[K_{j_i^*-1}, K_{j_i^*+1}]$.

Step 3.3: *Set* $\mathfrak{g}_i^1 = \tilde{K}$ *and* $\mathfrak{v}_i^1 = \hat{\phi}(\tilde{K})$, *where* $\hat{\phi}$ *is defined in* (1.20).

Step 4: *Check for convergence: If* $\|\mathfrak{v}^0 - \mathfrak{v}^1\|_\infty < \epsilon(1-\beta)$, $\epsilon \in \mathbb{R}_{++}$ *stop, else set* $\mathfrak{v}^0 = \mathfrak{v}^1$ *and* $\mathfrak{g}^0 = \mathfrak{g}^1$ *and return to step 3.*

This algorithm delivers the policy function $\mathfrak{g}$ not in terms of indices but of capital stocks that are optimal to choose if the current capital stock is point i in the grid $\mathscr{G}$. Thus, if we want to construct the time path of the capital stock from an arbitrary initial point K_0, we must use linear interpolation. For instance, the optimal stock of capital for period $t = 1$, say $\tilde{K}(K_0)$, is found from

$$\tilde{K}(K_0) = \mathfrak{g}_i + \frac{\mathfrak{g}_{i+1} - \mathfrak{g}_i}{K_{i+1} - K_i}(K_0 - K_i),$$

where $K_i < K_0 < K_{i+1}, K_i, K_{i+1} \in \mathscr{G}$.

It is next to impossible to program Algorithm 1.2.3 in GAUSS as a stand alone procedure that requires no more than a pointer to the utility function, the value of β, and an initial guess of the value function. The problem is that it is impossible to code $\hat{\phi}$ with a

pointer to u. One is bound to use global variables and procedures. For this reason we have decided not to include Algorithm 1.2.3 as a procedure into `DGE.src`. Instead, we coded it in the program `Ramsey2a.g` using global symbols, i.e., variables and procedure names that are known to the main program as well as to all other subprograms.

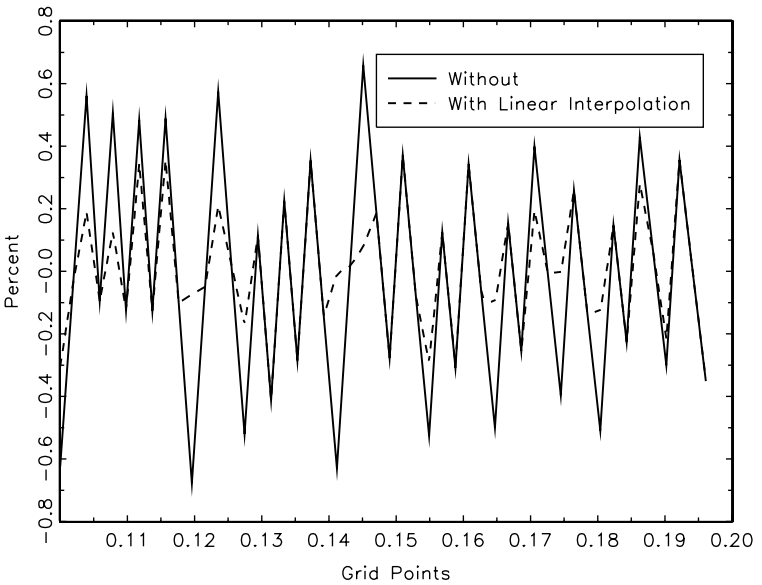

Figure 1.5: Deviation of the Policy Function from the Exact Solution

Figure 1.5 compares the performance of Algorithm 1.2.2, value function iteration over the grid points only, and Algorithm 1.2.3, value function iteration with linear interpolation between grid points but without binary search. Both computations used the same grid size $n = 50$. Of course, with linear interpolation, computing time increases (in our example from 0.4 to 12 seconds). However, the accuracy as measured by the deviation from the true solution is only slightly affected: the maximum absolute error decreases from 0.67 percent to 0.41 percent. As can be seen from Figure 1.5, there is almost no difference between values of the

capital stock near the steady state.[18] In the present application, linear interpolation between grid points without binary search is not an improvement over the standard value function iteration if we restrict the computing time to be the same in the two programs `SolveVIDc` (implementation of Algorithm 1.2.2) and `SolveVIDd` (implementation of Algorithm 1.2.3) and, hence, are able to increase the number of grid points in the former case. In more complex models which you will also encounter in the second part of this book, the state space may have three or four dimensions. In these applications linear interpolations with binary search will be the favorite method, since we get a given accuracy with less grid points.

1.3 The Stochastic Ramsey Model

1.3.1 Stochastic Output

In the Ramsey problem (1.9) everything is under the farmer's control. Yet, this is an overly optimistic picture of farming. Less rain during the summer causes harvest failure, whereas the right balance between rainfall and sunshine boosts crop growth. The amount of rainfall is outside the control of the farmer and, usually, he is unable to predict it accurately. The ensuing uncertainty turns the crop and, hence, consumption into stochastic variables. As a consequence, we must restate the farmer's decision problem in the framework of expected utility maximization. We illustrate the points that are involved in this task in Example 1.3.1. Since an in-depth treatment of the analytical framework that underlies stochastic control is beyond the scope of this book we refer the interested reader to Stokey and Lucas (1989).

[18] The policy and the value functions are more curved for low values of the capital stock so that the approximation is less accurate in this range. As one solution to this problem, one might choose an unequally-spaced grid with more points in the lower interval of the state space, e.g. $K_i = K_1 + \zeta(i-1)^2, \zeta = (K_n - K_1)/(n-1)^2$ or choose the distance between the logarithms of the neighboring grid points, $\ln K_i - \ln K_{i+1}$, to be constant. However, one can show that neither grid type dominates uniformly across applications.

Example 1.3.1

Assume the farmer's planing horizon is $T = 1$. His one-period utility function $u(C_t)$ is strictly increasing in consumption C_t. Output in period $t = 0$ is given by $f(K_0)$ and in period $t = 1$ by $Z_1 f(K_1)$, where $Z_1 = \underline{Z}$ with probability π and $Z_1 = \bar{Z} > \underline{Z}$ with probability $1 - \pi$. $f(K_t)$ is strictly increasing in the capital stock K_t. K_0 is given. Since the farmer does not plan beyond $t = 1$, we already know that he will choose $C_1 = Z_1 f(K_1)$. Given his investment decision in the current period K_1 his future consumption is a random variable with realizations $C_1(\underline{Z}) = \underline{Z} f(K_1)$ and $C_1(\bar{Z}) = \bar{Z} f(K_1)$. Hence, the farmer's expected life-time utility is

$$E_0 \left[u(C_0) + \beta u(C_1)\right] := u(f(K_0) - K_1) + \beta \left[\pi u(\underline{Z} f(K_1)) + (1 - \pi) u(\bar{Z} f(K_1))\right].$$

The farmer chooses K_1 to maximize this expression. Differentiating with respect to K_1 and setting to zero the resulting expression yields the following first-order condition:

$$\begin{aligned} u'(C_0) &= \beta \left[u'(\underline{Z} f(K_1)) \underline{Z} f'(K_1) \pi + u'(\bar{Z} f(K_1)) \bar{Z} f'(K_1)(1 - \pi)\right], \\ &= \beta E_0 \left[u'(C_1) Z_1 f'(K_1)\right]. \end{aligned}$$

This equation is the stochastic analog to the respective Euler equation in the deterministic case. It states that the utility loss from increased savings in the current period, $u'(C_0)$, must be compensated by the expected future utility increase.

We will consider the following stochastic infinite horizon Ramsey model, which is also known as the stochastic growth model:[19] The farmer solves

$$\max_{C_0} \quad E_0 \left[\sum_{t=0}^{\infty} \beta^t u(C_t)\right]$$

s.t. (1.21)

[19] Note, from here on $f(K)$ denotes gross value added and we consider depreciation explicitly. We need to do so, since using our specification of the production function from (1.2), $Z_t f(K_t)$ would imply stochastic depreciation otherwise.

$$\left.\begin{array}{rcl} K_{t+1} + C_t & \leq & Z_t f(K_t) + (1-\delta)K_t, \\ 0 & \leq & C_t, \\ 0 & \leq & K_{t+1}, \end{array}\right\} \; t = 0, 1, \ldots$$

K_0, Z_0 given.

It differs from the deterministic model in two respects: firstly, output at each period t depends not only on the amount of capital K_t but also on the realization of a stochastic variable Z_t capturing weather conditions. We assume that the farmer knows the amount of rainfall Z_t at harvest time, when he must decide about consumption. Secondly, and as a consequence of this assumption, in the current period $t = 0$ the farmer chooses only current consumption C_0. In the deterministic case, he gets no new information when the future unfolds. Therefore, he can safely determine consumption from the present to the very distant future. In technical terms, his decision problem is open-loop control, as opposed to close-loop control in the stochastic case. Here, as in Example 1.3.1, future consumption is a stochastic variable from the perspective of the current period. Thus, the farmer does better if he postpones the decision on period t consumption until this period t. As a consequence of the uncertainty with respect to consumption, the farmer aims at maximizing the expected value of his life-time utility. More specifically, the notation $E_0[\cdot]$ denotes expectations conditional with respect to the probability distribution of the sequence of random variables $\{C_t\}_{t=1}^{\infty}$ conditional on information available at $t = 0$.

A remarkable feature of the deterministic Ramsey problem (1.9) is its solution in terms of a time-invariant policy function g that relates the current choice variable K_{t+1} to the current state of the system K_t.[20] Would't it be nice if the solution to the stochastic model (1.21) turned out to be $K_{t+1} = g(Z_t, K_t)$? Indeed, we can get this result, if we are willing to restrict the class of probability laws governing Z_t to the class of stochastic processes that have the Markov property.

[20] Remember, we arrive at this results irrespective of whether we use the system of Euler equations or a dynamic programming argument to characterize the optimal solution.

If you are unfamiliar with Markov processes we ask you to consult Section 9.2, where we sketch the necessary definitions and tools. The important thing about these processes is that they preserve the recursive structure of the Ramsey model, so that we can employ dynamic programming techniques. Before we turn to this point, we consider the extension of the Euler equations approach to the stochastic Ramsey model.

1.3.2 Stochastic Euler Equations

First order conditions for the stochastic Ramsey model (1.21) can be derived in a manner analogous to the deterministic case. Consider the following Lagrangean function:

$$\mathscr{L} = E_0\Bigg\{\sum_{t=0}^{\infty}\beta^t\Big[u(C_t)+\mu_t C_t+\omega_{t+1}K_{t+1} + \lambda_t\left(Z_t f(K_t)+(1-\delta)K_t-C_t-K_{t+1}\right)\Big]\Bigg\}.$$

Since the expectations operator is a summation operator we can differentiate the expression in curly brackets with respect to C_0 and K_1 (see Example 1.3.1). This delivers

$$\begin{aligned}
\frac{\partial\mathscr{L}}{\partial C_0} &= E_0\{u'(C_0)-\lambda_0+\mu_0\}=0,\\
\frac{\partial\mathscr{L}}{\partial K_1} &= \beta E_0\{-\lambda_0+\omega_1+\beta\lambda_1(1-\delta+Z_1 f'(K_1))\}=0,\\
0 &= \lambda_0(Z_0 f(K_0)+(1-\delta)K_0-C_0-K_1),\\
0 &= \mu_0 C_0,\\
0 &= \omega_1 K_1.
\end{aligned}$$

Since, as in Example 1.3.1, C_0, K_1, and hence the multipliers λ_0, μ_0, and ω_1 are non-stochastic, we can replace the first condition with

$$u'(C_0)=\lambda_0-\mu_0$$

and the second with

$$\lambda_0 = \beta E_0 \lambda_1 \{1 - \delta + Z_1 f'(K_1)\} + \omega_1.$$

Now, consider the problem from $t = 1$ onwards, when Z_1 is known and K_1 given. The Lagrangean for this problem is

$$\mathscr{L} = E_1 \Bigg\{ \sum_{t=1}^{\infty} \beta^{t-1} \Big[u(C_t) + \mu_t C_t + \omega_{t+1} K_{t+1} + \lambda_t \left(Z_t f(K_t) + (1-\delta) K_t - C_t - K_{t+1} \right) \Big] \Bigg\}.$$

Proceeding as before, we find

$$\begin{aligned} u'(C_1) &= \lambda_1 - \mu_1, \\ \lambda_1 &= \beta E_1 \lambda_2 \{1 - \delta + Z_2 f'(K_2)\} + \omega_2. \end{aligned}$$

Continuing in this way, we find, since K_t must be optimal at t, that the plan for choosing $C_0, C_1, \ldots$ and $K_1, K_2, \ldots$ must solve the system:

$$u'(C_t) = \lambda_t - \mu_t, \tag{1.22a}$$

$$\lambda_t = \beta E_t \lambda_{t+1} \left[1 - \delta + Z_{t+1} f'(K_{t+1}) \right] + \omega_{t+1}, \tag{1.22b}$$

$$0 = \lambda_t (Z_t f(K_t) + (1-\delta) K_t - C_t - K_{t+1}), \tag{1.22c}$$

$$0 = \mu_t C_t, \tag{1.22d}$$

$$0 = \omega_{t+1} K_{t+1}. \tag{1.22e}$$

Thus, an interior solution with strictly positive consumption and capital at all dates t (i.e., $\forall t : \mu_t = \omega_{t+1} = 0$) must satisfy the stochastic analog to the Euler equation (1.11)

$$1 = \beta E_t \frac{u'(Z_{t+1} f(K_{t+1}) + (1-\delta) K_{t+1} - K_{t+2})}{u'(Z_t f(K_t) + (1-\delta) K_t - K_{t+1})} \times (1 - \delta + Z_{t+1} f'(K_{t+1})). \tag{1.23}$$

In addition to the stochastic Euler equation (1.23) there is also the stochastic analog of the transversality condition (1.12), namely

$$\lim_{t \to \infty} \beta^t E_t \lambda_t K_{t+1} = 0 \tag{1.24}$$

that provides a boundary condition for the solution to (1.23).

1.3.3 Stochastic Dynamic Programming

As in the deterministic Ramsey model there is a dynamic programming approach to characterize solutions of the stochastic Ramsey model (1.21). The value function $v(K, Z)$ is now defined as the solution to the following stochastic functional equation:[21]

$$v(K,Z) = \max_{0 \le K' \le Zf(K)+(1-\delta)K} u(Zf(K) + (1-\delta)K - K') + \beta E\left[v(K', Z')|Z\right],$$

where expectations are conditional on the given realization of Z. In the case of a Markov chain with realizations $[z_1, z_2, \ldots, z_n]$ and transition matrix $P = (p_{ij})$ the expression $E\left[v(K', Z')|Z\right]$ is given by

$$E\left[v(K', Z')|z_i\right] = \sum_{j=1}^{n} p_{ij} v(K', z_j)$$

and in the case of the continuous valued Markov process with conditional probability density function $\pi(z, Z')$ over the interval $[a, b]$ it is

$$E\left[v(K', Z')|z\right] = \int_a^b v(K', Z')\pi(z, Z')dZ'.$$

It requires some sophisticated mathematics to prove the existence and to find the properties of the value function and the associated policy function $K' = g(K, Z)$. We refer the interested reader to Stokey and Lucas (1989) and proceed under the assumption that both the value and the policy function exist and are sufficiently differentiable with respect to K. Under this assumption it is easy to use the steps taken on page 16 to show that the dynamic programming approach also delivers the stochastic Euler equation (1.23). We leave this as an exercise to the reader.

The solution to the deterministic Ramsey model is a time path for the capital stock. In the stochastic case $K' = g(K, Z)$ is a

[21] As explained in Section 9.1, a functional equation is an equation whose unknown is a function and not a point in $\mathbb{R}^n$.

stochastic variable, since Z is random. Thus, the policy function induces a probability distribution over the space of admissible capital stocks, and the solution we look for is a stochastic process $\{K_t\}_{t=0}^{\infty}$. Nevertheless, it is possible to adapt the techniques from the previous section to compute approximations to the realization of this process. We illustrate these techniques in the next subsection using the following stochastic version of Example 1.2.1.

Example 1.3.2

Let the one-period utility function u and the production function f be given by

$$u(C) := \ln C,$$
$$f(K) := K^{\alpha}, \quad \alpha \in (0,1),$$

respectively.

In Example 1.2.1 we find that K' is directly proportional to K^{α}. So let us try

$$K_{t+1} = g(K_t, Z_t) := AZ_tK_t^{\alpha}$$

with the unknown parameter A as policy function. If this function solves the problem, it must satisfy the stochastic Euler equation (1.23). To prove this assertion, we replace K_{t+1} in equation (1.23) by the rhs of the previous equation. This gives

$$1 = \beta E_t \left[\frac{(1-A)Z_tK_t^{\alpha}}{(1-A)Z_{t+1}[AZ_tK_t^{\alpha}]^{\alpha}} \alpha Z_{t+1}[AZ_tK_t^{\alpha}]^{\alpha-1} \right] = \frac{\alpha\beta}{A}.$$

If we put $A = \alpha\beta$ the function $g(Z_t, K_t) = \alpha\beta Z_tK_t^{\alpha}$ indeed satisfies the Euler equation, and thus is the policy function we look for.

1.3.4 Numerical Solutions

Extension of Deterministic Path. The solution method that we present in this paragraph fits well into the framework of our forward iteration algorithm presented in section 1.2.5. It dates back to Fair and Taylor (1983). Gagnon (1990) applied it to solve the stochastic Ramsey model (1.21).

Assume that at date t the productivity shock Z_t is observed. If from date t onwards there were no additional shocks, $Z_{t+s} = Z \ \forall s = 1, 2, \ldots$, the Ramsey model would behave as the deterministic model and approach a stationary solution. Thus, we could approximate the model's time path by solving the respective set of Euler equations for a given horizon T which we consider necessary to get close to the long-run equilibrium. The solution for the capital stock for date $t+1$, thus, represents the solution to the stochastic model under the assumption of perfect foresight and the assumption of no further shocks. Yet, if the productivity shock were identically and independently distributed over time with mean Z this would be the correct forecast. Given the solution for K_{t+1} we can subject the model to another shock Z_{t+1} and compute the deterministic path from there on to find K_{t+2}. Continuing in this way we can derive the approximate dynamics of the stochastic model for an arbitrary number of periods.

This approach also works if the productivity shock evolves according to a Markov process. In the following, we assume that the log of Z_t from (1.21) follows a first-order autoregressive process

$$\begin{aligned} &\ln Z_t = \varrho \ln Z_{t-1} + \epsilon_t \ \Leftrightarrow \ Z_t = Z_{t-1}^{\varrho} e^{\epsilon_t}, \\ &\quad \varrho \in (0,1), \quad \epsilon_t \sim N(0, \sigma^2). \end{aligned} \tag{1.25}$$

The farmer knows this law and observes ϵ_t prior to his decision on the future capital stock K_{t+1}. Thus, for the periods after t he expects

$$E_t(Z_{t+s}) = z_t^{\varrho^s}, \quad z_t = Z_t^{\varrho} e^{\epsilon_t}, \ s = 1, 2, \ldots.$$

To find K_{t+1} given K_t, we must solve the system of $T-1$ equations in $(K_{t+1}, K_{t+2}, \ldots, K_{t+T-1})$:

$$\begin{aligned} \frac{1}{\beta} &= \frac{u'(z_t^{\varrho} f(K_{t+1}) + (1-\delta)K_{t+1} - K_{t+2})}{u'(z_t f(K_t) + (1-\delta)K_t - K_{t+1})}(1 - \delta + z_t^{\varrho} f'(K_{t+1})), \\ \frac{1}{\beta} &= \frac{u'(z_t^{\varrho^2} f(K_{t+2}) + (1-\delta)K_{t+2} - K_{t+3})}{u'(z_t^{\varrho} f(K_{t+1}) + (1-\delta)K_{t+1} - K_{t+2})}(1 - \delta + z_t^{\varrho^2} f'(K_{t+2})), \\ \vdots &= \qquad \vdots, \end{aligned}$$

$$1 = \beta \frac{u'(z_t^{\varrho^{T-1}} f(K_{t+T-1}) + (1-\delta)K_{t+T-1} - K^*)}{u'(z_t^{\varrho^{T-2}} f(K_{t+T-2} + (1-\delta)K_{t+T-2} - K_{t+T-1})} \times (1 - \delta + z_t^{\varrho^{T-1}} f'(K_{t+T-1}).$$

The GAUSS program `Ramsey3a.g` implements this solution technique for the stochastic Ramsey model specified in Example 1.3.3. In this example we use a one-period utility function with a constant elasticity of the marginal utility $\eta > 0$:

$$u(c) := \begin{cases} \frac{c^{1-\eta}}{1-\eta} & \text{if} \quad \eta > 0 \wedge \eta \neq 1, \\ \ln c & \text{if} \quad \eta = 1. \end{cases} \tag{1.27}$$

In the literature on choice under uncertainty the parameter η is known as the coefficient of relative risk aversion.[22]

Example 1.3.3

Consider the following version of the stochastic Ramsey model (1.21). The farmer solves:

$$\max_{C_0} \quad E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\eta}}{1-\eta} \right] \qquad \eta > 0,\ \beta \in (0,1),$$

s.t.

$$\left.\begin{array}{rcl} K_{t+1} + C_t & \leq & Z_t K_t^{\alpha} + (1-\delta)K_t,\ \alpha \in (0,1), \\ Z_t & = & Z_{t-1}^{\varrho} e^{\epsilon_t},\ \varrho \in (0,1),\ \epsilon_t \sim N(0,\sigma^2), \\ 0 & \leq & C_t, \\ 0 & \leq & K_{t+1}. \end{array}\right\} \ t = 0, 1, \ldots$$

$K_0,\ Z_0$ given.

Figure 1.6 plots the time paths of the capital stock and of consumption in response to a negative productivity shock in period $t = 1$. The underlying parameter values are: $\alpha = 0.35$, $\beta = 0.984$, $\eta = 2$, $\delta = 0.01$, and $\varrho = 0.95$.[23] The shock itself results from a draw from a normal distribution with mean 0 and standard deviation $\sigma_\epsilon = 0.01$. Since the process for Z_t is highly autocorrelated,

[22] See, e.g., Mas-Colell, Whinston, and Green (1995), p. 194.

[23] In Section 1.5 we deal with the question of parameter choice. Here, you should simply note, that these figures are a reasonable choice.

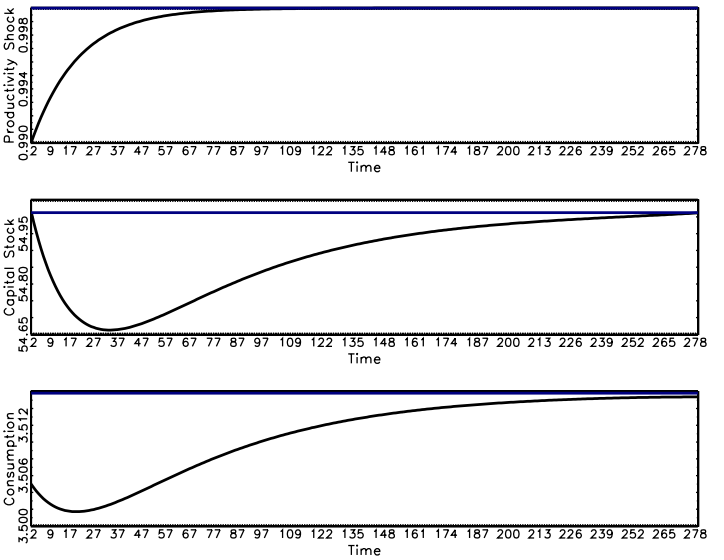

Figure 1.6: Example 1.3.3: Response to a Transitory Productivity Shock

it takes about 100 periods for Z_t to be near its stationary level of $Z \equiv 1$. This can be seen from the first panel of Figure 1.6. The farmer's response to this shock can be traced to two different motives. The crop failure reduces the farmer's income. As an immediate consequence, he curbs his current consumption, but by less than his income has dropped. Since the marginal utility of consumption is decreasing, he is better off when he spreads the loss over several periods. Therefore, the farmer invests less and next period's capital stock is below its long-run value (the straight line in the second panel of Figure 1.6). There is a second reason for the farmer to reduce investment: since Z_t is below average for quite a while, the return to investment will be low. This substitution effect accounts for the hump-shaped response of the capital stock and of consumption.

Figure 1.7 displays the farmer's response to recurring productivity shocks, produced by the GAUSS random number generator `rndn`. In this case, we need to solve the system of Euler equations of each time period considered. With $T = 200$ and 20 periods

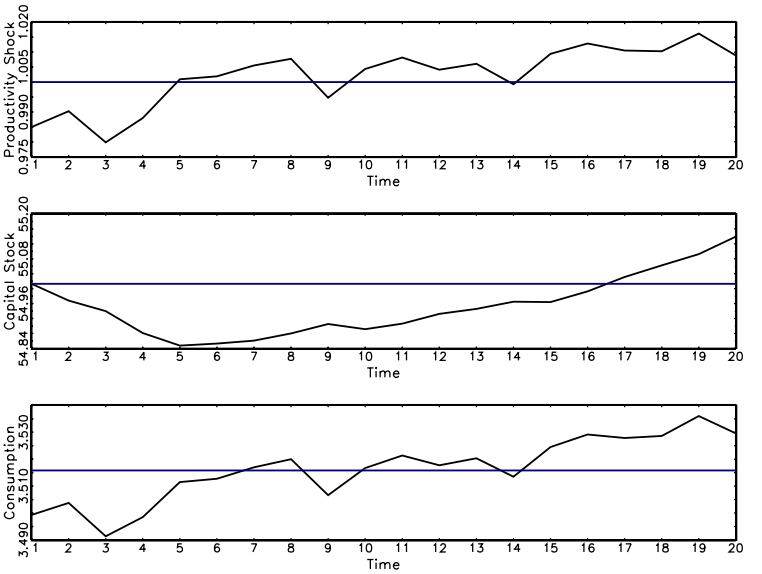

Figure 1.7: Example 1.3.3: Response to Recurring Productivity Shocks

the modified Newton-Raphson algorithm takes about 41 seconds to compute the time paths plotted in Figure 1.7. The horizontal lines represent the long-run equilibrium of the respective variable. The negative productivity shocks in the first four periods cause a recession: in his desire to smooth consumption the farmer runs down his capital stock. It takes a number of periods where the level of technology Z_t is around its mean $Z = 1$ and further positive shocks before the capital stock recovers.

Before we proceed, we summarize the extended path algorithm in the following steps:

Algorithm 1.3.1 (Deterministic Extended Path)

Purpose: *Numerical approximation of the solution of the stochastic Ramsey model* (1.21) *with a first order Markov process*

Steps:

Step 1: Initialize: Let n denote the number of periods to consider. Use a random number generator and draw a sequence of shocks $(\epsilon_t)_{t=0}^{n}$. Compute the time path of the productivity

shock $(Z_t)_{t=0}^n$ *from* $Z_t = Z_{t-1}^{\varrho} e^{\epsilon_t}$. *Select the initial capital stock* K_0. *Choose* T *large enough for the deterministic model to get close to its stationary solution* K^* *within* T *periods. (Iterate over* T *to see whether this condition holds.)*

Step 2: For $t = 0, 1, \ldots, n$ *repeat these steps:*

Step 2.1: Compute the expected time path of $(Z_{t+s})_{s=0}^{t+T}$ *from* $Z_{t+s} = Z_t^{\varrho^s}$.

Step 2.2: Set the initial capital stock to K_t *and the terminal capital stock to* $K_{t+T} = K^*$. *Solve the system of Euler equations*

$$\begin{aligned} &\frac{u'(Z_t^{\varrho^s} f(K_{t+s}) + (1-\delta)K_{t+s} - K_{t+s+1})}{u'(Z_t^{\varrho^{s+1}} f(K_{t+s+1} + (1-\delta)K_{t+s+1} - K_{t+s+2})} \\ &= \beta(1 - \delta + Z_t^{\varrho^{s+1}} f'(K_{t+s+1}), \\ &s = 0, 1, \ldots, T-2 \end{aligned}$$

and keep the solution for K_{t+1}.

In the next paragraph you will learn to compute a policy function that solves a discrete approximation of the Ramsey problem (1.21).

Value Function Iteration. In this paragraph we extend Algorithm 1.2.2 to compute an approximate policy function for the stochastic Ramsey model (1.21) when the productivity shock is a Markov chain with m different realizations $\mathbf{z} = [z_1, z_2, \ldots, z_m]$ and transition matrix $P = (p_{ij})$. The procedure is built on the result that the following sequence of iterations over the value function converges linearly at rate β:

$$\begin{aligned} v^{s+1}(K, z_i) = &\max_{K'} u(z_i f(K) + (1-\delta)K - K') \\ &+ \beta \sum_{j=1}^{m} p_{ij} v^s(K', z_j). \end{aligned}$$

The algorithm requires a discretization of the space of admissible capital stocks. An obvious choice would be the entire interval

between zero and the maximum sustainable capital stock $\bar{K}$. In the context of m different realizations of the productivity shock, the largest element in $\mathbf{z}$, $\bar{z}$ for short, determines $\bar{K}$ as solution to $\bar{z}f(\bar{K}) = \delta\bar{K}$, since, (as shown on p. 12) each $K \in [0, \bar{K}]$ will remain in this interval. Yet, with a reasonable number of grid points, say 500 or 1000, the resulting grid is by far too wide-meshed. For instance, a sensible choice of parameters like $\alpha = 0.35$, $\delta = 0.01$, and $\bar{z} = 1$, yields a maximum sustainable capital stock of almost 1200, and reliable results require a grid of at least one million points. As we shall see, however, the algorithm must loop m times over the steps performed in the deterministic case. Thus, to save on time, we must reduce the state space as far as possible. Recall that for a constant shift parameter z_j of the production function the capital stock approaches the associated stationary equilibrium K_j^* which solves $1 = \beta(1 - \delta + z_j f'(K_j^*))$. Define $\underline{K}^*$ ($\bar{K}^*$) as the stationary capital stock associated with the smallest (largest) element in $\mathbf{z}$. Any optimal sequence of the continuous valued problem originating in $[\underline{K}^*, \bar{K}^*]$ will remain there. Again, due to the discrete approximation, $\underline{K}^*$ and $\bar{K}^*$ are not identical to the stationary solutions of the approximate problem. They provide a first guess of the lower and upper grid points, say $\underline{K}$ and $\bar{K}$. We choose $\underline{K}$ as a fraction $\underline{x} < 1$ of $\underline{K}^*$ and $\bar{K}$ as a multiple $\bar{x} > 1$ of $\bar{K}^*$ and use K_j^*, to approximate the stationary value function. Let $\mathscr{G}$ denote a grid on $[\underline{K}, \bar{K}]$, and let $\mathfrak{v}_{ij}^0$ denote the expected life-time utility obtained under an optimal sequence of decisions when the initial capital stock and the initial productivity shock are given by the pair (K_i, z_j). An estimate of the stationary elements of $\mathfrak{v}$ is given by

$$\mathfrak{v}_{ij} = \underbrace{u(z_j f(K_j^*) + (1-\delta)K_j^* - K_j^*)}_{u_{ij}} + \beta \sum_{l=1}^{m} p_{jl}\mathfrak{v}_{il}$$

or in matrix notation

$$\mathfrak{v} = U + \beta\mathfrak{v}P',$$

where U is the matrix with typical element u_{ij}. Since P is a probability matrix, it has an inverse and

$$\mathfrak{v} = U(I - \beta P')^{-1}$$

is our guess of the stationary value function.

The policy function is a $n \times m$ matrix $\mathfrak{g}^0$. In its (ij)–th element we store the pointer to the k–th element of $\mathscr{G}$ that solves

$$\begin{aligned} k^* := &\operatorname*{argmax}_{1 \le k \le n} \big\{ u(z_j f(K_i) + (1-\delta)K_i - K_k) \\ &+ \beta \sum_{l=1}^{m} p_{jl} v(K_k, z_l) \big\}. \end{aligned} \tag{1.28}$$

For given z_j the policy function is still monotonically increasing in K and the value function is still strictly concave in K, if the one-period utility function u and the production function f are strictly increasing and strictly concave. Thus, when we solve the maximization problem for K_i we need only consider indices $k \ge k^*_{i-1}$ from the set $\{1, 2, \ldots n\}$, where k^*_{i-1} is the solution found for K_{i-1}. Furthermore, we can still employ Algorithm 1.2.1 to solve the maximization problem (1.28). To extend Algorithm 1.2.2 to the stochastic case, we must include an additional loop over the indices $j \in \{1, 2, \ldots m\}$ and must replace $\beta v(K')$ by the conditional expectation $E[v(K', z_l)|z_j]$. This gives:

Algorithm 1.3.2 (Value Function Iteration 2)

Purpose: *Compute an approximation to the policy function of the stochastic Ramsey model* (1.21) *with m–state Markov chain for the productivity shock.*

Steps:

Step 1: Choose a grid $\mathscr{G}$ of n equally spaced points over $[\underline{K}, \overline{K}]$:

$$\begin{aligned} &\mathscr{G} = [K_1, K_2, \ldots, K_n], \\ &K_i < K_{i+1}\ \forall i = 1, 2, \ldots n-1,\ K_1 = \underline{K}, K_n = \overline{K}. \end{aligned}$$

Step 2: Initialize the value function as follows: Let $U = (u_{ij})$. For all $j = 1, 2, \ldots, n$ compute

$$u_{ij} := u(z_j f(K^*_j) + (1-\delta)K^*_j - K^*_j),$$

where K_j^* *solves* $1 = \beta(1 - \delta + z_j f'(K_j^*))$ *and* $i = 1, 2, \ldots, n$. *Put*

$$\mathfrak{v}^0 = U(I - \beta P')^{-1}.$$

Step 3: Compute a new value function $\mathfrak{v}^1$ *and associated policy function* $\mathfrak{g}^1$*: For each* $j = 1, 2, \ldots, m$ *repeat these steps: Put* $k_0^* = 1$ *and for* $i = 1, 2, \ldots, n$ *find the index* k_i^* *that maximizes*

$$\mathfrak{w}_k = u(z_j f(K_i) + (1 - \delta)K_i - K_k) + \beta \sum_{l=1}^{m} p_{jl} \mathfrak{v}_{kl}^0$$

in the set of indices $k = k_{i-1}^*, k_{i-1}^* + 1, \ldots, n$. *Set* $\mathfrak{v}_{ij}^1 = \mathfrak{w}_{k_i^*}$ *and* $\mathfrak{g}_{ij}^1 = k_i^*$.

Step 4: Check for convergence: if

$$\max_{\substack{i=1,\ldots n \\ j=1,\ldots m}} |v_{ij}^1 - v_{ij}^0| \leq \epsilon(1 - \beta), \quad \epsilon \in \mathbb{R}_{++}$$

(or if the policy function has remained unchanged for a number of consecutive iterations) stop, else set $\mathfrak{v}^0 = \mathfrak{v}^1$ *and* $\mathfrak{g}^0 = \mathfrak{g}^1$ *and return to Step 3.*

The GAUSS procedure `SolveVISb` in `DGE.src` implements this algorithm as well as the Fortran subprogram `SolveVI` in `DGE.for`. Both use $\epsilon = 0.01$ and stop if either the largest element in $\mathfrak{v}^1 - \mathfrak{v}^0$ is smaller than $\epsilon(1 - \beta)$ or the policy function has remained unchanged in 30 consecutive iterations. We applied both routines to solve Example 1.3.4 using a grid of 1000 points on $[\underline{x}K^*, \bar{x}\overline{K}^*]$. Initially, we use $\underline{x}$ and $\bar{x}$ near one and adapt them when the program reports that the algorithm hits either the lower or the upper bound of the grid.

Example 1.3.4

Consider the following stochastic Ramsey model

$$\max_{C_0} \quad E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\eta}}{1 - \eta} \right] \quad \eta > 0, \, \beta \in (0, 1),$$

s.t.

$$\left.\begin{array}{rcll} K_{t+1} + C_t & \leq & Z_t K_t^{\alpha} + (1-\delta)K_t, & \alpha \in (0,1), \\ 0 & \leq & C_t, & \\ 0 & \leq & K_{t+1}, & \end{array}\right\} t = 0, 1, \ldots,$$

K_0, Z_0 given.

The Markov chain for the productivity shock Z_t has three states,

$$\mathbf{z} = [0.99, 1.00, 1.01]',$$

and the associated transition matrix is given by

$$P := \begin{bmatrix} 0.61 & 0.34 & 0.05 \\ 0.24 & 0.52 & 0.24 \\ 0.05 & 0.34 & 0.61 \end{bmatrix}.$$

The remaining parameters are calibrated as follows: $\alpha = 0.27$, $\beta = 0.984$, $\delta = 0.01$, and $\eta = 2$.

Run on a Pentium III, 800 MHz personal computer, both programs terminate after 226 iterations when the largest element in $\mathfrak{v}^1 - \mathfrak{v}^0$ is less than 0.0001546. The Fortran program takes about 17 seconds. Under GAUSS 4.0 the program `Ramsey3b.g` needs 177 seconds. The first panel of Figure 1.8 plots a realization of the productivity shock. The second and third panel display the time path of the capital stock and the time path of consumption, respectively. The dashed lines in panel two and three show the stationary levels of the stock of capital and of consumption associated with the intermediate level of the productivity shock of $z_2 = 1$. The jumps in the level of productivity are reflected in the pattern of consumption. Compared to the magnitude of the former, the magnitude of the latter is small. For instance, when the level of productivity drops by 1 percent in period $t = 2$, consumption falls by about 0.08 percent. Since investment is only a small fraction of the capital stock, the same shock reduces the stock of capital in period $t = 3$ by about 0.09 percent. In the first half of the simulation period productivity is either at its intermediate or at its lowest level. The farmer's desire to smooth consumption is

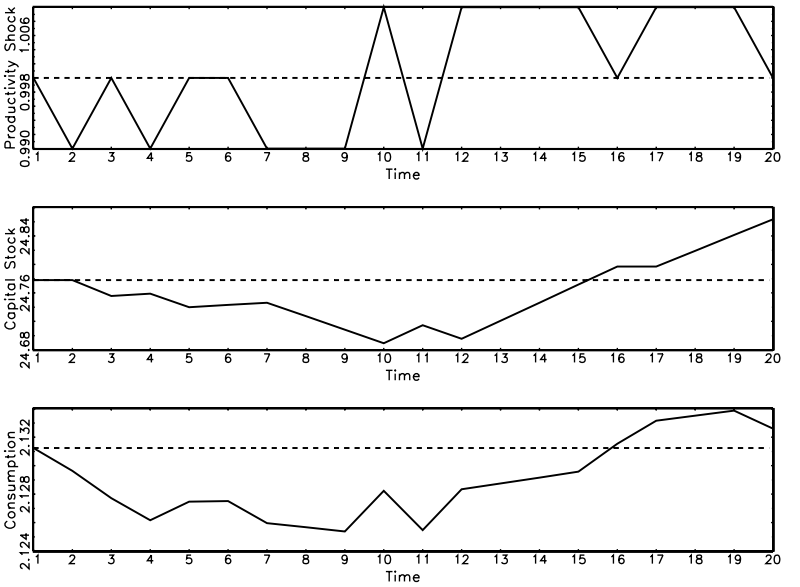

Figure 1.8: Example 1.3.4: Response to Recurring Productivity Shocks

reflected in the steady decline of the capital stock during this period. It requires a number of very favorable shocks in the second half of the simulation period to reverse this pattern.

1.4 Labor Supply, Growth, and the Decentralized Economy

1.4.1 Substitution of Leisure

So far we have taken labor supply as exogenous. Yet, as is well known, there are considerable employment fluctuations over the business cycles. In the context of our farming example, variations in labor input may arise from shocks to labor productivity, if the farmer values both consumption and leisure. To allow for that case we include leisure in the one-period utility function. Leisure L is the farmer's time endowment, which we normalize to 1, minus his working hours N. Thus we may state the one-period utility function now as

$$u(C, 1 - N). \tag{1.30}$$

In the following subsection we will ask what kinds of restrictions we must place on u besides the usual assumptions with respect to concavity and monotonicity when we deal with a growing economy. Before we proceed, we consider briefly what we can expect in general from including leisure into the one-period utility function.

Assume that the farmer observes an increase in today's marginal product of labor that he considers short-lived. How will he react?

In the current period the shock increases the farmer's opportunity set, since at any given level of labor input his harvest will be higher than before the shock. At the same time the shock changes the relative price of leisure: the farmer loses more output for each additional unit of leisure he desires. The overall effect of the shock on the intra-temporal substitution between labor and consumption depends upon the relative size of the associated income and substitution effect. If leisure and consumption are normal goods, the farmer wants both more consumption and more leisure (income effect). Yet, since leisure is more costly than before the shock, he also wants to substitute consumption against leisure (substitution effect).

In the intertemporal setting we are considering here, there is an additional, inter-temporal substitution effect. The shock raises the current reward for an additional hour of work vis-à-vis the future return. Consequently, the farmer will want to work more now and less in the future. He can achieve this goal by increasing today's savings and spending the proceeds in subsequent periods. Thus, investment serves as vehicle to the intertemporal substitution of consumption and leisure.

1.4.2 Growth and Restrictions on Technology and Preferences

Labor Augmenting Technical Progress. When we refer to economic growth we think of increases in output at given levels of input brought about by increases in technological knowledge. This

kind of technological progress is called disembodied as opposed to embodied progress that operates via improvements in the quality of the factors of production. Disembodied technological progress simply shifts the production function outward. Equivalently, we may think of it as if it redoubled the available physical units of labor and capital. For instance, if N is the amount of physical or raw labor and A its efficiency level, effective labor is AN. Using this concept, output is given by

$$Y_t = Z_t F(A_t N_t, B_t K_t),$$

where the efficiency factors A_t and B_t as well as the productivity shock Z_t are exogenously given time series or stochastic processes. The production function F has the usual properties, namely:[24]

1) $F_i > 0$ and $F_{ii} < 0$ (positive but diminishing marginal products),
2) $F(AN, 0) = 0$ and $F(0, BN) = 0$ (both factors of production are essential),
3) $\lambda Y = F(\lambda AN, \lambda BK)$ (constant returns to scale).

In Section 1.2.4 we have seen that the solution to the deterministic, infinite horizon Ramsey model approaches a stationary equilibrium. There is an appropriate concept of stationarity in models of growth, the so-called balanced growth path. Referring to Solow (1988), p. 4, we define a balanced growth path by two requirements:

1) output per working hour grows at a constant rate,
2) and the share of net savings in output is constant.

The motivation for this definition has two different sources. Firstly, from the empirical perspective, the balanced growth path replicates the broad facts about growth of advanced industrial economies.[25] Secondly, from the theoretical perspective, the balanced growth path allows to define variables relative to their trend path

[24] Here, and in the following, for any function $F(x_1, \ldots, x_n)$ the expression F_i denotes the first partial derivative of F with respect to x_i, and F_{ij} denotes the derivative of $F_i(x_1, \ldots, x_n)$ with respect to x_j.

[25] See, Solow (1988), p. 3ff.

that are stationary like the unscaled variables in no-growth models. Therefore, the techniques used to study stationary economies remain valid.

In Appendix 2 we show that for a balanced growth path to exist technical progress must be of the labor augmenting type, i.e., $B_t \equiv 1 \; \forall t$. As a consequence, we specify the production function as

$$Y_t = Z_t F(A_t N_t, K_t). \tag{1.31}$$

Trend versus Difference Stationary Growth. The specification (1.31) leaves two possible modelling choices for the process governing the evolution of the efficiency factor of raw labor. If we consider growth a deterministic process, the efficiency factor A_t grows at a given and constant growth factor $a > 1$:

$$A_{t+1} = aA_t. \tag{1.32}$$

Temporary variations around the long-run path are induced by the stochastic process $(Z_t)_{t=0}^{\infty}$, which we require to be covariance stationary.[26] To find the long-run behavior of output assume Z_t is equal to its unconditional mean $Z \equiv 1$. Since F has constant returns to scale we may write

$$Y_t = A_t F(N_t, K_t/A_t).$$

Note that according to our utility function (1.30) labor supply N_t is bounded above by 1. Since A_t grows at the constant rate $a - 1$, output will grow at the same constant rate, if both labor input and the quantity K_t/A_t are constant. Therefore, capital must grow at the same rate as output.

The assumption of deterministic growth has obvious empirical implications: output is a trend stationary stochastic process, i.e.,

[26] A stochastic process $(Z_t)_{t=0}^{\infty}$ is covariance stationary, if its unconditional mean is independent of time t and if the covariance between Z_t and Z_{t+s}, depends upon the time lag s but not on time itself.

when we subtract a linear trend from log-output, the resulting time series is a covariance stationary stochastic process.

In an influential paper NELSON and PLOSSER (1982) question this implication. They provide evidence that major macroeconomic aggregates are better modeled as difference stationary stochastic processes. A time series (x_t) is difference stationary if $(x_{t+1} - x_t)_{t=-\infty}^{\infty}$ is a covariance stationary stochastic process. In the context of our neoclassical production function we get this result, if we set $Z_t \equiv 1$ and let a difference stationary Markov process govern the evolution of the efficiency level of labor. For instance, we may assume A_t to follow the process

$$A_{t+1} = A_t e^{a+\epsilon_t}, \quad \epsilon_t \sim N(0, \sigma^2), a > 0. \tag{1.33}$$

Under this process the growth factor of the efficiency level of labor, A_{t+1}/A_t fluctuates around its long-run mean of e^a and the first difference of log-output, $\ln Y_{t+1} - \ln Y_t$, is covariance stationary. To see this, use

$$Y_t = A_t F(N_t, K_t/A_t)$$

and set $F(\cdot)$ equal to its long-run value $\bar{F} := F(N, K/A)$. Using (1.33), we get

$$\ln Y_{t+1} - \ln Y_t = \ln A_{t+1} - \ln A_t = a + \epsilon_t,$$

which is a white noise process.

Restrictions on Preferences. The restriction to labor augmenting technical progress is not sufficient to guarantee the existence of a balanced growth path when labor supply is endogenous. To see this, we restrict attention to the deterministic case and put $Z \equiv 1$ in (1.31). Using the one-period utility function (1.30), the farmer's maximization problem is

$$\max_{\{C_t, N_t\}_{t=0}^{\infty}} \quad \sum_{t=0}^{\infty} \beta^t u(C_t, 1 - N_t)$$

s.t. (1.34)

$$\left.\begin{array}{rcl} K_{t+1} + C_t & \leq & F(A_t N_t, K_t) + (1-\delta)K_t, \\ 0 & \leq & C_t, \\ 1 & \geq & N_t \geq 0, \\ 0 & \leq & K_{t+1}, \end{array}\right\} t = 0, 1, \ldots,$$

K_0 given.

Since we are interested in a long-run solution with positive consumption and leisure, we will ignore the non-negativity restrictions and the upper bound on labor in setting up the respective Lagrangean:

$$\mathscr{L} = \sum_{t=0}^{\infty} \beta^t \Big[u(C_t, 1-N_t) + \Lambda_t \big(F(A_t N_t, K_t) + (1-\delta)K_t - C_t - K_{t+1} \big) \Big].$$

Differentiating this expression with respect to C_t, N_t, and K_{t+1} provides the following set of first-order conditions:

$$0 = u_1(C_t, 1-N_t) - \Lambda_t, \tag{1.35a}$$

$$0 = -u_2(C_t, 1-N_t) + \Lambda_t F_1(A_t N_t, K_t) A_t, \tag{1.35b}$$

$$0 = -\Lambda_t + \beta \Lambda_{t+1}(1 - \delta + F_2(A_{t+1} N_{t+1}, K_{t+1})). \tag{1.35c}$$

Conditions (1.35a) and (1.35b) imply that the marginal rate of substitution between consumption and leisure, u_2/u_1, equals the marginal product of labor:

$$\frac{u_2(C_t, 1-N_t)}{u_1(C_t, 1-N_t)} = A_t F_1(A_t N_t, K_t). \tag{1.36}$$

Conditions (1.35a) and (1.35c) yield

$$\frac{u_1(C_t, 1-N_t)}{u_1(C_{t+1}, 1-N_{t+1})} = \beta(1 - \delta + F_2(A_{t+1} N_{t+1}, K_{t+1})). \tag{1.37}$$

Consider the right hand side of this equation. Since F is homogenous of degree one F_2 is homogenous of degree zero, i.e.,

$$F_2(A_{t+1} N_{t+1}, K_{t+1}) = F_2(N_{t+1}, K_{t+1}/A_{t+1}).$$

We have already seen that on a balanced growth path both N_{t+1} and K_{t+1}/A_{t+1} are constants. Thus, in the long run, the right hand side of equation (1.37) is constant and the left hand side must be, too. Now consider the resource constraint

$$K_{t+1} = Y_t - C_t + (1-\delta)K_t.$$

If capital and output grow at the common rate $a-1$, consumption must grow at the same rate, since otherwise the growth factor of capital g_K,

$$g_K := \frac{K_{t+1}}{K_t} = \frac{Y_t}{K_t} - \frac{C_t}{K_t} + (1-\delta),$$

is not constant. If consumption grows at the rate $a-1$ the marginal utility of consumption must fall at a constant rate. As we show in Appendix 2 this restricts the one-period utility function u to the class of constant-elasticity functions with respect to consumption. Further restrictions derive from condition (1.36). Since the marginal product of labor increases in the long run at the rate $a-1$ there must be exactly off-setting income and substitution effects with respect to the static labor supply decision. As we demonstrate in Appendix 2 we must restrict the one-period utility function (1.30) to

$$u(C, 1-N) = \begin{cases} C^{1-\eta} v(1-N) \text{ if } \eta \neq 1, \\ \ln C + v(1-N) \text{ if } \eta = 1. \end{cases} \tag{1.38}$$

The function v must be chosen so that $u(C, 1-N)$ is concave. Remember, that a function is concave, if and only if $u_{ii} \leq 0$ and $(u_{11}u_{22} - u_{12}^2) \geq 0$, and that it is strictly concave, if $u_{11} < 0$ and $(u_{11}u_{22} - u_{12}^2) > 0$.[27]. For example, in the parameterization of u that we use in Example 1.4.1 below, the restriction of η to $\eta > \theta/(1+\theta)$ implies that u is strictly concave.

Transformation to Stationary Variables. Given the restrictions on technology and preferences it is always possible to choose new variables that are constant in the long run. As an example,

[27] See, e.g., Takayama (1985), Theorem 1.E.13.

consider the deterministic Ramsey model (1.34). Assume $\eta \neq 1$ in (1.38) and deterministic growth of the efficiency level of labor according to (1.32). The static labor supply condition (1.36) can then be written as

$$\frac{v'(1-N_t)}{(1-\eta)v(1-N_t)}\frac{C_t}{A_t} = F_1(N_t, K_t/A_t) \tag{1.39}$$

and the intertemporal condition (1.37) is:

$$\begin{aligned}
&\frac{C_t^{-\eta}v(1-N_t)}{C_{t+1}^{-\eta}v(1-N_{t+1})} = \frac{(aA_t)^{\eta}C_t^{-\eta}v(1-N_t)}{A_{t+1}^{\eta}C_{t+1}^{-\eta}v(1-N_{t+1})} \\
&= \frac{a^{\eta}(C_t/A_t)^{-\eta}v(1-N_t)}{(C_{t+1}/A_{t+1})^{-\eta}v(1-N_{t+1})} \\
&= \beta(1-\delta+F_2(N_{t+1}, K_{t+1}/A_{t+1})).
\end{aligned} \tag{1.40}$$

Since F is homogenous of degree one, we can transform the resource constraint to

$$\frac{K_{t+1}}{A_{t+1}/a} = F(N_t, K_t/A_t) + (1-\delta)(K_t/A_t) - (C_t/A_t). \tag{1.41}$$

Equations (1.39) through (1.41) constitute a dynamic system in the new variables N_t, $c_t := C_t/A_t$, and $k_t = K_t/A_t$. Their stationary values N, c and k are found as solution to the system of three equations

$$\begin{aligned}
c &= \frac{(1-\eta)v(1-N)}{v'(1-N)}F_1(N,k), \\
1 &= \beta a^{-\eta}(1-\delta+F_2(N,k)), \\
0 &= F(N,k) - (1-\delta-a)k - c.
\end{aligned}$$

Note, that we can derive the efficiency conditions (1.39) through (1.41) from solving the problem

$$\max_{\{c_t, N_t\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \tilde{\beta}^t c_t^{1-\eta} v(1-N_t)$$

s.t.

$$\begin{aligned}
c_t &\leq F(N_t, k_t) + (1-\delta)k_t - ak_{t+1}, \\
k_0 &\quad \text{given},
\end{aligned}$$

with discount factor $\tilde{\beta} := \beta a^{1-\eta}$ in the stationary decision variables $c_t := C_t/A_t$ and $k_{t+1} := K_{t+1}/A_{t+1}$.

1.4.3 The Decentralized Economy

So far we have considered a single agent for ease of exposition. For each of the Ramsey models considered above, it is, however, straightforward to develop a model of a decentralized economy whose equilibrium allocation coincides with the equilibrium allocation of the respective Ramsey model. Since the latter is a utility maximizing allocation, the decentralized equilibrium is optimal in the sense of Pareto efficiency. In the static theory of general equilibrium with a finite dimensional commodity space the correspondence between a competitive equilibrium and a Pareto efficient allocation of resources is stated in the Two Fundamental Theorems of Welfare Economics.[28] The infinite horizon Ramsey model has infinitely many commodities. Nevertheless, as shown by Debreu (1954), it is possible to extend the correspondence between competitive equilibrium and Pareto efficiency to infinite dimensional commodity spaces.

We illustrate the relation between efficiency and intertemporal equilibrium by means of a simple example.

Firms. The production side of the economy consists of a large number of identical firms $i = 1, 2, \ldots, n$. Each firm uses labor N_i and capital K_i to produce a single output Y_i. The production function $ZF(AN_i, K_i)$ has the usual properties, in particular, it is homogenous of degree one (see page 51). Each firm hires labor and capital services on the respective markets. Let w and r denote the rental rates of labor and capital, respectively, in units of the final good. Since there is no link between successive periods, maximization of the firm's present value is equivalent to maximizing one-period profits

$$\Pi_i := ZF(AN_i, K_i) - wN_i - rK_i.$$

The first-order conditions imply

[28] For a statement, see, e.g., Starr (1997), pp. 144ff.

$$w = ZAF_1(AN_i, K_i) = ZAF_1(A, K_i/N_i),$$
$$r = ZF_2(AN_i, K_i) = ZF_2(A, K_i/N_i),$$

due to the homogeneity of degree zero of F_i. Since all firms face the same factor prices, they choose the same capital-labor ratio $k := K_i/N_i$ from the solution to the above equations. Therefore, output per unit of labor $y_i = Y_i/N_i = ZF(A, K_i/N_i)$ is the same for all firms: $y_i = y = ZF(A, k)$. These results imply the existence of an aggregate production function

$$Y = \sum_i Y_i = \sum_i N_i y = Ny = NZF(A, k) = ZF(AN, K),$$

where $N = \sum_i N_i$ and $K = \sum_i K_i$. In terms of this function, equilibrium on the markets for labor and capital services is given by

$$\begin{aligned} w &= ZAF_1(AN, K), \\ r &= ZF_2(AN, K), \end{aligned} \tag{1.42}$$

and the profits of all firms are zero:[29]

$$\Pi_i := Y_i - wN_i - rK_i = ZF(AN_i, K_i) \underbrace{-ZAF_1 N_i - ZF_2 K_i}_{-ZF(AN_i, K_i)}.$$

Households. Our example economy is populated by a continuum of households of mass 1, i.e., each individual household is assigned a unique real number h from the interval $[0, 1]$. All households have the same one-period utility function and the same time $t = 0$ capital stock. When they face a given path of output and factor prices they choose identical sequences of consumption and labor supply. Let $x(h)$ denote an arbitrary decision variable of household $h \in [0, 1]$ and put

$$x(h) = \bar{x}\, \forall h \in [0, 1].$$

[29] This is just Euler's theorem. For a general statement of this theorem, see, e.g., Sydsæter, Strøm and Berck (1999), p.28.

Since

$$\bar{x} = \int_0^1 x(h)dh = \int_0^1 \bar{x}dh$$

aggregate and individual variables are identical. As a consequence, we can consicer a representative member from $[0, 1]$ without explicit reference to his index h.

This representative household supplies labor services N_t with efficiency factor A_t and capital services K_t at the given real wage w_t and rental rate of capital r_t, respectively. He saves in terms of capital which depreciates at the rate $\delta \in (0, 1]$. Thus, his budget constraint reads:[30]

$$K_{t+1} - K_t \leq w_t N_t + (r - \delta)K_t - C_t. \tag{1.43}$$

The household seeks time paths of consumption and labor supply that maximize its life-time utility

$$\sum_{t=0}^{\infty} \beta^t u(C_t, 1 - N_t), \quad \beta \in (0, 1) \tag{1.44}$$

subject to (1.43) and the given initial stock of capital K_0. From the Lagrangean of this problem,

$$\begin{aligned}\mathscr{L} = \sum_{t=0}^{\infty} \beta^t \Big[& u(C_t, 1 - N_t) + \Lambda_t(w_t N_t + (1 - \delta + r_t)K_t \\ & - C_t - K_{t+1})\Big]\end{aligned}$$

we derive the following first-order conditions:

$$u_1(C_t, 1 - N_t) = \Lambda_t, \tag{1.45a}$$
$$u_2(C_t, 1 - N_t) = \Lambda_t w_t, \tag{1.45b}$$
$$\Lambda_t = \beta\Lambda_{t+1}(1 - \delta + r_{t+1}). \tag{1.45c}$$

[30] Here we use the fact that firms' profits are zero. In general, we must include the profits that firms distribute to their shareholders.

Using the factor market equilibrium conditions (1.42) to substitute for w_t and r_{t+1} and applying the Euler theorem to F,

$$Y_t = ZF(A_tN_t, K_t) = ZA_tF_1(A_tN_t, K_t)N_t + ZF_2(A_tN_t, K_t)K_t$$

equations (1.45) reduce to

$$\frac{u_2(C_t, 1-N_t)}{u_1(C_t, 1-N_t)} = ZA_tF_1(A_tN_t, K_t), \tag{1.46a}$$

$$\frac{u_1(C_t, 1-N_t)}{u_1(C_{t+1}, 1-N_{t+1})} = \beta(1-\delta+ZF_2(A_{t+1}N_{t+1}, K_{t+1})), \tag{1.46b}$$

$$K_{t+1} = ZF(A_tN_t, K_t) + (1-\delta)K_t - C_t. \tag{1.46c}$$

This system is identical to the first-order conditions that we derived for the Ramsey model (1.34) in equations (1.36) and (1.37) with the resource constraint being equal to (1.46c). Thus, the equilibrium time path of the decentralized economy is optimal in the sense that it maximizes the utility of all households given the resource constraint of the economy. On the other hand, a benevolent planer who solved the Ramsey problem (1.34) could implement this solution in terms of a competitive equilibrium. He simply has to choose time paths of wages and rental rates equal to the equilibrium sequences of the respective marginal products.

1.4.4 Numerical Solutions

The algorithms that we have employed to solve the stochastic Ramsey model (1.21) also work in the more general model with endogenous labor supply and technological progress. We use Example 1.4.1 to illustrate the basic steps. We will refer to this model as the benchmark model. More or less similar models appear amongst others in the papers by Hansen (1985), by King, Plosser, and Rebelo (1988a), and by Plosser (1989). In later chapters we will solve this model with various other methods. It thus serves as point of reference to compare the performance of different algorithms.

Example 1.4.1

Consider the following stochastic Ramsey model (1.21). The farmer solves:

$$\max_{C_0,N_0} \quad E_0\left[\sum_{t=0}^{\infty}\beta^t\frac{C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta}\right],$$
$$\beta\in(0,1),\theta\geq 0,\eta>\theta/(1+\theta),$$

s.t.

$$\left.\begin{array}{rcl} K_{t+1}+C_t &\leq& Z_t(A_tN_t)^{1-\alpha}K_t^{\alpha}+(1-\delta)K_t,\ \alpha\in(0,1),\\ A_{t+1} &=& aA_t,\ a\geq 1,\\ Z_{t+1} &=& Z_t^{\varrho}e^{\epsilon_t},\ \varrho\in(0,1),\ \epsilon_t\sim N(0,\sigma^2),\\ 0 &\leq& C_t,\\ 0 &\leq& K_{t+1},\end{array}\right\}\forall t,$$

$K_0,\ Z_0$ given.

From the Lagrangean

$$\mathscr{L}:=E_0\Big\{\sum_{t=0}^{\infty}\beta^t\Big[\frac{C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta}$$
$$+\Lambda_t\left(Z_t(A_tN_t)^{1-\alpha}K_t^{\alpha}+(1-\delta)K_t-C_t-K_{t+1}\right)\Big]\Big\}$$

we derive the following first-order conditions:

$$\frac{\theta C_t}{1-N_t}=(1-\alpha)Z_tA_t(A_tN_t)^{-\alpha}K_t^{\alpha},$$
$$C_t=Z_t(A_tN_t)^{1-\alpha}K_t^{\alpha}+(1-\delta)K_t-K_{t+1},$$
$$1=\beta E_t\left(\frac{C_t}{C_{t+1}}\right)^{\eta}\left(\frac{1-N_{t+1}}{1-N_t}\right)^{\theta(1-\eta)}$$
$$\times\left(1-\delta+\alpha Z_{t+1}(A_{t+1}N_{t+1})^{1-\alpha}K_{t+1}^{\alpha-1}\right).$$

In terms of stationary variables $c_t:=C_t/A_t$ and $k_t:=K_t/A_t$ this system is:

$$\frac{\theta c_t}{1-N_t}=(1-\alpha)Z_tN_t^{-\alpha}k_t^{\alpha},\tag{1.48a}$$

$$c_t = Z_t N_t^{1-\alpha} k_t^{\alpha} + (1-\delta)k_t - ak_{t+1}, \tag{1.48b}$$

$$1 = \beta a^{-\eta} E_t \left(\frac{c_t}{c_{t+1}}\right)^{\eta} \left(\frac{1-N_{t+1}}{1-N_t}\right)^{\theta(1-\eta)} \times \left(1-\delta+\alpha Z_{t+1} N_{t+1}^{1-\alpha} k_{t+1}^{\alpha-1}\right). \tag{1.48c}$$

Deterministic Extended Path. We are now in the position to derive the set of Euler equations that must be solved in Step 2.2 of Algorithm 1.3.1. For that purpose assume that the productivity shock equals its unconditional mean $Z \equiv 1$ forever so that the expectations operator can be dropped from (1.48). Since in the long run c_t as well as k_t and N_t are constant, we find the stationary solution if we put $c_t = c_{t+1} = c$ and $k_t = k_{t+1} = k$ in (1.48) and rearrange, where necessary. This yields:

$$\frac{N}{1-N} = \frac{1-\alpha}{\theta}\frac{N^{1-\alpha}k^{\alpha}}{c} = \frac{1-\alpha}{\theta}\frac{y}{c}, \tag{1.49a}$$

$$1 = \beta a^{-\eta}(1-\delta+\alpha N^{1-\alpha}k^{\alpha-1}) = \beta a^{-\eta}(1-\delta+\alpha(y/k)), \tag{1.49b}$$

$$\frac{y}{k} = \frac{N^{1-\alpha}k^{\alpha}}{k} = \frac{c}{k} + (a+\delta-1). \tag{1.49c}$$

Given the model's parameters, the second equation can be solved for y/k. We can use this result to get c/k and hence $(y/c) = (y/k)/(c/k)$ from the third equation, so that we can solve the first equation for N. Since

$$\frac{y}{k} = N^{1-\alpha}k^{\alpha-1}$$

we can solve this expression for k.

Now, choose T so that it is reasonable to assume that from the current period t the capital stock is close to k, i.e., $k_{t+T} \simeq k$. Furthermore, use (1.48b) to substitute for c_t and c_{t+1} in (1.48a) and (1.48c), replace Z_{t+s} with its expected value $Z_t^{\varrho^s}$, and put $k_{t+T} = k$. The result is a system with $2T-1$ equations in the $2T-1$ variables $(N_{t+s})_{s=0}^{T-1}, (K_{t+s})_{s=1}^{T-1}$:

$$
\begin{aligned}
0 &= Z_t^{\varrho^s} N_{t+s}^{1-\alpha} k_{t+s}^{\alpha} + (1-\delta)k_{t+s} - ak_{t+s+1} \\
&\quad - \frac{1-\alpha}{\theta}(1-N_{t+s}) Z_t^{\varrho^s} N_{t+s}^{-\alpha} k_{t+s}^{\alpha}, \\
0 &= \left(\frac{Z_t^{\varrho^{s+1}} N_{t+s+1}^{1-\alpha} k_{t+s+1}^{\alpha} + (1-\delta)k_{t+s+1} - ak_{t+s+2}}{Z_t^{\varrho^s} N_{t+s}^{1-\alpha} k_{t+s}^{\alpha} + (1-\delta)k_{t+s} - ak_{t+s+1}} \right)^{-\eta} \\
&\quad \times \left(\frac{1-N_{t+s}}{1-N_{t+s+1}} \right)^{\theta(1-\eta)} \\
&\quad - \beta a^{-\eta} \left(1 - \delta + \alpha Z_t^{\varrho^{s+1}} N_{t+s+1}^{1-\alpha} k_{t+s+1}^{\alpha-1} \right), s = 0, \ldots, T-2, \\
0 &= Z_t^{\varrho^{T-1}} N_{t+T-1}^{1-\alpha} k_{t+T-1}^{\alpha} + (1-\delta)k_{t+T-1} - ak \\
&\quad - \frac{1-\alpha}{\theta}(1-N_{t+T-1}) Z_t^{\varrho^{T-1}} N_{t+T-1}^{-\alpha} k_{t+T-1}^{\alpha}.
\end{aligned}
$$

In the GAUSS program `Ramsey4a.g` and the Fortran program `Ramsey4a.for` we use the modified Newton-Raphson method described in Algorithm 8.5.1 to solve this system.

Value Function Iteration. It requires only few additional work to use the value function algorithm to solve Example 1.4.1. In a first step we approximate the continuous valued Markov process by a Markov chain using Algorithm 9.2.1. Given the realizations $\mathbf{z} = [z_1, z_2, ..., z_m]$ and the associated transition matrix $P = (p_{ij})$ we must iterate over

$$
\begin{aligned}
& v^{s+1}(k, z_j) \\
&= \max_{N,k'} \frac{(z_j N^{1-\alpha} k^{\alpha} + (1-\delta)k - ak')^{1-\eta}(1-N)^{\theta(1-\eta)}}{1-\eta} \\
&\quad + \tilde{\beta} \sum_{l=1}^{m} p_{jl} v^s(k', z_l), \ \tilde{\beta} := \beta a^{1-\eta}.
\end{aligned}
$$

To find reasonable boundaries for the capital stock, we solve the system that defines the stationary solution for k using the minimal and maximal element in $\mathbf{z}$, respectively. Since N depends upon y/c which is independent of Z, the stationary solution for N is independent of Z, too. Thus

$$
k^*(z) = z^{1/(1-\alpha)} N (y/k)^{1/(\alpha-1)},
$$

where y/k is the solution of equation (1.49b). We use $k^*(z)$ together with N to initialize the value function in Step 2 of Algorithm 1.3.2. Thus, the element in the i–th row and j–th column of the Matrix U is given by:

$$u(z_j F(N, k^*(z_j)) + (1-\delta)k^*(z_j) - ak^*(z_j), 1-N).$$

We choose an interval $[\underline{k}, \bar{k}]$ that embeds $[k^*(z_1), k^*(z_m)]$, and define a grid $\mathscr{G}$ of n equally spaced points over this interval. The maximization problem is split into two steps. Given a pair (k, k') from $\mathscr{G}$, we first solve for the optimal N using the first-order condition (1.48a):

$$\theta \frac{z_i N^{1-\alpha} k^\alpha + (1-\delta)k - ak'}{1-N} = (1-\alpha) z_i N^{-\alpha} k^\alpha. \tag{1.51}$$

We use the single variable modified Newton-Raphson method to perform this step and code it in a subroutine that returns the utility derived from the triple (z, k, k'), when N is chosen to solve (1.51). The maximization over k' in the second step is done in the usual fashion within the subroutine that iterates over the value function. Since this routine only requires a pointer that returns the utility derived from (z, k, k'), there is no need to change that procedure.[31]

Before we present numerical results from this approach we digress and deal with the problem of selecting appropriate values for the model's parameters in the next section.

1.5 Model Calibration and Evaluation

The Ramsey model from Example 1.4.1, which is taken from King, Plosser, and Rebelo (1989a), pp. 198ff., presents an integrated framework for studying economic fluctuations in a growing economy. Since it excludes money altogether it is a real business cycle model. Models like this one have been used to demon-

[31] Note that $u(c, 1-N)$ with $c = zN^{1-\alpha}k^\alpha + (1-\delta)k - ak'$ and $N(z, k, k')$ as solution to (1.51) is strictly concave in k'.

strate that elementary economic principles may account for a substantial part of observed economic fluctuations. The common procedure in the studies of, e.g., KYDLAND and PRESCOTT (1982), LONG and PLOSSER (1983), and PLOSSER (1989) is to select reasonable values of the model's parameters (the calibration step), to simulate the model on a computer, and to compare the outcome in terms of second moments of the artificially created time series to those of real economic data. In this section we illustrate this approach using the model from Example 1.4.1 and seasonally adjusted quarterly economic data for the West German economy over the period 1975.i through 1989.iv.[32] We limit our attention to this time period for two reasons. Firstly, between 1960.i and 1975.i the West German average propensity to consume, c/y, has a clear upward trend. Had the German economy been on a balanced growth path this relation had been constant. Yet, the calibration step requires the steady state assumption to be approximately true. Secondly, the German unification in the fall of 1990 is certainly a structural break that violates the steady state assumption for the period after 1989.

Calibration. The basic idea behind calibration is the assumption that the real economic data were produced by the model at hand. To account for the representative agent nature of the model it is common to scale the data by the size of the population if appropriate. Since the model from Example 1.4.1 displays fluctuations around a stationary state, a valid procedure to select the model's key parameters is to use long-run time series averages.

In the stationary equilibrium of our model, output per household grows at the rate of labor augmenting technical progress $a-1$. Thus, we can infer a from fitting a linear time trend to gross domestic product at factor prices per capita. This gives $a = 1.005$, implying a quarterly growth rate of 0.5 percent. The second parameter of the production technology, α, equals the average wage share in gross domestic product at factor prices. The national accounts present no data on the wage income of self-employed

[32] Usually, the U.S. economy is taken for this purpose. But since this economy has been the focus of numerous real business cycle models we think it is interesting to use an economy that differs in a number of respects.

persons. Yet, from the viewpoint of economic theory, this group of households also receives wage income as well as capital income. To account for that fact we assume that the self-employed earn wages equal to the average wage of employees. Therefore, we get a higher wage share of $1 - \alpha = 0.73$ than the commonly used number of 0.65. The third parameter that describes the economy's production technology is the rate of depreciation δ. We compute this rate as the average ratio of quarterly real depreciation to the quarterly capital stock.[33] As compared to the number of 0.025 commonly used for the U.S. economy[34] our figure of $\delta = 0.011$ is much smaller. With these parameters at hand we can infer the productivity shock Z_t from the production function using the time series on the gross domestic product at factor prices Y_t, on hours H_t and on the stock of capital K_t:

$$Z_t = \frac{Y_t}{((1.005)^t H_t)^{0.73} K_t^{0.27}}.$$

Since our specification of the Markov process for Z_t implies

$$\ln Z_t = \varrho \ln Z_{t-1} + \epsilon_t,$$

where $\ln Z_t \approx (Z_t - Z)/Z$, we fit an AR(1) process to the percentage deviation of Z_t from its mean $Z \equiv 1$. This delivers our estimate of $\varrho = 0.90$ and of $\sigma = 0.0072$.

It is not possible to determine all of the parameters that describe the preferences of the representative household from aggregate time series alone. The critical parameter in this respect is the elasticity of the marginal utility of consumption $-\eta$. Microeconomic studies provide evidence that this elasticity varies both with observable demographic characteristics and with the level of wealth. Browning, Hansen, and Heckman (1999) argue that if constancy of this parameter across the population is imposed

[33] For this purpose we construct a quarterly series of the capital stock from yearly data on the stock of capital and quarterly data on investment and depreciation using the perpetual inventory method. The details of this approach can be found in the GAUSS program `GetPar1.g`.

[34] See, e.g., King, Plosser, and Rebelo (1988a), p. 214 and Plosser (1989), p. 75.

there is no strong evidence against η being slightly above one. We use $\eta = 2$ which implies that the household desires a smoother consumption profile than in the case of $\eta = 1$, i.e., the case of logarithmic preferences, which has been used in many studies. The reason for this choice is that a larger η reduces the variability of output, working hours, and investment, and, thus, provides a better match between the model and the respective German macroeconomic variables.[35] Once the choice of η is made there are several possibilities to select to value of the discount factor β. The first alternative uses the observed average (quarterly) capital-output ratio k/y to solve for β from equation (1.49b). In our case this violates the restriction $\beta < 1$. King, Plosser, and Rebelo (1988a), p. 207, equate the average rate of return on equity to $\alpha(y/k) - \delta$ in (1.49b) and solve for $\beta a^{-\eta}$. Other studies, e.g., Lucke (1998), p. 102, take the ex post real interest rate on short term bonds as estimators of $\alpha(y/k) - \delta$ in equation (1.49b). The average yearly return on the West German stock index DAX was about 8.5 percent, on the FAZ index 11.5 percent, and the ex post real interest rate on three month money market bonds about 2.7 percent. Given a and η, we use $\beta = 0.994$, which implies a yearly return of slightly above 6.5 percent. The final choice concerns the preference parameter θ. We use condition (1.49a) and choose θ so that $N = 0.13$, which is the average quarterly fraction of 1440 ($=16 \times 90$) hours spend on work by the typical German employee. Note that many other studies put $N = 1/3$ arguing that individuals devote about 8 hours a day to market activities.[36] However, we consider the typical individual to be an average over the total population, including children and retired persons. Therefore, we find a much smaller fraction of a sixteen hours day engaged in income earning activities. Table 1.1 summarizes our choice of parameters.

Model Evaluation. Although more formal testing procedures have been proposed (see Lucke (1998) for an overview) the typi-

[35] We strongly encourage you to use the executable `Ramsey4a.exe` to undertake a sensitivity analysis by changing the values of the parameters in the file `Parameters.txt`.

[36] See, e.g., Hansen (1985), p. 319f.

Table 1.1

Preferences	Production	
β=0.994	a=1.005	α=0.27
η=2.0	δ=0.011	ϱ=0.90
N=0.13	σ=0.0072	

cal way to evaluate small scale computable macroeconomic models is to compare the second moments of the time series obtained from simulations of the model to those of the respective macroeconomic aggregates. Most of these aggregates have an upward trend that must be removed to render the time series stationary. Most applications subject the logs of these aggregates to the Hodrick-Prescott or – for short – HP-filter that we describe in more detail in Section 9.4. The cyclical component of a time series that the filter returns is then the percentage deviation of the original series from its HP-trend component. The solution of our model are time paths of stationary variables $x_t := X_t/A_t$, where X_t denotes the level of the respective variable. Therefore, given our specification of the evolution of labor augmenting technical progress,

$$A_{t+1} = aA_t \quad \Leftrightarrow A_t = A_0 a^t,$$

we can recover the time paths of the logs of the levels from

$$\ln X_t = \ln x_t + \ln A_t = \ln x_t + \ln A_0 + at.$$

To get comparable results, we must apply the HP-filter to $\ln X_t$. Yet, we can bypass this step, since, as we demonstrate in 9.4, the cyclical component of $\ln x_t$ is equal to the cyclical component of $\ln X_t$.

Table 2.1 displays the results from solving and simulating the model from Example 1.4.1 using the two procedures described in the previous subsection. The second moments from the model are averages over 500 simulations. The length of the simulated time series is equal to the number of quarterly observations from 1975.i

through 1989.iv. At the beginning of the first quarter our model economy is on its balanced growth path. For the next 60 quarters it is hit by productivity shocks that drive the business cycle.

Table 1.2

	Extended Path			Value Function		
Variable	s_x	r_{xy}	r_x	s_x	r_{xy}	r_x
Output	1.45	1.00	0.64	1.41	1.00	0.62
	(1.14)	(1.00)	(0.80)	(1.14)	(1.00)	(0.80)
Investment	6.18	1.00	0.64	6.17	1.00	0.61
	(2.59)	(0.75)	(0.79)	(2.59)	(0.75)	(0.79)
Consumption	0.57	0.99	0.66	0.51	0.99	0.63
	(1.18)	(0.79)	(0.84)	(1.18)	(0.79)	(0.84)
Hours	0.78	1.00	0.64	0.79	1.00	0.61
	(0.78)	(0.40)	(0.31)	(0.78)	(0.40)	(0.31)
Real Wage	0.68	0.99	0.65	0.63	0.99	0.63
	(1.17)	(0.41)	(0.91)	(1.17)	(0.41)	(0.91)

Notes: Empirical values from HP-filtered German data in parenthesis. s_x:=standard deviation of HP-filtered simulated series of variable x, r_{xy}:=cross correlation of variable x with output, r_x:=first order autocorrelation of variable x.

For the extended path algorithm we assume that it takes 100 quarters to get close to the balanced growth path. Thus, for each of the 60 quarters we must solve a system of 199 non-linear equations to determine consumption and labor supply. Coded in Fortran it takes almost three hours on a Pentium III, 860 MHz machine to compute 500 artificial time series for output, investment, consumption, hours, and the real wage. For the value function iteration approach we used a grid of 2000 points and approximated the AR(1)-process of the productivity shock by a Markov chain with 9 elements. The difference between the largest and the smallest shock is equal to 3 times the standard deviation of the AR(1)-process. The Fortran program needs about 53 minutes to solve and simulate the model (the GAUSS program is much slower).

The results in Table 2.1 show small numeric differences between the time series moments generated by the extended path and the value function iteration algorithm. Those differences should be expected, since the approximation of the continuous valued AR(1)-process by a Markov chain is clearly not perfect. Experiments with the chosen parameters reveal that fitting an AR(1)-process to the realizations of the Markov process imply an AR(1) parameter of about 0.8 (instead of 0.9) and a standard deviation of the innovations of slightly below 0.007. To get close to the true parameters it requires a considerably larger interval (of about 20 times the size of the standard deviation of the AR(1) process) and much more grid points (about 71).

Now, consider the match between the data and the model's time series. The numbers in Table 2.1 reveal well known results. The model is able to reproduce the fact that investment is more volatile than output and consumption. Yet, it exaggerates this stylized fact of the business cycle. Furthermore, consumption is too smooth as compared to its empirical counterpart. Whereas the autocorrelations are quite in line with the data, the cross correlations between output and the other variables is almost perfect in the model, quite in contrast to the cross-correlations found in the data.

The quite obvious mismatch between the data and the artificial time series may be traced to two different sources. Firstly, we have not attempted to construct aggregates from the national income and product accounts (NIPA) that are consistent with the definition of output, capital, and labor in our model. Secondly, the benchmark model may be too simple to give an adequate account of the empirical facts.

Cooley and Prescott (1995) present a nice account of consistent measurement. Let us just consider two examples. The first one relates to consumption. In our model it is the flow of non-durables, whereas the NIPA report the quarterly expenditures on consumer durables and non-durables. From the viewpoint of our model, consumer durables are capital goods, and their purchases represent investment expenditures. Since the model predicts the latter to be more volatile than consumption, it should come as

no surprise that the consumer aggregate taken from the NIPA is more volatile than the consumption series from our model. As a second example take the capital stock. Since our model gives no explicit account of the government sector our measure of the capital stock includes the public stock of capital. Yet, the NIPA provide no data on depreciation for the public infrastructure. As a consequence, our measure of the rate of capital depreciation is biased downwards. Yet, with lower user costs of capital, the household's incentive for intertemporal substitution increases and investment becomes more volatile. For instance, if we increase δ from 0.0011 to 0.025, the ratio between the standard deviations of investment and output declines from about 4.3 to 3.5, which is much closer to the empirical ratio of 2.3.

To understand in what respects our benchmark model may be too simple, return to the household's first-order conditions with respect to consumption and labor supply in (1.48). For the present purpose we state them as follows:

$$w_t := (1-\alpha) Z_t N_t^{-\alpha} k_t^{\alpha},$$
$$w_t = \theta \lambda_t^{-1/\eta} (1-N_t)^{\frac{\theta(1-\eta)}{\eta}-1}.$$

The first line posits that the real wage per efficiency unit of labor w_t equals the marginal product of effective labor N_t. For a given capital stock k_t this relation defines a downward sloping labor demand schedule (see Figure 1.9) that is linear in the logs of the real wage and working hours, respectively. The second line defines an upward sloping labor supply schedule for a fixed multiplier λ_t.[37] A productivity shock raising Z from Z_1 to Z_2 shifts the labor demand schedule outward. Equilibrium in the labor market requires higher wages, and, as a result, the representative household supplies more hours. Thus, the immediate impact of the shock is to raise the real wage, hours, and output. Since current consumption is a normal good, it increases as a consequence of the higher current income. Investment increases for several reasons: Firstly, future consumption as well as future leisure are normal goods.

[37] Note, that we restricted η to $\eta > \theta/(1+\theta)$ so that the one-period utility function is strictly concave.

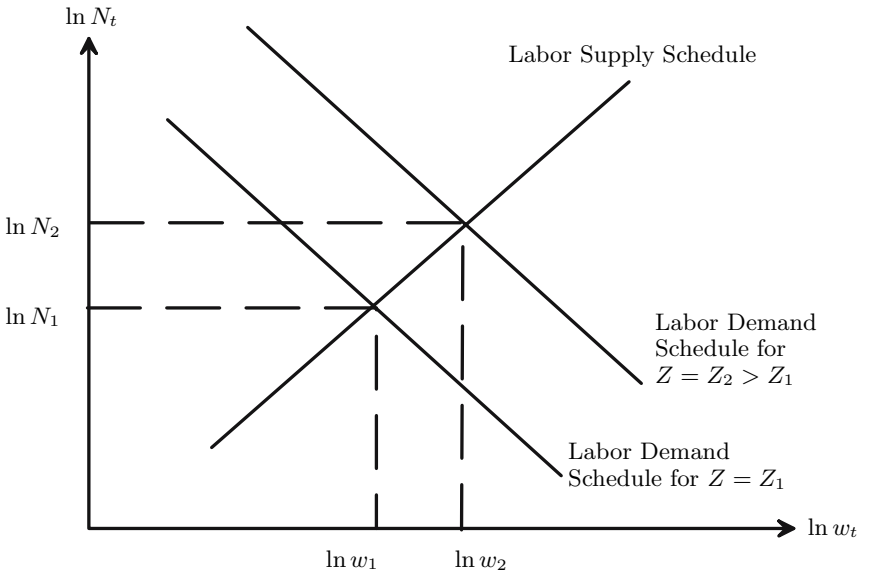

Figure 1.9: Productivity Shock in the Benchmark Business Cycle Model

Thus, the household wants to spend part of his higher current income on future consumption and future leisure. He builds up his stock of capital so that future production is potentially higher. Secondly, since the productivity shock is highly autocorrelated, the household expects above normal returns to capital. Thus, all variables in the model move closely together with income which, in turn, is driven by a single shock. In reality, however, there may be additional shocks. For instance, think of a preference shock that shift the labor supply curve to the left. That shock increases the real wage and reduces employment and output. As a consequence, the tight positive correlation between output, hours, and the real wage loosens.

In the problems to this chapter as well as in later chapters you will see how these and other extensions help to bring artificial and empirical data closer together. Yet, before we will construct more elaborate models, we proceed to introduce further solution techniques in the next chapter.

Appendix 1 Solution to Example 1.2.1

We derive the solution to Example 1.2.1 using iterations over the value function. Thus, letting $v^0 = 0$ we solve

$$v^1 = \max \ln(K^\alpha - K')$$

yielding $K' = 0$ and $v^1 = \alpha \ln K$. In the next step we seek K' that solves

$$v^2 = \max \ln(K^\alpha - K') + \beta\alpha \ln K'.$$

From the first order condition

$$\frac{1}{K^\alpha - K'} = \frac{\alpha\beta}{K'}$$

we get

$$\begin{aligned} K' &= \frac{\alpha\beta}{1+\alpha\beta} K^\alpha, \\ v^2 &= \alpha(1+\alpha\beta)\ln K + A_1, \\ A_1 &:= \ln(1/(1+\alpha\beta)) + \alpha\beta \ln(\alpha\beta/(1+\alpha\beta)). \end{aligned}$$

The value function in step $s = 3$ is given by

$$v^3 = \max \ln(K^\alpha - K') + \beta\alpha(1+\alpha\beta)\ln K' + \beta A_1$$

yielding

$$\begin{aligned} K' &= \frac{\alpha\beta + (\alpha\beta)^2}{1+\alpha\beta+(\alpha\beta)^2} K^\alpha, \\ v^3 &= \alpha(1+\alpha\beta+(\alpha\beta)^2)\ln K + A_2, \\ A_2 &= \ln\left[\frac{1}{1+\alpha\beta+(\alpha\beta)^2}\right] + \left(\alpha\beta+(\alpha\beta)^2\right)\ln\left[\frac{\alpha\beta+(\alpha\beta)^2}{1+\alpha\beta+(\alpha\beta)^2}\right] + \beta A_1. \end{aligned}$$

Continuing in this fashion we find the policy function in step s given by

$$K' = \frac{\sum_{i=1}^{s-1}(\alpha\beta)^s}{\sum_{i=0}^{s-1}(\alpha\beta)^s} K^\alpha$$

with limit $s \to \infty$ equal to

$$K' = \alpha\beta K^{\alpha}.$$

Obviously, from the first two steps, the value function is a linear function of $\ln K$. To infer the parameters of $v := \lim_{s\to\infty} v^s$, we use the method of undetermined coefficients.

Assume $v = a + b \ln K$. Solving

$$\max_{K'} \quad \ln(K^{\alpha} - K') + \beta(a + b \ln K')$$

yields

$$K' = \frac{\beta b}{1 + \beta b} K^{\alpha}.$$

Therefore

$$v = \underbrace{\alpha(1 + \beta b)}_{b} \ln K + \underbrace{\beta a + \ln\left[\frac{1}{1 + \beta b}\right] + \beta b \ln\left[\frac{\beta b}{1 + \beta b}\right]}_{a}.$$

Equating the constant on the right hand side of this equation to a and the slope parameter to b, we get:

$$\begin{aligned} b &= \alpha(1 + \beta b) \Rightarrow b = \frac{\alpha}{1 - \alpha\beta}, \\ a &= \beta a + \ln\left[\frac{1}{1 + \beta b}\right] + \beta b \ln\left[\frac{\beta b}{1 + \beta b}\right], \\ &\Rightarrow a = \frac{1}{1 - \beta}\left[\ln(1 - \alpha\beta) + \frac{\alpha\beta}{1 - \alpha\beta} \ln \alpha\beta\right]. \end{aligned}$$

Appendix 2 Restrictions on Technology and Preferences

Here we derive formally the restrictions that we must place on technology and preferences to ensure the existence of a balanced growth path. We draw heavily on the Appendix to King, Plosser, and Rebelo (1988) and, like Solow (1988), p. 35f., define a balanced growth path as an equilibrium that features (see page 51)

- a constant rate of output growth
- and a constant share of savings in output.

Technology. The constant share of savings S_t in output implies that output and capital must grow at the same rate: using the economy's resource constraint, we find:[38]

$$g_K = \frac{K_{t+1}}{K_t} = \frac{\overbrace{Y_t - C_t}^{S_t} + (1-\delta)K_t}{K_t} = \frac{S_t}{Y_t}\frac{Y_t}{K_t} + (1-\delta).$$

So, if S_t/Y_t is constant, so must be Y_t/K_t, and, hence, output and capital must grow at the same rate.

Now, consider the general case of labor and capital augmenting technical progress:

$$Y_t = F(A_t N_t, B_t K_t), \quad A_t = A_0 a^t,\ B_t = B_0 b^t.$$

Since F is linear homogenous, the growth factor of output, g_Y, can be factored as follows

$$g_y = \frac{Y_{t+1}}{Y_t} = \frac{B_{t+1}K_{t+1}}{B_t K_t}\frac{F(X_{t+1},1)}{F(X_t,1)} = b g_K g_F, \tag{A.2.1a}$$

$$X_t := (A_0/B_0)(a/b)^t(N_t/K_t) \Rightarrow g_X = \frac{a g_N}{b g_K}. \tag{A.2.1b}$$

Since $g_Y = g_K$ we get from (A.2.1a)

$$1 = b g_F.$$

There are two cases to consider:

1) $b = g_F \equiv 1$

[38] In the following the symbol g_X denotes the growth factor of the variable X.

2) and $g_F = 1/b, b > 1$.

In the first case technical progress is purely labor augmenting and for $g_F \equiv 1$ we must have $g_X = 1$, implying $g_K = ag_N$. Now, in our representative agent framework with a constant population size, N is bounded between zero and one. Thus, a constant rate of capital and output growth requires $g_N = 1$ (otherwise $N \to 1$). Therefore, output and capital grow at the rate of labor augmenting technical progress $a - 1$. For the share of savings to remain constant, consumption must also grow at this rate.

Now consider the second case. For

$$g_F := \frac{F(X_{t+1}, 1)}{F(X_t, 1)} = \text{constant} > 1$$

X_t must grow at the constant rate

$$g_X = \frac{ag_N}{bg_K},$$

which again implies $g_N = 1$. Let

$$X_t = X_0 c^t, \quad c = \frac{a}{bg_K},$$

and define $f(X_t) := F(X_t, 1)$ so that the condition reads

$$\frac{f(X_0 c^{t+1})}{f(X_0 c^t)} = \text{constant}.$$

Since this must hold for arbitrary given initial conditions X_0, differentiation with respect to X_0 implies

$$0 = \frac{1}{f(X_t)^2} \left\{ f(X_t) f'(X_{t+1}) \frac{X_{t+1}}{X_0} - f(X_{t+1}) f'(X_t) \frac{X_t}{X_0} \right\} dX_0,$$
$$0 = \left\{ \frac{f'(X_{t+1}) X_{t+1}}{f(X_{t+1})} - \frac{f'(X_t) X_t}{f(X_t)} \right\} \frac{f(X_{t+1})}{f(X_t)} \frac{dX_0}{X_0}.$$

For the term in curly brackets to be zero, the elasticity of f with respect to X_t must be a constant, say α:

$$\frac{f'(X_t) X_t}{f(X_t)} = \alpha.$$

Yet, the only functional form with constant elasticity is

$$f(X) = ZX^{\alpha}$$

with Z an arbitrary constant of integration. Thus, output must be given by a Cobb-Douglas function

$$\begin{aligned} Y &= F(AN, BK) = BK(f(AN/BK)) = BKZ(AN/BK)^{\alpha} \\ &= Z(AN)^{\alpha}(BK)^{1-\alpha}. \end{aligned}$$

Yet, if F is Cobb-Douglas, technical progress can always be written as purely labor-augmenting, since

$$Y_t = Z(A_t N_t)^{\alpha}(B_t K_t)^{1-\alpha} = Z(\tilde{A}_t N_t)^{\alpha} K_t^{1-\alpha}, \quad \tilde{A}_t := A_t B_t^{(1-\alpha)/\alpha}.$$

Preferences. Consider equation (1.37) which determines the farmer's savings decision. We reproduce it here for convenience:

$$\frac{u_1(C_t, 1-N_t)}{u_1(C_{t+1}, 1-N_{t+1})} = \beta(1 - \delta + F_2(A_{t+1}N_{t+1}, K_{t+1})). \tag{A.2.2}$$

On a balanced growth path with constant supply of labor the right hand side of this equation is a constant, since A_t and K_t grow at the same rate and $F_2(AN, K) = F_2(N, K/A)$. On that path the resource constraint is given by

$$\begin{aligned} C_t &= F(A_t N_t, K_t) + (1-\delta)K_t - K_{t+1} \\ &= A_t\big[F(N, K/A) + (1-\delta)(K/A) - a(K/A)\big]. \end{aligned}$$

Since the term in brackets is constant, consumption grows at the rate $a - 1$, and we may write:

$$C_t = C_0 a^t.$$

On the balanced growth path equation (A.2.2), thus, may be written as:

$$\frac{u_1(C_0 a^t, 1-N)}{u_1(C_0 a^{t+1}, 1-N)} = \Delta = \text{constant}.$$

This must hold irrespective of the arbitrary constant C_0. Differentiating with respect to C_0 yields:

$$\Delta \frac{dC_0}{C_0} \left\{ \frac{u_{11}(C_t, 1-N)}{u_1(C_t, 1-N)} C_t - \frac{u_{11}(C_{t+1}, 1-N)}{u_1(C_{t+1}, 1-N)} C_{t+1} \right\} = 0.$$

The term in curly brackets is zero, if the elasticity of the marginal utility of consumption $(u_{11}/u_1)C$, is a constant, say $-\eta$. Integrating

$$\frac{du_1(C, 1-N)}{u_1(C, 1-N)} = -\eta \frac{dC}{C}$$

on both sides gives

$$\ln u_1(\cdot) = -\eta \ln C + \ln v_1(1-N), \Rightarrow u_1(\cdot) = C^{-\eta} v_1(1-N)$$

where v_1 is an arbitrary function of leisure $1-N$. Integrating once more with respect to C yields

$$u(C, 1-N) = \begin{cases} C^{1-\eta} v_1(1-N) + v_2(1-N) \text{ if } \eta \neq 1, \\ v_1(1-N) \ln C + v_2(1-N) \text{ if } \eta = 1. \end{cases} \tag{A.2.3}$$

Restrictions on the functions v_1 and v_2 derive from the static condition on labor supply in equation (1.36). Remember, this condition is

$$\frac{u_2(C_t, 1-N_t)}{u_1(C_t, 1-N_t)} = A_t F_1(A_t N_t, K_t),$$

in general, and

$$\frac{u_2(C, 1-N)}{u_1(C, 1-N)} = A F_1(N, K/A)$$

along the balanced growth path. Write this as

$$\ln[u_2(C, 1-N)] = \ln[u_1(C, 1-N)] + \ln A + \ln[F_1(N, K/A)],$$

and differentiate with respect to C and A. The result is

$$\underbrace{\frac{u_{21}(\cdot)}{u_2(\cdot)} C}_{\xi} \frac{dC}{C} = \underbrace{\frac{u_{11}(\cdot)}{u_1(\cdot)} C \frac{dC}{C}}_{-\eta} + \frac{dA}{A},$$

where ξ denotes the elasticity of the marginal utility of leisure with respect to consumption. Since $dC/C = dA/A$ in the long-run, this condition restricts ξ to

$$\xi = 1 - \eta.$$

Using (A.2.3) this implies that

- $v_2(1-N)$ is a constant in the case of $\eta \neq 1$,
- $v_1(1-N)$ is a constant in the case of $\eta = 1$.

Setting the respective constants equal to zero and 1, respectively, yields the functional forms of the one-period utility function given in (1.38).

Problems

1.1 **Transition Dynamics and Endogenous Growth.** The following endogenous growth model is based upon Lucas (2000). The description of the dynamics is adapted from Grüner and Heer (2000).
Consider the following deterministic Ramsey problem that is augmented by a human capital sector. Households live inifinitely maximizing intertemporal utility:

$$\sum_{t=0}^{\infty} \beta^t \frac{\left(c_t l_t^{\theta}\right)^{1-\eta}}{1-\eta}, \quad 0 < \beta < 1,\ 0 < \theta,$$

where c_t and l_t denote consumption and leisure in period t. The individual can allocate his time endowment B to work n, learning v and leisure l:

$$B = n_t + v_t + l_t.$$

The human capital of the representative individual h is determined by the time v he allocates to learning according to:

$$h_{t+1} = h_t \left(1 + D v_t^{\gamma}\right).$$

Physical capital k_t accumulates according to:

$$k_{t+1} = (1-\tau_w) n_t h_t w_t + (1 + (1-\tau_r) r_t)\, k_t + b_t - c_t,$$

where wage income and interest income are taxes at the rates τ_w and τ_r respectively. Pre-tax wage income is given by the product of the wage rate w_t, the time people work n_t, and the human capital h_t. r_t and b_t denote the real interest rate and government transfers, respectively.
Production per capita y is a function of capital k and effective labor nh. Output is produced with a CES technology:

$$y_t = F(k, nh) = a_0 \left(a_1 k^{\sigma_p} + a_2 (nh)^{\sigma_p}\right)^{\frac{1}{\sigma_p}},$$

where σ_p denotes the elasticity of substitution in production. Define the state variable $z \equiv \frac{k}{nh}$. The production per effective labor is defined by

$f(z) \equiv F(z, 1)$. In a factor market equilibrium, factors are rewarded with their marginal products:

$$w = f(z) - zf'(z),$$
$$r = f'(z).$$

The government receives revenues from taxing labor income and capital income. The government budget is balanced so that government consumption g and transfers b equal tax revenues in any period:

$$g_t + b_t = \tau_w n_t h_t w_t + \tau_r r_t k_t.$$

Periods t correspond to years. The model is calibrated as follows: $\eta = 2.0$, $\theta = 0.5$, $\eta = 0.97$, $B = 2.13$, $D = 0.035$, $\gamma = 0.8$, $\sigma_p = -2/3$, $a_0 = 0.77$, $a_1 = 0.36$, $a_2 = 0.64$, $\tau_w = 0.36$, $\tau_r = 0.40$. The share of government consumption in output is $g/y = 0.21$.

a) Derive the first-order conditions of the household and the equilibrium conditions of the model.
b) On a balanced growth path, consumption, output, physical capital, and human capital grow at a constant rate μ, while the time allocation is constant. Derive the equations that characterize the balanced growth equilibrium. For this reason, express the equations with the help of stationary variables. For example, divide the government budget constraint by y_t.
c) Use our non-linear equation solver to compute the stationary equilibrium.
d) How does the growth rate react to a reduction of the capital income tax rate τ_r from 40% to 25% that is financed i) by a reduction in transfers b_t and ii) by an increase in the wage income tax rate τ_w? Explain why the growth rate decreases in the latter case.
e) Compute the dynamics for the transition between the old steady state that is characterized by a capital income tax rate $\tau_r = 40\%$ and the new steady state that is characterized by $\tau_r = 25\%$. Assume that during the transition and in the new steady state, g/y and b/y are constant and that the wage income tax rate τ_w adjusts in order to balance the government budget. Use backward iteration to compute the dynamics. (difficult)

1.2 **Stochastic Growth.** In the benchmark model of Example 1.4.1 labor-augmenting technical progress grows deterministically. Suppose instead the following production function

$$Y_t = (A_t N_t)^{1-\alpha} K_t^{\alpha},$$

where the log of labor augmenting technical progress A_t is a random walk with drift μ:

$$\ln A_t = \mu + \ln A_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma^2).$$

The household's preferences are the same as those presented in Example 1.4.1. Use $\mu = 0.00465$ and $\sigma = 0.006$ and the parameter values given in Table 1.1 to calibrate the model. Except for the stock of capital define stationary variables as in Section 1.4.2. For the capital stock define $k_t := K_t/A_{t-1}$. This ensures that the stationary capital stock k_t is still a predetermined variable at the beginning of period t.

a) Derive the first-order conditions for the planers problem:

$$\max_{C_0, N_0} \quad E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta} \right\}$$

s.t.

$$\left.\begin{array}{rcl} K_{t+1} + C_t & \leq & (A_t N_t)^{1-\alpha} K_t^{\alpha} + (1-\delta)K_t, \\ A_t & = & A_t e^{\mu+\epsilon_t}, \\ 0 & \leq & C_t, \\ 1 & \geq & N_t \geq 0, \\ 0 & \leq & K_{t+1}, \end{array}\right\} t = 0, 1, \ldots,$$

$K_0,\ A_0$ given.

b) State this set of equations in terms of stationary variables and compute the balanced growth path.

c) Place a grid over k_t and solve the model via value function iteration. (Hint: Don't forget to use the correct discount factor!).

d) Use a random number generator and simulate the model. Compute second moments from first-differences of the logged variables (why?) and compare their empirical analogs obtained from German data (see the following table).

Variable	s_x	r_{xy}	r_x
Output	0.75	1.00	0.24
Consumption	0.76	0.56	0.04
Investment	1.99	0.68	0.25
Hours	0.97	0.59	−0.26
Real Wage	1.01	−0.14	−0.23

Notes: Second moments from first differences of logged German data, 70.i to 89.iv. s_x:= standard deviation of variable x, s_{xy}:=cross correlation of x with output y, r_x:=first order autocorrelation.

1.3 **Productivity and Preference Shocks.** Empirically the correlation between working hours and the real wage is close to zero. The benchmark model, however, predicts a strong positive correlation. In the following model, which is adapted from HOLLAND and SCOTT (1998), we introduce

a preference shock in the benchmark model of Example 1.4.1. Specifically, we assume that the parameter θ in the momentary utility function of the representative household is not a constant but a random variable θ_t that is governed by a first-order autoregressive process:

$$\theta_t = \theta^{1-\gamma}\theta_{t-1}^{\gamma}e^{\xi_t}, \quad \gamma \in [0,1],\ \xi_t \sim N(0,\sigma_\xi^2).$$

The innovations ξ_t induce shifts of the labor supply schedule along a given labor demand schedule. By this, they counteract the positive correlation between the real wage and working hours introduced by shocks to total factor productivity Z_t. The planers problem is a follows:

$$\max_{C_0,N_0} \quad E_0\left\{\sum_{t=0}^{\infty}\beta^t\frac{C_t^{1-\eta}(1-N_t)^{\theta_t(1-\eta)}}{1-\eta}\right\}$$

s.t.

$$\left.\begin{array}{rcl} K_{t+1}+C_t & \leq & Z_t(A_tN_t)^{1-\alpha}K_t^{\alpha}+(1-\delta)K_t, \\ A_t & = & aA_{t-1}, \quad a \geq 1, \\ Z_t & = & Z_{t-1}^{\varrho}e^{\epsilon_t}, \quad \epsilon_t \sim N(0,\sigma_\epsilon^2), \\ 0 & \leq & C_t, \\ 1 & \geq & N_t \leq 0, \\ 0 & \leq & K_{t+1}, \end{array}\right\} t = 0,1,\ldots,$$

$K_0,\ A_0\ Z_0$ given.

Use the parameter values given in Table 1.1 to calibrate this model. In addition, put $\gamma = 0.9$ and $\sigma_\gamma = 0.01$ and calibrate θ so that the stationary fraction of working hours equals $N = 0.13$.

a) Derive the first-order conditions for the planer's problem and write it down in terms of stationary variables. Modify the extended path algorithm 1.3.1 to suit this model.
b) Simulate the model several hundred times. Pass the time series for working hours and the real wage to the HP-filter and compute the average cross-correlation between those two variables.
c) Repeat this exercise for a value of $\sigma_x i$ close to zero.

1.4 **Business Cycle Fluctuations and Home Production.**
In the US economy, hours worked fluctuate considerably more than productivity, and the correlation is close to zero. The standard real business cycle model presented in Section 1.4 has considerable difficulties to replicate this fact. For our German calibration, for example, hours worked and productivity have approximately equal standard deviations (0.77% and 0.72%, respectively). The following extension of the stochastic growth model is based on Benhabib, Rogerson, and Wright (1991). In their model, agents work in the production of both a market-produced good M and a home-produced good H.

Households maximize intertemporal utility

$$E_0\Big\{\sum_{t=0}^{\infty}\beta^t\Big[\frac{C_t^{1-\eta}(L_t)^{\theta(1-\eta)}}{1-\eta}\Big]\Big\}$$

where C_t is the following composite of the consumptions of good M and H:

$$C_t=\Big(aC_{Mt}^{\phi}+(1-a)C_{Ht}^{\phi}\Big)^{\frac{1}{\phi}}.$$

The time endowment of one unit is allocated to market and home production according to:

$$1=L_t-N_{Mt}-N_{Ht}.$$

Notice that the two types of work are assumed to be perfect substitutes, while the two consumption goods are combined by an aggregator that implies a constant elasticity of substitution equal to $\phi/(1-\phi)$.
The model has two technologies:

$$Y_{Mt}=F\left(Z_{Mt},K_{Mt},N_{Mt}\right)=Z_{Mt}K_{Mt}^{\alpha}N_{Mt}^{1-\alpha},$$
$$Y_{Ht}=G\left(Z_{Ht},K_{Ht},N_{Ht}\right)=Z_{Ht}K_{Ht}^{\gamma}N_{Ht}^{1-\gamma}.$$

The technology shocks follows the processes:

$$Z_{M,t+1}=\rho Z_{Mt}+\epsilon_{Mt},$$
$$Z_{H,t+1}=\rho Z_{Ht}+\epsilon_{Ht},$$

where $\epsilon_{it}\sim N(0,\sigma_i^2)$ are normally i.i.d. for $i=M,H$ and have a contemporaneous correlation $r_{MH}=cor(\epsilon_{Mt},\epsilon_{Ht})$.
Total capital $K_t=K_{Mt}+K_{Ht}$ accumulates according to

$$K_{t+1}=(1-\delta)K_t+I_t.$$

New capital is produced only in the market sector implying the constraints:

$$C_{Mt}+I_t=Y_{Mt},$$
$$C_{Ht}=Y_{Ht}.$$

Model periods correspond to quarters. The model is calibrated as follows: $\beta=0.99$, $\alpha=0.36$, $\delta=0.025$, $\eta=1.5$, $\phi=0.8$, $\gamma=0.08$, $r_{MH}=0.66$, $\rho=0.9$, $\sigma_M=\sigma_H=0.007$. The steady state leisure $\bar{L}=0.7$ is used to calibrate θ. a is set so that $C_H/C_M=1/4$.

a) Derive the first-order conditions of the model.
b) Compute the steady state and calibrate the parameters a and θ.
c) Compute the standard deviation of hours worked in the market activity, N_{Mt} , and productivity, Z_{Mt}, as well as the correlation of N_{Mt} and Z_{Mt}. Apply the HP-filter to the simulated time series. Use the techniques applied to the solution of the stochastic growth model with endogenous labor as presented in section 1.4. Explain why the variance of hours worked has increased. Vary ϕ and analyze the sensitivity of your result with regard to this parameter. Explain your result.

Chapter 2

Linear Quadratic and Linear Approximation Methods

Overview. This chapter describes algorithms that derive from the properties of the stochastic linear quadratic optimal control problem; also referred to as the optimal linear regulator problem, or LQ problem for short. This class of problems features a quadratic current period return function and a linear law of motion. The policy function is linear and there are several methods to compute it numerically. One of them iterates over the value function and another one solves the related system of stochastic linear difference equations using methods from linear algebra.

Two different approximation methods derive from the LQ problem. The first approach approximates a given model so that its return function is quadratic and the law of motion is linear and solves the approximate model by value function iterations. The second approach relies on a linear approximation of the model's Euler equations and law of motion and solves the ensuing system of linear stochastic difference equations.

In the next section we state the LQ problem and derive some important properties of its solution. Section 2.2 presents an algorithm based on the linear quadratic approximation. This algorithm works well for models where the equilibrium allocation is identical to the solution of a central planing problem. It is, however, cumbersome to adapt to models where the equilibrium of the decentralized economy differs from the planing solution. In this case, it is much easier to solve the approximate linear model by the method presented in Section 2.3. Various applications will illustrate these points in Section 2.4.

2.1 The Linear Quadratic Model

This section presents the linear quadratic model and derives some of its important properties. Since its main purpose is to provide a framework for both linear quadratic and linear approximation methods, we postpone detailed algorithms for the computation of the policy function until Section 2.2 and Section 2.3, respectively.

2.1.1 Description

Consider an economy governed by the following stochastic linear law of motion:

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + \boldsymbol{\epsilon}_{t+1}. \tag{2.1}$$

The n-dimensional column vector $\mathbf{x}_t$ holds those variables that are predetermined at period t. A fictitious social planner sets the values of the variables stacked in in the m-dimensional column vector $\mathbf{u}_t$. We refer to $\mathbf{x}$ as the state vector and to $\mathbf{u}$ as the control vector. $A \in \mathbb{R}^{n\times n}$ and $B \in \mathbb{R}^{n\times m}$ are matrices. Due to the presence of shocks, the planner cannot control this economy perfectly. The n vector of shocks $\boldsymbol{\epsilon}$ has a multivariate normal distribution with $E(\boldsymbol{\epsilon}) = \mathbf{0}$ and covariance matrix[1] $E(\boldsymbol{\epsilon}\boldsymbol{\epsilon}') = \Sigma$. The planner must choose $\mathbf{u}_t$ before he can realize the size of the shocks.

Given $\mathbf{x}_0$ the planner's objective is to maximize

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\mathbf{x}_t' Q \mathbf{x}_t + \mathbf{u}_t' R \mathbf{u}_t + 2\mathbf{u}_t' S \mathbf{x}_t\right], \quad \beta \in (0,1), \tag{2.2}$$

subject to (2.1). The current period objective function

$$g(\mathbf{x}_t, \mathbf{u}_t) := \begin{bmatrix}\mathbf{x}_t', & \mathbf{u}_t'\end{bmatrix} \begin{bmatrix} Q & S' \\ S & R \end{bmatrix} \begin{bmatrix}\mathbf{x}_t \\ \mathbf{u}_t\end{bmatrix} \tag{2.3}$$

is quadratic and concave in $(\mathbf{x}_t', \mathbf{u}_t')$. This requires that both the symmetric $n \times n$ matrix Q and the symmetric $m \times m$ matrix R are negativ semidefinite.

Note that this specification encompasses non-stochastic state variables and first-order (vector) autoregressive processes.

[1] Remember that a prime denotes transposition, i.e., $\boldsymbol{\epsilon}'$ is a row vector and $\boldsymbol{\epsilon}$ a column vector.

2.1.2 Derivation of the Policy Function

The Bellman equation for the stochastic LQ problem is given by

$$v(\mathbf{x}) := \max_{\mathbf{u}} \mathbf{x}'Q\mathbf{x} + 2\mathbf{u}S\mathbf{x} + \mathbf{u}'R\mathbf{u} + \beta Ev(A\mathbf{x} + B\mathbf{u} + \boldsymbol{\epsilon}), \quad (2.4)$$

where we used (2.1) to replace next period's state variables in $Ev(\cdot)$. Expectations are taken conditional on the information contained in the current state $\mathbf{x}$. We guess that the value function is given by $v(\mathbf{x}) := \mathbf{x}'P\mathbf{x} + d$, P being a n dimensional symmetric, negative semidefinite square matrix and $d \in \mathbb{R}$ an unknown constant.[2] we may write (2.4) as follows:[3]

$$\begin{aligned}\mathbf{x}'P\mathbf{x} + d = \max_{\mathbf{u}} \Big[&\mathbf{x}'Q\mathbf{x} + 2\mathbf{u}S\mathbf{x} + \mathbf{u}'R\mathbf{u} \\ &+ \beta E\Big((A\mathbf{x} + B\mathbf{u} + \boldsymbol{\epsilon})'P(A\mathbf{x} + B\mathbf{u} + \boldsymbol{\epsilon}) + d\Big)\Big].\end{aligned} \quad (2.5)$$

Evaluating the conditional expectations on the rhs. of (2.5) yields:

$$\begin{aligned}\mathbf{x}'P\mathbf{x} + d = \max_{\mathbf{u}} \Big[&\mathbf{x}'Q\mathbf{x} + 2\mathbf{u}S\mathbf{x} + \mathbf{u}'R\mathbf{u} + \beta\mathbf{x}'A'PA\mathbf{x} \\ &+ 2\beta\mathbf{x}'A'PB\mathbf{u} + \beta\mathbf{u}'B'PB\mathbf{u} + \beta\mathrm{tr}\,(P\Sigma) + \beta d\Big].\end{aligned} \quad (2.6)$$

In the next step we differentiate the rhs. of (2.6) with respect to the control vector $\mathbf{u}$, set the result equal to the zero vector, and solve for $\mathbf{u}$. This provides the solution for the policy function:

$$\mathbf{u} = -\underbrace{(R + \beta B'PB)^{-1}(S + \beta B'PA)}_{F}\mathbf{x}. \quad (2.7)$$

To find the solution for the matrix P and the constant d, we eliminate $\mathbf{u}$ from the Bellman equation (2.6) and compare the

[2] Note, since $\mathbf{x}_t'P\mathbf{x}_t$ is a quadratic form, it is not restrictive to assume that P is symmetric. Furthermore, since the value function of a well defined dynamic programming problem is strictly concave, P must be negative semidefinite.

[3] If you are unfamiliar with matrix algebra, you may find it helpful to consult Section 8.1. We present the details of the derivation of the policy function in Appendix 3.

quadratic forms and the constant terms on both sides. It turns out that P must satisfy the following implicit equation, known as algebraic matrix Riccati equation :

$$\begin{aligned} P = Q + \beta A'PA \\ - (S + \beta B'PA)'(R + \beta B'PB)^{-1}(S + \beta B'PA), \end{aligned} \tag{2.8}$$

and that d is given by:

$$d = \frac{\beta}{1-\beta}\text{tr}\,(P\Sigma).$$

The solution of (2.8) can be obtained by iterating on the matrix Riccati difference equation

$$\begin{aligned} P_{s+1} = Q + A'(P_s)A \\ - (S + \beta B'(P_s)A)'(R + \beta B'(P_s)B)^{-1}(S + \beta B'(P_s)A) \end{aligned}$$

starting with some initial negative definite matrix P_0.[4] Other methods to solve (2.8) rely on eigenvalue-eigenvector decompositions. Since we will use iterations over the value function later on, we will not explore these methods further. Once the solution for P has been computed, the dynamics of the model is governed by

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + \boldsymbol{\epsilon}_{t+1} = (A - FB)\mathbf{x}_t + \boldsymbol{\epsilon}_{t+1}.$$

2.1.3 Certainty Equivalence

The solution of the stochastic LQ problem has a remarkable feature. Since the covariance matrix of the shocks Σ appears neither in equation (2.7) nor in equation (2.8), the optimal control is independent of the stochastic properties of the model summarized by Σ. Had we considered a deterministic linear quadratic problem by assuming $\boldsymbol{\epsilon}_t = \mathbf{0}\forall t$, we would have found the same feedback

[4] For example $P_0 = -0.01I_n$.

rule (2.7). You may want to verify this claim by solving Problem 2.1. This property of the stochastic LQ problem is called *certainty equivalence principle*. It is important to note that if we use the LQ approximation to solve an arbitrary economic model we enforce the certainty equivalence principle on this solution. This may hide important properties of the model. It is, for instance, hard to imagine that agents in two otherwise identical economies A and B use the same feedback rules if the variance of productivity shocks in economy A is much larger than in economy B.

2.1.4 Derivation of the Euler Equations

As we have seen in Chapter 1 an alternative way to derive the dynamic path of an optimizing model is to consider the model's Euler equations. It will be helpful for the approach taken in Section 2.3 to separate the state variables into two categories. Variables that have a given initial condition but are otherwise determined endogenously are stacked in the n vector $\mathbf{x}$. Purely exogenous shocks are summarized in the l vector $\mathbf{z}$. As in the previous subsection $\mathbf{u}$ is the m vector of controls. The planner's current period return function is the following quadratic form:

$$\begin{aligned} g(\mathbf{x}_t, \mathbf{u}_t, \mathbf{z}_t) := \mathbf{x}_t' A_{xx} \mathbf{x}_t + \mathbf{u}_t' A_{uu} \mathbf{u}_t + \mathbf{z}_t A_{zz} \mathbf{z}_t \\ + 2\mathbf{u}_t' A_{ux} \mathbf{x}_t + 2\mathbf{u}_t' A_{uz} \mathbf{z}_t + 2\mathbf{x}_t' A_{xz} \mathbf{z}_t. \end{aligned} \tag{2.9}$$

$A_{ij}, i, j \in \{x, u, z\}$ are given matrices. The transition law of the endogenous state variables is

$$\mathbf{x}_{t+1} = B_x \mathbf{x}_t + B_u \mathbf{u}_t + B_z \mathbf{z}_t, \tag{2.10}$$

where $B_x \in \mathbb{R}^{n \times n}$, $B_u \in \mathbb{R}^{n \times m}$, and $B_z \in \mathbb{R}^{n \times l}$ are given matrices. The shocks follow a first-order vector autoregressive process

$$\mathbf{z}_{t+1} = \Pi \mathbf{z}_t + \boldsymbol{\epsilon}_{t+1}, \quad \boldsymbol{\epsilon} \sim N(\mathbf{0}, \Sigma). \tag{2.11}$$

The eigenvalues of $\Pi \in \mathbb{R}^{l \times l}$ lie inside the unit circle. The planner maximizes

$$E_0 \sum_{t=0}^{\infty} \beta^t g(\mathbf{x}_t, \mathbf{u}_t, \mathbf{z}_t) \tag{2.12}$$

subject to (2.10) and (2.11).

Let $\boldsymbol{\lambda_t}$ denote the n vector of Lagrange multipliers. The Lagrangian of this LQ problem is

$$\mathscr{L} = E_0 \sum_{t=0}^{\infty} \beta^t \Big[g(\mathbf{x}_t, \mathbf{u}_t, \mathbf{z}_t) + 2\boldsymbol{\lambda}_t'(B_x\mathbf{x}_t + B_u\mathbf{u}_t + B_z\mathbf{z}_t - \mathbf{x}_{t+1}) \Big].$$

Differentiating this expression with respect to $\mathbf{u}_t$ and $\mathbf{x}_{t+1}$ provides the following set of first-order conditions:

$$\begin{aligned} 0 &= A_{uu}\mathbf{u}_t + A_{ux}\mathbf{x}_t + A_{uz}\mathbf{z}_t + B_u'\boldsymbol{\lambda}_t, \\ \boldsymbol{\lambda}_t &= \beta E_t \left[A_{xx}\mathbf{x}_{t+1} + A_{xz}\mathbf{z}_{t+1} + A_{ux}'\mathbf{u}_{t+1} + B_x'\boldsymbol{\lambda}_{t+1} \right]. \end{aligned}$$

The first of these equations may be rewritten as:

$$C_u\mathbf{u}_t = C_{x\lambda} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} + C_z\mathbf{z}_t, \tag{2.13a}$$

whereas the second equation and the transition law (2.10) can be summarized in the following matrix difference equation:

$$\begin{aligned} D_{x\lambda}E_t \begin{bmatrix} \mathbf{x}_{t+1} \\ \boldsymbol{\lambda}_{t+1} \end{bmatrix} + F_{x\lambda} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} &= D_uE_t\mathbf{u}_{t+1} + F_u\mathbf{u}_t \\ &\quad + D_zE_t\mathbf{z}_{t+1} + F_z\mathbf{z}_t. \end{aligned} \tag{2.13b}$$

The matrices in these equations relate to those of the original problem as follows:

$$\begin{aligned} C_u &:= A_{uu}, & C_{x\lambda} &:= -\left[A_{ux}, B_u'\right], \\ C_z &:= -A_{uz}, \end{aligned}$$

$$D_{x\lambda} := \begin{bmatrix} \beta A_{xx} & \beta B_x' \\ I_n & 0_{n\times n} \end{bmatrix}, \quad F_{x\lambda} := \begin{bmatrix} 0_{n\times n} & -I_n \\ -B_x & 0_{n\times n} \end{bmatrix},$$

$$D_u := \begin{bmatrix} -\beta A_{ux}' \\ 0_{n\times m} \end{bmatrix}, \qquad F_u := \begin{bmatrix} 0_{n\times m} \\ B_u \end{bmatrix},$$

$$D_z := \begin{bmatrix} -\beta A_{xz} \\ 0_{n\times l} \end{bmatrix}, \qquad F_z := \begin{bmatrix} 0_{n\times l} \\ B_z \end{bmatrix},$$

where I_n and $0_{n\times m}$ denote the n dimensional identity matrix and the $n \times m$ zero matrix, respectively.

Equations (2.13) describe a system of stochastic linear difference equations in two parts. The first part (2.13a) determines the control variables as linear functions of the model's state variables, $\mathbf{x}_t$, exogenous shocks $\mathbf{z}_t$, and the vector of Lagrange multipliers, often referred to as the vector of costate variables. The second part (2.13b) determines the dynamics of the vector of state and costate variables. In Section 2.3 equations (2.13) will serve as framework to study the approximate dynamics of non-linear models. Before we explore this subject and discuss the solution of (2.13), we consider the computation of the policy function via value function iterations in the next section.

2.2 LQ Approximation

This section provides the details of an algorithm proposed by Hansen and Prescott (1994). Their approach rests on a linear quadratic approximation of a given model and they device a simple to program iterative procedure to compute the policy function of the approximate model. In Section 2.2.1, we use the deterministic Ramsey model deterministic from Example 1.2.1 to illustrate the various steps. Section 2.2.2 outlines the general approach and its implementation in the GAUSS program `SolveLQA`.

2.2.1 An Illustrative Example

The Model. The previous subsection shows that the policy function of the LQ problem is independent of the second moments (and, a fortiori, of any higher moments) of the shocks. Therefore, nothing is lost but much is gained in notational simplicity, if we use the deterministic Ramsey model from example 1.2.1 to illustrate the approach of Hansen and Prescott (1994). In this example the farmer solves

$$\begin{aligned} \max_{\{C_t\}_{t=0}^{\infty}} \quad & \sum_{t=0}^{\infty} \beta^t \ln C_t \quad \beta \in (0,1), \\ \text{s.t. } & K_{t+1} + C_t \leq \quad K_t^{\alpha},\ \alpha \in (0,1),\ t = 0, 1, \ldots \\ & K_0 \text{ given.} \end{aligned}$$

C_t denotes consumption at time t, and K_t is the stock of capital. The dynamics of this model is determined by two equations:

$$1 = \beta \frac{C_t}{C_{t+1}} \alpha K_{t+1}^{\alpha-1}, \tag{2.14a}$$

$$K_{t+1} = K_t^{\alpha} - C_t. \tag{2.14b}$$

The first equation is a special case of the Euler equation (1.11) in the case of logarithmic preferences and a Cobb-Douglas production function. The second equation is the economy's resource constraint.

Approximation Step. We want to approximate this model by a linear quadratic problem. Towards this end we must look for a linear law of motion and put all remaining nonlinear relations into the current period return function $\ln C_t$. We achieve this using investment expenditures $I_t = K_t^{\alpha} - C_t$ instead of consumption as control variable. Remember, this model assumes 100 percent depreciation (i.e., $\delta = 1$), so that the linear transition law is:

$$K_{t+1} = I_t. \tag{2.15}$$

Let $g(K_t, I_t) := \ln(K_t^{\alpha} - I_t)$ denote the current period utility function. We approximate this function by a quadratic function in (K, I) at the point of the stationary solution of the model. This solution derives from equations (2.14) and (2.15) for $K_{t+1} = K_t = \bar{K}$ and $C_{t+1} = C_t = \bar{C}$. Thus,

$$\bar{K} = (\alpha\beta)^{(1/(1-\alpha))}, \tag{2.16a}$$

$$\bar{I} = \bar{K}. \tag{2.16b}$$

A second order Taylor series approximation of g yields:

$$\begin{aligned} g(K,I) \simeq\ & g(\bar{K},\bar{I}) + g_K(K-\bar{K}) + g_I(I-\bar{I}) \\ & + (1/2)g_{KK}(K-\bar{K})^2 + (1/2)g_{II}(I-\bar{I})^2 \\ & + (1/2)(g_{KI}+g_{IK})(K-\bar{K})(I-\bar{I}). \end{aligned} \tag{2.17}$$

For latter purposes, we want to write the rhs. of this equation using matrix notation.[5] To take care of the constant and the linear terms we define the vector $(1,K,I)^T$ and the 3×3 matrix $Q=(q_{ij})$ and equate the r.h.s. of (2.17) to the product

$$[1,K,I]Q\begin{bmatrix}1\\K\\I\end{bmatrix}.$$

Comparing terms on both sides of the resulting expression and using the symmetry of the second order mixed partial derivatives ($g_{KI}=g_{IK}$) yields the elements of Q:

$$\begin{aligned} q_{11}&=g-g_K\bar{K}-g_I\bar{I}+(1/2)g_{KK}\bar{K}^2+g_{KI}\bar{K}\bar{I}+(1/2)g_{II}\bar{I}^2,\\ q_{12}&=q_{21}=(1/2)(g_K-g_{KK}\bar{K}-g_{KI}\bar{I}),\\ q_{13}&=q_{31}=(1/2)(g_I-g_{II}\bar{I}-g_{KI}\bar{K}),\\ q_{23}&=q_{32}=(1/2)g_{KI},\\ q_{22}&=(1/2)g_{KK},\\ q_{33}&=(1/2)g_{II}. \end{aligned}$$

In the next step we use Q and the even larger vector $\mathbf{w}=[1,K,I,1,K']$ (where K' denotes next period's stock of capital) to write the r.h.s. of the Bellman equation, $g(K,I)+\beta v(K')$, in matrix notation. This gives:

$$[1,K,I,1,K']\underbrace{\begin{bmatrix}Q & 0_{3\times 2}\\ 0_{2\times 3} & \beta V^0_{2\times 2}\end{bmatrix}}_{R_{5\times 5}}\begin{bmatrix}1\\K\\I\\1\\K'\end{bmatrix},\ V^0:=\begin{bmatrix}v^0_{11} & v^0_{12}\\ v^0_{21} & v^0_{22}\end{bmatrix}. \tag{2.18}$$

We initialize V^0 with a negative definite matrix, e.g., $V^0=-0.001I_2$. Our goal is to eliminate all future variables (here it

[5] To prevent confusion, we depart from our usual notation temporarily and let the superscript T denote the transpose operator. As usual in dynamic programming, the prime $'$ denotes next period's variables.

is just K') using the linear law of motion. Next, we perform the maximization step that allows us to eliminate the controls (here it is just I). After that step we have a new guess for the value function, say V^1. We use this guess as input in a new round of iterations until V^0 and V^1 are sufficiently close together.

Reduction Step. We begin to eliminate K' and the constant from (2.18) so that the resulting quadratic form is reduced to a function of the current state K and the current control I. Note that $K' = I$ can be written as dot product:

$$K' = [0, 0, 1, 0] \begin{bmatrix} 1 \\ K \\ I \\ 1 \end{bmatrix},$$

and observe that

$$\begin{bmatrix} 1 \\ K \\ I \\ 1 \\ K' \end{bmatrix} = \begin{bmatrix} I_4 \\ 0\ 0\ 1\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ K \\ I \\ 1 \end{bmatrix}.$$

Thus, we may express (2.18) equivalently as:

$$[1, K, I, 1, K'] R_{5\times 5} \begin{bmatrix} 1 \\ K \\ I \\ 1 \\ K' \end{bmatrix} = [1, K, I, 1] R_{4\times 4} \begin{bmatrix} 1 \\ K \\ I \\ 1 \end{bmatrix},$$

where

$$R_{4\times 4} = \begin{bmatrix} I_4 \\ 0\ 0\ 1\ 0 \end{bmatrix}^T R_{5\times 5} \begin{bmatrix} I_4 \\ 0\ 0\ 1\ 0 \end{bmatrix}.$$

So what was the trick? In words: use the rightmost variable in $\mathbf{w}^T = [1, K, I, 1, K']$ and write it as linear function of the remaining variables. This gives a row vector with 4 elements. Append

this vector to the identity matrix of dimension 4 to get the transformation matrix S. The matrix of the Bellman equation with K' eliminated is $R_{4\times4} = S^T R_{5\times5} S$.

In the same way we can eliminate the second constant. The constant in terms of the remaining variables $[1, K, I]$ is determined by the dot product:

$$1 = [1, 0, 0] \begin{bmatrix} 1 \\ K \\ I \end{bmatrix}.$$

Thus, the matrix S is now

$$S = \begin{bmatrix} I_3 \\ 1\ 0\ 0 \end{bmatrix},$$

and the rhs. of the Bellman equation in terms of $[1, K, I]$ is

$$g(K, I) + \beta v(I) = [1, K, I] R_{3\times3} \begin{bmatrix} 1 \\ K \\ I \end{bmatrix}, \quad R_{3\times3} = S' R_{4\times4} S.$$

Maximization Step. In this last step we eliminate I from the r.h.s. of the Bellman equation to find

$$[1, K] R_{2\times2} \begin{bmatrix} 1 \\ K \end{bmatrix}.$$

The matrix $R_{2\times2}$ will be our new guess of the value function.

After the last reduction step, the quadratic form is:

$$\begin{aligned} &[1, K, I] \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} 1 \\ K \\ I \end{bmatrix} \\ &= r_{11} + (r_{12} + r_{21})K + (r_{13} + r_{31})I + (r_{23} + r_{32})KI \\ &+ r_{22}K^2 + r_{33}I^2. \end{aligned}$$

Setting the derivative of this expression with respect to I equal to zero and solving for I gives:

$$I = -\underbrace{\frac{r_{13}+r_{31}}{2r_{33}}}_{i_1} - \underbrace{\frac{r_{23}+r_{32}}{2r_{33}}}_{i_2} K = -\frac{r_{13}}{r_{33}} - \frac{r_{23}}{r_{33}}K,$$

where the last equality follows from the symmetry of R. Thus, we can use

$$S = \begin{bmatrix} I_2 \\ -i_1 \quad -i_2 \end{bmatrix}$$

to reduce $R_{3\times 3}$ to the new guess of the value function:

$$V^1 = S^T R_{3\times 3} S.$$

We stop iterations, if the maximal element in $|V^1 - V^0|$ is smaller than $\epsilon(1-\beta)$ for some small positive ϵ (see (8.72) in Section 8.4 on this choice).

2.2.2 The General Method

Notation. Consider the following framework: There is a n vector of state variables $\mathbf{x}$, a m vector of control variables $\mathbf{u}$, a current period return function $g(\mathbf{x}, \mathbf{u})$, and a discount factor $\beta \in (0, 1)$. As you will see in a moment, it will be helpful to put $x_1 = 1$. All non-linear relations of the model are part of the specification of g, and the remaining linear relations define the following law of motion:

$$\mathbf{x}' = A\mathbf{x} + B\mathbf{u}. \tag{2.19}$$

Furthermore, there is a point $[\bar{\mathbf{x}}^T, \bar{\mathbf{u}}^T]^T$. Usually, this will be the stationary solution of the deterministic counterpart of the model under consideration.

Approximation Step. Let $Q \in \mathbb{R}^{l\times l}$, $l = n + m$ denote the matrix of the linear quadratic approximation of the current period return function $g(\cdot)$, and define the $n + m$ column vector $\mathbf{y} = [\mathbf{x}^T, \mathbf{u}^T]^T$. From a Taylor series expansion of g at $\bar{\mathbf{y}}$, we get:

$$\mathbf{y}^T Q \mathbf{y} = g(\bar{\mathbf{y}}) + \sum_{i=1}^{n+m} g_i(y_i - \bar{y}_i) + \frac{1}{2}\sum_{i=1}^{n+m}\sum_{j=1}^{n+m} g_{ij}(y_i - \bar{y}_i)(y_j - \bar{y}_j),$$

where g_i and g_{ij} are first and second partial derivatives of g at $\bar{\mathbf{y}}$, respectively.[6] Comparing terms on both sides of this expression delivers the elements of $Q = (q_{ij})$:

$$\begin{aligned} q_{11} &= g(\bar{\mathbf{y}}) + \textstyle\sum_{i=1}^{n+m} g_i \bar{y}_i + \frac{1}{2} \sum_{i=1}^{n+m} \sum_{j=1}^{n+m} g_{ij} \bar{y}_i \bar{y}_j, \\ q_{1i} = q_{i1} &= \tfrac{1}{2} g_i - \tfrac{1}{2} \textstyle\sum_{j=1}^{n+m} g_{ij} \bar{y}_j, \ i = 2, 3, \ldots, n+m, \\ q_{ij} = q_{ji} &= \tfrac{1}{2} g_{ij}, \ i, j = 2, 3, \ldots, n+m. \end{aligned} \tag{2.20}$$

Except in very rare cases, where g_i and g_{ij} are given by simple analytic expressions, one will use numeric differentiation (see Section 8.3.1). For instance, to use our program `SolveLQA`, the user must supply a procedure `gproc` that returns the value of g at an arbitrary point $[\mathbf{x}^T, \mathbf{u}^T]^T$. Note that you must pass $(1, x_2, \ldots, x_n, u_1, \ldots, u_n)^T$ to that procedure, even if the 1 is not used in `gproc`. This ensures that any procedure that computes the gradient of g returns a vector with l elements and that any procedure that returns the Hesse matrix returns a $l \times l$ matrix. Given this procedure, our GAUSS programs `CDJac` and `CDHesse` compute the gradient vector $\nabla g = [0, g_2, g_3, \ldots, g_{n+m}]$ and the Hesse matrix $H := (h_{ij}) \equiv (g_{ij})$, $i, j = 1, 2, \ldots, n+m$ from which `SolveLQA` builds Q using the above formulas. All of this is done without any further intervention of the user. If higher accuracy in the computation of the Hesse matrix is desired, the user can supply a routine `MyGrad` that returns the gradient vector of g. He must then set the flag `_MyGrad=1` to let the program know that an analytic gradient is available. `SolveLQA` will then use `MyGrad` to compute the Hesse matrix using the forward difference Jacobian programmed in `CDJac`.

Reduction Steps. Let R^s denote the matrix that represents the quadratic form on the rhs of the Bellman equation at reduction step s, where

$$R^1 := \begin{bmatrix} Q_{n+m \times (n+m)} & 0_{(n+m) \times n} \\ 0_{n \times (n+m)} & \beta V^0_{n \times n} \end{bmatrix}.$$

In addition, let $\mathbf{c}_s^T$ denote the $n+1-s$-th row of the matrix

[6] Note, since $x_1 = 1$, we have $g_1 = 0$ and $g_{1i} = g_{i1} = 0$ for $i = 1, 2, \ldots, l$.

$$C_s = \begin{bmatrix} A & B & 0_{n\times(n-s)} \end{bmatrix}.$$

Then, for $s = 1, 2, \ldots, n$ iterate over

$$R^{s+1} = \begin{bmatrix} I_{2n+m-s} \\ \mathbf{c}_s^T \end{bmatrix}^T R^s \begin{bmatrix} I_{2n+m-s} \\ \mathbf{c}_s^T \end{bmatrix}.$$

Maximization Steps. After the last reduction step the matrix R is reduced to a square matrix of size $n + m$. There are m maximization steps to be taken until R is reduced further to a square matrix of size n, which is our new guess of the value function. At step $s = 1, 2, \ldots, m$ the optimal choice of the control variable u_{m+1-s} as linear function of the variables $[x_1, \ldots, x_n, u_1, \ldots, u_{m-s}]$ is given by the row vector

$$\mathbf{d}_s^T = \left[-\frac{r_{1k}}{r_{kk}}, -\frac{r_{2k}}{r_{kk}}, \ldots, -\frac{r_{k-1,k}}{r_{kk}} \right], \; k = n + m - s.$$

Therefore, we iterate over

$$R^{s+1} = \begin{bmatrix} I_{n+m-s} \\ \mathbf{d}_s^T \end{bmatrix}^T R^s \begin{bmatrix} I_{n+m-s} \\ \mathbf{d}_s^T \end{bmatrix}, \; s = 1, 2, \ldots, m.$$

If R is reduced to size n, we have found a new guess of the value function $V^1 = R^{m+1}$, and we compare its elements to those of V^0. If they are close together,

$$\max_{ij} |v_{ij}^0 - v_{ij}^1| < \epsilon(1 - \beta),$$

we stop iterations. Otherwise we replace V^0 with V^1 and restart.

Computation of the Policy Function. It is good idea to store the vectors $\mathbf{d}_s$ in a $m\times(n+m-1)$ matrix D. After convergence, we can use $D = (d_{ij})$ to derive the policy matrix $F \in \mathbb{R}^{m\times n} = (f_{ij})$ that defines the controls as functions of the states. This works as follows: The policy vector $\mathbf{d}_m$ (i.e., the last row of D) holds the coefficients that determine the first control variable u_1 as function of the n state variables:

$$u_1 = \sum_{i=1}^{n} d_{mi} x_i \;\Rightarrow\; f_{1i} = d_{mi}.$$

The second control is given by

$$u_2 = \sum_{i=1}^{n} d_{m-1,i} x_i + d_{m-1,n+1} u_1$$
$$\Rightarrow \; f_{2i} = d_{m-1,i} + d_{m-1,n+1} f_{1i}.$$

Therefore, we may compute the coefficients of F recursively from:

$$f_{ji} = d_{m+1-j,i} + \sum_{k=1}^{j-1} d_{m+1-j,n+k} f_{ki},$$
$$j = 1, \ldots, m, \; i = 1, \ldots, n.$$

As a final check of the solution, we can use

$$|\bar{\mathbf{u}} - F\bar{\mathbf{x}}|.$$

i.e. the discrepancy between the stationary solution of the controls from the original model and those computed using the linear policy function.

2.3 Loglinear Approximation

In this section we return to the system of stochastic difference equations.(2.13) Remember, this system is one way to characterize the solution of the linear quadratic problem. However, we are by no means restricted to this interpretation. More generally, we may consider this system as approximation of an arbitrary non-linear model. In the next subsection we explain this approximation by means of the stochastic growth model from Example 1.3.2 and use a matrix factorization to solve the ensuing system of two linear, stochastic difference equations. Section 2.3.2 presents the solution method for the general case (2.13) and explains the use of our GAUSS program `SolveLA` that implements this method.

2.3.1 An Illustrative Example

The Model. We use the stochastic Ramsey (or growth) model from Example 1.3.2. The problem is to solve

$$\begin{aligned}
&\max_{C_0} \; E_0 \sum_{t=0}^{\infty} \beta^t \ln(C_t), \; \beta \in (0,1), \\
&\text{s.t. } K_{t+1} = Z_t K_t^{\alpha} - C_t, \; \alpha \in (0,1), \; t = 0, 1, \ldots, \\
&\qquad K_0, Z_0 \text{ given.}
\end{aligned}$$

C_t denotes consumption at time t, K_t is the stock of capital and Z_t the productivity shock.

We know from the previous chapter[7] that any solution satisfies the following equations:

$$\frac{1}{C_t} = \lambda_t, \tag{2.21a}$$

$$\lambda_t = \beta E_t \lambda_{t+1} \alpha Z_{t+1} K_{t+1}^{\alpha-1}, \tag{2.21b}$$

$$K_{t+1} = Z_t K_t^{\alpha} - C_t. \tag{2.21c}$$

The Lagrange multiplier λ_t of the budget constraint (2.21c) is the model's costate variable. The control variable is C_t and the state variable with given initial condition is the capital stock K_t. The model's second state variable is the productivity shock Z_t.

Loglinear Approximation. Since the solution to (2.21) is a time invariant distribution for the capital stock, there is no obvious point at which we could approximate the stochastic model. The way out is to use the deterministic counterpart of the model that results when the productivity shock equals its unconditional mean of $Z = 1$ for ever. As we know from the previous chapter, this deterministic model features a unique stationary solution $(\bar{K}, \bar{C})$ determined by the following equations:[8]

$$\bar{K} = (\alpha\beta)^{1/(1-\alpha)}, \tag{2.22a}$$

$$\bar{C} = \bar{K}^{\alpha} - \bar{K}. \tag{2.22b}$$

Differentiating (2.21a) on the left hand side with respect to C_t and on the right hand side with respect to λ_t gives:

[7] Equations (2.21) are a special case of conditions (1.22) for an interior solution.

[8] Put $\bar{Z} = 1$, $K_t = K_{t+1} = \bar{K}$, and $C_t = C_{t+1} = \bar{C}$ and ignore the conditional expectation in (2.21) to get this result.

$$-\frac{1}{\bar{C}^2}dC_t = d\lambda_t,$$

where the partial derivatives are evaluated at the stationary solution. Note that for any variable x, the expression

$$\hat{x} := (x - \bar{x})/\bar{x} \equiv \frac{dx}{\bar{x}} \simeq \ln(x/\bar{x}) \tag{2.23}$$

is the percentage deviation of x from the point $\bar{x}$.[9] Multiplying the above equation by $1/\bar{C} = \bar{\lambda}$ yields condition (2.21a) in terms of deviations of C_t and λ_t from their respective stationary values:

$$-\hat{c}_t = \hat{\lambda}_t. \tag{2.24a}$$

Performing the same exercise with respect to the Euler equation (2.21b) and the budget constraint (2.21c) completes the approximation step. After some algebra, we get:[10]

$$\hat{\lambda}_t = \beta E_t\hat{\lambda}_{t+1} - (1-\alpha)E_t\hat{K}_{t+1} + E_t\hat{Z}_{t+1}, \tag{2.24b}$$

$$\hat{K}_{t+1} = \frac{1}{\beta}\hat{K}_t - \frac{1-\alpha\beta}{\alpha\beta}\hat{C}_t + \frac{1}{\alpha\beta}\hat{Z}_t. \tag{2.24c}$$

Or, using the matrix notation of (2.13):

$$C_u{=}-1, \qquad C_{x\lambda}{=}[0\ 1], \qquad C_z{=}0,$$

$$D_{x\lambda}{=}\begin{bmatrix}1-\alpha & -1\\ 1 & 0\end{bmatrix}, \quad F_{x\lambda}{=}\begin{bmatrix}0 & 1\\ -\frac{1}{\beta} & 0\end{bmatrix}, \quad D_u{=}\begin{bmatrix}0\\ 0\end{bmatrix},$$

[9] That is where the term loglinear as opposed to linear approximation comes from.

[10] The steps involved in getting (2.24b) are:

$$d\lambda_t = \alpha\bar{K}^{\alpha-1}E_t d\lambda_{t+1} + \beta\alpha(\alpha-1)\bar{\lambda}\bar{K}^{\alpha-2}E_t dK_{t+1} + \beta\alpha\bar{\lambda}\bar{K}^{\alpha-1}E_t dZ_{t+1},$$

where we used $\bar{Z} = 1$. Since $\beta\alpha\bar{K}^{\alpha-1} = 1$ (see (2.22a)), we find

$$\frac{d\lambda_t}{\bar{\lambda}} = E_t\frac{d\lambda_{t+1}}{\bar{\lambda}} + (\alpha-1)E_t\frac{dK_{t+1}}{\bar{K}} + E_t\frac{dZ_{t+1}}{\bar{Z}}.$$

Using (2.23) yields (2.24b).

$$F_u = \begin{bmatrix} 0 \\ -\frac{1-\alpha\beta}{\alpha\beta} \end{bmatrix}, \quad D_z = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad F_z = \begin{bmatrix} 0 \\ \frac{1}{\alpha\beta} \end{bmatrix}.$$

In the next step we reduce (2.24) to a system of two difference equations. In general, we will use matrix algebra to do so. Here, it is much simpler. First, we assume that the percentage deviation of the productivity shock from its mean $\bar{Z} = 1$ follows an AR(1) process:

$$\hat{Z}_t = \varrho \hat{Z}_{t-1} + \epsilon_t, \quad \varrho \in (0,1),\ \epsilon_t \sim N(0, \sigma^2). \tag{2.25}$$

This implies

$$E_t \hat{Z}_{t+1} = \varrho \hat{Z}_t.$$

Second, we replace $\hat{C}_t$ in equation (2.24c) by $-\hat{\lambda}_t$, and, thirdly, we use this new equation to eliminate $\hat{K}_{t+1}$ from (2.24b). After a modest amount of algebra involved in collecting terms, we may write (2.24b) and (2.24c) as follows:

$$E_t \begin{bmatrix} \hat{K}_{t+1} \\ \hat{\lambda}_{t+1} \end{bmatrix} = \begin{bmatrix} \frac{1}{\beta} & \frac{1-\alpha\beta}{\alpha\beta} \\ \frac{1-\alpha}{\beta} & \frac{1-\alpha(1-\alpha\beta)}{\alpha\beta} \end{bmatrix} \begin{bmatrix} \hat{K}_t \\ \hat{\lambda}_t \end{bmatrix} + \begin{bmatrix} \frac{1}{\alpha\beta} \\ \frac{1-\alpha}{\alpha\beta} - \varrho \end{bmatrix} \hat{Z}_t,$$

$$\Leftrightarrow \ E_t \begin{bmatrix} \hat{K}_{t+1} \\ \hat{\lambda}_{t+1} \end{bmatrix} = W \begin{bmatrix} \hat{K}_t \\ \hat{\lambda}_t \end{bmatrix} + R \hat{Z}_t. \tag{2.26}$$

Solving (2.26)**.** We apply the technique proposed by King and Watson (2002). They use the Schur factorization of the matrix W to split the system into two parts. The same thing can be achieved using the Jordan canonical form, as advocated by Blanchard and Kahn (1980). Yet, the numerical derivation of the Schur factorization is more reliable.[11]

Remember from Section 8.1 that a real matrix W can be factored into

$$W = TST^{-1},$$

[11] See Golub and Van Loan (1996), p. 317 on this point.

where S is a (possibly complex) upper triangular matrix, whose diagonal elements are the eigenvalues of W. As you will see in a moment, W from equation (2.26) has two real eigenvalues so that $S \in \mathbb{R}^{2\times 2}$ is real, as is the transformation matrix T satisfying $T' = T^{-1}$.

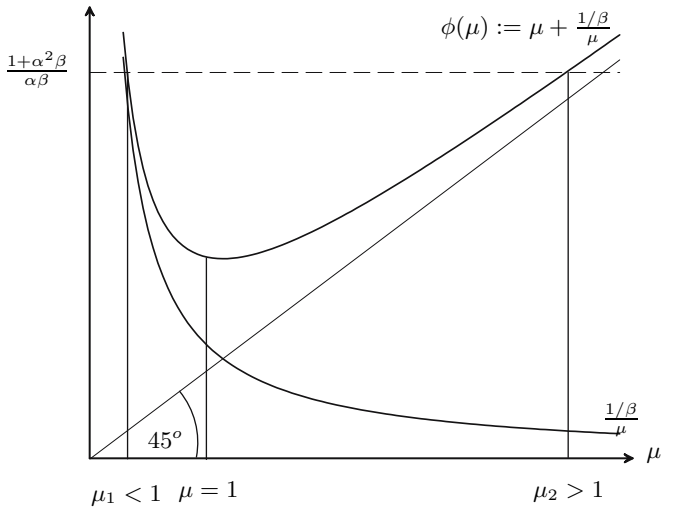

Figure 2.1: Roots of W

It is not difficult to verify that W has two real roots $0 < \mu_1 < 1$ and $\mu_2 > 1$. Consider the following relations (see (8.19)):

1) the determinant of W equals the product of the roots, i.e.,

$$\mu_1\mu_2 = |W| = \frac{1-\alpha(1-\alpha\beta)}{\alpha\beta^2} - \frac{(1-\alpha)(1-\alpha\beta)}{\alpha\beta^2} = (1/\beta).$$

2) the trace of W equals the sum of the roots, i.e.,

$$\mu_1 + \mu_2 = (1/\beta) + \frac{1-\alpha(1-\alpha\beta)}{\alpha\beta} = \frac{1+\alpha^2\beta}{\alpha\beta}.$$

Therefore, the roots of W solve

$$\phi(\mu) := \mu + \frac{1/\beta}{\mu} = \frac{1+\alpha^2\beta}{\alpha\beta}.$$

The solutions are the points of intersection between the horizontal line through $(1+\alpha^2\beta)/(\alpha\beta)$ and the graph of $\phi(\mu)$ (see Figure 2.1). The graph of $\phi(\mu)$ is u-shaped. Its asymptotes are the 45-degree line and the hyperbola $1/(\beta\mu)$. Since

$$\phi(1) = 1 + \frac{1}{\beta} < \frac{1+\alpha^2\beta}{\alpha\beta},$$

the first root μ_1 is located to the left of 1 and the second root must lie to the right of 1.

We use this result to write

$$T^{-1}WT = S = \begin{bmatrix} \mu_1 & s_{12} \\ 0 & \mu_2 \end{bmatrix} \tag{2.27}$$

for a suitable transformation matrix T. In practice, we will use appropriate software to compute S and T. Hence, suppose we know T. In the new variables

$$\begin{bmatrix} \tilde{K}_t \\ \tilde{\lambda}_t \end{bmatrix} := T^{-1} \begin{bmatrix} \hat{K}_t \\ \hat{\lambda}_t \end{bmatrix} \Leftrightarrow T \begin{bmatrix} \tilde{K}_t \\ \tilde{\lambda}_t \end{bmatrix} := \begin{bmatrix} \hat{K}_t \\ \hat{\lambda}_t \end{bmatrix} \tag{2.28}$$

the system of difference equations may be written as:[12]

$$E_t \begin{bmatrix} \tilde{K}_{t+1} \\ \tilde{\lambda}_{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} \mu_1 & s_{12} \\ 0 & \mu_2 \end{bmatrix}}_{S} \begin{bmatrix} \tilde{K}_t \\ \tilde{\lambda}_t \end{bmatrix} + \underbrace{\begin{bmatrix} q_1 \\ q_2 \end{bmatrix}}_{Q=T^{-1}R} \hat{Z}_t. \tag{2.29}$$

Consider the second equation of this system, which is a relation in the new variable $\tilde{\lambda}_t$ and the exogenous shock:

$$E_t\tilde{\lambda}_{t+1} = \mu_2\tilde{\lambda}_t + q_2\hat{Z}_t. \tag{2.30}$$

We can solve this equation for $\tilde{\lambda}_t$ via repeated substitution: from (2.30) we get

$$\tilde{\lambda}_t = \frac{1}{\mu_2}E_t\tilde{\lambda}_{t+1} - \frac{q_2}{\mu_2}\hat{Z}_t. \tag{2.31}$$

[12] Pre-multiply (2.26) by T^{-1} and use the definitions in (2.28) to arrive at this representation.

Shifting the time index one period into the future and taking expectations conditional on information as of period t yields:

$$E_t\tilde{\lambda}_{t+1} = \frac{1}{\mu_2}E_t\tilde{\lambda}_{t+2} - \frac{q_2}{\mu_2}E_t\hat{Z}_{t+1} = \frac{1}{\mu_2}E_t\tilde{\lambda}_{t+2} - \frac{q_2}{\mu_2}\varrho\hat{Z}_t, \quad (2.32)$$

due to (2.25). Substitution into (2.31) gives:

$$\tilde{\lambda}_t = \frac{1}{\mu_2^2}E_t\tilde{\lambda}_{t+2} - \left[\frac{q_2}{\mu_2} + \frac{q_2}{\mu_2}\frac{\varrho}{\mu_2}\right]\hat{Z}_t.$$

We can use (2.32) to get an expression for $\tilde{\lambda}_{t+3}$ and so on up to period $t+\tau$:

$$\tilde{\lambda}_t = \left[\frac{1}{\mu_2}\right]^\tau E_t\tilde{\lambda}_{t+\tau} - \frac{q_2}{\mu_2}\sum_{i=0}^{\tau-1}\left[\frac{\varrho}{\mu_2}\right]^i\hat{Z}_t. \quad (2.33)$$

Suppose that the sequence

$$\left\{\frac{1}{\mu_2^\tau}E_t\tilde{\lambda}_{t+\tau}\right\}_{\tau=0}^{\infty}$$

converges towards zero for $\tau \to \infty$. This is not very restrictive: since $1/\mu_2 < 1$, it is sufficient to assume that $E_t\lambda_{t+\tau}$ is bounded. Intuitively, this assumption rules out speculative Bubbles along explosive paths and renders the solution unique. In addition, it guarantees that the transversality condition (1.24) is met. In this case we can compute the limit of (2.33) for $\tau \to \infty$:

$$\tilde{\lambda}_t = -\frac{q_2/\mu_2}{1-(\varrho/\mu_2)}\hat{Z}_t. \quad (2.34)$$

We substitute this solution into the second equation of (2.28),[13]

$$\tilde{\lambda}_t = t^{21}\hat{K}_t + t^{22}\hat{\lambda}_t,$$

to get the solution for $\hat{\lambda}_t$ in terms of $\hat{K}_t$ and $\hat{Z}_t$:

[13] We denote the elements of T^{-1} by (t^{ij}).

$$\hat{\lambda}_t = -\underbrace{\frac{t^{21}}{t^{22}}}_{=:L_{\lambda x}} \hat{K}_t - \underbrace{\frac{q_2/\mu_2}{t^{22}(1-(\varrho/\mu_2))}}_{=:L_{\lambda z}} \hat{Z}_t. \tag{2.35}$$

Thus, the policy function that gives $\hat{C}_t$ in terms of the (deviations) of the state variables is

$$\hat{C}_t = L_{ux}\hat{K}_t + L_{uz}\hat{Z}_t.$$

From the first equation of (2.26),

$$\hat{K}_{t+1} = \frac{1}{\beta}\hat{K}_t + \frac{1-\alpha\beta}{\alpha\beta}\hat{\lambda}_t + \frac{1}{\alpha\beta}\hat{Z}_t,$$

we can derive the solution for $\hat{K}_t$:

$$\begin{aligned}\hat{K}_{t+1} = &\underbrace{\left(\frac{1}{\beta} - \frac{1-\alpha\beta}{\alpha\beta}\frac{t^{21}}{t^{22}}\right)}_{=:L_{xx}} \hat{K}_t \\ &+ \underbrace{\left(\frac{1}{\alpha\beta} - \frac{1-\alpha\beta}{\alpha\beta}\frac{q_2/\mu_2}{t^{22}(1-(\varrho/\mu_2))}\right)}_{=:L_{xz}} \hat{Z}_t.\end{aligned}$$

Thus, given a sequence of shocks $\{\hat{Z}_t\}_{t=0}^T$ and an initial $\hat{K}_0$ we may compute the entire time path of consumption and the stock of capital by iteration over

$$\begin{aligned}\hat{C}_t &= L_{ux}\hat{K}_t + L_{uz}\hat{Z}_t, \\ \hat{K}_{t+1} &= L_{xx}\hat{K}_t + L_{xz}\hat{Z}_t.\end{aligned}$$

In the next subsection we basically use the same steps to derive the policy function in terms of the matrices L_{ij} for the general system.(2.13) If you dislike linear algebra , you may skip this section and note that the GAUSS program `SolveLA` performs the above computations for the general case. The program requires the matrices from (2.13) as input and outputs the matrices L_{ij}.

2.3.2 The General Method

In this subsection we consider the solution of a system of linear stochastic difference equations given in the form of (2.13), which derive from the LQ problem. There are related ways to state and solve such systems. The list of references includes the classical paper by Blanchard and Kahn (1980), Chapter 3 of the book by Farmer (1993), the papers of King and Watson (1998), (2002), and the approach proposed by Uhlig (1999). Our statement of the problem is that proposed by Burnside (1999), but we solve it along the lines of King and Watson (2002).

The Problem. Consider the system of stochastic difference equations (2.36):

$$C_u \mathbf{u}_t = C_{x\lambda} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} + C_z \mathbf{z}_t, \tag{2.36a}$$

$$D_{x\lambda} E_t \begin{bmatrix} \mathbf{x}_{t+1} \\ \boldsymbol{\lambda}_{t+1} \end{bmatrix} + F_{x\lambda} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} = D_u E_t \mathbf{u}_{t+1} + F_u \mathbf{u}_t + D_z E_t \mathbf{z}_{t+1} + F_z \mathbf{z}_t. \tag{2.36b}$$

To ease notation we use $n(x)$ do denote the dimension (i.e., the number of elements) of the vector $\mathbf{x}$. We think of the $n(u)$ vector $\mathbf{u}_t$ as the collection of variables that are determined within period t as linear functions of the model's state variables. We distinguish between three kinds of state variables: those with given initial conditions build the $n(x)$ vector $\mathbf{x}_t$; the $n(\lambda)$ vector $\boldsymbol{\lambda}_t$ collects those variables, whose initial values may be chosen freely. In the LQ problem these are the costate variables. In the stochastic growth model it is just the Lagrange multiplier of the budget constraint. Purely exogenous stochastic shocks are stacked in the $n(z)$ vector $\mathbf{z}_t$. We assume that $\mathbf{z}_t$ is governed by a stable vector autoregressive process of first-order with normally distributed innovations $\boldsymbol{\epsilon}_t$:

$$\mathbf{z}_t = \Pi \mathbf{z}_{t-1} + \boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \sim N(\mathbf{0}, \Sigma). \tag{2.37}$$

Stability requires that the eigenvalues of the matrix Π lie within the unit circle.

System Reduction. We assume that the first equation can be solved for the vector $\mathbf{u}_t$:

$$\mathbf{u}_t = C_u^{-1} C_{x\lambda} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} + C_u^{-1} C_z \mathbf{z}_t. \tag{2.38}$$

Shifting the time index one period into the future and taking expectations conditional on information as of period t yields:

$$E_t \mathbf{u}_{t+1} = C_u^{-1} C_{x\lambda} E_t \begin{bmatrix} \mathbf{x}_{t+1} \\ \boldsymbol{\lambda}_{t+1} \end{bmatrix} + C_u^{-1} C_z E_t \mathbf{z}_{t+1}. \tag{2.39}$$

The solutions (2.38) and (2.39) allow us to eliminate $\mathbf{u}_t$ and $E_t \mathbf{u}_{t+1}$ from (2.36b):

$$\begin{aligned} \left(D_{x\lambda} - D_u C_u^{-1} C_{x\lambda}\right) E_t \begin{bmatrix} \mathbf{x}_{t+1} \\ \boldsymbol{\lambda}_{t+1} \end{bmatrix} = &- \left(F_{x\lambda} - F_u C_u^{-1} C_{x\lambda}\right) \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} \\ &+ \left(D_z + D_u C_u^{-1} C_z\right) E_t \mathbf{z}_{t+1} \\ &+ \left(F_z + F_u C_u^{-1} C_z\right) \mathbf{z}_t. \end{aligned}$$

Assume that this system can be solved for $E_t(\mathbf{x}_{t+1}, \boldsymbol{\lambda}_{t+1})'$. In other words, the matrix $D_{x\lambda} - D_u C_u^{-1} C_{x\lambda}$ must be invertible. Using $E_t \mathbf{z}_{t+1} = \Pi \mathbf{z}_t$, which is implied by (2.37), we get the following reduced dynamic system:

$$\begin{aligned} E_t \begin{bmatrix} \mathbf{x}_{t+1} \\ \boldsymbol{\lambda}_{t+1} \end{bmatrix} &= W \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} + R \mathbf{z}_t, \\ W &= -\left(D_{x\lambda} - D_u C_u^{-1} C_{x\lambda}\right)^{-1} \left(F_{x\lambda} - F_u C_u^{-1} C_{x\lambda}\right), \\ R &= \left(D_{x\lambda} - D_u C_u^{-1} C_{x\lambda}\right)^{-1} \\ &\quad \times \left[\left(D_z + D_u C_u^{-1} C_z\right) \Pi + \left(F_z + F_u C_u^{-1} C_z\right)\right]. \end{aligned} \tag{2.40}$$

Change of Variables. Consider the Schur factorization of the matrix W:

$$S = T^{-1} W T,$$

which gives raise to the following partitioned matrices :

$$\begin{aligned} S &= \begin{bmatrix} S_{xx} & S_{x\lambda} \\ 0 & S_{\lambda\lambda} \end{bmatrix} \\ &= \begin{bmatrix} T^{xx} & T^{x\lambda} \\ T^{\lambda x} & T^{\lambda\lambda} \end{bmatrix} \begin{bmatrix} W_{xx} & W_{x\lambda} \\ W_{\lambda x} & W_{\lambda\lambda} \end{bmatrix} \begin{bmatrix} T_{xx} & T_{x\lambda} \\ T_{\lambda x} & T_{\lambda\lambda} \end{bmatrix}. \end{aligned} \tag{2.41}$$

We assume that the eigenvalues of W appear in ascending order on the main diagonal of S (see 8.1). To find a unique solution, $n(x)$ eigenvalues must lie within the unit circle and $n(\lambda)$ eigenvalues must have modulus greater than one. In the new variables

$$\begin{bmatrix} \tilde{\mathbf{x}}_t \\ \tilde{\boldsymbol{\lambda}}_t \end{bmatrix} := \begin{bmatrix} T^{xx} & T^{x\lambda} \\ T^{\lambda x} & T^{\lambda\lambda} \end{bmatrix} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} \tag{2.42}$$

the dynamic system (2.40) can be rewritten as

$$\begin{aligned} E_t \begin{bmatrix} \tilde{\mathbf{x}}_{t+1} \\ \tilde{\boldsymbol{\lambda}}_{t+1} \end{bmatrix} &= \begin{bmatrix} S_{xx} & S_{x\lambda} \\ 0 & S_{\lambda\lambda} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{x}}_t \\ \tilde{\boldsymbol{\lambda}}_t \end{bmatrix} + \begin{bmatrix} Q_x \\ Q_\lambda \end{bmatrix} \mathbf{z}_t, \\ Q &= [Q_x, Q_\lambda]' = T^{-1}R. \end{aligned} \tag{2.43}$$

Solving the Second Part of (2.43)**.** Consider the second line of (2.43), which is a linear system in $\tilde{\boldsymbol{\lambda}}$ alone:

$$E_t \tilde{\boldsymbol{\lambda}}_{t+1} = S_{\lambda\lambda} \tilde{\boldsymbol{\lambda}}_t + Q_\lambda \mathbf{z}_t. \tag{2.44}$$

Its solution is given by:

$$\tilde{\boldsymbol{\lambda}}_t = \Phi \mathbf{z}_t. \tag{2.45}$$

The rows of the matrix Φ are computed in the following steps.

The matrix $S_{\lambda\lambda}$ is upper triangular with all of its eigenvalues μ_i on the main diagonal being larger than one in absolute value:

$$S_{\lambda\lambda} = \begin{bmatrix} \mu_1 & s_{12} & \dots & s_{1n(\lambda)} \\ 0 & \mu_2 & \dots & s_{2n(\lambda)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mu_{n(\lambda)} \end{bmatrix}.$$

Therefore, the last line of (2.44) is a stochastic difference equation in the single variable $\tilde{\lambda}_{n(\lambda)}$, just like equation (2.30):

$$E_t\tilde{\lambda}_{n(\lambda)\,t+1} = \mu_{n(\lambda)}\tilde{\lambda}_{n(\lambda)\,t} + \mathbf{q}'_{n(\lambda)}\mathbf{z}_t, \tag{2.46}$$

where $\mathbf{q}'_{n(\lambda)}$ denotes the last row of the matrix Q_λ. Note, that $\tilde{\lambda}_{n(\lambda)\,t}$ – as every other component of $\tilde{\boldsymbol{\lambda}}_t$ – may be a complex variable. Yet, since the modulus (i.e., the absolute value) of the complex number $\mu_{n(\lambda)}$ is larger than one, the sequence

$$\left\{\frac{1}{\mu^{\tau}_{n(\lambda)}}E_t\tilde{\lambda}_{n(\lambda)\,t+\tau}\right\}_{\tau=0}^{\infty}$$

will converge to zero if the sequence

$$\left\{E_t\tilde{\lambda}_{n(\lambda)\,t+\tau}\right\}_{\tau=0}^{\infty}$$

is bounded (see 9.1). Given this assumption, we know from equation (2.34) that the solution to (2.46) is a linear function of $\mathbf{z}_t$:

$$\tilde{\lambda}_{n(\lambda)\,t} = \underbrace{\left(\phi_{n(\lambda)\,1}, \phi_{n(\lambda)\,2}, \ldots, \phi_{n(\lambda),n(z)}\right)'}_{\boldsymbol{\phi}'_{n(\lambda)}}\mathbf{z}_t.$$

To determine the yet unknown coefficients of this function, i.e., the elements of the row vector $\boldsymbol{\phi}'_{n(\lambda)}$, we proceed as follows: we substitute this solution into equation (2.46). This yields:

$$\begin{aligned}
\boldsymbol{\phi}'_{n(\lambda)}E_t\mathbf{z}_{t+1} &= \mu_{n(\lambda)}\boldsymbol{\phi}'_{n(\lambda)}\mathbf{z}_t + \mathbf{q}'_{n(\lambda)}\mathbf{z}_t,\\
\left(\boldsymbol{\phi}'_{n(\lambda)}\Pi - \boldsymbol{\phi}'_{n(\lambda)}\mu_{n(\lambda)}\right)\mathbf{z}_t &= \mathbf{q}'_{n(\lambda)}\mathbf{z}_t,\\
\boldsymbol{\phi}'_{n(\lambda)}\left(\Pi - \mu_{n(\lambda)}I_{n(z)}\right)\mathbf{z}_t &= \mathbf{q}'_{n(\lambda)}\mathbf{z}_t,
\end{aligned}$$

where the second line follows from (2.37). Equating the coefficients on both sides of the last line of the preceding expression gives the solution for the unknown vector $\boldsymbol{\phi}_{n(\lambda)}$:

$$\boldsymbol{\phi}'_{n(\lambda)} = \mathbf{q}'_{n(\lambda)}\left(\Pi - \mu_{n(\lambda)}I_{n(z)}\right)^{-1}. \tag{2.47}$$

Since the eigenvalues of Π are inside the unit circle, this solution exists.

Now, consider the next to last line of (2.44):

$$E_t\tilde{\lambda}_{n(\lambda)-1\,t+1} = \mu_{n(\lambda)-1}\tilde{\lambda}_{n(\lambda)-1\,t} + s_{n(\lambda)-1,n(\lambda)}\tilde{\lambda}_{n(\lambda)\,t} + \mathbf{q}'_{n(\lambda)-1}\mathbf{z}_t,$$
$$E_t\tilde{\lambda}_{n(\lambda)-1\,t+1} = \mu_{n(\lambda)-1}\tilde{\lambda}_{n(\lambda)-1\,t} + s_{n(\lambda)-1,n(\lambda)}\boldsymbol{\phi}'_{n(\lambda)}\mathbf{z}_t + \mathbf{q}'_{n(\lambda)-1}\mathbf{z}_t.$$

The solution to this equation is given by the row vector $\boldsymbol{\phi}'_{n(\lambda)-1}$. Repeating the steps from above, we find:

$$\boldsymbol{\phi}'_{n(\lambda)-1} = \left(\mathbf{q}'_{n(\lambda)-1} + s_{n(\lambda)-1\,n(\lambda)}\boldsymbol{\phi}'_{n(\lambda)}\right)\left(\Pi - \mu_{n(\lambda)-1}I_{n(z)}\right)^{-1}. \tag{2.48}$$

Proceeding from line $n(\lambda)-1$ to line $n(\lambda)-2$ and so forth until the first line of (2.44) we are able to compute all of the rows $\boldsymbol{\phi}'_i$ of the matrix Φ. The respective formula is:

$$\begin{aligned}\boldsymbol{\phi}'_i &= \left[\mathbf{q}'_i + \sum_{j=i+1}^{n(\lambda)} s_{i,j}\boldsymbol{\phi}'_j\right]\left(\Pi - \mu_i I_{n(z)}\right)^{-1}, \\ i &= n(\lambda), n(\lambda)-1, \ldots, 1.\end{aligned} \tag{2.49}$$

Policy Function for $\boldsymbol{\lambda}_t$. Given the solution for $\tilde{\boldsymbol{\lambda}}_t$ we can use (2.42) to find the solution for $\boldsymbol{\lambda}_t$ in terms of $\mathbf{x}_t$ and $\mathbf{z}_t$. The second part of (2.42) is:

$$\tilde{\boldsymbol{\lambda}}_t = T^{\lambda x}\mathbf{x}_t + T^{\lambda\lambda}\boldsymbol{\lambda}_t.$$

Together with (2.45) this gives:

$$\boldsymbol{\lambda}_t = \underbrace{-\left(T^{\lambda\lambda}\right)^{-1}T^{\lambda x}}_{L_{\lambda x}}\mathbf{x}_t + \underbrace{\left(T^{\lambda\lambda}\right)^{-1}\Phi}_{L_{\lambda z}}\mathbf{z}_t. \tag{2.50}$$

Policy Function for $\mathbf{x}_{t+1}$. In obvious notation the first part of (2.40) may be written as:

$$\mathbf{x}_{t+1} = W_{xx}\mathbf{x}_t + W_{x\lambda}\boldsymbol{\lambda}_t + R_x\mathbf{z}_t.$$

Substitution for $\boldsymbol{\lambda}_t$ from (2.50) gives:

$$\begin{aligned}\mathbf{x}_{t+1} = &\underbrace{\left(W_{xx} - W_{x\lambda}\left(T^{\lambda\lambda}\right)^{-1}T^{\lambda x}\right)}_{L_{xx}}\mathbf{x}_t \\ &+ \underbrace{\left(W_{xx}\left(T^{\lambda\lambda}\right)^{-1}\Phi + R_x\right)}_{L_{xz}}\mathbf{z}_t.\end{aligned} \tag{2.51}$$

The expression for L_{xx} may be considerably simplified. In terms of partitioned matrices the expression $W = TST^{-1}$ may be written as:

$$\begin{bmatrix} W_{xx} & W_{x\lambda} \\ W_{\lambda x} & W_{\lambda\lambda} \end{bmatrix} = \begin{bmatrix} T_{xx} & T_{x\lambda} \\ T_{\lambda x} & T_{\lambda\lambda} \end{bmatrix} \begin{bmatrix} S_{xx} & S_{x\lambda} \\ 0 & S_{\lambda\lambda} \end{bmatrix} \begin{bmatrix} T^{xx} & T^{x\lambda} \\ T^{\lambda x} & T^{\lambda\lambda} \end{bmatrix},$$

which implies:

$$\begin{aligned} W_{xx} &= T_{xx}S_{xx}T^{xx} + T_{xx}S_{x\lambda}T^{\lambda x} + T_{x\lambda}S_{\lambda\lambda}T^{\lambda x}, \\ W_{x\lambda} &= T_{xx}S_{xx}T^{x\lambda} + T_{xx}S_{x\lambda}T^{\lambda\lambda} + T_{x\lambda}S_{\lambda\lambda}T^{\lambda\lambda}. \end{aligned}$$

Substituting the right hand sides of these equations into the expression for L_{xx} from (2.51) gives:

$$L_{xx} = T_{xx}S_{xx}\left(T^{xx} - T^{x\lambda}\left(T^{\lambda\lambda}\right)^{-1}T^{\lambda x}\right).$$

Since

$$\begin{bmatrix} T_{xx} & T_{x\lambda} \\ T_{\lambda x} & T_{\lambda\lambda} \end{bmatrix} = \begin{bmatrix} T^{xx} & T^{x\lambda} \\ T^{\lambda x} & T^{\lambda\lambda} \end{bmatrix}^{-1}$$

the formula for the inverse of a partitioned matrix (see, e.g., Murata (1977), p.9) implies:

$$(T_{xx})^{-1} = T^{xx} - T^{x\lambda}\left(T^{\lambda\lambda}\right)^{-1}T^{\lambda x}.$$

Putting all pieces together, we find:

$$L_{xx} = T_{xx}S_{xx}T_{xx}^{-1}.$$

Policy Function for $\mathbf{u}_t$. Using equation (2.38) the solutions for $\mathbf{x}_t$ and $\boldsymbol{\lambda}_t$ imply the following policy function for the vector $\mathbf{u}_t$:

$$\mathbf{u}_t = \underbrace{C_u^{-1}C_{x\lambda}\begin{bmatrix} I_{n(x)} \\ L_{\lambda x} \end{bmatrix}}_{L_{ux}}\mathbf{x}_t + \underbrace{\left(C_u^{-1}C_{x\lambda}\begin{bmatrix} 0_{n(x)\times n(z)} \\ L_{\lambda z} \end{bmatrix} + C_u^{-1}C_z\right)}_{L_{uz}}\mathbf{z}_t. \tag{2.52}$$

Implementation. Our GAUSS program `SolveLA` performs the computation of the policy matrices according to the formulas given above. It uses the GAUSS intrinsic command

```
{S,T}=Schtoc(Schur(W))
```

to get the matrices S and T. However, the eigenvalues on the main diagonal of S are not ordered. We use the Givens rotation described in Section 8.1 to sort the eigenvalues in ascending order. The program's input are the matrices from (2.36), the matrix Π from (2.37), and the number of elements of $\mathbf{x}_t$. The program checks whether $n(x)$ of the eigenvalues of W are inside the unit circle. If not, it stops with an error message. Otherwise it returns the matrices L_{xx}, L_{xz}, $L_{\lambda x}$, $L_{\lambda z}$, L_{ux}, and L_{uz}. A Fortran version of this program is also available. The program uses the LAPACK subroutine `ZGEES` to get the Schur decomposition. The advantage of this subroutine is an option to sort the eigenvalues on the main diagonal of S.

2.4 Applications

2.4.1 The Benchmark Model

In Chapter 1, Example 1.4.1, we present the benchmark model. In this model, a representative agent chooses feed-back rules for consumption and labor supply that maximize his expected live time utility over an infinite time horizon. This section shows how to solve this model using the linear-quadratic and the loglinear approximation methods.

The Model. Briefly, the problem is to find policy functions for consumption and working hours that solve the following problem:

$$\max_{C_0,N_0} \quad E_0\left[\sum_{t=0}^{\infty}\beta^t\frac{C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta}\right], \qquad (2.53)$$

$$\beta\in(0,1),\ \theta\geq 0, \eta>\theta/(1+\theta),$$

s.t.

$$\left.\begin{array}{rcl} K_{t+1} + C_t &\leq& Z_t(A_tN_t)^{1-\alpha}K_t^{\alpha} + (1-\delta)K_t,\ \alpha \in (0,1), \\ A_{t+1} &=& aA_t,\ a \geq 1, \\ Z_{t+1} &=& Z_t^{\varrho}e^{\epsilon_t},\ \varrho \in (0,1),\ \epsilon_t \sim N(0,\sigma^2), \\ 0 &\leq& C_t, \\ 0 &\leq& K_{t+1}, \end{array}\right\} \forall t,$$

$K_0,\ Z_0$ given.

As in Chapter 1, C_t is consumption in period t, N_t are working hours, K_t is the stock of capital available for production in period t, A_t is the level of labor augmenting technical progress, and Z_t is a shock to total factor productivity, whose log, $z_t := \ln Z_t$, follows an AR(1) process with normally distributed innovations ϵ_t and autocorrelation parameter ϱ. $a \geq 1$ is the growth factor of technical progress, $1-\alpha$ is labor's share in output, the parameters β, η, and θ describe the agent's preferences, and δ is the rate of capital depreciation.

This model depicts a growing economy. To apply our methods, we must choose variables that are stationary. First, consider the Lagrangean:

$$\begin{aligned} \mathscr{L} &:= E_0\Bigg\{\sum_{t=0}^{\infty}\beta^t\Bigg[\frac{C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta} \\ &\quad + \Lambda_t\big(Z_t(A_tN_t)^{1-\alpha}K_t^{\alpha} + (1-\delta)K_t - C_t - K_{t+1}\big)\Bigg]\Bigg\} \\ &= E_0\Bigg\{\sum_{t=0}^{\infty}\beta^t A_t^{1-\eta}\Bigg[\frac{(C_t/A_t)^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta} \\ &\quad + \Lambda_t A_t^{\eta-1}\big(Z_t(A_tN_t)^{1-\alpha}K_t^{\alpha} + (1-\delta)K_t - C_t - K_{t+1}\big)\Bigg]\Bigg\} \\ &= E_0\Bigg\{\sum_{t=0}^{\infty}\beta^t A_t^{1-\eta}\Bigg[\frac{(C_t/A_t)^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta} \\ &\quad + \Lambda_t A_t^{\eta}\big(Z_tN_t^{1-\alpha}(K_t/A_t)^{\alpha} + (1-\delta)(K_t/A_t) - (C_t/A_t) \\ &\quad - (K_{t+1}/A_t)\big)\Bigg]\Bigg\}. \end{aligned} \tag{2.54}$$

Second, note that

$$\beta^t A_t^{1-\eta} = (\beta a^{1-\eta})^t,$$

since (by repeated substitution)

$$A_t = a^t A_0$$

and since we are free to choose units so that $A_0 \equiv 1$. As a consequence, the first-order conditions of problem (2.53) are the same as those of

$$\begin{aligned} &\max_{c_0, N_0} \quad \sum_{t=0}^{\infty} \tilde{\beta}^t \Big[\frac{c_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta} \\ &+ \lambda_t \left(Z_t N_t^{1-\alpha} k_t^{\alpha} + (1-\delta)k_t - c_t - a k_{t+1} \right) \Big] \Big\}, \\ &\tilde{\beta} := \beta a^{1-\eta}, c_t := C_t/A_t, k_t := K_t/A_t, \lambda_t := \Lambda_t A_t^{\eta}. \end{aligned} \tag{2.55}$$

First-Order Conditions and Stationary Solution. The first-order conditions of problem (2.55) are given in (1.48). For the present purpose it will be convenient to state these conditions a bit differently. By differentiating (2.54) with respect to c_t, N_t, k_{t+1}, and λ_t we find:

$$\lambda_t = c_t^{-\eta}(1-N_t)^{\theta(1-\eta)}, \tag{2.56a}$$

$$0 = \theta c_t^{1-\eta}(1-N_t)^{\theta(1-\eta)-1} - (1-\alpha)\lambda_t Z_t N_t^{-\alpha} k_t^{\alpha}, \tag{2.56b}$$

$$\lambda_t = \beta a^{-\eta} E_t \lambda_{t+1} \big(1 - \delta + \alpha Z_{t+1} N_{t+1}^{1-\alpha} k_{t+1}^{\alpha-1}\big), \tag{2.56c}$$

$$a k_{t+1} = Z_t N_t^{1-\alpha} k_t^{\alpha} + (1-\delta)k_t - c_t. \tag{2.56d}$$

From this set of equations we get the stationary solution of the deterministic counterpart of the model by dropping the time indices and the expectations operator and by setting Z equal to its unconditional mean of unity. Although we derive this solution in Section 1.4.4, we repeat the resulting equations for your convenience. Equation (2.56c) delivers the solution for the output-capital ratio $y/k = (N^{1-\alpha}k^{\alpha})/k$:

$$\frac{y}{k} = \frac{a^{\eta} - \beta(1-\delta)}{\alpha\beta}. \tag{2.57a}$$

We can use this solution in the stationary version of the budget constraint (2.56d) to solve for the consumption-capital ratio c/k:

$$\frac{c}{k} = \frac{y}{k} + (1 - a - \delta). \tag{2.57b}$$

Combing (2.56a) and (2.56b), we find

$$\frac{N}{1-N} = \frac{1-\alpha}{\theta}\frac{y}{c}. \tag{2.57c}$$

Since the share of consumption in output c/y equals $(c/k)/(y/k)$, we can solve this equation for the stationary level of working hours N. Finally, since

$$\frac{y}{k} = N^{1-\alpha}k^{\alpha-1} \Rightarrow k = N(y/k)^{1/(\alpha-1)}, \tag{2.57d}$$

we get the solutions for k, $c = k(c/k)$ $y = k(y/k)$ and investment expenditures $i = (a + \delta - 1)k$. Given these solutions, we can move on to implement the linear quadratic approximation of the benchmark model.

Linear Quadratic Approximation. At first sight, it seems that there are no linear equations in the benchmark model at all. Yet, this is just a matter of definitions. First, let us define

$$z_t := \ln Z_t$$

so that the law of motion of the productivity shock becomes

$$z_t = \varrho z_{t-1} + \epsilon_t. \tag{2.58a}$$

Second, we use investment expenditures i_t instead of consumption c_t as a control variable. Put

$$i_t = Z_t N_t^{1-\alpha} k_t^{\alpha} - c_t \tag{2.58b}$$

so that equation (2.56d) can be written as:

$$k_{t+1} = \frac{1}{a} i_t + \frac{1-\delta}{a} k_t, \tag{2.58c}$$

which is linear in k_{t+1}, k_t, and i_t. Let $\mathbf{x}_t := [1, k_t, z_t]'$ denote the vector of states and $\mathbf{u}_t := [i_t, N_t]'$ the vector of controls. Then, equations (2.58a) and (2.58b) may be written in matrix notation as follows:

$$\mathbf{x}_{t+1} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & (1-\delta)/a & 0 \\ 0 & 0 & \varrho \end{bmatrix}}_{A} \mathbf{x}_t + \underbrace{\begin{bmatrix} 0 & 0 \\ 1/a & 0 \\ 0 & 0 \end{bmatrix}}_{B} \mathbf{u}_t + \begin{bmatrix} 0 \\ 0 \\ \epsilon_t \end{bmatrix}. \tag{2.59}$$

The remaining non-linearities are handled by the algorithm. The current period return function in the scaled variables is given by:

$$g(\mathbf{x}, \mathbf{u}) := \frac{1}{1-\eta} \left(e^{z_t} N_t^{1-\alpha} k_t^{\alpha} - i_t \right)^{1-\eta} (1 - N_t)^{\theta(1-\eta)}.$$

You must write a subroutine, say `GProc`, that takes the vector `ybar`$=[1, k, z, i, N]'$ as input and returns the value of g at this point. As shown in the previous paragraph, the appropriate discount factor is $\tilde{\beta} = \beta a^{1-\eta}$. The command

```
{fmat,prec}=SolveLQA(&GProc,beta,ybar,A,B)
```

returns the 2×3 matrix `fmat`, whose first (second) row stores the coefficients of the linear policy function for investment expenditures (working hours).

The GAUSS program `Ramsey4c.g` solves the benchmark model for the parameter values from Table 1.1. Given the coefficients of the policy function, we compute artificial time series for output, investment, consumption, working hours, and the real wage. The length of each series is 60 quarters.[14] We use the linear law of motion (2.59), the definitions

$$\begin{aligned} y_t &= e^{z_t} N_t^{1-\alpha} k_t^{\alpha}, \\ c_t &= y_t - i_t, \\ w_t &= (1-\alpha) y_t / N_t, \end{aligned} \tag{2.60}$$

and the GAUSS random number generator `rndn` to compute these time series. The logs of the results are passed to the HP-Filter to

[14] See Section 1.4.4 on the issues of parameter choice and model evaluation.

get percentage deviations from the stationary solutions.[15] Table 2.1 displays second moments from the filtered series. They are averages of 500 simulations.

Table 2.1

	Linear Quadratic			Loglinear		
Variable	s_x	r_{xy}	r_x	s_x	r_{xy}	r_x
Output	1.43	1.00	0.63	1.43	1.00	0.63
Investment	6.19	0.99	0.62	6.10	1.00	0.62
Consumption	0.56	0.99	0.64	0.56	0.99	0.64
Hours	0.77	1.00	0.62	0.77	1.00	0.62
Real Wage	0.67	0.99	0.64	0.67	0.99	0.64

Notes: s_x:=standard deviation of HP-filtered simulated series of variable x, r_{xy}:=cross correlation of variable x with output, r_x:=first order autocorrelation of variable x.

You may want to compare these numbers to those in Table 1.2 presenting the results from the extended path and the value function iteration algorithm, respectively.

Loglinear Approximation. Our starting point are the first-order conditions stated in (2.56) and the associated stationary solution given in (2.57). At the point of the stationary solution the loglinearized equations (2.56) are:

$$\begin{aligned}
\hat{\lambda}_t &= -\eta\hat{c}_t - \theta(1-\eta)\frac{N}{1-N}\hat{N}_t,\\
\hat{\lambda}_t + \alpha\hat{k}_t + \hat{Z}_t &= (1-\eta)\hat{c}_t + \xi_1\hat{N}_t,\\
(1-\alpha)\xi_2 E_t\hat{k}_{t+1} - E_t\hat{\lambda}_{t+1} + \lambda_t &= (1-\alpha)\xi_2 E_t\hat{N}_{t+1} + \xi_2 E_t\hat{Z}_{t+1},\\
aE_t\hat{k}_{t+1} - (a^\eta/\beta)\hat{k}_t &= -(c/k)\hat{c}_t + (1-\alpha)(y/k)\hat{N}_t\\
&\quad + (y/k)\hat{Z}_t,
\end{aligned}$$

where

$$\begin{aligned}
\xi_1 &:= \big(1-\theta(1-\eta)\big)\frac{N}{1-N} + \alpha,\\
\xi_2 &:= 1 - \beta a^{-\eta}(1-\delta).
\end{aligned}$$

[15] See Section 9.4 on the HP-Filter.

It is easy to rewrite this linear system in the matrix notation used in (2.13). Put $[\mathbf{x}_t, \boldsymbol{\lambda}_t]' := [\hat{k}_t, \hat{\lambda}_t]'$ and $\mathbf{u}_t := [\hat{c}_t, \hat{N}_t]'$. The matrices C_u, $C_{x\lambda}$, and C_z from the static equations (2.13a) are:

$$C_u := \begin{bmatrix} -\eta & -\theta((1-\eta)\frac{N}{1-N} \\ 1-\eta & \xi_1 \end{bmatrix}, \quad C_{x\lambda} := \begin{bmatrix} 0 & 1 \\ \alpha & 1 \end{bmatrix}, \quad C_z := \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

The matrices of the dynamic part of the system are:

$$D_{x\lambda} := \begin{bmatrix} (1-\alpha)\xi_2 & -1 \\ a & 0 \end{bmatrix}, \quad F_{x\lambda} := \begin{bmatrix} 0 & 1 \\ -\beta a^{-\eta} & 0 \end{bmatrix},$$
$$D_u := \begin{bmatrix} 0 & (1-\alpha)\xi_2 \\ 0 & 0 \end{bmatrix}, \qquad F_u := \begin{bmatrix} 0 & 0 \\ -(c/k) & (1-\alpha)(y/k) \end{bmatrix},$$
$$D_z := \begin{bmatrix} \xi_2 \\ 0 \end{bmatrix}, \qquad F_z := \begin{bmatrix} 0 \\ y/k \end{bmatrix}.$$

From these matrices, the GAUSS program `SolveLA` computes the policy functions L_{xx}, L_{xz}, L_{ux}, and L_{uz}. We can then iterate over

$$\begin{aligned} \hat{x}_{t+1} &= L_{xx}\hat{x}_t + L_{xz}\hat{z}_t, \\ \hat{z}_{t+1} &= \varrho\hat{z}_t + \epsilon_{t+1}, \\ \begin{bmatrix} \hat{c}_t \\ \hat{N}_t \end{bmatrix} &= L_{ux}\hat{x}_t + L_{uz}\hat{z}_t \end{aligned}$$

to get time paths for the deviations of the stock of capital, of consumption, and of working hours from their respective stationary solutions. To compute the time series for production, investment expenditures, and the real wage, we use the loglinear versions of equations (2.60):

$$\begin{aligned} \hat{y}_t &= (1-\alpha)\hat{N}_t + \alpha\hat{k}_t + \hat{Z}_t, \\ \hat{i}_t &= (y/i)\hat{y}_t - [(y/i)-1]\hat{c}_t, \quad y/i = (y/k)/(a+\delta-1), \\ \hat{w}_t &= \hat{y}_t - \hat{N}_t. \end{aligned}$$

Columns five through seven of Table 2.1 show the second moments obtained from 500 simulations. The GAUSS program `Ramsey4d.g` makes use of the same sequence of productivity

shocks that underlies the linear quadratic solution. Except for investment expenditures, which are a little more volatile in the linear quadratic solution, we obtain virtually the same results. This difference may be traced to the fact that investment rather than consumption is the control variable in the linear quadratic approximation. In the loglinear approximation, the equation for $\hat{i}_t$ implies the following formula for the variance of investment:

$$\text{var}\,(\hat{i}_t) = a^2\text{var}\,(\hat{y}_t) + b^2\text{var}\,(\hat{c}_t) + 2ab\,\text{cov}\,(\hat{y}_t, \hat{c}_t),$$

where $a = y/i$ and $b = (y/i) - 1$. If we compute the variance of investment expenditures in the linear quadratic model according to this formula, we get about the same standard deviation of 6.1 as we find in the loglinear case.

2.4.2 Time to Build

Gestation Period. In the benchmark model investment projects require one quarter to complete. In their classic article KYDLAND and PRESCOTT (1982) use a more realistic gestation period. Based on published studies of investment projects they assume that it takes four quarters for an investment project to be finished. The investment costs are spread out evenly over this period. Yet, the business cycle in this extended model is similar to the business cycle in their benchmark model with a one quarter lag. We introduce the time-to-build assumption into the benchmark model of the previous section. Our results confirm their findings. Nevertheless, we think this venture is worth the while, since it nicely demonstrates the ease of applying the linear quadratic solution algorithm to a rather tricky dynamic model.

The model that we consider uses the same specification of the household's preferences and production technology as the model in the previous section. The timing of investment expenditures differs from this model in the following way. In each quarter t the representative household launches a new investment project. After four quarters this project is finished and adds to the capital stock. The investment costs are spread out over the entire gestation period. More formally, let S_{it}, $i = 1, 2, \ldots, 4$ denote an investment

project that is finished after i periods and that requires the household to pay the fraction ω_i of its total costs. At any period, there are four unfinished projects so that total investment expenditures I_t amount to

$$I_t = \sum_{i=1}^{4} \omega_i S_{it}, \quad \sum_{i=1}^{4} \omega_i = 1. \tag{2.62}$$

Obviously, the S_{it} are related to each other in the following way:

$$\begin{aligned} S_{1t+1} &= S_{2t}, \\ S_{2t+1} &= S_{3t}, \\ S_{3t+1} &= S_{4t}, \end{aligned} \tag{2.63}$$

and the capital stock evolves according to

$$K_{t+1} = (1-\delta)K_t + S_{1t}. \tag{2.64}$$

First-Order Conditions. Since the model exhibits growth, we transform it to a stationary problem. As in Section 2.4.1 we put $c_t := C_t/A_t$, $i_t := I_t/A_t$, $k_t := K_t/A_t$, $\lambda_t := \Lambda_t A_t^{\eta}$, $s_{it} = S_{it}/A_t$, and $\tilde{\beta} := \beta a^{1-\eta}$. In this model, the vector of states is $\mathbf{x}_t = [1, k_t, s_{1t}, s_{2t}, s_{3t}, \ln Z_t]'$ and the vector of controls is $\mathbf{u} = [s_{4t}, N_t]'$. From (2.63) and (2.64) we derive the following law of motion of the stationary variables:

$$\mathbf{x}_{t+1} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1-\delta}{a} & \frac{1}{a} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{a} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{a} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho \end{bmatrix}}_{:=A} \mathbf{x}_t + \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \frac{1}{a} & 0 \\ 0 & 0 \end{bmatrix}}_{=:B} \mathbf{u}_t + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ \epsilon_t \end{bmatrix}. \tag{2.65}$$

The remaining task is to compute the stationary equilibrium. Consider the Lagrangean of the stationary problem:

$$\mathscr{L} = E_0 \sum_{t=0}^{\infty} \tilde{\beta}^t \Bigg\{ \frac{c_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta} + \lambda_t\big(Z_t N_t^{1-\alpha} k_t^{\alpha} - \sum_{i=1}^{4} \omega_i s_{it} - c_t\big) + \gamma_t((1-\delta)k_t + s_{1t} - a k_{t+1}) \Bigg\},$$

where γ_t is the Lagrange multiplier of the transformed constraint (2.64). Differentiating this expression with respect to c_t, N_t, s_{4t} and k_{t+4} provides the following conditions:[16]

$$\lambda_t = c_t^{-\eta}(1-N_t)^{\theta(1-\eta)}, \tag{2.66a}$$

$$\frac{\theta c_t}{1-N_t} = (1-\alpha) Z_t N_t^{-\alpha} k_t^{\alpha}, \tag{2.66b}$$

$$0 = E_t\big[-\omega_4\lambda_t - (\tilde{\beta}/a)\omega_3\lambda_{t+1} - (\tilde{\beta}/a)^2\omega_2\lambda_{t+2} - (\tilde{\beta}/a)^3\omega_1\lambda_{t+3} + (\tilde{\beta}/a)^3\gamma_{t+3}\big], \tag{2.66c}$$

$$0 = E_t\big[-(\tilde{\beta}/a)^3\gamma_{t+3} + (\tilde{\beta}/a)^4(1-\delta)\gamma_{t+4} + (\tilde{\beta}/a)^4\lambda_{t+4}\alpha Z_{t+4} N_{t+4}^{1-\alpha} k_{t+4}^{\alpha-1}\big]. \tag{2.66d}$$

The first and the second condition are standard and need no comment. The third and the fourth condition imply the following Euler equation in the shadow price of capital:

$$0 = E_t\Big\{ \omega_4[(\tilde{\beta}/a)(1-\delta)\lambda_{t+1} - \lambda_t] + \omega_3(\tilde{\beta}/a)[(\tilde{\beta}/a)(1-\delta)\lambda_{t+2} - \lambda_{t+1}] + \omega_2(\tilde{\beta}/a)^2[(\tilde{\beta}/a)(1-\delta)\lambda_{t+3} - \lambda_{t+2}] + \omega_1(\tilde{\beta}/a)^3[(\tilde{\beta}/a)(1-\delta)\lambda_{t+4} - \lambda_{t+3}] + (\tilde{\beta}/a)^4\alpha\lambda_{t+4} Z_{t+4} N_{t+4}^{1-\alpha} k_{t+4}^{\alpha-1} \Big\}.$$

Stationary Equilibrium. On a balanced growth path, where $Z_t = 1$ and $\lambda_t = \lambda_{t+1}$ for all t, this expression reduces to

[16] To keep track of the various terms that involve s_{4t} and k_{t+4}, it is helpful to write out the sum for $t = 0, 1, 2, 3, 4$.

$$\frac{y}{k} = \frac{a - \tilde{\beta}(1-\delta)}{\alpha\tilde{\beta}} \left[\omega_1 + (a/\tilde{\beta})\omega_2 + (a/\tilde{\beta})^2\omega_3 + (a/\tilde{\beta})^3\omega_4\right]. \tag{2.67}$$

Given a, β, δ, and η, we can solve this equation for the output-capital ratio y/k. From $(1-\delta)k+s_1 = ak$ we find $s_1 = (a+\delta-1)k$, the stationary level of new investment projects started in each period. Total investment per unit of capital is then given by

$$\frac{i}{k} = \frac{1}{k}\sum_{i=1}^{4}\omega_i s_i = (a+\delta-1)\sum_{i=1}^{4} a^{i-1}\omega_i.$$

Using this, we can solve for

$$\frac{c}{k} = \frac{y}{k} - \frac{i}{k}.$$

Since $y/c = (y/k)/(c/k)$, we can finally solve the stationary version of (2.66b) for N. This solution in turn provides $k = N(y/k)^{1/(\alpha-1)}$, which allows us to solve for i and c. The final step is to write a procedure that returns the current period utility as a function of $\mathbf{x}$ and $\mathbf{u}$. The latter is given by:

$$g(\mathbf{x},\mathbf{u}) := \frac{1}{1-\eta}\left(e^{\ln Z_t} N_t^{1-\alpha} k_t^{\alpha} - \sum_{i=1}^{4} s_{it}\right)^{1-\eta} (1-N_t)^{\theta(1-\eta)}.$$

Results. The GAUSS program `ToB.g` implements the solution. We use the parameter values from Table 1.1 and assume $\omega_i = 0.25, i = 1,\ldots,4$. Table 2.2 displays the averages of 500 time series moments computed from the simulated model. We used the same random numbers in both the simulations of the benchmark model and the simulations of the time-to-build model. Thus, the differences revealed in Table 2.2 are systematic and not random.

In the time-to-build economy output, investment, and hours are a little less volatile than in the benchmark economy. The intuition behind this result is straightforward. When a positive technological shock hits the benchmark economy the household takes the chance, works more at the higher real wage and transfers part

Table 2.2

	Benchmark			Time to Build		
Variable	s_x	r_{xy}	r_x	s_x	r_{xy}	r_x
Output	1.45	1.00	0.63	1.37	1.00	0.63
Investment	6.31	0.99	0.63	5.85	0.99	0.65
Consumption	0.57	0.99	0.65	0.58	0.97	0.56
Hours	0.78	1.00	0.63	0.71	0.98	0.65
Real Wage	0.68	0.99	0.64	0.68	0.98	0.58

Notes: s_x:=standard deviation of HP-filtered simulated series of variable x, r_{xy}:=cross correlation of variable x with output, r_x:=first order autocorrelation of variable x.

of the increased income via capital accumulation into future periods. Since the shock is highly autocorrelated, the household can profit from the still above average marginal product of capital in the next quarter. Yet in the time-to-build economy intertemporal substitution is not that easy. Income spent on additional investment projects will not pay out in terms of more capital income until the fourth quarter after the shock. However, at this time a substantial part of the shock has faded. This reduces the incentive to invest and, therefore, the incentive to work more.

Lawrence Christiano and Richard Todd (1996) embed the time-to-build structure in a model where labor augmenting technical progress follows a random walk. They use a different parameterization of the weights ω_i. Their argument is that investment projects typically begin with a lengthy planning phase. The overwhelming part of the project's costs are spent in the construction phase. As a consequence, they set $\omega_1 = 0.01$ and $\omega_2 = \omega_3 = \omega_4 = 0.33$. This model is able to account for the positive autocorrelation in output growth, whereas the Kydland and Prescott (1992) parameterization of the same model – $\omega_i = 0.25$, $i = 1, \ldots, 4$ – is not able to replicate this empirical finding. However, the random walk assumption does not lent itself to the linear quadratic approach, and, therefore we will not pursue this matter any further.

2.4.3 New Keynesian Phillips Curve

Money in General Equilibrium. So far we have restricted our attention to non-monetary economies. In this subsection we focus on the interaction of real and monetary shocks to explain the business cycle.

Introducing money into a dynamic general equilibrium model is not an easy task. As a store of value money is dominated by other interest bearing assets like corporate and government bonds or stocks, and in the basically one-good Ramsey model there is no true need for a means of exchange. So how do we guarantee a positive value of pure fiat outside money in equilibrium?

Monetary theory has pursed three approaches (see, e.g., Walsh (2003)). The first device is to assume that money yields direct utility, the second strand of the literature imposes transaction costs, and the third way is to guarantee an exclusive role for money as a store of value. We will pursue the second approach in what follows and assume transaction costs to be proportional to the volume of trade. Moreover, a larger stock of real money balances relative to the volume of trade reduces transaction costs (see Leeper and Sims (1994)). Different from other approaches, as, e.g., the cash-in-advance assumption, our particular specification implies the neutrality of monetary shocks in the log-linear model solution. This allows us to focus on other deviations from the standard model that are required to explain why money has real effects.

The most prominent explanation for the real effects of money that has been pursued in the recent literature are nominal rigidities that arise from sticky wages and/or prices. Among the various models probably the Calvo (1983) model has gained the most widespread attention.[17] For this reason we use the discrete time version of his assumption on price setting to introduce nominal frictions into the monetary economy that we consider in the following paragraphs.

[17] A non-exhaustive list of models of nominal rigidities includes Bergin and Feenstra (2000), Chari, Kehoe, and McGrattan (2000), Cho and Cooley (1995), Cooley and Hansen (1995,1998), ChristianoChristiano, Eichenbaum, and Evans (1997), Hairault and Portier (1995).

The Calvo (1983) hypothesis provides a first-order condition for the optimal relative price of a monopolistically competitive firm that is able to adjust its price optimally whereas a fraction of other firms is not permitted to do so. The log-linear version of this condition (see equation (A.4.11e) in Appendix 4) relates the current inflation rate to the expected inflation rate and a measure of labor market tension. It thus provides solid microfoundations for the well-known Phillips curve that appears in many textbooks. This curve is derived from the short-run aggregate supply function and relates inflation to expected inflation and cyclical unemployment.[18] In the Calvo (1983) model the deviation of marginal costs from their average level measures labor market tension. Since this equation resembles the traditional Phillips curve it is sometimes referred to as the New Keynesian Phillips curve.

The Household Sector. The representative household has the usual instantaneous utility function u defined over consumption C_t and leisure $1 - N_t$, where N_t are working hours:

$$u(C_t, 1 - N_t) := \frac{C_t^{1-\eta}(1 - N_t)^{\theta(1-\eta)}}{1 - \eta}. \tag{2.68}$$

The parameters of this function are non-negative and satisfy $\eta > \theta/(1 + \theta)$. The household receives wages, rental income from capital services, dividends D_t and a lump-sum transfer from the government T_t. We use P_t to denote the aggregate price level. The wage rate in terms of money is W_t and the rental rate in terms of consumption goods is r_t. The household allocates its income net of transaction costs TC_t to consumption, additional holdings of physical capital K_t and real money balances M_t/P_t, where M_t is the beginning-of-period stock of money. This produces the following budget constraint:

$$\frac{W_t}{P_t}N_t + r_t K_t + D_t + T_t - TC_t - C_t \geq \frac{M_{t+1} - M_t}{P_t} + K_{t+1} - (1-\delta)K_t. \tag{2.69}$$

Transactions costs are given by the following function

[18] See, e.g., Mankiw (2000), pp. 364.

$$TC_t = \gamma \left(\frac{C_t}{M_{t+1}/P_t} \right)^{\kappa} C_t, \quad \gamma,\ \kappa > 0. \tag{2.70}$$

Importantly, the assumption that the costs TC_t depend upon the ratio of consumption to real end-of-period money holdings M_{t+1}/P_t is responsible for the neutrality of money in our model. The household maximizes the expected discounted stream of future utility

$$E_0 \sum_{t=0}^{\infty} \beta^t u(C_t, 1 - N_t)$$

subject to (2.69) and (2.70).

Money Supply. The government sector finances the transfers to the household sector from money creation. Thus,

$$T_t = \frac{M_{t+1} - M_t}{P_t}. \tag{2.71}$$

We assume that the monetary authority is not able to monitor the growth rate of money supply perfectly. In particular, we posit the following stochastic process for the growth factor of money supply $\mu_t := M_{t+1}/M_t$:

$$\hat{\mu}_{t+1} = \rho^{\mu} \hat{\mu}_t + \epsilon_t^{\mu}, \quad \hat{\mu}_t := \ln \mu_t - \ln \mu,\ \epsilon_t^{\mu} \sim N(0, \sigma^{\mu}). \tag{2.72}$$

In the stationary equilibrium money grows at the rate $\mu - 1$.

Price Setting. To motivate price setting by individual firms we assume that there is a final goods sector that assembles the output of a large number J_t of intermediary producers to the single good Y_t according to

$$Y_t = \left[J_t^{-1/\epsilon} \sum_{j=1}^{J_t} Y_{jt}^{(\epsilon-1)/\epsilon} \right]^{\epsilon/(\epsilon-1)}, \quad \epsilon > 1. \tag{2.73}$$

The money price of intermediary product j is P_{jt} and final output sells at the price P_t. The representative firm in the final sector

takes all prices as given. Maximizing its profits $P_t Y_t - \sum_{j=1}^{J_t} P_{jt} Y_j$ subject to (2.73) produces the following demand for good j:

$$Y_{jt} = \left(\frac{P_{jt}}{P_t} \right)^{-\epsilon} \frac{Y_t}{J_t}. \tag{2.74}$$

Accordingly, ϵ is the price elasticity of demand for good j. It is easy to demonstrate that the final goods producers earn no profits if the aggregate price index P_t is given by the following function:

$$P_t = \left[\frac{1}{J_t} \sum_{j=1}^{J_t} P_{jt}^{1-\epsilon} \right]^{1/(1-\epsilon)}. \tag{2.75}$$

An intermediary producer j combines labor N_{jt} and capital services K_{jt} according to the following production function:

$$Y_{jt} = Z_t (A_t N_{jt})^{1-\alpha} K_{jt}^{\alpha} - F, \quad \alpha \in (0,1), F > 0. \tag{2.76}$$

F is a fixed cost in terms of forgone output. We will use F to determine the number of firms on a balanced growth path from the zero profit condition. As in all our other models A_t is an exogenous, deterministic process for labor augmenting technical progress,

$$A_{t+1} = a A_t, \quad a \geq 1,$$

and Z_t is a stationary, stochastic process for total factor productivity that follows

$$\hat{Z}_{t+1} = \rho^Z \hat{Z}_t + \epsilon_t^Z, \quad \hat{Z}_t = \ln Z_t, \ \epsilon_t^Z \sim N(0, \sigma^Z).$$

Note that α, F, A_t, and Z_t are common to all intermediary producers, who also face the same price elasticity ϵ.

From now on we must distinguish between two types of firms, which we label A and N, respectively. Type A firms are allowed to set their price P_{At} optimally, whereas type N firms are not. To prevent their relative price P_{Nt}/P_t from falling short of the aggregate price level, type N firms are permitted to increase their price

according to the average inflation factor π. This is the inflation factor on a balanced growth path without any uncertainty. Thus

$$P_{Nt} = \pi P_{Nt-1}. \tag{2.77}$$

To which type an individual firm j belongs is random. At each period $(1-\varphi)J_t$ of firms receive the signal to choose their optimal relative price P_{At}/P_t. The fraction $\varphi \in [0,1]$ must apply the rule (2.77). Those firms that are free to adjust their price solve the following problem:

$$\begin{aligned} \max_{P_{At}} \quad & E_t \sum_{\tau=t}^{\infty} \varphi^{\tau-t} \varrho_\tau \left[\left(\frac{\pi^{\tau-t} P_{At}}{P_\tau} \right) Y_{A\tau} - g_\tau(Y_{A\tau} + F) \right] \\ \text{s.t.} \quad & Y_{A\tau} = \left(\frac{\pi^{\tau-t} P_{At}}{P_\tau} \right)^{-\epsilon} \frac{Y_\tau}{J_\tau}. \end{aligned} \tag{2.78}$$

The sum to the right of the expectations operator E_t is the discounted flow of real profits earned until the firm will be able to reset its price optimally again. Real profits are given by the value of sales in units of the final good $[(\pi^{\tau-t}P_{At})/P_\tau]Y_\tau$ minus production cost $g_\tau(Y_{A\tau}+F)$, where g_τ are the firm's variable unit costs.[19] The term $\varphi^{\tau-t}$ captures the probability that in period τ the firm is still a type N producer. ϱ_τ is the discount factor for time τ profits. We show in Section 4.3.3 that this factor is related to the household's discount factor β and marginal utility of wealth Λ_τ by the following formula:

$$\varrho_\tau = \beta^{\tau-t} \frac{\Lambda_\tau}{\Lambda_t}. \tag{2.79}$$

Intermediary producers distribute their profits to the household sector. Thus,

$$D_t := \sum_{j=1}^{J_t} \frac{P_{jt}}{P_t} Y_{jt} - \frac{W_t}{P_t} N_{jt} - r_t K_{jt}. \tag{2.80}$$

[19] Note that equation (2.76) implies that the firm must produce the amount $Y_{jt} + F$ in order to sell Y_{jt}.

This equation closes the model. To streamline the presentation we restrict ourselves to the properties of the stationary equilibrium and the simulation results. Appendix 4 provides the mathematical details of the analysis and the loglinear model used for the simulation.

Stationary Equilibrium. The model of this section depicts a growing economy. For this reason we must scale the variables so that they are stationary on a balanced growth path. As previously, we use the following definitions: $c_t := C_t/A_t$, $y_t := Y_t/A_t$, $k_t := K_t/A_t$, $\lambda_t := \Lambda_t A_t^{\eta}$. In addition, we define the inflation factor $\pi_t := P_t/P_{t-1}$ and real end-of-period money balances $m_t := M_{t+1}/(A_t P_t)$. The stationary equilibrium of the deterministic model has the following properties:

(1) The productivity shock and the money supply shock equal their respective means $Z_t = Z \equiv 1$ and $\mu_t = \mu$ for all t.
(2) Inflation is constant: $\pi = \frac{P_t}{P_{t-1}}$ for all t.
(3) All (scaled) variables are constant.
(4) All firms in the intermediary sector earn zero profits.

There are two immediate consequences of these assumptions. First, inflation is directly proportional to the growth rate of money supply $\mu - 1$:[20]

$$\pi = \frac{\mu}{a}.$$

Second, the optimal relative price of type A firms satisfies

$$\frac{P_A}{P} = \frac{\epsilon}{\epsilon - 1} g,$$

i.e., it is determined as a markup on the firm's marginal costs g. Furthermore, the formula for the price index given in equation (A.4.5) implies $P_A = P$ so that $g = (\epsilon - 1)/\epsilon$ and $P_N = P$. Since all firms charge the same price, the market share of each producer is Y/J. Therefore, working hours and capital services are equal across firms, $N_j = N/J$, and $K_j = K/J$, and profits amount to

[20] See equation (A.4.2c) for $m_t = m_{t+1}$.

$$D_j = \frac{Y}{J} - g\left(\frac{Y}{J} + F\right).$$

Imposing $D_j = 0$ for all j and using $Y/J = (AN/J)^{1-\alpha}(K/J)^{\alpha} - (F/J)$ yields

$$j := \frac{J_t}{A_t} = \frac{N^{1-\alpha}k^{\alpha}}{\epsilon F}.$$

Thus, to keep profits at zero, the number of firms must increase at the rate $a - 1$ on the balanced growth path. The production function (2.76) thus implies

$$y = \frac{\epsilon - 1}{\epsilon} N^{1-\alpha}k^{\alpha}.$$

Using this in the first-order condition for cost minimization with respect to capital services (see equation (A.4.3b)) implies

$$r = \alpha(y/k).$$

Eliminating r from the Euler equation for capital delivers the well known relation between the output-capital ratio and the household's discount factor β:

$$\frac{y}{k} = \frac{a^{\eta} - \beta(1-\delta)}{\alpha\beta}. \tag{2.81a}$$

This result allows to solve for the consumption-output ratio via the economy's resource constraint (see (A.4.9)):

$$\frac{c}{y} = \left(1 + \frac{1 - a - \delta}{y/k}\right)\left[1 + \gamma\left(\frac{C}{\mu(M/P)}\right)^{\kappa}\right]^{-1}.$$

The stationary version of the Euler condition for money balances (see equation (A.4.2e)) delivers:

$$\frac{\beta a^{1-\eta}}{\mu} = 1 - \kappa\gamma\left(\frac{C}{\mu(M/P)}\right)^{1+\kappa}. \tag{2.81b}$$

We need a final equation to determine the stationary level of working hours. Using the results obtained so far we derive this relation

from the household's first-order condition with respect to labor supply (see equation (A.4.2b)):

$$\frac{N}{1-N} = \frac{1-\alpha}{\theta}\left(1+\frac{1-a-\delta}{y/k}\right)^{-1} h(c/x), \tag{2.81c}$$
$$h(c/x) := \frac{1+\gamma(c/x)^{\kappa}}{1+\gamma(1+\kappa)(c/x)^{\kappa}}, \quad \frac{c}{x} := \frac{PC}{\mu M}.$$

It is obvious from equation (2.81a) that the output-capital ratio and therefore also the capital-labor ratio k/N and labor productivity y/N are independent of the money growth rate. As can be seen from (2.81b), the velocity of end-of-period money balances $c/x \equiv C/(\mu(M/P))$ is an increasing function of the money growth rate. In the benchmark model of Section 2.4.1 working hours are determined by the first two terms on the rhs of (2.81c). The presence of money adds the factor $h(c/x)$. It is easy to show that $h(c/x)$ is an decreasing function of the velocity of money (c/x). Since $N/(1-N)$ increases with N, steady-state working hours are a decreasing function of the money growth rate.

Calibration. We do not need to assign new values to the standard parameters of the model. The steady state relations presented in the previous paragraph show that the usual procedure to calibrate β, α, a, and δ is still valid. We will also use the empirical value of N to infer θ from (2.81c). This implies a slightly smaller value of θ as compared to the value of this parameter in the benchmark model. Nothing is really affected from this choice.

Unfortunately, there is no easy way to determine the parameters of the productivity shock, since there is no simple aggregate production function that we could use to identify Z_t. The problem becomes apparent from the following equation, which we derive in Appendix 4:

$$\hat{y}_t = \vartheta(1-\alpha)\hat{N}_t + \vartheta\alpha\hat{k}_t + \vartheta\hat{Z}_t(1-\vartheta)\hat{j}_t, \quad \vartheta = \frac{\epsilon}{\epsilon-1}. \tag{2.82}$$

This equation is the model's analog to the log-linear aggregate production function in the benchmark model given by

$$\hat{y}_t = (1-\alpha)\hat{N}_t + \alpha\hat{k}_t + \hat{Z}_t.$$

Since $\vartheta > 1$ we overstate the size of $\hat{Z}_t$, when we use this latter equation to estimate the size of the technology shock from data on output, hours, and the capital stock. Furthermore, in as much as the entry of new firms measured by $\hat{j}_t$ depends upon the state of the business cycle, the usual measure of $\hat{Z}_t$ is further spoiled. We do not consider this book to be the right place to develop this matter further. Possible remedies have been suggested for instance by Rotemberg and Woodford (1995) and Hairault and Portier (1995). Instead, we continue to use the parameters from the benchmark model so that we are able to compare our results to those obtained in the Section 2.4.1 and Section 2.4.2.

What we need further are the parameters of the money supply process, of the transaction costs function, and of the structure of the monopolistic intermediary goods sector.

Our measure of money supply is the West-German monetary aggregate M1 per capita. As in Section 1.5 we focus on the period 1975.i through 1989.iv. The average quarterly growth rate of this aggregate was 1.67 percent. We fitted an AR(1) process to the deviations of μ_t from this value. The autocorrelation parameter from this estimation is not significantly different from zero and the estimated standard deviation of the innovations is $\sigma^\mu = 0.0173$. We use the average velocity of M1 with respect to consumption of 0.84 to determine γ from (2.81b). Finally, we can use the following observation to find an appropriate value of κ: The lhs of equation (2.81b) is equal to

$$\frac{1}{\pi(1-\delta+r)}.$$

The term in the denominator is the nominal interest rate factor, i.e., one plus the nominal interest rate q, say. This implies the following long run interest rate elasticity of the demand for real money balances:

$$\frac{d(M/P)/(M/P)}{dq/q} = \frac{-1}{(1+\kappa)\pi(1-\delta+r)}.$$

The estimate of this elasticity provided by Hoffman, Rasche, and Tieslau (1995) is about -0.2. Since $1/R \approx 1$ we use $\kappa = 4$.

Table 2.3

Preferences	Production	
β=0.994	a=1.005	α=0.73
η=2.0	δ=0.011	ρ^Z=0.90
N=0.13	σ^Z=0.0072	
Money Supply	**Transactions Costs**	**Market Structure**
μ=1.0167	$C/(M/P)$=0.84	φ=0.25
ρ^μ=0.0	κ=4.0	ϵ=6.0
σ^μ=0.0173		

The degree of nominal rigidity in our model is determined by the parameter φ. According to the estimates found in ROTEMBERG (1987) it takes about four quarters to achieve full price adjustment. Therefore, we use $\varphi = 0.25$. LINNEMANN (1999) presents estimates of markups for Germany, which imply a price elasticity of $\epsilon = 6$. Table summarizes this choice of parameters.

Results. The GAUSS program `NKPK.g` implements the solution. To understand the mechanics of the model, we consider the case without nominal frictions first. Figure 2.2 displays the time paths of several variables after a one-time shock to the money supply process (2.72) in period $t = 3$ of size σ^μ. Before this shock the economy was on its balanced growth path, after this shock the growth factor of money follows (2.72) with $\epsilon_t^\mu = 0$. For this reason time paths like those in Figure 2.2 are known as impulse responses.

The case $\rho^\mu = 0$ highlights the unanticipated effect of the shock, since after period 3 the money growth rate is back on its stationary path. The money transfer in period 3 raises the household's income unexpectedly. Since both consumption and leisure are normal goods the household's demand for consumption increases and its labor supply decreases. The latter raises the real wage so that marginal costs increase. Higher costs and excess demand raise inflation. This increase just offsets the extra amount of money so that the real stock of money does not change. Therefore, none of the real variables really changes. Money is neutral.

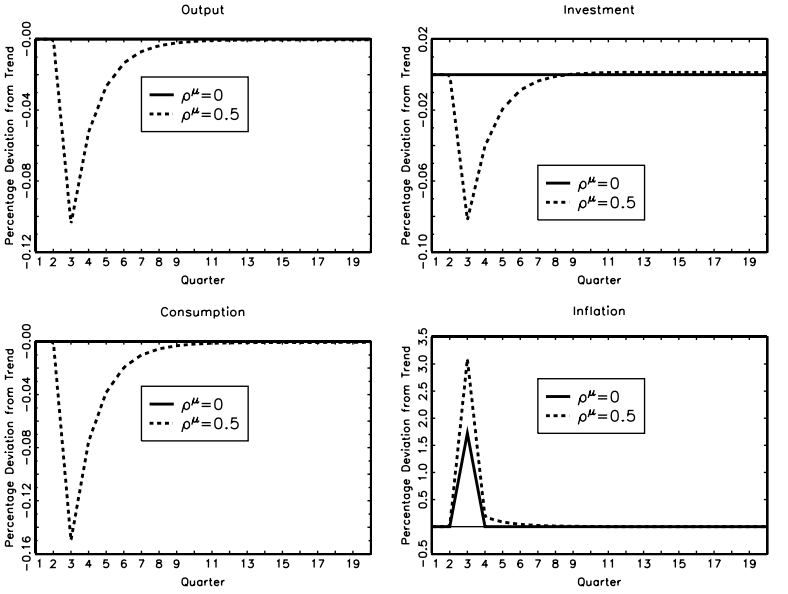

Figure 2.2: Real Effects of a Monetary Shock in the Model Without Nominal Rigidities

This can bee seen in Figure 2.2 since the impulse responses of output, consumption, and investment coincide with the zero line.

Things are different when the shock is autocorrelated. In this case there is also an anticipated effect. Households know that money growth will remain above average for several periods and expect above average inflation. This in turn increases the expected costs of money holdings and households reduce their cash holdings. As a consequence, the velocity of money with respect to consumption increases. To offset this negative effect on transaction costs the households reduce consumption. Their desire to smooth consumption finally entails less investment. Note however that these effects are very small. For instance, consumption in period 3 is 0.16 percent below its stationary value, and investment is 0.08 percent below its steady state level.

We find very different impulse responses, if nominal rigidities are present. This can be seen in Figure 2.3. Since inflation cannot adjust fully, households expect above average inflation even in the

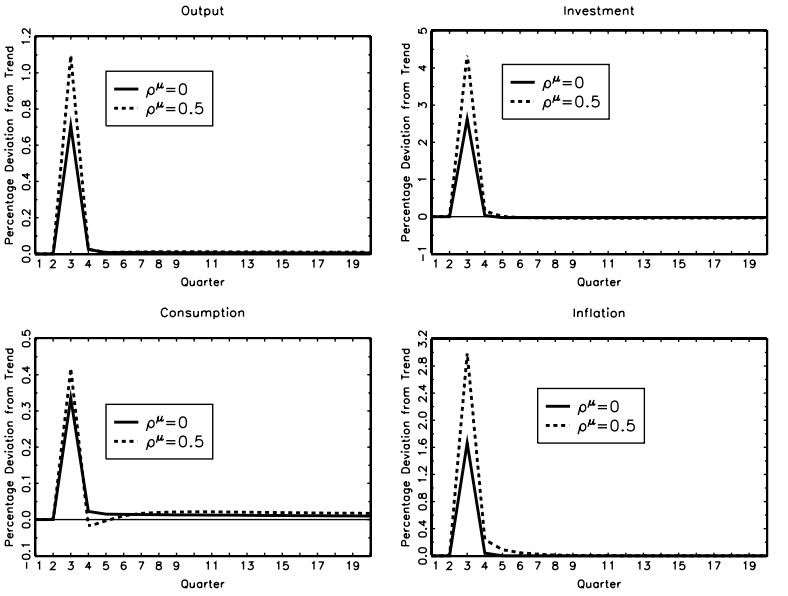

Figure 2.3: Impulse Responses to a Monetary Shock in the New Keynesian Phillips Curve Model

case of $\rho^{\mu} = 0$. This creates a desire to shift consumption to the current period so that there is excess demand. Monopolistically competitive firms are willing to satisfy this demand since their price exceeds their marginal costs. Thus output increases. The household's desire to spread the extra income over several periods spurs investment into physical capital.

There is another noteworthy property of the model: The spike-like shape of the impulse responses. Consumption, hours, output, and investment are almost back on their respective growth paths after period 3, irrespective of whether or not the monetary shock is autocorrelated. This is in stark contrast to the findings of empirical studies. For instance, according to the impulse responses estimated by Cochrane (1998) the response of output is hump shaped and peaks after eight quarters. The apparent failure of the model to explain the persistence of a monetary shock has let many researches to question the usefulness of the New Keynesian Phillips curve. In a recent paper Eichenbaum and Fisher

(2004) argue that the Calvo (1983) model is able to explain persistent effects of monetary shocks if one abandons the convenient but arbitrary assumption of a constant price elasticity.

Table 2.4

	$\vartheta = 1, \varphi = 0, \sigma^{\mu} = 0$			$\sigma^{\mu} = 0$			$\sigma^{\mu} = 0.0173$		
Variable	s_x	r_{xy}	r_x	s_x	r_{xy}	r_x	s_x	r_{xy}	r_x
Output	1.43	1.00	0.63	1.55	1.00	0.68	1.69	1.00	0.56
	(1.14)	(1.00)	(0.80)	(1.14)	(1.00)	(0.80)	(1.14)	(1.00)	(0.80)
Consump-	0.53	0.99	0.65	0.55	0.98	0.72	0.64	0.98	0.52
tion	(1.18)	(0.79)	(0.84)	(1.18)	(0.79)	(0.84)	(1.18)	(0.79)	(0.84)
Invest-	6.16	1.00	0.63	6.87	1.00	0.67	7.31	1.00	0.58
ment	(2.59)	(0.75)	(0.79)	(2.59)	(0.75)	(0.79)	(2.59)	(0.75)	(0.79)
Hours	0.76	1.00	0.63	0.59	0.99	0.75	0.97	0.86	0.23
	(0.78)	(0.40)	(0.31)	(0.78)	(0.40)	(0.31)	(0.78)	(0.40)	(0.31)
Real	0.67	0.99	0.65	0.66	0.99	0.72	0.81	0.97	0.45
Wage	(1.17)	(0.41)	(0.91)	(1.17)	(0.41)	(0.91)	(1.17)	(0.41)	(0.91)
Inflation	0.27	−0.53	−0.07	0.31	−0.48	−0.05	1.62	0.30	−0.06
	(0.28)	(0.04)	(−0.03)	(0.28)	(0.04)	(−0.03)	(0.28)	(0.04)	(−0.03)

Notes: s_x:=standard deviation of HP-filtered simulated series of variable x, r_{xy}:=cross correlation of variable x with output, r_x:=first order autocorrelation of variable x. Empirical magnitudes in parenthesis.

Table 2.4 reveals the contribution of monetary shocks to the business cycle. To fully understand the model we must disentangle several mechanisms that work simultaneously. For this reason, columns 2 to 4 present simulations, where neither monetary shocks, nor nominal rigidities, nor monopolistic elements are present. This requires to set $\vartheta = 1$, $\varphi = 0$, and $\sigma^{\mu} = 0$ in the program `NKPK.g`. Obviously, this model behaves almost like the benchmark model (see Tables 2.1 and 2.2).

Next consider columns 5 to 7. In this model, there are no monetary shocks, but there are monopolistic price setters facing nominal rigidities. The most immediate differences are: output is more volatile and hours are less volatile than in the benchmark model.

How can this happen? Note that under monopolistic price setting the marginal product of labor is larger than it is under perfect competition. The same is true for the marginal product of capital. Thus, a technology shock that shifts the production function outward boosts output more than it would do in a competitive environment. Due to the fixed costs of production, the shock also raises profits and thus dividend payments to the household. This in turn increases the household's demand for leisure. Since prices do not fully adjust, these effects are a bit smaller than they are in a purely real model without nominal frictions.[21]

Columns 8 to 10 present the results from simulations where both technology shocks and monetary shocks are present. The most noteworthy effect concerns working hours. The standard deviation of this variable increases by 64 percent. The wealth effect that we identified above now works in the opposite direction: A monetary shock squeezes the profits of firms, since marginal costs rise and prices cannot fully adjust. As a consequence, the household's demand for leisure falls. But note, most of the shock is absorbed by inflation, which increases substantially.

[21] A detailed comparison between a real and a monetary model of monopolistic price setting appears in Maussner (1999).

Appendix 3 Solution of the Stochastic LQ problem

In this Appendix we provide the details of the solution of the stochastic linear quadratic (LQ) problem. If you are unfamiliar with matrix algebra, you should consult 8.1 before proceeding.

Using matrix algebra we may write the Bellman equation (2.5) as follows:

$$\begin{aligned}\mathbf{x}'P\mathbf{x}+d=\max_{\mathbf{u}}\ \Big[&\mathbf{x}'Q\mathbf{x}+\mathbf{u}'R\mathbf{u}+2\mathbf{u}'S\mathbf{x}\\&+\beta E\Big(\mathbf{x}'A'PA\mathbf{x}+\mathbf{u}'B'PA\mathbf{x}+\boldsymbol{\epsilon}'PA\mathbf{x}\\&+\mathbf{x}'A'PB\mathbf{u}+\mathbf{u}'B'PB\mathbf{u}+\boldsymbol{\epsilon}'PB\mathbf{u}\\&+\mathbf{x}'A'P\boldsymbol{\epsilon}+\mathbf{u}'B'P\boldsymbol{\epsilon}+\boldsymbol{\epsilon}'P\boldsymbol{\epsilon}+d\Big)\Big].\end{aligned}\tag{A.3.1}$$

Since $E(\boldsymbol{\epsilon})=\mathbf{0}$ the expectation of all linear forms involving the vector of shocks $\boldsymbol{\epsilon}$ evaluate to zero. The expectation of the quadratic form $\boldsymbol{\epsilon}'P\boldsymbol{\epsilon}$ is:

$$E(\sum_{i=1}^{n}\sum_{i=1}^{n}p_{ij}\epsilon_i\epsilon_j)=\sum_{i=1}^{n}\sum_{j=1}^{n}p_{ij}\sigma_{ij},$$

where σ_{ij} (σ_{ii}) denotes the covariance (variance) between ϵ_i and ϵ_j (of ϵ_i). It is not difficult to see that this expression equals $\mathrm{tr}\,(P\Sigma)$. Furthermore, since $P=P'$ and

$$z:=\mathbf{u}'B'PA\mathbf{x}=z'=(\mathbf{x}'A'PB'\mathbf{u})'$$

we may write the Bellman equation as

$$\begin{aligned}\mathbf{x}'P\mathbf{x}+\mathbf{d}=\max_{\mathbf{u}}\ \Big[&\mathbf{x}'Q\mathbf{x}+2\mathbf{u}S\mathbf{x}+\mathbf{u}'R\mathbf{u}+\beta\mathbf{x}'A'PA\mathbf{x}\\&+2\beta\mathbf{x}'A'PB\mathbf{u}+\beta\mathbf{u}'B'PB\mathbf{u}+\beta\mathrm{tr}\,(P\Sigma)+\beta d\Big].\end{aligned}\tag{A.3.2}$$

This is equation (2.6) in the main text. Differentiation of the rhs of this expression with respect to $\mathbf{u}$ yields

$$2S\mathbf{x}+2R\mathbf{u}+2\beta(\mathbf{x}'A'PB)'+2\beta(B'PB)\mathbf{u}.$$

Setting this equal to the zero vector and solving for $\mathbf{u}$ gives

$$\begin{aligned} &\underbrace{(R+\beta B'PB)}_{C^{-1}}\mathbf{u} = -\underbrace{(S+\beta B'PA)}_{D}\mathbf{x} \\ \Rightarrow \quad &\mathbf{u} = -CD\mathbf{x}. \end{aligned} \tag{A.3.3}$$

If we substitute this solution back into (A.3.2), we get:

$$\begin{aligned} \mathbf{x}'P\mathbf{x}+d = {} & \mathbf{x}'Q\mathbf{x} - 2(CD\mathbf{x})'S\mathbf{x} + (CD\mathbf{x})'RCD\mathbf{x} + \beta\mathbf{x}'A'PA\mathbf{x} \\ & - 2\beta\mathbf{x}'A'PBCD\mathbf{x} + \beta(CD\mathbf{x})'B'PBCD\mathbf{x} + \beta\mathrm{tr}\,(P\Sigma) + \beta d \\ = {} & \mathbf{x}'Q\mathbf{x} + \beta\mathbf{x}'A'PA\mathbf{x} \\ & - 2\mathbf{x}'D'C'S\mathbf{x} - 2\beta\mathbf{x}'A'PBCD\mathbf{x} \\ & + \mathbf{x}'D'C'RCD\mathbf{x} + \beta\mathbf{x}'D'C'B'PBCD\mathbf{x} \\ & + \beta\mathrm{tr}\,(P\Sigma) + \beta d. \end{aligned}$$

The expression on the fourth line can be simplified to

$$\begin{aligned} &- 2\mathbf{x}'D'C'S\mathbf{x} - \underbrace{2\beta\mathbf{x}'A'PBCD\mathbf{x}}_{=2\beta\mathbf{x}'D'C'B'PA\mathbf{x}} \\ &= -2\mathbf{x}'D'C'\underbrace{(S+\beta B'PA)}_{D}\mathbf{x} = -2\mathbf{x}'D'C'D\mathbf{x}. \end{aligned}$$

The terms on the fifth line add to

$$\mathbf{x}'D'C'\underbrace{(R+\beta B'PB)C}_{I}D\mathbf{x} = \mathbf{x}'D'C'D.$$

Therefore,

$$\mathbf{x}'P\mathbf{x}+d = \mathbf{x}'Q\mathbf{x}+\beta\mathbf{x}'A'PA\mathbf{x}-\mathbf{x}'D'C'D\mathbf{x}+\beta\mathrm{tr}\,(P\Sigma)+\beta d. \tag{A.3.4}$$

For this expression to hold, the coefficient matrices of the various quadratic forms on both sides of equation (A.3.4) must satisfy the matrix equation

$$P = Q + \beta A'PA + D'C'D,$$

and the constant d must be given by

$$d = \frac{\beta}{1-\beta}\mathrm{tr}\,(P\Sigma).$$

This finishes the derivation of the solution of LQ the problem.

Appendix 4 Derivation of the Loglinear Model of the New Keynesian Phillips Curve

In this Appendix we provide the details of the solution of the model from Section 2.4.3.

The Household's Problem. The Lagrangean of the household's problem is:

$$\begin{aligned}\mathscr{L} = E_0 \sum_{t=0}^{\infty} \beta^t \Bigg\{ & \frac{C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta} \\ & + \Lambda_t \Bigg[\frac{W_t}{P_t} N_t + (r_t - \delta) K_t + D_t + T_t \\ & - \gamma \left(\frac{C_t}{M_{t+1}/P_t} \right)^{\kappa} C_t - C_t - (K_{t+1} - K_t) - \frac{M_{t+1} - M_t}{P_t} \Bigg] \Bigg\}.\end{aligned}$$

Differentiating this expression with respect to C_t, N_t, K_{t+1} and M_{t+1} provides the following first-order conditions:

$$\begin{aligned}0 &= C_t^{-\eta}(1-N_t)^{\theta(1-\eta)} - E_t \Lambda_t \left[1 + \gamma(1+\kappa) \left(\frac{C_t}{M_{t+1}/P_t} \right)^{\kappa} \right], \\ 0 &= \theta C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)-1} - \Lambda_t \frac{W_t}{P_t}, \\ 0 &= \Lambda_t - \beta E_t \Lambda_{t+1}(1 - \delta + r_{t+1}), \\ 0 &= E_t \left\{ -\frac{\Lambda_t}{P_t} + \kappa\gamma \left(\frac{C_t}{M_{t+1}/P_t} \right)^{\kappa+1} \frac{\Lambda_t}{P_t} + \beta \frac{\Lambda_{t+1}}{P_{t+1}} \right\}.\end{aligned} \tag{A.4.1}$$

As usual, we must define variables that are stationary. We choose $c_t := C_t/A_t$, $k_t := K_t/A_t$, $\lambda_t := \Lambda_t A_t^{\eta}$, $w_t := W_t/(P_t A_t)$, $m_{t+1} := M_{t+1}/(A_t P_t)$, and $j_t := J_t/A_t$. The inflation factor is $\pi_t := P_t/P_{t-1}$. Since the price level is determined in period t, this variable is also a period t variable. The growth factor of money supply, also determined in period t, is given by $\mu_t := M_{t+1}/M_t$, where M_t is the beginning-of-period money stock and M_{t+1} the end-of-period money stock. In these variables, we can rewrite the system (A.4.1) as follows:

$$c_t^{-\eta}(1-N_t)^{\theta(1-\eta)} = \lambda_t\left(1+\gamma(1+\kappa)\left(\frac{c_t}{m_{t+1}}\right)^{\kappa}\right), \tag{A.4.2a}$$

$$\lambda_t w_t = \theta c_t^{1-\eta}(1-N_t)^{\theta(1-\eta)-1}, \tag{A.4.2b}$$

$$m_{t+1} = \frac{\mu_t}{a\pi_t}m_t, \tag{A.4.2c}$$

$$\lambda_t = \beta a^{-\eta}E_t\lambda_{t+1}\left(1-\delta+r_{t+1}\right), \tag{A.4.2d}$$

$$\beta a^{-\eta}E_t\frac{\lambda_{t+1}}{\pi_{t+1}} = \lambda_t\left(1-\kappa\gamma\left(\frac{c_t}{m_{t+1}}\right)^{\kappa+1}\right). \tag{A.4.2e}$$

Price Setting. To study the price setting behavior, it is convenient to first solve the firm's cost minimization problem

$$\min_{N_{jt},K_{jt}} \quad \frac{W_t}{P_t}N_{jt} + r_t K_{jt} \quad \text{s.t. (2.76)}.$$

The first-order conditions for this problem are easy to derive. They are:

$$w_t = g_t(1-\alpha)Z_t N_{jt}^{-\alpha}(K_{jt}/A_t)^{\alpha} = g_t(1-\alpha)Z_t(k_t/N_t)^{\alpha}, \tag{A.4.3a}$$

$$r_t = g_t\alpha Z_t N_{jt}^{1-\alpha}(K_{jt}/A_t)^{\alpha-1} = g_t\alpha Z_t(k_t/N_t)^{\alpha-1}, \tag{A.4.3b}$$

where g_t is the Lagrange multiplier of the constraint (2.76), and $w_t := W_t/(P_t A_t)$ is the real wage rate per unit of effective labor. It is well known from elementary production theory that g_t equals the marginal costs of production. Furthermore, the constant scale assumption with respect to $Y_{jt}+F$ also implies that g_t are the variable unit costs of production:

$$g_t = \frac{(W_t/P_t)N_{jt} + r_t K_{jt}}{Y_{jt}+F}.$$

Marginal costs as well as the capital-output ratio are the same in all intermediary firms due to the symmetry that is inherent in the specification of the demand and production function. For later use we note the factor demand functions that are associated with this solution:

$$N_{jt} = \frac{Y_{jt}+F}{A_t Z_t}\left(\frac{1-\alpha}{\alpha}\right)^{\alpha}\left(\frac{w_t}{r_t}\right)^{-\alpha}, \tag{A.4.4a}$$

$$K_{jt} = \frac{Y_{jt}+F}{Z_t}\left(\frac{1-\alpha}{\alpha}\right)^{\alpha-1}\left(\frac{w_t}{r_t}\right)^{1-\alpha}. \tag{A.4.4b}$$

In each period $(1-\varphi)J_t$ firms choose their optimal money price P_{At} and φJ_t firms increase their price according to average inflation,

$$P_{Nt} = \pi P_{Nt-1}.$$

Therefore, the aggregate price level given in equation (2.75) is:

$$P_t = \left[(1-\varphi)P_{At}^{1-\epsilon} + \varphi(\pi P_{Nt-1})^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}.$$

Now observe that the pool of firms that are not allowed to choose their price optimally consists itself of firms that were able to set their optimal price in the previous period and those unlucky once that were not allowed to do so. Thus, P_{Nt-1} is in turn the following index:

$$P_{N\,t-1} = \left[(1-\varphi)P_{A\,t-1}^{1-\epsilon} + \varphi(\pi P_{N\,t-2})^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}.$$

Using this formula recursively establishes:

$$P_t = \left[(1-\varphi)\left\{P_{A\,t}^{1-\epsilon} + \varphi(\pi P_{A\,t-1})^{1-\epsilon} + \varphi^2(\pi^2 P_{A\,t-2})^{1-\epsilon} + \dots\right\}\right]^{\frac{1}{1-\epsilon}},$$

which implies

$$\varphi(\pi P_{t-1})^{1-\epsilon} = \left[(1-\varphi)\{\varphi(\pi P_{A\,t-1})^{1-\epsilon} + \varphi^2(\pi^2 P_{A\,t-2})^{1-\epsilon} + \dots\}\right].$$

Thus, the aggregate price level can equivalently be written as

$$P_t = \left[(1-\varphi)P_{At}^{1-\epsilon} + \varphi(\pi P_{t-1})^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}. \tag{A.4.5}$$

We turn now to the first-order conditions that determine the optimal price of type A firms. Maximizing the expression in (2.78) with respect to P_{At} provides the following condition:

$$\underbrace{\frac{\epsilon-1}{\epsilon}}_{=:1/\vartheta} P_{At} E_t \sum_{\tau=t}^{\infty} \varphi^{\tau-t} \varrho_\tau \left(\frac{\pi^{\tau-t}}{P_\tau}\right)^{(1-\epsilon)} \frac{Y_\tau}{J_\tau}$$
$$= E_t \sum_{\tau=t}^{\infty} \varphi^{\tau-t} \varrho_\tau \left(\frac{\pi^{\tau-t}}{P_\tau}\right)^{-\epsilon} g_\tau \frac{Y_\tau}{J_\tau}.$$

We multiply both sides by $P_t^{-\epsilon}$ and substitute the rhs of equation (2.79) for ϱ_τ the rhs. This delivers the following condition:

$$\begin{aligned}&\frac{1}{\vartheta}\left(\frac{P_{At}}{P_t}\right)E_t\sum_{\tau=t}^{\infty}(\varphi\beta a^{-\eta})^{\tau-t}\frac{\lambda_\tau}{\lambda_t}\pi^{(1-\epsilon)(\tau-t)}\left(\frac{P_\tau}{P_t}\right)^{\epsilon-1}\frac{Y_\tau}{J_\tau}\\&=E_t\sum_{\tau=t}^{\infty}(\varphi\beta a^{-\eta})^{\tau-t}\frac{\lambda_\tau}{\lambda_t}\pi^{-\epsilon(\tau-t)}g_\tau\left(\frac{P_\tau}{P_t}\right)^{\epsilon}\frac{Y_\tau}{J_\tau}.\end{aligned}\tag{A.4.6}$$

Our next task is to determine aggregate output and employment. Note from (2.74) that final goods producers use different amounts of type A and N goods since the prices of these inputs differ. Therefore, aggregate output is:

$$\begin{aligned}Y_t&=(1-\varphi)J_t\frac{P_{At}}{P_t}Y_{At}+\varphi J_t\frac{\pi}{\pi_t}Y_{Nt}\\&=(1-\varphi)J_t\left[\frac{P_{At}}{P_t}\left(Z_tA_tN_{At}(K_{At}/A_tN_{At})^{1-\alpha}-F\right)\right]\\&\quad+\varphi J_t\left[\frac{\pi}{\pi_t}\left(Z_tA_tN_{Nt}(K_{Nt}/A_tN_{Nt})^{1-\alpha}-F\right)\right].\end{aligned}$$

Using the fact that all producers choose the same capital-labor ratio k_t/N_t provides:

$$\begin{aligned}Y_t=A_t&\left[\frac{P_{At}}{P_t}Z_t\underbrace{(1-\varphi)J_tN_{At}}_{n_tN_t}(k_t/N_t)^{1-\alpha}+\frac{\pi}{\pi_t}Z_t\underbrace{\varphi J_tN_{Nt}}_{(1-n_t)N_t}(k_t/N_t)^{1-\alpha}\right]\\&-J_tF\left[(1-\varphi)\frac{P_{At}}{P_t}+\varphi\frac{\pi}{\pi_t}\right],\end{aligned}$$

where the fraction of workers employed by type A firms n_t is given by:

$$n_t:=\frac{(1-\varphi)J_tN_{At}}{N_t}.\tag{A.4.7}$$

From this we derive the following equation in terms of aggregate output per efficiency unit A_t:

$$\begin{aligned}y_t:=\frac{Y_t}{A_t}&=Z_tN_t^{\alpha}k_t^{1-\alpha}\left[n_t\frac{P_{At}}{P_t}+(1-n_t)\frac{\pi}{\pi_t}\right]\\&\quad-j_tF\left[(1-\varphi)\frac{P_{At}}{P_t}+\varphi\frac{\pi}{\pi_t}\right].\end{aligned}\tag{A.4.8}$$

In the loglinear version of this equation the variable n_t drops out. Thus, there is no need to derive the equation that determines this variable.

Finally, consider the household's budget constraint (2.69). In equilibrium it holds with equality. Using the government's budget constraint (2.71) and the definition of dividends (2.80), we end up with the following resource constraint:

$$ak_{t+1} = y_t + (1-\delta)k_t - \gamma\left(\frac{c_t}{m_{t+1}}\right)^{\kappa} c_t - c_t. \tag{A.4.9}$$

The Loglinear Model. The dynamic model consists of equations (A.4.2), (A.4.3), (A.4.5),(A.4.6), (A.4.8), and (A.4.9). The stationary equilibrium of this system is considered in the main text so that we can focus on the derivation of the loglinear equations. First, consider the variables that play the role of the control variables in the system (2.36). These are the deviations of consumption, working hours, output, the inflation factor, the real wage rate, and the rental rate of capital from their respective steady state levels:

$$\mathbf{u}_t := [\hat{c}_t, \hat{N}_t, \hat{y}_t, \hat{\pi}_t, \hat{w}_t, \hat{r}_t]'.$$

The state variables with predetermined initial conditions are the stock of capital and beginning-of-period money real money balances. Thus, in terms of (2.36):

$$\mathbf{x}_t = [\hat{k}_t, \hat{m}_t]'.$$

Purely exogenous are the technological shock $\hat{Z}_t$, the monetary shock $\hat{\mu}_t$, and the entrance rate of firms $\hat{j}_t$ into the intermediary goods sector. For the latter we will assume it is independent of the state of the business cycle so that $\hat{j}_t = 0$ for all t.[22] Thus,

$$\mathbf{z}_t = [\hat{Z}_t, \hat{\mu}_t]'.$$

The remaining variables are the shadow price of capital λ_t, firms' marginal costs g_t, and real end-of-period money balances m_{t+1}. Note, that we cannot determine the latter from equation (A.4.2c), since we need this equation to determine $\mathbf{u}_t$. Thus, in addition to λ_t and g_t, this variable is a costate. To keep to the dating convention in (2.36) we define the auxiliary variable $x_t \equiv m_{t+1}$. Hence, our vector of costate variables comprises:

[22] For instance, Rotemberg and Woodford (1995) link $\hat{j}_t$ to the technological shock.

$\lambda_t = [\hat{\lambda}_t, \hat{g}_t, \hat{x}_t]'$.

We first present the static equations that relate control variables to state and costate variables. The log-linear versions of equations (A.4.2a) through (A.4.2c) are

$$-(\eta + \xi_1)\hat{c}_t - \xi_2 \hat{N}_t = \hat{\lambda}_t - \xi_1 \hat{x}_t, \tag{A.4.10a}$$

$$(1-\eta)\hat{c}_t - \xi_3 \hat{N}_t - \hat{w}_t = \hat{\lambda}_t, \tag{A.4.10b}$$

$$\hat{\pi}_t = \hat{m}_t - \hat{x}_t + \hat{\mu}_t, \tag{A.4.10c}$$

$$\xi_1 := \frac{\kappa\gamma(1+\kappa)(c/x)^\kappa}{1+\gamma(1+\kappa)(c/x)^\kappa}, \quad \frac{c}{x} = \frac{C}{\mu(M/P)},$$

$$\xi_2 := \theta(1-\eta)\frac{N}{1-N},$$

$$\xi_3 := [\theta(1-\eta)-1]\frac{N}{1-N}.$$

The loglinear cost-minimizing conditions (A.4.3) deliver two further equations:

$$(1-\alpha)\hat{N}_t + \hat{w}_t = (1-\alpha)\hat{k}_t + \hat{g}_t + \hat{Z}_t, \tag{A.4.10d}$$

$$-\alpha \hat{N}_t + \hat{r}_t = -\alpha \hat{k}_t + \hat{g}_t + \hat{Z}_t. \tag{A.4.10e}$$

To derive the sixth equation we use the formula for the price level to write

$$\pi_t = \frac{P_t}{P_{t-1}} = \left[(1-\varphi)\left(\frac{P_{At}}{P_t}\underbrace{\frac{P_t}{P_{t-1}}}_{\pi_t}\right)^{1-\epsilon} + \varphi\pi^{1-\epsilon}\right]^{\frac{1}{1-\epsilon}}.$$

Log-linearizing at $P_A/P = 1$ provides:

$$\hat{\pi}_t = \frac{1-\varphi}{\varphi}\widehat{P_{At}/P}_t.$$

We use this relation to derive

$$\hat{y}_t - \vartheta\alpha\hat{N}_t = \vartheta(1-\alpha)\hat{k}_t + \vartheta\hat{Z}_t + (1-\vartheta)\hat{j}_t. \tag{A.4.10f}$$

from equation (A.4.8). The six equations (2.24a) through (A.4.10f) determine the control variables. We turn now to the dynamic equations

that determine the time paths of $\hat{k}_t$, $\hat{m}_t$, $\hat{x}_t \equiv \hat{m}_{t+1}$, $\hat{\lambda}_t$, and $\hat{g}_t$. The log-linear versions of the resource constraint (A.4.9), the Euler equations for capital and money balances (A.4.2d) and (A.4.2e), and the definition $x_t := m_{t+1}$ are:

$$aE_t\hat{k}_{t+1} - (1-\delta)\hat{k}_t - \xi_4\hat{x}_t = \frac{y}{k}\hat{y}_t - \xi_5\hat{c}_t, \tag{A.4.11a}$$

$$-E_t\hat{\lambda}_{t+1} + \hat{\lambda}_t = \xi_6 E_t\hat{r}_{t+1}, \tag{A.4.11b}$$

$$E_t\hat{\lambda}_{t+1} - \hat{\lambda}_t - \xi_7\hat{x}_t = -\xi_7\hat{c}_t + E_t\hat{\pi}_{t+1}, \tag{A.4.11c}$$

$$E_t\hat{m}_{t+1} - \hat{x}_t = 0, \tag{A.4.11d}$$

$$\begin{aligned}
\xi_4 &:= \kappa\gamma(c/x)^\kappa(c/k),\\
\xi_5 &:= (1+\gamma(1+\kappa)(c/x)^\kappa)(c/k),\\
\xi_6 &:= 1-\beta a^{-\eta}(1-\delta),\\
\xi_7 &:= \frac{\kappa\gamma(1+\kappa)(c/x)^{1+\kappa}}{1-\gamma\kappa(c/x)^{1+\kappa}}.
\end{aligned}$$

The remaining fifth equation is the loglinear condition for the firms' optimal price:

$$\frac{(1-\varphi)(1-\varphi\beta a^{-\eta})}{\varphi}\hat{g}_t = -\beta a^{-\eta}E_t\hat{\pi}_{t+1} + \hat{\pi}_t. \tag{A.4.11e}$$

This looks nice and resembles a Phillips curve since it relates the current inflation rate to the expected future rate of inflation and a measure of labor market tension, which is here given by the deviation of marginal costs from their steady state level. It requires a substantial amount of algebra to get this relation and it is this task to which we turn next. Considering (A.4.6) we find:

$$\begin{aligned}
&(\widehat{P_{At}/P_t})\frac{1}{\vartheta}\frac{y}{j}\underbrace{\left(1+\varphi\beta a^{-\eta}+(\varphi\beta a^{-\eta})^2+\dots\right)}_{(1-\varphi\beta a^{-\eta})^{-1}}\\
&\quad+\frac{1}{\vartheta}\frac{y}{j}\sum_{\tau=t}^{\infty}(\varphi\beta a^{-\eta})^{\tau-t}E_t\left[(\widehat{\lambda_\tau/\lambda_t})+(\epsilon-1)(\widehat{P_\tau/P_t})+(\widehat{y_\tau/j_\tau})\right]\\
&=g\frac{y}{j}\sum_{\tau=t}^{\infty}(\varphi\beta a^{-\eta})^{\tau-t}E_t\left[(\widehat{\lambda_\tau/\lambda_t})+\epsilon(\widehat{P_\tau/P_t})+(\widehat{y_\tau/j_\tau})+\hat{g}_\tau\right].
\end{aligned}$$

Since $\vartheta g = 1$, $\widehat{\lambda_{\tau=t}/\lambda_t} = 0$, $\widehat{P_{\tau=t}/P_t} = 0$, and $\widehat{P_{At}/P_t} = [\varphi/(1-\varphi)]\hat{\pi}_t$ (see above), we can simplify this expression to

$$\frac{\varphi}{(1-\varphi)(1-\varphi\beta a^{-\eta})}\hat{\pi}_t = \sum_{\tau=t}^{\infty}(\varphi\beta a^{-\eta})^{\tau-t}E_t\left[\widehat{(P_\tau/P_t)} + \hat{g}_\tau\right]. \tag{A.4.12}$$

Next, we shift the time index one period into the future, multiply through by $\varphi\beta a^{-\eta}$, and compute the conditional expectation of the ensuing expression:

$$\begin{aligned}
&\left(\frac{\varphi}{1-\varphi}\right)\left(\frac{\varphi\beta a^{-\eta}}{1-\varphi\beta a^{-\eta}}\right)E_t\hat{\pi}_{t+1} \\
&= E_t\Bigg[\left(\varphi\beta a^{-\eta}\right)^2\widehat{\left(\frac{P_{t+2}}{P_{t+1}}\right)} + \left(\varphi\beta a^{-\eta}\right)^3\widehat{\left(\frac{P_{t+3}}{P_{t+1}}\right)} + \cdots + \varphi\beta a^{-\eta}\hat{g}_{t+1} \\
&\quad + (\varphi\beta a^{-\eta})^2\hat{g}_{t+2} + \ldots\Bigg].
\end{aligned}$$

We subtract this equation from (A.4.12) to arrive at:

$$\begin{aligned}
&\frac{\varphi}{(1-\varphi)(1-\varphi\beta a^{-\eta})}\left(\hat{\pi}_t - \varphi\beta a^{-\eta}E_t\hat{\pi}_{t+1}\right) \\
&= \hat{g}_t + E_t\Bigg[\varphi\beta a^{-\eta}\widehat{\left(\frac{P_{t+1}}{P_t}\right)} + \left(\varphi\beta a^{-\eta}\right)^2\left\{\widehat{\left(\frac{P_{t+2}}{P_t)}\right)} - \widehat{\left(\frac{P_{t+2}}{P_{t+1}}\right)}\right\} \\
&\quad + \left(\varphi\beta a^{-\eta}\right)^3\left\{\widehat{\left(\frac{P_{t+3}}{P_t}\right)} - \widehat{\left(\frac{P_{t+3}}{P_{t+1}}\right)}\right\} + \ldots\Bigg].
\end{aligned} \tag{A.4.13}$$

Since

$$\widehat{\left(\frac{P_\tau}{P_t}\right)} = \sum_{s=t+1}^{\tau}\hat{\pi}_s,$$

the terms in curly brackets reduce to $\hat{\pi}_{t+1}$ so that the sum in brackets equals

$$\hat{\pi}_{t+1}\underbrace{\left[\varphi\beta a^{-\eta} + (\varphi\beta a^{-\eta})^2 + \ldots\right]}_{(\varphi\beta a^{-\eta})/(1-\varphi\beta a^{-\eta})}.$$

Substituting these results back into (A.4.13) delivers equation (A.4.11e).

To determine the time path of investment, we start from

$$i_t = y_t - \left(1 + \gamma \left(\frac{c_t}{x_t}\right)^{\kappa}\right) c_t, \quad x_t \equiv m_{t+1}.$$

The loglinearized version of this equation is:

$$\hat{i}_t = \iota_1 \hat{y}_t - \iota_2 \hat{c}_t + \iota_3 \hat{x}_t,$$
$$\iota_1 := (y/i) = \frac{y/k}{a + \delta - 1}, \quad \iota_2 := \left(1 + (1+\kappa)\gamma \left(\frac{c}{x}\right)^{\kappa}\right) \frac{c}{i},$$
$$\iota_3 := \kappa\gamma \left(\frac{C}{\mu(M/P)}\right)^{\kappa} \frac{c}{i}, \quad \frac{c}{i} = \frac{y}{i} - 1.$$

Problems

2.1 Consider the deterministic linear quadratic optimal control problem of maximizing

$$\sum_{t=0}^{\infty} \beta^t \left[\mathbf{x}_t' Q \mathbf{x} + \mathbf{u}_t' R \mathbf{u}_t + 2\mathbf{u}_t' S \mathbf{x}_t\right]$$

subject to the linear law of motion

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t.$$

Adapt the steps followed in Section 2.1 and Appendix 3 to this problem and show that the optimal control as well as the matrix P are the solutions to equations (2.7) and (2.8), respectively.

2.2 Show that the linear quadratic problem with the current period return function

$$g(\mathbf{x}_t, \mathbf{u}_t, \mathbf{z}_t) := \mathbf{x}_t' A_{xx} \mathbf{x}_t + \mathbf{u}_t' A_{uu} \mathbf{u}_t + \mathbf{z}_t A_{zz} \mathbf{z}_t + 2\mathbf{u}_t' A_{ux} \mathbf{x}_t + 2\mathbf{u}_t' A_{uz} \mathbf{z}_t + 2\mathbf{x}_t A_{xz} \mathbf{z}_t$$

and the law of motion

$$\mathbf{x}_{t+1} = B_x \mathbf{x}_t + B_u \mathbf{u}_t + B_z \mathbf{z}_t$$

is a special case of the problem stated in equations (2.2) and (2.1). Toward that purpose define

$$\tilde{\mathbf{x}}_t = \begin{bmatrix} \mathbf{x}_t \\ \mathbf{z}_t \end{bmatrix}, \quad \tilde{\boldsymbol{\epsilon}}_t = \begin{bmatrix} 0_{n\times 1} \\ \boldsymbol{\epsilon}_t \end{bmatrix}$$

and show how the matrices A, B, Q, R, and S must be chosen so that both problems coincide.

2.3 Write a program (in your favorite language) that solves for the policy function of the simple deterministic Ramsey problem considered in Section 2.2.1. Use analytic first and second derivatives to form the quadratic approximation of the farmer's utility function.

2.4 Instead of the Calvo (1983) model, consider the following model of price setting introduced in Hairault and Portier (1995). Intermediate producers face convex costs of adjusting their price given by

$$PC_{jt} := (\psi/2)\left(\frac{P_{jt}}{P_{jt-1}} - \pi\right)^2.$$

Thus they solve the following problem:

$$\max \quad E_0 \sum_{t=0}^{\infty} \varrho_t \left[(P_{jt}/P_t)Y_{jt} - (W_t/P_t)N_{jt} - r_t K_{jt} - PC_{jt}\right],$$

s.t.

$$Y_{jt} = (P_{jt}/P_t)^{-\epsilon}(Y_t/J_t),$$
$$Y_{jt} = Z_t(A_t N_{jt})^{\alpha} K_{jt}^{1-\alpha} - F.$$

Calibrate the parameter ψ so that a one percent deviation of the firm's inflation factor P_{jt}/P_{jt-1} from average the average inflation factor entails costs of 0.01 percent of the firm's value added. Do you find more persistence of a money supply shock with this alternative specification of nominal rigidities? What happens, if you increase ψ?

2.5 In most OECD countries, wages and labor productivity are acyclic or even negatively correlated with output and working hours, while, in the stochastic Ramsey model, however, these correlations are positive and close to one (please compare table 2.1). One possible remedy for this shortcoming of the stochastic growth model is the introduction of a government spending shock. The following model is adapted from Baxter and King (1993) and Ambler and Paquet (1996).
Consider the stochastic growth model where the number of agents is normalized to one. Assume that utility is also a function of government consumption, where due to our normalization per capita government spending G_t is also equal to total government spending G_t. In particular, government consumption substitutes for private consumption C_t^p:

$$C_t = C_t^p + \vartheta G_t,$$

with $\vartheta < 1$ as some forms of government spending, for example military spending, do not provide utility for private consumption. The household maximizes her intertemporal utility:

$$\max_{C_0^p, N_0} \quad E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta} \right],$$
$$\beta \in (0,1), \eta \geq 0, \theta \geq 0, \eta > \theta/(1+\theta),$$

subject to the budget constraint

$$C_t^p + I_t^p = (1-\tau)(w_t N_t + r_t K_t^p) + Tr_t.$$

Both wage income $w_t N_t$ and interest income $r_t K_t$ are taxed at the constant rate τ. The household also receives lump-sum transfers Tr_t from the government. The private capital stock evolves according to:

$$K_{t+1}^p = (1-\delta)K_t^P + I_t^p,$$

where δ denotes the depreciation rate. Production is described by a Cobb-Douglas Production Function, $Y_t = Z_t N_t^{\alpha} K_t^{1-\alpha}$, where the productivity Z_t follows an AR(1) process, $Z_{t+1} = Z_t^{\varrho} e^{\epsilon_t}$, with $\epsilon_t \sim N(0, \sigma^2)$ and $\varrho = 0.90$ and $\sigma = 0.007$. Factors are rewarded by their marginal products. Government consumption $G_t = g_t \bar{G}$ follows a stochastic process:

$$\ln g_t = \rho_g \ln g_{t-1} + \epsilon_t^g,$$

with $\epsilon_t^g \sim N(0, \sigma_g^2)$ and $\rho_g = 0.95$ and $\sigma_g = 0.01$. In the steady state, government consumption is constant and equal to 20% of output, $\bar{G} = 0.2\bar{Y}$. In equilibrium, the government budget is balanced:

$$\tau(w_t N_t + r_t K_t^p) = G_t + Tr_t.$$

The model is calibrated as follows: $\beta = 0.99$, $\eta = 2.0$, $\psi = 0.5$, $\alpha = 0.6$, $\delta = 0.02$. θ and τ are chosen so that the steady state labor supply $\bar{N}$ and transfers $\overline{Tr}$ are equal to 0.30 and 0, respectively.

a) Compute the steady state.
b) Compute the log-linear solution. Simulate the model and assume that ϵ_t and ϵ_t^g are uncorrelated. What happens to the correlation of labor productivity and wages with output and employment?
c) Assume that transfers are zero, $Tr_t = 0$, and that the income tax τ_t always adjusts in order to balance the budget. How are your results affected?
d) Assume now that the government expenditures are split evenly on government consumption G_t and government investment I_t^G. Government capital K_t^G evolves accordingly

$$K_{t+1}^G = (1-\delta)K_t^G + I_t^G,$$

and production is now given by

$$Y_t = Z_t = Z_t N_t^{\alpha} K_t^{1-\gamma} \left(K_t^G\right)^{1-\alpha-\gamma}$$

with $\alpha = 0.6$ and $\gamma = 0.3$. Recompute the model.

2.6 In the previous problem, you have learned about the 'wealth effect' of government demand. An increase in government expenditures results in a reduction of transfers and, hence, wealth of the households is decreased. Consequently, the households increase their labor supply and both employment and output increase. In this problem, you will learn about the traditional Keynesian IS-LM effect. Expansionary fiscal policy increases aggregate demand and demand-constrained firms increases their output as prices are fixed in the short run. The model follows Linnemann and Schabert (2003).
Households maximize the expected value of a discounted stream of instantaneous utility:

$$\max_{C_0,N_0} \quad E_0\left[\sum_{t=0}^{\infty}\beta^t\frac{C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta}\right],$$
$$\beta\in(0,1),\eta\geq 0,\theta\geq 0,\eta>\theta/(1+\theta).$$

A role for money is introduced into the model with the help of a cash-in-advance constraint:

$$P_tC_t\leq M_t+P_tTr_t,$$

Nominal consumption purchases P_tC_t are constrained by nominal beginning-of period money balances M_t and nominal government transfers P_tTr_t.[23] The household holds two kinds of assets, nominal money M_t and nominal bonds, B_t. Bonds yield a gross nominal return R_t. In addition, agents receive income from labor, $P_tw_tN_t$, government transfers, P_tTr_t, and from firm profits, $\int_0^1\Omega_{it}\,di$. The budget constraint is given by:

$$M_{t+1}+B_{t+1}+P_tc_t=P_tw_tN_t+R_tB_t+M_t+P_tTr_t+\int_0^1\Omega_{it}di.$$

The number of firms i is one, $i\in(0,1)$. Firms are monopolistically competitive and set their prices in a staggered way as in the model of Section 2.4.3. Accordingly, profit maximization of the firms implies the New Keynesian Phillips curve:

$$\hat{\pi}_t=\psi\widehat{mc}_t+\beta E_t\left\{\hat{\pi}_{t+1}\right\},\quad \psi=(1-\varphi)(1-\beta\varphi)\varphi^{-1},$$

[23] Government transfers are included in this cash-in-advance specification in order to avoid the following: an expansionary monetary policy consisting in a rise of M_{t+1} already increases prices P_t due to the expected inflation effect. Accordingly, real money balances M_t/P_t fall and so does real consumption C_t if government transfers do not enter the cash-in-advance constraint. This, however, contradicts empirical evidence.

where mc_t denotes marginal costs (compare (A.4.11e)).
Firms produce with labor only:

$$y_{it} = N_{it}.$$

Cost minimization implies that the real wage is equal to marginal costs:

$$w_t = mc_t.$$

The government issues money and nominal riskless one-period bonds and spends its revenues on government spending, G_t, and lump-sum transfers:

$$P_t Tr_t + P_t G_t + M_t + R_t B_t = B_{t+1} + M_{t+1}.$$

Real government expenditures follow an AR(1)-process:

$$\ln G_t = \rho \ln G_{t-1} + (1-\rho) \ln G + \epsilon_t$$

with $\epsilon_t \sim N(0, \sigma^2)$ and $\rho = 0.90$ and $\sigma = 0.007$.
Monetary policy is characterized by a forward-looking interest-rate rule:

$$\hat{R}_{t+1} = \rho_\pi E - t\hat{\pi}_{t+1} + \rho_y E_t \hat{y}_{t+1}, \quad \rho_{pi} > 1.$$

The restriction ρ_π is imposed in order to ensure uniqueness of the equilibrium.

a) Compute the first-order conditions of the household.
b) Compute the stationary equilibrium that is characterized by a zero-supply of bonds, $B_t = 0$,[24] and $R > 1$ (in this case, the cash-in-advance constraint is always binding). Furthermore, in equilibrium, the aggregate resource constraint is given by $y_t = c_t + G_t$ and firms are identical, $y_{it} = y_t = N_t = N_{it}$. Define the equilibrium with the help of the stationary variables $\{\pi_t, w_t, m_t \equiv \frac{M_t}{P_{t-1}}, R_t, y_t, G_t\}$.
c) Compute the steady-state.
d) Calibrate the model as in the previous problem. In addition, set $\rho_\pi = 1.5$, $\rho_y \in \{0, 0.1, 0.5\}$, $\pi = 1$, and $\varphi = 0.75$.
e) Log-linearize the model and compute the dynamics. How does consumption react to an expansionary fiscal policy? Does it increase (as IS-LM implies) or decrease (due to the wealth effect)?
f) Assume now that the interest-rate rule is subject to an exogenous autocorrelated shock with autoregressive parameter $\rho_R \in \{0, 0.5\}$. How does a shock affect the economy?
g) Assume that monetary policy is described by a money-growth rule that is subject to an autoregressive shock. Recompute the model for an autoregressive parameter $\rho_\mu \in \{0, 0.5\}$ and compare the impulse responses to those implied by an interest-rate rule.

[24] Why can we set the nominal bonds supply equal to zero?

Chapter 3

Parameterized Expectations

Overview. We know from Chapter 1 that there are two ways to characterize the solution of a Ramsey problem or, more generally, of a recursive dynamic general equilibrium (DGE) model: (1) in terms of a policy function that relates the model's decision or control variables to the model's state variables or (2) in terms of a system of stochastic difference equations that determines the time paths of the model's endogenous variables. The method presented in this chapter rests on yet a third solution concept. In the rational expectations equilibrium of a recursive DGE model agents' conditional expectations are time invariant functions of the model's state variables. The parameterized expectations approach (PEA) applies methods from function approximation (see Section 8.2) to these unknown functions. In particular, it uses simple functions in place of the true but unknown expectations and employs Monte Carlo techniques to determine their parameters.

The PEA has several advantages vis-a-vis both the value function iteration approach and the extended path algorithm. In contrast to the former it does not suffer from the curse of dimensionality and, therefore, can be applied to models with many endogenous state variables. Unlike the latter, it deals easily with binding constraints. Our applications in Section 3.3 illustrate these issues.

We describe the PEA in two steps. (1) In the next section we look at the solution of a Ramsey problem from a different angle. Instead of focusing on agents' policy functions we consider their conditional expectations of future prices, quantities, and shocks. So far we have tacitly assumed that these expectations are consistent with the model. It is now the time to make this more obvious. The ensuing concept of a rational expectations equilibrium is yet

another way to describe the solution of a DGE model. Its feature is a time invariant (possibly vector-valued) function $\mathscr{E}$ used by agents to predict their future economic environment. An approximate solution is a simple function $\hat{\mathscr{E}}$, for example a finite degree polynomial, that approximates $\mathscr{E}$ sufficiently well. We will see that the definition of conditional expectations provides the clue to compute the parameters of $\hat{\mathscr{E}}$. This naturally implies a general framework for the PEA.

(2) Section 3.2 considers the single steps of this algorithm in more detail. Specifically, we will deal with two approaches to solve the fixed-point problem that defines the PEA solution. The first approach is iterative, the second solves a non-linear equations problem. Both approaches require good starting values and so we will consider this problem subsequently. The final problem that we consider is accuracy. Function approximation in general and the PEA in particular provide a number of ways to increase the accuracy of a given solution. Yet, since we do not know the true solution, we need measures of accuracy that help us to decide whether a given approximate solution is sufficiently close to the true but unknown function.

Having read this chapter you will have seen applications of almost every tool from the collection of numerical methods presented in Chapter 8.

3.1 Characterization of Approximate Solutions

This section provides a general description of the parameterized expectations approach (PEA). In the first subsection we use the stochastic growth model from Example 1.3.3 to illustrate the basic idea. The second subsection provides the general framework and the third subsection highlights the relation between the PEA and models of adaptive learning.

3.1.1 An Illustrative Example

The Model. Consider the stochastic Ramsey model from Example 1.3.3. We derive the first-order conditions that characterize its

solution in Section 1.3.2 (see (1.23)). We repeat it for your convenience:

$$K_{t+1} = Z_t K_t^{\alpha} + (1-\delta)K_t - C_t, \tag{3.1a}$$

$$C_t^{-\eta} = \beta E_t\left[C_{t+1}^{-\eta}\left(1-\delta+\alpha Z_{t+1}K_{t+1}^{\alpha-1}\right)\right]. \tag{3.1b}$$

Remember, C_t denotes consumption, K_t the stock of capital, and Z_t is the size of the shock to total factor productivity that evolves over time according to

$$Z_t = Z_{t-1}^{\varrho} e^{\epsilon_t}, \quad \epsilon_t \sim N(0, \sigma^2). \tag{3.1c}$$

Equation (3.1a) is the economy's resource constraint. Implicit in equation (3.1b) is the statement that the expected marginal rate of substitution between current and future consumption must equal the expected gross return on investment, i.e., one plus the marginal product of capital net of depreciation δ.

Conditional Expectations. We know from Chapter 1 that the solution to this set of equations can be written in terms of a time invariant policy function $K_{t+1} = g(K_t, Z_t)$. This implies that the conditional expectation on the rhs of (3.1b) is also a time invariant function $\mathscr{E}$ of the model's state variables K_t and Z_t. To see this, let

$$C(K_t, Z_t) := Z_t K_t^{\alpha} + (1-\delta)K_t - g(K_t, Z_t)$$

denote the solution for consumption given K_t and Z_t. Therefore, $C_{t+1} = C(g(K_t, Z_t), Z_{t+1})$. Using (3.1c) to replace Z_{t+1}, we may summarize the expression inside the expectations operator E_t in a function $\phi(K_t, Z_t, \epsilon_{t+1})$:

$$\begin{aligned}\phi(K_t, Z_t, \epsilon_{t+1}) :=& C(g(K_t, Z_t), Z_t^{\varrho} e^{\epsilon_{t+1}})^{-\eta} \\ &\times \left(1-\delta+\alpha Z_t^{\varrho} e^{\epsilon_{t+1}} g(K_t, Z_t)^{\alpha-1}\right).\end{aligned}$$

Since the innovations to (3.1c) are normal variates with density $\pi(\epsilon)$, we get $\mathscr{E}$ via integration of $\phi(\cdot)\pi(\epsilon)$:

$$\mathscr{E}(K_t, Z_t) := \int_{-\infty}^{\infty} \phi(K_t, Z_t, \epsilon)\pi(\epsilon)d\epsilon.$$

Approximation of $\mathscr{E}$. Suppose we knew $\mathscr{E}$. Then, given an arbitrary initial capital stock K_0 and an arbitrary initial level of the total factor productivity Z_0 we can compute the rhs of equation (3.1b) and solve the system of two equations (3.1) for K_1 and C_0. With K_1 at hand, Z_1 derived from Z_0, and a draw from the $N(0, \sigma^2)$-distribution, we can use $\mathscr{E}(K_1, Z_1)$ and (3.1) again to solve for (K_2, C_1). Repeating these steps over and over we can trace out an entire time path for the variables of our model economy.

As with the policy function g, there is in general no analytic solution for $\mathscr{E}$. The idea behind the PEA is to approximate the unknown function $\mathscr{E}$ by a simple function ψ, say. For instance, DEN HAAN and MARCET (1990) use $\psi(\gamma_1, \gamma_2, \gamma_3, K_t, Z_t) = \gamma_1 K_t^{\gamma_2} Z_t^{\gamma_3}$ to approximate the solution of the stochastic difference equation (3.1).

Given the choice of the approximating function, the remaining task is to fix its parameters. Remember the definition of conditional expectations: let y denote a random variable that we wish to forecast using observations on $(x_1, x_2, \ldots, x_n)$. We seek a function h that minimizes the expected mean quadratic error

$$E[(y - h(x_1, x_2, \ldots, x_n))^2].$$

The solution to this problem is the conditional expectation:[1]

$$E[y|(x_1, x_2, \ldots, x_n)] := \arg\min_h E[(y - h(x_1, x_2, \ldots, x_n))^2].$$

The parameter choice mimics this definition. We need some additional notation to describe this procedure. For simplicity, we stack the model's variables in the vector $\mathbf{u}_t := (C_t, K_t, K_{t+1}, Z_t)$ and collect the model's states in the vector $\mathbf{x}_t := (K_t, Z_t)$. We use $\psi(\boldsymbol{\gamma}, \mathbf{x}_t)$ to denote the function approximating $\mathscr{E}$ for a given p-vector of parameters $\boldsymbol{\gamma} := (\gamma_1, \gamma_2, \ldots, \gamma_p)$ and assume that a time path $\{\mathbf{u}_t\}_{t=0}^T$ of length $T + 1$ has been computed based on a given (K_0, Z_0) and T draws from the $N(0, \sigma^2)$-distribution. To emphasize the dependence on $\boldsymbol{\gamma}$ we write $\mathbf{u}_t(\boldsymbol{\gamma})$ and $\mathbf{x}_t(\boldsymbol{\gamma})$. Given this, let

[1] See, e.g., Sargent (1987), p. 224.

$$\phi(\mathbf{u}_{t+1}(\boldsymbol{\gamma})) := C_{t+1}^{-\eta}(1 - \delta + \alpha Z_{t+1} K_{t+1}^{\alpha-1})$$

denote the (ex post) rhs of equation (3.1b) associated with this time path so that $\phi(\mathbf{u}_{t+1}(\boldsymbol{\gamma})) - \psi(\boldsymbol{\gamma}, \mathbf{x}_t(\boldsymbol{\gamma}))$ is the time $t+1$ prediction error. Next, define the map $\Gamma : \mathbb{R}^p \to \mathbb{R}^p$ by

$$\Gamma(\boldsymbol{\gamma}) := \arg\min_{\boldsymbol{\xi}} \sum_{t=0}^{T-1} \left[\phi(\mathbf{u}_{t+1}(\boldsymbol{\gamma})) - \psi(\boldsymbol{\xi}, \mathbf{x}_t(\boldsymbol{\gamma}))\right]^2 .$$

Thus, $\Gamma(\boldsymbol{\gamma})$ is the parameter vector $\boldsymbol{\xi}$ that minimizes the sum of squared prediction errors associated with the time path that results from predictions of the rhs of equation (3.1b) using the function $\psi(\boldsymbol{\gamma}, \cdot)$. The fixed point $\boldsymbol{\gamma}_{p,T}$ of this mapping,

$$\boldsymbol{\gamma}_{p,T} = \Gamma(\boldsymbol{\gamma}_{p,T}),$$

is the approximate model solution. It depends on the length of the time path T and the function $\psi(\cdot)$.

3.1.2 A General Framework

This section describes the parameterized expectations approach in more general terms. Each of the representative agent models that you will encounter in this book fits the following framework.

Let $\mathbf{u}_t$ denote an l-dimensional vector that collects all of the model's variables. This vector belongs to some subset U of $\mathbb{R}^l$.[2] It is convenient to consider two further subsets of the variables in $\mathbf{u}_t$. The first subset, the m-vector $\mathbf{s}_t$, includes all exogenous stochastic processes with the Markov property that drive the model.[3] The second subset collects the model's state variables in the n-dimensional vector $\mathbf{x}_t \in X \subset \mathbb{R}^n$. Note that $\mathbf{x}_t$ includes

[2] Many of the variables of an economic model are restricted to belong to a given subinterval of the real line. For instance, output, consumption, investment, and the stock of capital cannot be negative. For this reason, we restrict $\mathbf{u}_t$ (as well as the vectors $\mathbf{s}_t$ and $\mathbf{x}_t$) to a subset U of $\mathbb{R}^l$. The set U implicitly imposes the restrictions on the values which the variables can take.

[3] See Section 9.2 on this property.

$\mathbf{s}_t$ and additional variables from the vector $\mathbf{u}_t$. The state variables summarize the information that is relevant to predict the future economic environment. In addition, there are two vector-valued functions that govern the model's dynamics. The function $\boldsymbol{\phi}$ with argument $\mathbf{u}_{t+1}$ maps U to a V, a subset of $\mathbb{R}^k$. The function $\mathbf{g}$ with arguments $E_t[\boldsymbol{\phi}(\mathbf{u}_{t+1})]$, $\mathbf{u}_t$, and $\mathbf{s}_t$ collects the model's Euler equations, definitions, resource constraints, and so forth. In the example from the previous subsection $\boldsymbol{\phi}$ equals the single-valued expression to the right of the conditional expectations operator E_t, $\mathbf{g}$ is given by equations (3.1), and V is the one-dimensional space of positive real numbers $\mathbb{R}_+$. Accordingly, the system of stochastic difference equations that drive the model can be written as follows:

$$\mathbf{g}(E_t[\phi(\mathbf{u}_{t+1})], \mathbf{u}_t, \mathbf{s}_t) = \mathbf{0} \text{ for all } t = 0, 1, \ldots, \infty. \tag{3.2}$$

Due to the recursive nature of the model (that allows for its solution in terms of a time invariant policy function) there is a time invariant conditional expectations function $\mathscr{E}$ given as the solution to

$$\mathscr{E} := \arg \min_{\mathbf{h}: X \to V} E[\boldsymbol{\phi}(\mathbf{u}_{t+1}) - \mathbf{h}(\mathbf{x}_t)]^2$$

that solves (3.2), i.e.,

$$\mathbf{g}(\mathscr{E}(\mathbf{x}_t), \mathbf{u}_t, \mathbf{s}_t) = \mathbf{0} \text{ for all } t = 0, 1, \ldots, \infty. \tag{3.3}$$

The parameterized expectations approach approximates this solution in the following steps:

Algorithm 3.1.1 (PEA)

Purpose: *Approximate the solution to* (3.3)

Steps:

Step 1: Choose a function $\boldsymbol{\psi}(\boldsymbol{\gamma}, \cdot) : X \to V$ *that depends on the vector of parameters* $\boldsymbol{\gamma} \in \mathbb{R}^p$.

Step 2: Draw a sequence of shocks $\{\mathbf{s}_t\}_{t=0}^T$.

Step 3: Iterate on

$$\mathbf{g}(\boldsymbol{\psi}(\boldsymbol{\gamma}, \mathbf{x}_t(\boldsymbol{\gamma})), \mathbf{u}_t(\boldsymbol{\gamma}), \mathbf{s}_t) = \mathbf{0}$$

to find the sequence $\{\mathbf{u}_t(\boldsymbol{\gamma}), \mathbf{s}_t\}_{t=0}^{T}$.

Step 4: Find the fixed point $\boldsymbol{\gamma}_{p,T} = \Gamma(\boldsymbol{\gamma}_{p,T})$ *of the map* Γ *defined by*

$$\Gamma(\boldsymbol{\gamma}) := \arg\min_{\boldsymbol{\xi}} \frac{1}{T} \sum_{t=0}^{T-1} \|\boldsymbol{\phi}(\mathbf{u}_{t+1}(\boldsymbol{\gamma})) - \boldsymbol{\psi}(\boldsymbol{\xi}, \mathbf{x}_t(\boldsymbol{\gamma}))\|^2,$$

where $\|\cdot\|$ *denotes the Euclidean norm.*

Step 5: Decide whether $\psi(\boldsymbol{\gamma}_{p,T}, \cdot)$ *is close to the true but unknown solution* $\mathscr{E}$. *If not, change either* T *or* $\boldsymbol{\psi}(\cdot)$ *and return to Step 1.*

Marcet and Marshall (1992, 1994) provide conditions on the functions $\mathbf{g}$, $\boldsymbol{\phi}$, $\boldsymbol{\psi}$ as well as on the process $\{\mathbf{s}_t\}_{t=0}^{\infty}$ that make the PEA a meaningful concept. Using a weaker definition of an approximate solution than that given in Step 4 they are able to show that the approximation can be made arbitrarily close to the true solution (3.3) by letting $T \to \infty$ and $p \to \infty$. Since we will be dealing with the computation of $\boldsymbol{\gamma}_{p,T}$ for a fixed T and p we can sidestep the involved technical details and can proceed with the definition given in Step 4.

3.1.3 Adaptive Learning

Models of Learning. There is an interesting relation between the approximate solution discussed in the previous section and attempts to formalize how agents learn about their environment.

The rational expectations equilibrium defined in (3.3) presupposes two requirements: individual rationality and mutual consistency of perceptions of the environment. The agents in the model use the true conditional expectations function for their forecasts. They have somehow solved estimation and inference problems that an econometrician must deal with. Models of learning depict economic agents as econometricians that use current

and past observations to estimate the parameters of the economy's law of motion. Since the actual law of motion depends upon the law perceived by agents, this is like chasing a moving target. Agents that act like econometricians are not as smart as those that populate the rational expectations equilibrium. For that reason, SARGENT (1993) refers to the former as agents with 'bounded rationality', a term coined by HERBERT SIMON (1957). Others use the term 'adaptive learning' to characterize this approach. EVANS and HONKAPOHJA (2001) provide an introduction into the related methods and present many applications. In the following paragraphs we will sketch an adaptive learning process whose stationary point is the approximate solution discussed in the previous subsection.

Recursive Least Squares. Assume you want to estimate the linear equation

$$y_i = \boldsymbol{\gamma}' \mathbf{x}_i + \epsilon_i, \quad i = 1, 2, ..., t,$$

where $\boldsymbol{\gamma}$ is a p-dimensional column vector[4] of parameters related to the observations of p independent variables collected in the column vector $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})'$. Put $\mathbf{y} = (y_1, y_2, \ldots, y_t)'$ and $X = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_t)'$. The well-known formula for the least squares estimator gives:[5]

$$\boldsymbol{\gamma}_t = (X'X)^{-1} X'\mathbf{y} = \left(\sum_{i=1}^{t} \mathbf{x}_i \mathbf{x}_i' \right)^{-1} \left(\sum_{i=1}^{t} \mathbf{x}_i y_i \right). \tag{3.4}$$

Suppose you have estimated $\boldsymbol{\gamma}$ from $t-1$ observations and now you are given one additional observation $(y_t, x_{t1}, x_{t2}, ..., x_{tp})$. There is a convenient formula that updates your estimate as follows:[6]

[4] The operator $'$ transposes rows and columns of a matrix. Thus, if $\mathbf{x}$ is a column vector $\mathbf{x}'$ is the corresponding row vector.

[5] This formula is derived in most introductory and advanced textbooks on econometrics. See, e.g., GREENE (2003), pp. 19ff. or JUDGE et al. (1988), p. 164ff.

[6] You can verify this formula by substituting the definitions of $\boldsymbol{\gamma}_t$ from (3.4) and of R_t into (3.5).

$$\begin{aligned}\boldsymbol{\gamma}_t &= \boldsymbol{\gamma}_{t-1} + \frac{1}{t} R_t^{-1} \mathbf{x}_t (y_t - \boldsymbol{\gamma}'_{t-1} \mathbf{x}_t), \\ R_t &= R_{t-1} + \frac{1}{t}(\mathbf{x}_t \mathbf{x}'_t - R_{t-1}),\end{aligned} \tag{3.5}$$

where

$$R_t := \frac{1}{t}\left(\sum_{i=1}^{t} \mathbf{x}_i \mathbf{x}'_i\right)$$

is a square matrix of dimension p and R_t^{-1} its inverse. The update of $\boldsymbol{\gamma}$ in the first line of (3.5) uses the most recent forecast error $y_t - \boldsymbol{\gamma}'_{t-1}\mathbf{x}_t$.

Learning Dynamics and the PEA. Suppose the agents in our model economy were not able to compute the true conditional expectations function $\mathscr{E}$. For ease of exposition assume that the range of $\boldsymbol{\phi} : X \to V$ is a subset of the real line (as in the Ramsey model of Section 3.1.1). Let $\psi(\boldsymbol{\gamma}_t, \cdot)$ denote the agents' forecast of $\phi(\cdot)$ using their most recent estimate of the parameter vector $\boldsymbol{\gamma}_t$. Since the entire history of the model economy depends upon the sequence of estimates $\{\boldsymbol{\gamma}_\tau\}_{\tau=0}^t$ the time sequence of the model's variables is different from the sequence $\{\mathbf{u}_\tau, \mathbf{s}_\tau\}_{\tau=0}^t$ obtained for a given and constant vector $\boldsymbol{\gamma}$. To emphasize this difference, we use $\tilde{\mathbf{u}}_t$ and $\tilde{\mathbf{x}}_t$ to denote the vector of variables and the vector of states, respectively, that are associated with a given sequence of estimates $\{\boldsymbol{\gamma}_\tau\}_{\tau=0}^t$. Assume agents use non-linear least squares to estimate $\boldsymbol{\gamma}$, i.e., at period t they choose $\boldsymbol{\gamma}_t$ to minimize

$$\frac{1}{t}\sum_{i=0}^{t-1}[\phi(\tilde{\mathbf{u}}_{i+1}) - \psi(\boldsymbol{\gamma}_t, \tilde{\mathbf{x}}_i)]^2.$$

A solution to this problem that fits into the framework of recursive least squares can be found as follows. Linearize $\psi(\boldsymbol{\gamma}, \cdot)$ at the previous estimate $\boldsymbol{\gamma}_{t-1}$:

$$\psi(\boldsymbol{\gamma}_t, \cdot) \simeq \psi(\boldsymbol{\gamma}_{t-1}, \cdot) + \nabla\psi(\boldsymbol{\gamma}_{t-1})(\boldsymbol{\gamma}_t - \boldsymbol{\gamma}_{t-1}),$$

where the symbol $\nabla\psi(\boldsymbol{\gamma}_{t-1})$ denotes the row vector of first derivatives of the function ψ evaluated at the point $\boldsymbol{\gamma}_{t-1}$. Put

$$\bar{y}_i := \phi(\tilde{\mathbf{u}}_{i+1}) - \psi(\boldsymbol{\gamma}_{t-1}, \tilde{\mathbf{x}}_i) + \nabla\psi(\boldsymbol{\gamma}_{t-1})\boldsymbol{\gamma}_{t-1},$$
$$\bar{\mathbf{x}}_i' := \nabla\psi(\boldsymbol{\gamma}_{t-1}),$$

and solve

$$\min_{\boldsymbol{\gamma}_t} \quad \frac{1}{t}\sum_{i=0}^{t-1}[\bar{y}_i - \boldsymbol{\gamma}_t'\bar{\mathbf{x}}_i]^2.$$

The solution is given by (3.4) with y_i and $\mathbf{x}_i$ replaced by $\bar{y}_i$ and $\bar{\mathbf{x}}_i$, respectively. Now, we are able to apply the recursive formula (3.5) to formulate the dynamics of our model under non-linear least squares learning:

$$\begin{aligned}
\boldsymbol{\gamma}_t &= \boldsymbol{\gamma}_{t-1} + \frac{1}{tR_t}\nabla\psi(\boldsymbol{\gamma}_{t-1})'(\phi(\tilde{\mathbf{u}}_t) - \psi(\boldsymbol{\gamma}_{t-1}, \tilde{\mathbf{x}}_{t-1})), \\
R_t &= R_{t-1} + \frac{1}{t}\left(\nabla\psi(\boldsymbol{\gamma}_{t-1})'\nabla\psi(\boldsymbol{\gamma}_{t-1}) - R_{t-1}\right), \\
0 &= g(\psi(\boldsymbol{\gamma}_t, \tilde{\mathbf{x}}_t), \tilde{\mathbf{u}}_t, \mathbf{s}_t).
\end{aligned} \tag{3.6}$$

Marcet and Marshall (1994) show, that the approximate solution defined in Step 4 of Algorithm 3.1.1 for $t \to \infty$, denoted by $\boldsymbol{\gamma}_p$, is a rest point of this process. Furthermore, if the absolute values of the eigenvalues of $\Gamma(\boldsymbol{\gamma})$ evaluated at $\boldsymbol{\gamma}_p$ are less than one, there is a neighborhood $\mathscr{N}(\boldsymbol{\gamma}_p)$ such that all $\boldsymbol{\gamma} \in \mathscr{N}(\boldsymbol{\gamma}_p)$ converge to this rest point.

3.2 Computation of the Approximate Solution

This section considers the single steps of the PEA Algorithm 3.1.1 in more detail. We start with the choice of the sample size T and the approximating function $\boldsymbol{\psi}$ in the next subsection.

3.2.1 Choice of T and $\boldsymbol{\psi}$

Sample Size. We note in Section 3.1.2 that the accuracy of the approximation increases with T. The underlying intuition is as follows. Suppose we would compute the time sequence of $\Omega_T := \{\mathbf{u}_t\}_{t=0}^T$ from the true function $\mathscr{E}$. In this case Ω_T is a sample

drawn from the ergodic distribution that is the solution of the system of stochastic difference equations defined in (3.3). As usual in sampling, the larger Ω_T is, the better does it represent the properties of the underlying distribution. In particular, those parts of the space U where the solution spends most of its time receive a high frequency count in Ω_T, whereas those parts of U which are visited very rarely appear hardly in Ω_T. As a consequence, the non-linear least squares estimator invoked in Step 4 of the PEA will be eager to keep the expectational errors small on those subsets of U, which we are most interested in. Of course, this property carries over to any sufficiently good approximation $\boldsymbol{\psi}$ of $\mathscr{E}$.

Applications of the PEA to solve the simple stochastic growth model from Example 1.3.3 therefore use large integer values of T. For instance, Duffy and McNelis (2001) use $T = 2,000$, den Haan and Marcet (1990) choose $T = 2,500$, the Fortran programs of Marcet and Lorenzoni (1999) allow for a maximum of 10,000 data points, and Christiano and Fisher (2000) even put $T = 100,000$. To eliminate the influence of the initial value $\mathbf{x}_0$ one can disregard the first 0.5 or 1.0 percent of the data points from the simulated time series and choose the parameter vector $\boldsymbol{\gamma}_{p,T}$ with respect to the remaining sample.

Function Approximation. More challenging is the choice of $\boldsymbol{\psi}$. Remember that this function is vector-valued, as it maps points from a subset of $\mathbb{R}^n$ to points in a subset of $\mathbb{R}^k$. If we think of the j-th coordinate of $\boldsymbol{\psi}$ as a map $\psi_j : X \subset \mathbb{R}^n \to \mathbb{R}$ we can reduce this problem to the simpler one of approximating a real-valued function. In Section 8.2 we present various ways to approximate a given function. In the following we use a complete set of polynomials of degree p in the n variables $(x_{1t}, \ldots, x_{nt})$ to build ψ_j. The members of the set are products of monomials $(x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n})$, where $\sum_{i=1}^n k_i = p$. These are easy to deal with in the PEA for the following reason: in many applications we do not know the boundaries of X in advance. However, the domain of orthogonal families of polynomials, as for instance the Hermite or Chebyshev polynomials, are certain compact intervals of the real line. When we use bases from members of these families, we must specify a

compact region X before we start the computations. This is not necessary in the case of monomials, since their domain is the entire real line. The drawback from using monomials that we will encounter later is the problem of multicollinearity. Very often, higher order terms of a variable x_i appear to be indistinguishable from one another on the computer. Hence, even if the theory tells us that we get a more accurate solution if we increase the degree of the polynomial we will not be able to achieve this on the computer.

Of course, the PEA is not restricted to a certain class of functions, and we encourage you to redo our examples using, e.g., Chebyshev polynomials or neural networks (see Section 8.2).

3.2.2 Iterative Computation of the Fixed Point

Convergence. There is a last step to be taken in order to implement the parameterized expectations approach: the actual computation of the parameters of the expectations function $\boldsymbol{\psi}(\boldsymbol{\gamma}, \cdot)$. Probably the most obvious thing to do is to iterate on the mapping Γ defined in Step 4 of Algorithm 3.1.1,

$$\boldsymbol{\gamma}_{s+1} = \Gamma(\boldsymbol{\gamma}_s), s = 0, 1, \ldots, \tag{3.7}$$

starting with an arbitrary $\boldsymbol{\gamma}_0$. However, since (3.7) is essentially a non-linear difference equation, this procedure need not converge, even if the fixed point exists. DEN HAAN and MARCET (1990) as well as MARCET and MARSHALL (1994) propose to iterate on

$$\boldsymbol{\gamma}_{s+1} = (1 - \lambda)\boldsymbol{\gamma}_s + \lambda\Gamma(\boldsymbol{\gamma}_s) \tag{3.8}$$

for some $\lambda \in (0, 1]$ to foster convergence. Indeed, if the related adaptive learning model is locally stable, there are starting values $\boldsymbol{\gamma}_0$ such that for a sufficiently small λ (3.8) will converge.

Non-Linear Least Squares. If we use this iterative procedure we have to solve

$$\min_{\boldsymbol{\xi}} \frac{1}{T} \sum_{t=0}^{T-1} \|\phi(\mathbf{u}_{t+1}(\boldsymbol{\gamma})) - \boldsymbol{\psi}(\boldsymbol{\xi}, \mathbf{x}_t(\boldsymbol{\gamma}))\|^2$$

at each step. This breaks down to solving k non-linear least squares problems. To see this let $\phi_j(\cdot)$ and $\psi_j(\cdot)$ denote the j-th component of $\boldsymbol{\phi}$ and $\boldsymbol{\psi}$, respectively, and partition the parameter vector $\boldsymbol{\gamma}$ so that $\boldsymbol{\gamma}_j := [\gamma_{1j}, \ldots, \gamma_{pj}]$, $j = 1, 2, \ldots, k$. With this notation the minimization problem can be rewritten as

$$\begin{aligned}
&\min_{\boldsymbol{\xi}} \frac{1}{T} \sum_{t=0}^{T-1} \sum_{j=1}^{k} [\phi_j(\mathbf{u}_{t+1}(\boldsymbol{\gamma})) - \psi_j(\boldsymbol{\xi}_j, \mathbf{x}_t(\boldsymbol{\gamma}))]^2, \\
&\equiv \min_{\boldsymbol{\xi}} \sum_{j=1}^{k} \frac{1}{T} \sum_{t=0}^{T-1} [\phi_j(\mathbf{u}_{t+1}(\boldsymbol{\gamma})) - \psi_j(\boldsymbol{\xi}_j, \mathbf{x}_t(\boldsymbol{\gamma}))]^2, \\
&\equiv \sum_{j=1}^{k} \min_{\boldsymbol{\xi}_j} \frac{1}{T} \sum_{t=0}^{T-1} [\phi_j(\mathbf{u}_{t+1}(\boldsymbol{\gamma})) - \psi_j(\boldsymbol{\xi}_j, \mathbf{x}_t(\boldsymbol{\gamma}))]^2.
\end{aligned}$$

In our applications we use the damped Gauss-Newton method explained in Section 8.6.2 to solve this problem. In the early stages of the iterations over (3.7) it is not necessary to compute the minimum with great accuracy. Thus one can make the algorithm faster by choosing very generous stopping criteria. For instance, the programs by Marcet and Lorenzoni (1999) bypass the convergence test (8.74) (see Section 8.4 on this criterion).

3.2.3 Direct Computation of the Fixed Point

In this subsection we consider the PEA as solution to a complicated system of $k \times p$ non-linear equations.

Remember the following notation used so far. $\psi_j(\boldsymbol{\gamma}_j, \mathbf{x}_t(\boldsymbol{\gamma}))$, $j = 1, 2, \ldots, k$ is the time t forecast of the j-th conditional expectation given the vector of states $\mathbf{x}_t$ and the parameter vector $\boldsymbol{\gamma} = [\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \ldots, \boldsymbol{\gamma}_k]$, where $\boldsymbol{\gamma}_j = [\gamma_{1j}, \gamma_{2j}, \ldots, \gamma_{pj}]$. Accordingly, $\phi_j(\mathbf{u}_{t+1}(\boldsymbol{\gamma}))$, is the time $t+1$ value of the expression to the right of the expectations operator that defines the j-th conditional expectation.

In this notation, the $k \times p$ first-order conditions for the minimization problem in Step 4 of Algorithm 3.1.1 may be written as follows:

$$0 = \frac{-2}{T}\sum_{t=0}^{T-1}[\phi_j(\mathbf{u}_{t+1}(\boldsymbol{\gamma})) - \psi_j(\boldsymbol{\xi}_j, \mathbf{x}_t(\boldsymbol{\gamma}))]\frac{\partial \psi_j}{\partial \xi_{ij}}(\boldsymbol{\xi}_j, \mathbf{x}_t(\boldsymbol{\gamma})),$$
$$\text{for all } i = 1, 2, \ldots, p, \text{ and } j = 1, 2, \ldots, k.$$

The iterative procedure of the previous subsection solves this problem for $\boldsymbol{\xi} = [\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \ldots, \boldsymbol{\xi}_k]$ given $\boldsymbol{\gamma}$ and stops if $\boldsymbol{\xi} \simeq \boldsymbol{\gamma}$. Here we replace $\boldsymbol{\xi}$ in the above system with $\boldsymbol{\gamma}$ to get:

$$0 = \underbrace{\frac{-2}{T}\sum_{t=0}^{T-1}[\phi_j(\mathbf{u}_{t+1}(\boldsymbol{\gamma})) - \psi_j(\boldsymbol{\gamma}_j, \mathbf{x}_t(\boldsymbol{\gamma}))]\frac{\partial \psi_j}{\partial \gamma_{ij}}(\boldsymbol{\gamma}_j, \mathbf{x}_t(\boldsymbol{\gamma}))}_{=:\varphi_{ij}(\boldsymbol{\gamma})}, \tag{3.9}$$
$$\text{for all } i = 1, 2, \ldots, p, \text{ and } j = 1, 2, \ldots, k.$$

The zero of this non-linear system of equations in $\boldsymbol{\gamma}$ is an equivalent characterization of the approximate model solution. Thus, instead of the iterative procedure outlined above, we can apply a non-linear equation solver to (3.9).

This sounds nice and easy! But think of the following issues. Routines that solve non-linear equations, as the modified Newton-Raphson method with line search, require a starting value. With an arbitrary $\boldsymbol{\gamma}_0$, however, it may not be possible to perform the simulations in Step 3 of Algorithm 3.1.1 at all. For instance, it may happen that at some t a non-negativity constraint implicit in the definition of $\mathbf{g}(\cdot)$ is violated so that it is impossible to compute $\{\mathbf{u}_t\}_{t=0}^T$. Even if this does not happen at the given $\boldsymbol{\gamma}_0$ the algorithm may want to try a vector where it is not possible to simulate the model for all T. For this reason any procedure that performs the simulation step must return an error flag that signals the calling program to stop. Otherwise your program will crash because of overflows, underflows or other run-time errors arising from undefined numerical operations. By the same token, the procedure that computes the rhs of (3.9) must return an error flag to the non-linear equations solver telling it to stop or to look for a different $\boldsymbol{\gamma}$ if it is not possible to evaluate all the $\varphi_{ij}(\boldsymbol{\gamma})$. Yet, standard software usually assumes that it is possible to evaluate a given non-linear system everywhere and there is no way to tell the

program to do otherwise. So, unless you write your own non-linear equations solver (or trust our routines) you are bound to find very good starting values. It is this issue that we turn to next.

3.2.4 Starting Points

Good starting values are essential to both the iterative and the direct approach to locate the PEA solution. The iterations over (3.8) may not converge if the initial point is outside of the basin of attraction of the respective learning algorithm, and non-linear equations solvers easily get stuck if the simulation step fails. There are several ways to handle this problem.

Homotopy. In mathematics two vector-valued functions $\mathbf{f} : X \to Y$ and $\mathbf{g} : X \to Y$ are said to be homotopic if $\mathbf{f}$ can be continuously deformed into $\mathbf{g}$. A function $\mathbf{h}(\mathbf{x}, s)$ that performs this task, i.e., that equals $\mathbf{f}$ for $s = 0$ and $\mathbf{g}$ for $s = 1$, is called a homotopy function. For instance,

$$\mathbf{h}(\mathbf{x}, s) := (1 - s)\mathbf{f}(\mathbf{x}) + s\mathbf{g}(\mathbf{x}) \tag{3.10}$$

is a homotopy function.

Suppose we want to solve $\mathbf{g}(\mathbf{x}) = \mathbf{0}$ and know the solution $\mathbf{x}_0$ of $\mathbf{f}(\mathbf{x}) = \mathbf{0}$. The idea behind homotopy methods is to construct a path in $X \times \mathbb{R}$ that takes us from the known solution to the solution of the problem of interest. Simple continuation methods use the linear homotopy (3.10), form an increasing sequence $0 < s_1 < s_2 < \cdots < 1$ and solve the related sequence of problems $\mathbf{h}(\mathbf{x}, s_i) = \mathbf{0}$. If the homotopy path in $X \times \mathbb{R}$ has peaks and troughs along the s dimension, simple continuation methods can fail. More advanced methods construct the homotopy path by solving a related system of differential equations.[7]

As regards DGE models the problem is to construct a simple model whose solution is either known or has been found in previous work and to move gradually from this model to the model of interest. This may be simple, as it is in the stochastic growth model, where an analytic solution exists for log-preferences and

[7] See, e.g., Judd (1998), pp. 176ff.

full depreciation of capital (see Example 1.3.2). As you will see in Section 3.3.1, in this case we can also derive an analytic expression for the conditional expectations function $\mathscr{E}(K_t, Z_t)$. We can then use small steps to move from $\delta = 1$ and $\eta = 1$ to a version of this model where the rate of capital depreciation δ is in the range of empirical estimates and where the elasticity of the marginal utility of consumption η is different from one. However, if we think of a model with capital and real balances of money as a second asset, it is less obvious from where to start. Moreover, if the model of interest departs significantly from the simple stochastic growth model it may be very cumbersome to trace out a sequence of more and more complicated models. For this reason, we consider search methods that are easy to implement for any kind of model and that have been found effective in quite different areas, such as automatic programming, machine learning, game theory, and numerical optimization.

Genetic Algorithms. In Section 8.6.4 we introduce genetic algorithms as a tool to minimize a given function. Here, our problem is to find the zeros of a set of non-linear functions $f^1(\mathbf{x}), f^2(\mathbf{x}), \ldots, f^n(\mathbf{x})$. But the solution to this problem is also a minimum of

$$g(\mathbf{x}) := \sum_{i=1}^{n} (f^i(\mathbf{x}))^2.$$

Though the converse is not true, a solution to this minimization problem might be a good starting point for a non-linear equations solver.

Using the Log-linear Model. In Section 2.3 we consider the log-linear approximation method. Very often it is possible to obtain the solution from this method with little additional effort. The extra work to be done is to obtain the log-linearized equations of the model. One can then use the linear policy functions to trace out a path for the vector $\mathbf{u}_t$ and solve the non-linear regression problem

$$\min_{\boldsymbol{\gamma}_0} \quad \frac{1}{T} \sum_{t=0}^{T-1} \left[\boldsymbol{\phi}(\mathbf{u}_{t+1}) - \boldsymbol{\psi}(\boldsymbol{\gamma}_0, \mathbf{x}_t) \right]^2 .$$

At this point one can also apply econometric tests to check whether the chosen degree of $\boldsymbol{\psi}$ is appropriate. For instance, if the t-ratio of a regressor is smaller than unity, one might exclude it from the regression.

3.2.5 Accuracy of Solutions

The last step of Algorithm 3.1.1 is to check whether the solution obtained in Step 4 is sufficiently close to the true but unknown solution $\mathscr{E}$. To guide this decision several measures of accuracy of the resulting approximate solution have been proposed in the literature.[8] Since many of them are related to the simple stochastic growth model, we will consider only those tests that are not tied to a specific model.

A simple and easy to compute measure is the sum of squared errors computed at the final solution $\boldsymbol{\gamma}_{p,T}$:

$$S(\boldsymbol{\gamma}_{p,T}) = \frac{1}{T}\sum_{t=0}^{T-1} \|\boldsymbol{\phi}(\mathbf{u}_{t+1}(\boldsymbol{\gamma}_{p,T})) - \boldsymbol{\psi}(\boldsymbol{\gamma}_{p,T}, \mathbf{x}_t(\boldsymbol{\gamma}_{p,T}))\|^2. \quad (3.11)$$

A smaller value of $S(\cdot)$ indicates a better approximation.

Den Haan and Marcet (1994) develop a formal test, which is based on the following property of a rational expectations equilibrium. The forecast errors of agents must be uncorrelated with past observations of the economic environment. Otherwise agents could use this correlation to obtain better forecasts. To test this proposition the authors consider the ex post residuals from the model's Euler equations. For instance, in the stochastic growth of Section 3.1.1 the expression

$$e_t = \beta C_{t+1}^{-\eta}\left(1 - \delta + \alpha Z_{t+1} K_{t+1}^{\alpha-1}\right) - C_t^{-\eta}$$

measures the extend to which the Euler equation (3.1b) is violated after the date $t+1$ variables have been revealed to agents. Den Haan and Marcet (1994) propose to regress e_t on lagged values of the model's variables and test whether the coefficients in

[8] See Taylor and Uhlig (1990) for an overview.

this regression are all equal to zero. We provide the details of this test in Section 9.3. The resulting DM-statistic is asymptotically distributed as a χ^2-variable. The authors propose to run a large number of simulations of the model and record the simulations where the DM-statistic is either smaller than the 2.5-percent critical value or larger than the 97.5-percent critical value. From a good solution we expect that about 5 percent of the simulations fall into these two regions.

3.3 Applications

3.3.1 Stochastic Growth with Non-negative Investment

The Model. Our first example application of the parameterized expectations approach (PEA) concerns the stochastic growth model. Since the PEA can easily handle constraints we consider a variant of this model, where it is impossible to convert capital goods back into consumption goods so that gross investment is bounded from below by zero. We present this model in Example 3.3.1.

Example 3.3.1

The representative agent in the economy solves the decision problem:

$$\max_{C_0} \quad E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\eta}}{1-\eta} \right], \quad \eta > 0,\ \beta \in (0,1),$$

s.t.

$$\left.\begin{array}{rcl} K_{t+1} + C_t & \leq & Z_t K_t^{\alpha} + (1-\delta)K_t,\ \alpha \in (0,1), \\ Z_t & = & Z_{t-1}^{\varrho} e^{\epsilon_t},\ \varrho \in (0,1),\ \epsilon_t \sim N(0,\sigma^2), \\ 0 & \leq & C_t, \\ 0 & \leq & K_{t+1} - (1-\delta)K_t, \end{array}\right\} \ t = 0, 1, \ldots$$

$K_0,\ Z_0$ given.

Here t is the time index, C denotes consumption, K is the stock of capital, and Z is the level of productivity.

First-Order Conditions. Using the techniques developed in Chapter 1, the first-order conditions for this problem are:

$$0 = C_t^{-\eta} - \mu_t - \beta E_t\big[C_{t+1}^{-\eta}(1 - \delta + \alpha Z_{t+1} K_{t+1}^{\alpha-1}) - \mu_{t+1}(1-\delta)\big], \tag{3.13a}$$

$$0 = Z_t K_t^{\alpha} + (1-\delta)K_t - C_t - K_{t+1}, \tag{3.13b}$$

$$0 = \mu_t[K_{t+1} - (1-\delta)K_t], \tag{3.13c}$$

$$0 \leq \mu_t, \tag{3.13d}$$

$$0 \leq K_{t+1} - (1-\delta)K_t. \tag{3.13e}$$

Here μ_t is the Lagrange multiplier related to the non-negativity constraint on investment

$$I_t := K_{t+1} - (1-\delta)K_t.$$

Lines (3.13c) to (3.13e) are the Kuhn-Tucker conditions associated with this constraint: either the constraint does not bind, in which case $\mu_t = 0$ (from (3.13c)) or gross investment is zero, in which case $\mu_t \geq 0$. Equation (3.13a) is the model's Euler equation and equation (3.13b) the economy's resource restriction.

Implementation. We already know from Example 1.3.2 that this model has an analytic solution for the policy function if $\eta = \delta = 1$, which is given by

$$K_{t+1} = \alpha\beta Z_t K_t^{\alpha}.$$

Since, in this case,

$$C_t = (1 - \alpha\beta) Z_t K_t^{\alpha},$$

the non-negativity constraint never binds, irrespective of the size of the productivity shock. Therefore, we can evaluate the rhs of (3.13a) to find the analytic solution for the conditional expectations function $\mathscr{E}(K_t, Z_t)$. We ask you to perform this task in Problem 2.1. Here, we just report the result:

$$\mathscr{E}(K_t, Z_t) = \frac{1}{(1-\alpha\beta)\beta} K_t^{-\alpha} Z_t^{-1}.$$

We use this information to approximate $\mathscr{E}$ in the general case $\eta > 0$ and $\delta \in [0,1]$ by an exponential polynomial in (K_t, Z_t) and compute solutions for the case of a first and a second degree complete polynomial:

$$\psi(\boldsymbol{\gamma}, K_t, Z_t) := \exp\bigl(\gamma_1 + \gamma_2 \ln K_t + \gamma_3 \ln Z_t\bigr),$$
$$\psi(\boldsymbol{\gamma}, K_t, Z_t) := \exp\bigl(\gamma_1 + \gamma_2 \ln K_t + \gamma_3 \ln Z_t + \gamma_4 (\ln K_t)^2 + \gamma_5 \ln K_t \ln Z_t + \gamma_6 (\ln Z_t)^2\bigr).$$

The Kuhn-Tucker conditions in the simulation are implemented as follows: given (K_t, Z_t) we first solve for

$$C_t = (\beta \psi(\boldsymbol{\gamma}, K_t, Z_t))^{-1/\eta},$$

and compute

$$K_{t+1} = Z_t K_t^{\alpha} + (1-\delta) K_t - C_t.$$

We then test the non-negativity constraint. If

$$K_{t+1} - (1-\delta) K_t < 0,$$

we set

$$\bar{C}_t = Z_t K_t^{\alpha} < C_t,$$

and

$$K_{t+1} = (1-\delta) K_t.$$

The Lagrange multiplier μ_t is found by solving

$$\mu_t = \bar{C}_t^{-\eta} - \beta \psi(\boldsymbol{\gamma}, K_t, Z_t).$$

This solution is always positive, if $C_t > \bar{C}_t$, i.e., if the constraint binds. Indeed, setting up the model with a non-negativity constraint is computationally easier than without this restriction. Otherwise we would have to check for $K_{t+1} < 0$ and terminate the computations with the given vector of parameters $\boldsymbol{\gamma}$. This introduces a discontinuity that must be handled explicitly in the non-linear equations solver. Nevertheless, it may happen that this routine tries a vector $\boldsymbol{\gamma}$ for which it is not possible to evaluate $\psi(\cdot)$, either because of an over- or an underflow. So the program must take this into account (see below).

Programs. The Fortran 95 routine `Ramsey3c1.for` implements fixed point iterations along with homotopy to find the approximate PEA solution. If you have access to a Fortran 95 compiler and the IMSL library you can build on our source files and assemble your own executable. Otherwise you must use the executable `Ramsey3c1.exe`. You can supply your own parameters in the file `Parameters.txt`. The program is not limited to a first or second degree polynomial. But you will see that the program is hardly able to find a solution for third and higher degree polynomials. The reason is that some of the higher order terms are almost collinear to others so that the Gauss-Newton step cannot be solved. If you desire higher precision, you must do some programming of your own and change `Ramsey3c1.for`. Gradually try higher order terms and keep those that work. In the file `Parameters.txt` you can give starting and ending values of η and δ along with step sizes for both parameters that determine the homotopy steps. Also, you can input your own starting values for the iterations over $\boldsymbol{\gamma}$ in the file `ParIni.txt`. We provide more details on how to use `Ramsey3c1.exe` in the file `Readme.txt`.

Our experiments with this program show that one can usually find the solution in one step starting at the vector that corresponds to the analytical solution of the model with loglinear preferences and full depreciation. This is owed to the constraint on investment. Without this piece of code, it is impossible to find the solution without using homotopy.

The Fortran 95 routine `Ramsey3c2.for` and the related source files implement the PEA solution by solving the system of nonlinear equations (3.9). Our non-linear equations solver is a variant of Algorithm 8.5.1 extended by the line search method of Algorithm 8.5.2. Since we do not know any bounds for the parameter vector $\boldsymbol{\gamma}$, we can not apply Algorithm 8.5.1 directly. Nevertheless, there are bounds for $\boldsymbol{\gamma}$. For instance, when we use double precision variables, the domain of the argument of the exponential function is bounded from above by about 710. To prevent the program from crashing, the subroutine that evaluates $\psi(\boldsymbol{\gamma}, \cdot)$ sets an error flag, if the argument of the exponential function exceeds this upper bound. The Newton step that caused that error

is then scaled down until the error flag clears. Despite this the algorithm may fail to find a solution due to roundoff errors that accumulate during the computation of the long sequences for the model's variables. As a consequence, the numerical approximation of the Jacobian may be rather imprecise so that the downhill movement ends before the zero is found. To reduce this problem, we approximate the Jacobian not by forward differences but by central differences that are more accurate.[9] We employ one of the genetic algorithms of Section 8.6.4 to find good starting values for the non-linear equation solver. If you use the executable `Ramsey3c2.exe` you can input the model's parameters in the file `Parameters.txt`. For more details on how to use this program see the `Readme.txt`.

Results. Table 3.1 presents statistics from four different simulations. Throughout, we used 30,000 observations to compute the parameters of the expectations function. The model's parameters are the same as those given in Table 1.1. Thus, the elasticity of production with respect to capital, α, equals 0.27, the discount factor β is 0.994, the elasticity of marginal utility of consumption η is 2.0, the rate of capital depreciation δ is 0.011, and the autocorrelation parameter of the productivity shock ϱ equals 0.90.

When we also use the same standard deviation of the innovations as in Table 1.1, $\sigma = 0.0072$, the non-negativity constraint never binds. The first two rows of Table 3.1 show that the second degree polynomial approximation provides a better approximation of the solution: it provides a slightly smaller sum of squared errors. More obviously, however, the number of DM-statistics (out of 500) below (above) the 2.5% (97.5%) percentile is closer to the theoretical 5%. The test rests on a regression of the Euler equation residual on a constant, five lags of consumption, and five lags of the productivity shock. Each of the 500 regressions has 3000 observations. This result is well in line with the findings of Den Haan and Marcet (1994), who conclude that the second degree

[9] See Dennis and Schnabel (1983), p. 77 ff. for a comparison of the two methods. Our implementation is the Fortran 95 subroutine `CDJac` in the file `MNR.for`.

Table 3.1

Simulation	$S(\boldsymbol{\gamma})$	σ_C	DM-Stat	
$\sigma = 0.0072$				
d=1	$1.366262E-7$	0.020729	2.0	4.6
d=2	$1.364703E-7$	0.020738	2.0	2.6
$\sigma = 0.05$				
d=1	$6.646593E-6$	0.144870	2.4	4.4
d=2	$6.575600E-6$	0.144750	1.2	3.6

Notes: σ is the standard deviation of the innovations of the productivity shock, d denotes the degree of the polynomial used to approximate $\psi(\cdot)$, $S(\boldsymbol{\gamma})$ is the sum of squared prediction errors at the final solution (see (3.11)), σ_C is the standard deviation of consumption, DM-Stat: the percentage of simulations (out of 500) whose DM-statistic is below (above) the 2.5 (97.5) percentile of the $\chi^2(11)$ distribution.

polynomial provides a good approximation of the solution of the stochastic growth model.[10]

In order to find binding constraints, we need larger productivity shocks. For $\sigma = 0.05$, we find that the non-negativity constraint binds in about 2 percent of the periods. The intuition should be clear: when total factor productivity Z_t is unexpectedly bad, the household would like to consume part of his capital stock to smooth his consumption stream. The overall picture remains the same and confirms the theory: the second degree polynomial gives a better approximation than the first degree polynomial.

With respect to run time, fixed point iterations outperform the direct solution method. On a machine with 1.6 GHz, Pentium 4 processor, `Ramsey3c1.exe` needs about 11 minutes to compute the solution for the first degree polynomial and about 30 minutes to find the solution for the second degree polynomial. In both runs we used $\lambda = 0.3$ in equation (3.8). The direct method was not able to find the solution using the parameters of the analytic solution in the case of log preferences and full depreciation of capital as starting values. To find acceptable starting values, the

[10] These authors consider the stochastic growth model without the non-negativity constraint on investment.

genetic algorithm must compute the long time sequences for each chromosome in each generation. Therefore, it takes much longer to find a solution. In the case of a first degree polynomial, we used a population size of 76 and stopped after generation 100. Starting from the best chromosome of generation 100 the non-linear equations solver refined the solution in a few steps (as in the fixed point iterations we required four good digits for each element of the parameter vector). Run time was about 17 minutes. In the case of a second degree polynomial (where imprecision in the computation of the Jacobian increases), we often needed to run the genetic algorithm more than once to get starting values from which the non-linear equations solver was able to find the final solution. In these cases the program ran for one to three hours before it found the solution. When the genetic algorithm returned a good starting point after the first run, it took no more time than fixed point iterations.

This sounds like bad news for the direct method. But note, the simple stochastic growth model with non-negative investment may be too simple a model to derive sound conclusions on the relative competitiveness of both methods. In the next section, where we report on the PEA solution of Example 1.4.1, this will become obvious.

3.3.2 The Benchmark Model

The Model. In Chapter 1 we solve the benchmark model by iterations over the value function and by the extended path method. Section 2.4.1 presents the results from the linear quadratic and the loglinear methods. Here we apply the PEA to the benchmark model. For your convenience, we present this model here again as Example 3.3.2.

Example 3.3.2

The representative household solves:

$$\max_{C_0,N_0} \quad E_0\left[\sum_{t=0}^{\infty}\beta^t\frac{C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}}{1-\eta}\right],$$
$$\beta\in(0,1),\theta\geq 0,\eta>\theta/(1+\theta),$$

s.t.

$$\left.\begin{array}{rcl} K_{t+1}+C_t & \leq & Z_t(A_tN_t)^{1-\alpha}K_t^{\alpha}+(1-\delta)K_t,\ \alpha\in(0,1), \\ A_{t+1} & = & aA_t,\ a\geq 1, \\ Z_{t+1} & = & Z_t^{\varrho}e^{\epsilon_t},\ \varrho\in(0,1),\ \epsilon_t\sim N(0,\sigma^2), \\ 0 & \leq & C_t, \\ 0 & \leq & K_{t+1}, \end{array}\right\}\ \forall t,$$

$K_0,\ Z_0$ given,

where t is the time index, C is consumption, N is labor supply, K is the stock (as well as the flow) of capital (services), A is the level of labor augmenting technical progress that increases deterministically at rate $a-1$, and Z is the level of total factor productivity. ———

First-Order Conditions. Our starting point are the first-order conditions as given in Section 2.4.1 and repeated in the following:

$$\lambda_t = c_t^{-\eta}(1-N_t)^{\theta(1-\eta)}, \tag{3.15a}$$

$$0 = \theta c_t^{1-\eta}(1-N_t)^{\theta(1-\eta)-1} - (1-\alpha)\lambda_t Z_t N_t^{-\alpha}k_t^{\alpha}, \tag{3.15b}$$

$$\lambda_t = \beta a^{-\eta}E_t\lambda_{t+1}(1-\delta+\alpha Z_{t+1}N_{t+1}^{1-\alpha}k_{t+1}^{\alpha-1}), \tag{3.15c}$$

$$ak_{t+1} = Z_tN_t^{1-\alpha}k_t^{\alpha}+(1-\delta)k_t-c_t. \tag{3.15d}$$

Remember the following definitions: $c_t := C_t/A_t$, $k_t := K_t/A_t$, and $\lambda_t = A_t^{\eta}\Lambda_t$, where Λ_t is the Lagrange multiplier of the budget constraint, which equals the marginal utility of consumption (see (3.15a)). The second line in (3.15) is the optimality condition with respect to labor supply, the third line is the Euler equation for the capital stock, and the fourth line is the budget constraint. We can eliminate consumption from the first two equations of (3.15). The result is an implicit equation for N_t:

$$((1-\alpha)/\theta)Z_tN_t^{-\alpha}k_t^{\alpha} = \lambda_t^{-1/\eta}(1-N_t)^{[\theta(1-\eta)/\eta]-1}. \tag{3.16}$$

We parameterize the conditional expectation on the rhs of (3.15c) by a second degree exponential polynomial in k_t and Z_t:

$$\psi(\boldsymbol{\gamma},k_t,Z_t) := \exp\big(\gamma_1+\gamma_2\ln k_t+\gamma_3\ln Z_t + \gamma_4(\ln k_t)^2+\gamma_5\ln k_t\ln Z_t+\gamma_6(\ln Z_t)^2\big).$$

Therefore, we can proceed as follows: given k_t and Z_t we get λ_t from $\lambda_t = \beta a^{-\eta}\psi(\boldsymbol{\gamma}, k_t, Z_t)$ and use this result to solve (3.16) for N_t.[11] Given λ_t and N_t we can solve (3.15a) for consumption c_t. Finally, we use the budget constraint (3.15d) to compute k_{t+1}.

Implementation. By now, you should be able to demonstrate that this model has an analytic solution in the case of logarithmic preferences and full depreciation of capital. As a consequence, there is also an analytic solution to the conditional expectations function on the rhs of equation (3.15c). Therefore, one can start from this vector using fixed point iterations. Note, however, that this model has no lower bound on investment that prevents the capital stock from becoming negative. Therefore, it requires small homotopy steps to approach the solution. We found that it is less cumbersome to use a genetic algorithm together with our nonlinear equations solver to find the solution.

It does not need much programming to modify the file used to solve the example of the previous subsection. The main changes are confined to the subroutine that returns the lhs of the system of equations whose zero we want to find. In addition, we must provide a piece of code that solves (3.16). You will find the details in the file `Ramsey4c.for`. The executable file `Ramsey4c.exe` uses the parameters in the file `Parameters.txt` to solve the model.

Results. We solve the model for the same set of parameters used in Chapter 1 (see Table 1.1) using 30,500 observations. To reduce the influence of the initial value of the capital stock (we equate k_0 with the stationary solution for the capital stock in the deterministic counterpart of the model) on the solution, we discard the first 500 observations. Depending on the details of the genetic algorithm it takes about 45 minutes to several hours on a 2.4 GHz Pentium 4 CPU to find the solution. Table 3.2 presents the second moments from the HP-filtered time series. As previously, the statistics are averages out of 500 simulations of the model. There

[11] As an exercise, you can demonstrate that (3.16) has a unique solution $N(k_t, Z_t, \lambda_t)$ for each $(k_t, Z_t, \lambda_t) > 0$. To find this solution, we employ Algorithm 8.5.1.

Table 3.2

Variable	s_x	r_{xy}	r_x
Output	1.43	1.00	0.63
Investment	6.15	1.00	0.63
Consumption	0.55	0.99	0.65
Hours	0.77	1.00	0.63
Real Wage	0.67	0.99	0.64

Notes: s_x:=standard deviation of HP-filtered simulated series of variable x, r_{xy}:=cross correlation of variable x with output, r_x:=first order autocorrelation of variable x.

is no noteworthy difference from the results obtained from the methods presented in Chapters 1 and 2 (see Tables 1.2 and 2.1).

The DM-test indicates a good solution: from 500 simulations with 3,000 observations each we obtain the DM-statistic by regressing the Euler equation residual on a constant, 5 lags of consumption, 5 lags of working hours, and 5 lags of the productivity shock. 5 percent of the simulations are either below the 2.5 percentile or above the 97.5 percentile of the $\chi^2(16)$-distribution.

3.3.3 Limited Participation in Financial Markets

Motivation. In the textbook IS-LM model an expansionary monetary shock lowers the nominal interest rate. Since inflationary expectations do not adjust immediately the real interest rate also declines. This spurs investment expenditures, which in turn raise aggregate spending. Given a sufficiently elastic short run supply function output and employment increase. This story is in line with the empirical evidence provided by vector autoregressions.[12] Yet, most monetary DGE models do not reproduce this liquidity effect. Consider, for instance, the model presented in Section 2.4.3. In this model there is only an anticipated inflation effect on the nominal interest rate: when agents learn about a temporarily

[12] See, e.g, Christiano, Eichenbaum, and Evans (1999).

high money growth rate, they expect a rise of future inflation and demand a higher nominal interest rate.

In this section we present a model of a monetary economy that is able to account for both the liquidity and the inflationary expectations effect.[13] The model includes a rudimentary banking sector. Households face a cash-in-advance constraint and can lend part of their financial wealth M_t to the banking sector at the gross nominal interest q_t (one plus the nominal interest rate). The firms in this model pay wages to the household sector before they sell their output. To finance their wage bill they borrow money from the banking sector. The government injects money into the economy via the banking sector. The crucial assumption is that banks receive the monetary transfer after households have decided about the volume of their banking deposits. Given the additional money, banks lower the nominal interest rate to increase their loans to firms. At the reduced credit costs firms hire more labor and increase production. The fact that households cannot trade on the market for deposits after the monetary shock has been observed has given the model its name: limited participation model.

The Banking Sector. At the beginning of period t banks receive deposits of size B_t from households. Government transfers amount to $M_{t+1} - M_t$, where M_t are beginning-of-period money balances. Banks, thus, are able to lend $B_t + (M_{t+1} - M_t)$ to firms. At the end of the period they pay interest and principal $q_t B_t$ to their creditors and distribute the remaining profits,

$$D_t^B = \frac{q_t(B_t + M_{t+1} - M_t)}{P_t} - \frac{q_t B_t}{P_t} = q_t \frac{M_{t+1} - M_t}{P_t} \tag{3.17}$$

to the household sector. As in Section 2.4.3 P_t denotes the money price of output.

Producers. The representative producer employs labor N_t and capital services K_t to produce output according to

$$Y_t = Z_t(A_t N_t)^{1-\alpha} K_t^{\alpha}, \quad \alpha \in (0,1). \tag{3.18}$$

[13] The model is based on a paper by Lawrence Christiano, Martin Eichenbaum, and Charles Evans (1997). Different from their model, we also include capital services as a factor of production.

As in the benchmark model A_t is the level of labor-augmenting technical progress that grows deterministically at the rate $a - 1 \geq 0$. Total factor productivity Z_t is governed by the stochastic process

$$Z_t = Z_{t-1}^{\rho^Z} e^{\epsilon_t^Z}, \quad \epsilon_t^Z \sim N(0, \sigma^Z). \tag{3.19}$$

Producers hire workers at the money wage rate W_t and capital services at the real rental rate r_t. Since they have to pay workers in advance, they borrow $W_t N_t$ at the nominal rate of interest $q_t - 1$ from banks. Hence, their profits are given by

$$D_t^P = Y_t - q_t \frac{W_t}{P_t} N_t - r_t K_t. \tag{3.20}$$

Maximizing (3.20) with respect to N_t and K_t provides the following first-order conditions:

$$q_t w_t = (1 - \alpha) Z_t N_t^{-\alpha} k_t^{\alpha}, \quad w_t := \frac{W_t}{A_t P_t}, \; k_t := \frac{K_t}{A_t}, \tag{3.21a}$$

$$r_t = \alpha Z_t N_t^{1-\alpha} k_t^{\alpha-1}. \tag{3.21b}$$

Consequently, profits in the production sector are zero.

Money Supply. Money supply is governed by the same process that we used in Section 2.4.3. Thus,

$$\mu_t := \frac{M_{t+1}}{M_t}, \quad \mu_t = \mu^{1-\rho^\mu} \mu_{t-1}^{\rho^\mu} e^{\epsilon_t^\mu}, \quad \epsilon_t^\mu \sim N(0, \sigma^\mu). \tag{3.22}$$

Households. The households' total financial wealth at the beginning of period t is given by $M_t = B_t + X_t$, where B_t is the amount deposited at banks and X_t are cash balances kept for the purchase of consumption goods. Since households receive wages before they go shopping, their cash-in-advance constraint is

$$C_t \leq \frac{X_t + W_t N_t}{P_t}. \tag{3.23}$$

The real income of households consists of wages $W_t N_t / P_t$, net rental income $(r_t - \delta) K_t$ from capital services (where capital depreciates at the rate δ), interest on banking deposits $(q_t - 1) B_t / P_t$,

and dividends from banks D_t^B. This income is split between consumption C_t and savings S_t. Savings are used to increase financial wealth M_t and the stock of physical capital K_t. Accordingly, the budget constraint is given by:

$$\begin{aligned} &K_{t+1} - K_t + \frac{(X_{t+1} - X_t) + (B_{t+1} - B_t)}{P_t} \\ &\leq \frac{W_t}{P_t} N_t + (r_t - \delta) K_t + (q_t - 1) \frac{B_t}{P_t} + D_t^B - C_t. \end{aligned} \tag{3.24}$$

In the following we depart from our usual specification of the household's preferences over consumption and leisure and use the following instantaneous utility function:

$$u(C_t, N_t) := \frac{1}{1-\eta} \left[(C_t - \theta A_t N_t^{\nu})^{1-\eta} - 1 \right], \quad \theta > 0,\ \nu > 1.$$

As you will see in a moment, this function implies a labor supply schedule that depends on the real wage only. In particular, labor supply does not depend on wealth. Technically, this makes it easy to solve for N_t given the real wage and to separate the role of the elasticity of labor supply $1/(\nu - 1)$ from other factors.[14]

The household maximizes the expected stream of discounted utility

$$E_0 \sum_{t=0}^{\infty} \beta^t u(C_t, N_t)$$

with respect to C_0, N_0, K_1, X_1, and B_1 subject to (3.23) and (3.24). Since the household must decide on the size of its nominal deposits before the monetary shock is observed, X_t and B_t are state variables of the model. The Lagrangean for this problem is:

$$\mathscr{L} = E_0 \sum_{t=0}^{\infty} \beta^t \left\{ \frac{1}{1-\eta} \left[(C_t - \theta A_t N_t^{\nu})^{1-\eta} - 1 \right] \right.$$

[14] This functional form dates back to a paper by Greenwood, Hercowitz, and Huffman (1988). It is also used by Lawrence Christiano, Martin Eichenbaum, and Charles Evans (1997), which is why we use it here.

$$+\Lambda_t\left[\frac{W_t}{P_t}N_t+(r_t-\delta)K_t+\frac{(q_t-1)B_t}{P_t}+D_t^B-C_t\right.$$
$$\left.-(K_{t+1}-K_t)-\frac{(X_{t+1}-X_t)+(B_{t+1}-B_t)}{P_t}\right]$$
$$\left.+\Xi_t\left[\frac{X_t+W_tN_t}{P_t}-C_t\right]\right\}.$$

From this expression we can derive the set of first order conditions that describes the household's decisions. In the following, we present these conditions in terms of the stationary variables $y_t := Y_t/A_t$, $c_t := C_t/A_t$, $k_t := K_t/A_t$, $w_t := W_t/(A_tP_t)$, $\pi_t := P_t/P_{t-1}$, $\lambda_t := \Lambda_t A_t^\eta$, $x_t := X_t/(A_{t-1}P_{t-1})$, $m_t := M_t/(A_{t-1}P_{t-1})$, and $\xi_t := \Xi_t A_t^\eta$. The definitions of x_t and m_t guarantees that these variables are pre-determined at the beginning of period t.

$$\lambda_t+\xi_t=(c_t-\theta N_t^\nu)^{-\eta}, \tag{3.26a}$$
$$N_t=\left(\frac{w_t}{\theta\nu}\right)^{\frac{1}{\nu-1}}, \tag{3.26b}$$
$$\lambda_t=\beta a^{-\eta}E_t\lambda_{t+1}\left(1-\delta+\alpha Z_{t+1}N_{t+1}^{1-\alpha}k_{t+1}^{\alpha-1}\right), \tag{3.26c}$$
$$\lambda_t=\beta a^{-\eta}E_t\left(\frac{\lambda_{t+1}q_{t+1}}{\pi_{t+1}}\right), \tag{3.26d}$$
$$\lambda_t=\beta a^{-\eta}E_t\left(\frac{\lambda_{t+1}+\xi_{t+1}}{\pi_{t+1}}\right), \tag{3.26e}$$
$$0=\xi_t(x_t/(a\pi_t)+w_tN_t-c_t). \tag{3.26f}$$

Equation (3.26a) shows that the marginal utility of consumption departs from the shadow price of wealth λ_t as long as the cash-in-advance constraint binds, i.e., if $\xi_t > 0$. The related Kuhn-Tucker condition is equation (3.26f). Equation (3.26b) is the labor supply schedule. The well-known Euler equation for capital is given in (3.26c). Together with equations (3.26d) and (3.26e) it implies equal expected rewards on the holdings of physical capital, of banking deposits, and of cash balances.

In addition to these equations the household's budget constraint is satisfied with the equality sign and the cash-in-advance

constraint holds. Since

$$w_t N_t = \frac{B_t + M_{t+1} - M_t}{A_t P_t} = m_{t+1} - x_t/(a\pi_t),$$

we may write the latter in the following way:

$$c_t = m_{t+1}, \quad \text{if } \xi_t > 0, \tag{3.27a}$$

$$c_t \leq m_{t+1}, \quad \text{if } \xi_t = 0, \tag{3.27b}$$

$$m_{t+1} = \frac{\mu_t m_t}{a\pi_t}, \tag{3.27c}$$

where the third equation is implied from the definition of m_t. In equilibrium, the household's budget constraint reduces to the well-known resource restriction:

$$ak_{t+1} = Z_t N_t^{1-\alpha} k_t^{\alpha} + (1-\delta)k_t - c_t. \tag{3.28}$$

Stationary Equilibrium. In a stationary equilibrium all shocks equal their unconditional means, $Z_t \equiv 1$ and $\mu_t \equiv \mu$ for all t, and all (scaled) variables are constant. Equation (3.27c) implies that the rate of inflation is proportional to the money growth rate:

$$\pi = \frac{\mu}{a}. \tag{3.29a}$$

The Euler equation for capital (3.26c) delivers

$$1 = \beta a^{-\eta} \underbrace{(1 - \delta + \alpha(y/k))}_{1-\delta+r} \Rightarrow \frac{y}{k} = \frac{a^{\eta} - \beta(1-\delta)}{\alpha\beta}. \tag{3.29b}$$

Together with (3.26d) this implies the Fisher equation, here written in terms of gross rates:

$$q = \pi(1 - \delta + r). \tag{3.29c}$$

Given this, the stationary version of (3.26e) implies:

$$\xi = \lambda(q - 1). \tag{3.29d}$$

Accordingly, the cash-in-advance constraint binds in equilibrium if the nominal interest rate is positive: $q - 1 > 0$. Combining (3.29a) and (3.29c), we find that this condition is satisfied, if the growth rate of money is not too small:

$$\mu > \beta a^{1-\eta}.$$

Finally note that equation (3.26b) and equation (3.21a) imply

$$N^{\nu-1} = \frac{1}{q}\left(\frac{1-\alpha}{\nu\theta}\frac{y}{N}\right). \tag{3.29e}$$

Since y/N is a function of y/k, it is independent of the money growth rate. Yet, according to (3.29c) and (3.29a) q is an increasing function of μ. Thus, steady-state working hours depend inversely on the rate of money growth. As in the model of Section 2.4.3 money is not super-neutral.

The PEA Solution. Our model has two exogenous shocks, Z_t and μ_t, and three variables with given initial conditions, k_t, m_t, and x_t. However, there are not enough equations to determine consumption, working hours, the rate of inflation, the nominal interest rate, and the Lagrange multiplier of the cash-in-advance constraint given the former variables. We must define additional co-state variables. However, there is no easy way to do so, since the three Euler equations (3.26c) through (3.26e) have the same lhs. Technically speaking, the system of stochastic difference equations $\mathbf{g}(\cdot)$ is not easily invertible. There are various possible ways to deal with this situation. The following is the solution that really works.

As in the applications above we parameterize the rhs of the Euler equation for capital:

$$\lambda_t = \beta a^{-\eta}\psi^1(\boldsymbol{\gamma}_1, k_t, m_t, x_t, Z_t, \mu_t). \tag{3.30a}$$

Since $m_t > 0$ in any solution where money has a positive value, we multiply the second Euler equation (3.26d) on both sides by m_{t+1} and parameterize the ensuing rhs of this equation:

$$m_{t+1}\lambda_t = \beta a^{-\eta}\psi^2(\boldsymbol{\gamma}_2, k_t, m_t, x_t, Z_t, \mu_t). \tag{3.30b}$$

Analogously, we multiply the third Euler equation by x_{t+1} and put

$$x_{t+1}\lambda_t = \beta a^{-\eta}\psi^3(\boldsymbol{\gamma}_3, k_t, m_t, x_t, Z_t, \mu_t). \tag{3.30c}$$

We are now able to trace out a time path as follows: Given the five-tuple $(k_t, m_t, x_t, Z_t, \mu_t)$ we use (3.30a) to solve for λ_t. We use this solution to infer m_{t+1} and x_{t+1} from (3.30b) and (3.30c), respectively. Given m_t and m_{t+1} equation (3.27c) delivers π_t. Since

$$w_t N_t = \frac{\mu_t m_t - x_t}{a\pi_t}.$$

we can solve for $w_t N_t$ and use this in (3.26b) to solve for N_t. In the next step we use the first-order condition for labor demand (3.21a) to solve for q_t. Finally we check the Kuhn-Tucker conditions: assume $\xi_t = 0$. This implies

$$\bar{c}_t = \lambda_t^{-1/\eta} + \theta N_t^{\nu},$$

from (3.26a). If $\bar{c}_t < m_{t+1}$ we accept this solution. Otherwise we put $\tilde{c}_t = m_{t+1}$ and solve for ξ_t from (3.26a):

$$\xi_t = (\tilde{c}_t - \theta N_t^{\nu})^{-\eta} - \lambda_t.$$

Since $\bar{c}_t > \tilde{c}_t$, we also have $\xi_t > 0$. In the last step we compute k_{t+1} from the resource constraint (3.28).

Implementation. The Fortran program `LP.for` implements the PEA solution. As in the previous applications we use exponentials of simple polynomials for ψ^i, $i = 1, 2, 3$. The program allows the user to find the solution either iteratively or in one step by solving the related system of non-linear equations. In both cases the program obtains starting values from the solution of the log-linearized model. We do this for the following reason. Since we have five state variables and three expectational equations the potential number of coefficients in the expectational equations is large. For instance, a complete second degree polynomial in five variables has 21 coefficients. Accordingly, the potential of multicollinearity among the 21 regressors is high and we do not consider

higher degree polynomials. Given the log-linear solution, we compute time paths for the relevant variables. In a first step we look at the correlation matrix between the potential regressors and exclude those that are highly correlated with others.[15] In a second step we regress the error terms from the log-linear solution on the remaining regressors. For this step, we use the Gauss-Newton method presented as Algorithm 8.6.2. Given these initial values we either invoke our non-linear equations solver or compute new time paths and estimates until the estimates converge. In a third step we reduce the set of regressors further: we exclude all regressors whose t-ratios from the solution of step 2 are smaller than one in absolute value. As it turns out, we get good results with a small number of coefficients.

Note also that the number of regressors depends on your assumptions with regard to monetary policy. If the monetary authority is able to control money supply perfectly, i.e., $\sigma^{\mu} = 0$, the vector $\boldsymbol{\mu} := [\mu_1, \ldots, \mu_T]'$ is a vector of constants. Neither $\boldsymbol{\mu}$ nor any of its integer powers or cross-products with other variables can be used as regressor. To see this, consider the case

$$\psi(k_t, \mu_t) := \exp(\gamma_1 + \gamma_2 \ln(k_t) + \gamma_3 \ln(\mu_t)).$$

The Jacobian matrix of ψ with respect to γ_i is given by:

$$\begin{bmatrix} \psi(k_1, \mu_1) & \psi(k_1, \mu_1)\ln(k_1) & \psi(k_1, \mu_1)\ln(\mu_1) \\ \psi(k_2, \mu_2) & \psi(k_2, \mu_2)\ln(k_2) & \psi(k_2, \mu_2)\ln(\mu_2) \\ \vdots & \vdots & \vdots \\ \psi(k_T, \mu_T) & \psi(k_T, \mu_T)\ln(k_T) & \psi(k_T, \mu_T)\ln(\mu_T) \end{bmatrix}.$$

Thus, if $\mu_t = \mu$ for all t, the third column of this matrix is a multiple $\ln(\mu)$ of the first and the Jacobian is singular. Accordingly, the Gauss-Newton step cannot be computed.

Concluding this paragraph, we strongly advice you to go through steps one to three from above for every parameter set that you wish to consider.

[15] The program allows you to write this matrix to a file without doing any further computations.

Table 3.3

Preferences	Production		Money Supply
β=0.994	a=1.005	α=0.27	μ=1.0167
η=2.0	δ=0.011	ρ^Z=0.90	ρ^μ=0.0
N=0.13	σ^Z=0.0072		σ^μ=0.0173
ν=5.0			

Results. If not mentioned otherwise the following results are computed for the calibration displayed in Table 3.3. The parameters for the production side and for money supply are the same as those used in Section 2.4.3 and are, therefore, reproduced from Table 2.3. The preference parameters β and η are the same as in the benchmark model. Furthermore, we choose θ so that stationary working hours are $N = 0.13$. The parameter that determines the labor supply elasticity ν is taken from Heer and Maussner (2004).

Table 3.4 displays the arguments and estimated coefficients of the functions that we use to parameterize expectations. They are the results of the steps described in the previous paragraph.

We will first consider the relative strength between the liquidity and the anticipated inflation effect. If the monetary shock is not autocorrelated – as our estimates of this process from German data indicate – there is no anticipated inflation effect. This effect gains importance, if the autocorrelation parameter ρ^μ increases. The impulse responses displayed in Figure 3.1 show this very clearly. The monetary shock hits the economy in period $t = 3$. The solid lines correspond to the case $\rho^\mu = 0$. The liquidity effect is obvious from the lower right panel of Figure 3.1. The additional supply of money lowers the nominal interest rate. The costs of hiring labor decrease, working hours and production increase. Part of the extra income is consumed and part is transferred to future periods via additional capital accumulation. The positive effect on consumption is very small, and, thus, not visible in Figure 3.1.

The dotted lines correspond to an autocorrelated money supply process. In addition to the liquidity effect, there is also an

Table 3.4

Regressors	$\sigma^\mu = 0.0173$			$\sigma^\mu = 0$		
	ψ^1	ψ^2	ψ^3	ψ^1	ψ^2	ψ^3
c	4.1457	2.4675	−8.5449	4.1161	2.3547	−3.7464
	(56.45)	(18.70)	(−23.03)	(54.70)	(87.23)	(−59.33)
k	−1.4201	−0.8764	1.2091	−1.2503	−0.6635	0.0492
	(−36.28)	(−3.68)	(6.17)	(−32.01)	(−13.80)	(1.51)
x	−0.1440		−4.8323	−0.1033	0.0016	−2.1977
	(−3.43)		(−22.76)	(−2.39)	(1.64)	(−60.54)
Z	−0.4868	−0.1099	−3.6561	−0.4859	−0.1046	−3.7700
	(−183.87)	(−14.31)	(−275.13)	(−183.33)	(−31.95)	(−1692.54)
μ	−0.0040	0.0140	−0.1237			
	(−2.21)	(2.06)	(−13.68)			
k^2	0.2333	0.1841	−0.5130	0.1570	0.0880	0.0225
	(13.20)	(1.71)	(−5.80)	(8.90)	(4.05)	(1.53)
x^2	−0.0227		−0.7571	−0.0164		−0.3410
	(−3.50)		(−22.96)	(−2.45)		(−60.36)
Z^2	−0.1063		−5.0438	−0.1023		−5.7785
	(−1.74)		(−16.58)	(−1.66)		(−112.73)
μ^2	0.0734		3.9122			
	(1.61)		(17.30)			

Notes: c refers to the intercept, t-ratios of estimated coefficients at final solution in parenthesis.

inflationary expectations effect. As can be seen from Figure 3.1 the latter dominates the former for our choice of parameters. Since households expect higher inflation, their costs of holding money balances increase. They substitute physical capital for financial wealth so that there is a stronger increase in investment. Since the cash-in-advance constraint binds, the reduced money holdings entail lower consumption. On the production side the increased nominal interest rate reduces working hours and output. This negative income effect puts additional pressure on consumption.

Table 3.5 presents second moments from a two different simulations of the model. The first run considers the case of steady money growth, i.e., $\sigma^\mu = 0$, the second simulation assumes monetary shocks of the size observed in the data.

First, consider columns 2 to 4. They show one obvious difference between the benchmark model and the present model (compare Table 3.2). The standard deviation of working hours

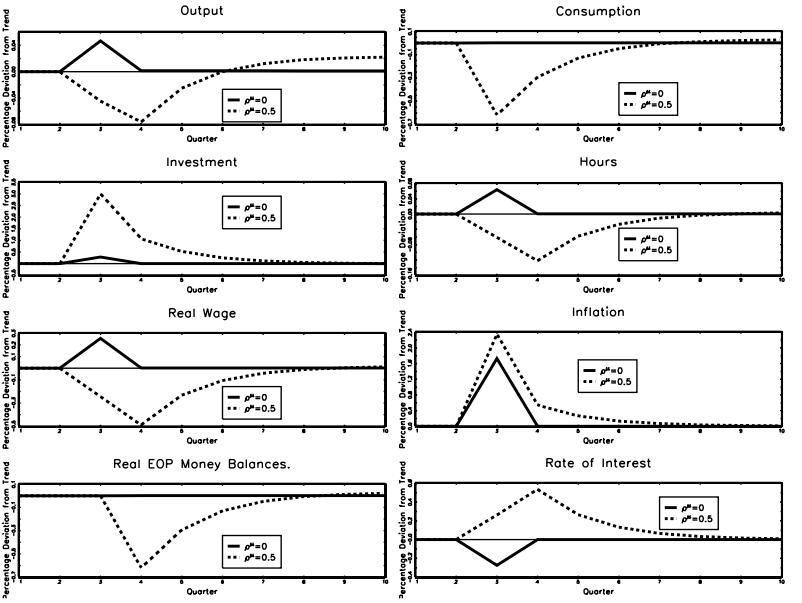

Figure 3.1: Impulse Responses to a Monetary Shock in the Limited Participation Model

in the benchmark model is almost four times larger, and, as a consequence, output fluctuations are more pronounced. This difference is easily traced to the small elasticity of labor supply of $1/(\nu - 1) = 0.25$. In the benchmark model the Frisch elasticity of labor supply is determined implicitly, and is about 1.7.[16]

Given our calibration, the real effects of monetary shocks in the limited participation model are quite small and negligible. The standard deviations in column 5 differ from those in column 2 only in two instances: due to the liquidity effect, the standard deviation of the real wage is about 7 percent higher, which translates into a greater variability of hours. Of course, inflation is

[16] The Frisch elasticity measures the relative change of working hours to a one-percent increase of the real wage, given the marginal utility of wealth λ. In the steady state of the benchmark model it is given by

$$\frac{dN/N}{dw/w} = \frac{1-N}{N}\left(\frac{1-\eta}{\eta}\theta - 1\right)^{-1}.$$

Table 3.5

Variable	PEA Solution						Loglinear Solution		
	$\sigma^\mu = 0$			$\sigma^\mu = 0.0173$			$\sigma^\mu = 0.0173$		
	s_x	r_{xy}	r_x	s_x	r_{xy}	r_x	s_x	r_{xy}	r_x
Output	0.98	1.00	0.68	0.98	1.00	0.68	0.98	1.00	0.68
Investment	4.47	1.00	0.69	4.47	1.00	0.69	4.45	1.00	0.69
Consumption	0.34	0.98	0.67	0.34	0.98	0.67	0.34	0.98	0.67
Hours	0.17	0.87	0.78	0.18	0.83	0.67	0.18	0.83	0.67
Real Wage	0.69	0.87	0.78	0.74	0.83	0.67	0.73	0.83	0.67
Inflation	0.27	−0.43	−0.09	1.69	−0.02	−0.07	1.70	−0.02	−0.09

Notes: s_x:=standard deviation of HP-filtered simulated series of variable x, r_{xy}:=cross correlation of variable x with output, r_x:=first order autocorrelation of variable x.

substantially more volatile if monetary shocks are present. The standard deviation of this variable is almost equal to σ^μ.

Finally, consider columns 7 to 9. They present the second moments obtained from the simulation of the loglinear solution of the model. Both the loglinear and the PEA simulations use the same random numbers, so that differences between the two solutions are non-random.[17] Compared to columns 4 to 6, no noteworthy difference is discernible.

[17] They are random only in so far as the PEA solution depends itself on a long sequence of random numbers.

Problems

3.1 Consider the stochastic Ramsey Model from Example 1.3.2. The problem is to solve

$$\max \sum_{t=0}^{\infty} \beta^t \ln C_t$$
$$\text{subject to} \qquad K_{t+1} = Z_t K_t^{\alpha} - C_t,$$
$$K_0, Z_0 \text{ given.}$$

We already know that the policy function g is given by

$$g(K_t, Z_t) = \alpha\beta Z_t K_t^{\alpha}.$$

Use this information to find the analytic solution for the expectations function $\mathscr{E}$.

3.2 Consider the stochastic growth model with irreversible investment from Example 3.3.1. Instead of using simple polynomials use
a) neural networks
b) and Chebyshev polynomials
to solve this model. Use the same number of parameters as we did in Section 3.3.1 and compare your results to ours. Use the sum of squared errors and the DM-statistic for this purpose.
(Hint: you can use the solution from a) to find suitable lower and upper bounds for K and Z, that are necessary when you use Chebyshev polynomials.)

3.3 A less complicated DGE model of a monetary economy than the limited participation model of Section 3.3.3 is the model of Cooley and Hansen (1989). This paper introduces money into the model of Hansen (1985) via a cash-in-advance constraint. The authors demonstrate that a policy of constant money growth does not alter the business cycle characteristics of the original model and that an erratic money supply resembling the US historical experience alters the behavior of real variables slightly.
Cooley and Hansen (1989) solve their model with a variant of the linear-quadratic method of Section 2.2. We ask you to employ the PEA to solve their model and to reproduce their results.
We use the same symbols for consumption, capital, working hours, money balances, and so forth as in the model of Section 3.3.3. The representative household solves the following problem:

$$\max_{C_0,N_0,K_1,M_1} \quad E_0 \sum_{t=0}^{\infty} \beta^t \left(\ln C_t - \theta N_t\right)$$

subject to

$$K_{t+1} - K_t + \frac{M_{t+1} - M_t}{P_t} \leq \frac{W_t}{P_t} N_t + (r_t - \delta) K_t + T_t - C_t,$$

$$C_t \leq \frac{M_t}{P_t} + T_t.$$

Money supply is determined by

$$T_t = \frac{M_{t+1} - M_t}{P_t}, \quad M_{t+1} = \mu_t M_t.$$

The policy of a constant money supply implies $\mu_t = \mu$ for all t, whereas

$$\hat{\mu}_t = (1 - \rho^{\mu})\mu + \rho^{\mu} \hat{\mu}_{t-1} + \epsilon_t^{\mu}, \quad \epsilon_t^{\mu} \sim N(0, \sigma^{\mu}), \ \hat{\mu}_t := \ln(\mu_t/\mu)$$

describes an erratic money supply.
The representative firm solves

$$\max_{N_t,K_t} Z_t N_t^{\alpha} K_t^{1-\alpha} - \frac{W_t}{P_t} N_t - r_t K_t,$$

where Z_t is governed by

$$\ln Z_t = \rho^Z \ln Z_{t-1} + \epsilon_t^Z, \quad \epsilon_t^Z \sim N(0, \sigma^Z).$$

a) Set up the Lagrangean of the household's problem and derive the first-order conditions for this problem.
b) Use the first-order conditions of the firm's problem to substitute for the wage rate and the rental rate in the household's optimality conditions and derive the system of stochastic difference equations that govern the model's dynamics.
c) Solve for the model's balanced growth path and show that working hours are a decreasing function of the steady-state growth rate of money.
d) Consult the Appendix of DEN HAAN and MARCET (1994) to find out how they solve this model using the PEA.
e) COOLEY and HANSEN (1989) calibrate their model as follows: $\beta = 0.99$, $\theta = 2.86$, $\alpha = 0.64$, $\rho^Z = 0.95$, $\sigma^Z = 0.00721$, $\rho^{\mu} = 0.48$, and $\sigma^{\mu} = 0.009$. Use a polynomial in K_t, Z_t, and μ_t to parameterize the conditional expectation appearing in the Euler equation of capital and solve the model.
f) Use the solution and simulate the model. As in the original paper use time series with 115 observations and compute seconds moments as averages over 50 simulations from the HP-filtered time series.

3.4 Consider the Cooley and Hansen (1989) model with current period utility given by

$$u(C_t, N_t) := \frac{1}{1-\eta} C_t^{1-\eta}(1-N_t)^{\theta(1-\eta)}, \quad \eta \neq 1,\ \theta > 0, \eta > \frac{\theta}{1+\theta}.$$

Put $\eta = 2$ and choose θ so that working hours in the steady-state equal $N = 0.33$.
With these preferences it is no longer possible to solve the model along the lines of den Haan and Marcet (1994). To solve the model you must parameterize the conditional expectations not only in the Euler equation for capital but also in the Euler equation for money balances.
Solve the model and compare your results to those of the original model. (Hint: Use $\Lambda_t m_{t+1} = \psi^2(K_t, Z_t, m_t, \mu_t)$ as the second parameterized equation.)

Chapter 4

Projection Methods

Overview. In this chapter we introduce projection methods or, as they are also referred to, weighted residual methods.[1] These methods are closely related to the parameterized expectations approach (PEA) that we presented in the previous chapter. Moreover, the conventional PEA can be interpreted as a special case of projection methods. In Chapter 3 we approximate the conditional expectations of the household with the exponential of a polynomial, therefore restricting the function to be positive. In order to find the best approximation, we use a Monte-Carlo simulation and non-linear regression methods. In this chapter, again, we use a parameterized polynomial to approximate a policy function. However, we will generalize the PEA along certain dimensions: (1) The functions that we approximate do not need to be the conditional expectations that characterize the first-order conditions of the agents in our model. Instead, we may approximate the agent's policy function, or the value function of the problem at hand. (2) In Chapter 3 we have seen that using higher order monomials to achieve a better approximation may not always be successful. Here we circumvent this problem, since we will choose Chebyshev polynomials, which belong to the class of families of orthogonal

[1] Early expositions of projection methods are provided by Judd (1992,1998) and Reddy (1993). McGrattan (1996) also consider so-called finite-element methods that approximate the solution over non-overlapping subdomains of the state-space. In these methods, low-polynomials are fitted on subdomains rather than high polynomials on the entire state space. Our piecewise linear approximation of the value function in Algorithm 1.2.3 can be interpreted as a finite-element method. In the following, we will not consider these methods and refer the interested reader to Reddy (1993) and McGrattan (1996).

polynomials.[2] (3) In Algorithm 3.1.1, a large sequence of shocks is generated in order to simulate the behavior of the economy. In this chapter, you will learn that Monte-Carlo simulation is only one possible solution technique in order to fit the parameterized function. Other methods are possible and often preferable.

This chapter is structured as follows. First, the general idea of projection methods is presented. Second, we consider the various steps that constitute this class of methods in more detail. There, it will become obvious that we need several numerical tools to implement a particular method. Among them are numerical integration and optimization as well as finding the zeros of a set of non-linear equations. Third, we apply projection methods to the deterministic and the simple stochastic growth model and compare our results to those of Chapter 1 and Chapter 3. As an additional application, we study the equity premium puzzle, i.e. the (arguably) missing explanation for the observation that the average return on equities has been so much higher than the one on bonds over the last century. For this reason, we introduce asset pricing into the stochastic growth model.

4.1 Characterization of Projection Methods

4.1.1 An Example

Projection methods derive approximate solutions to functional equations. The unknown of a functional equation is not a point in $\mathbb{R}^n$ but a function f that maps $\mathbb{R}^n$ to $\mathbb{R}^m$. Since an appropriately defined set of functions is itself a vector space, the problem is to pick an element from a function space. Different from $\mathbb{R}^n$, however, function spaces have infinite dimensions, and in many circumstances it is impossible to derive analytic solutions. Projection methods use a family of polynomials $\mathcal{P} := \{\varphi_i\}_{i=0}^{\infty}$ and approximate f by a finite sum of members of this family.

[2] If you are not already familiar with polynomial approximation in general and with Chebyshev polynomials in particular, it is now a good time to read Section 8.2.

To be concrete, consider the ordinary differential equation[3]

$$\dot{x}(t) + x(t) = 0, \quad x(0) = 1, \tag{4.1}$$

with solution

$$x(t) = e^{-t}.$$

Suppose we use the monomials $(1, t, t^2)$ to approximate the solution in the interval $[0, 2]$:[4]

$$\hat{x}(t) = 1 + \gamma_1 t + \gamma_2 t^2. \tag{4.2}$$

How shall we choose the unknown parameters γ_i, $i = 1, 2$? In econometrics we approximate a given set of points $(x_i, y_i) \in \mathbb{R}^2$ by drawing a line so that the sum of squared distances of (x_i, y_i) from this line attains a minimum. Something similar to this also works here. Let us define the residual function

$$R(\boldsymbol{\gamma}, t) := \underbrace{\gamma_1 + 2\gamma_2 t}_{d\hat{x}/dt} + \underbrace{1 + \gamma_1 t + \gamma_2 t^2}_{\hat{x}(t)}, \tag{4.3}$$

i.e. the function that we get if we substitute our guess of the solution into the functional equation (4.1). Then, our goal could be to choose the parameters so that over $[0, 2]$ $\hat{x}(t)$ is as close as possible to $x(t)$ in the sense of

$$\min_{\gamma_1, \gamma_2} \quad \int_0^2 R(\boldsymbol{\gamma}, t)^2 dt. \tag{4.4}$$

The first-order conditions for this problem are given by the following two equations:

$$0 = \int_0^2 R(\boldsymbol{\gamma}, t) \frac{\partial R(\boldsymbol{\gamma}, t)}{\partial \gamma_1} dt,$$
$$0 = \int_0^2 R(\boldsymbol{\gamma}, t) \frac{\partial R(\boldsymbol{\gamma}, t)}{\partial \gamma_2} dt.$$

[3] In the following, we draw on McGrattan (1999).
[4] Note that we set $\gamma_0 = 1$ to satisfy the boundary condition $x(0) = 1$.

Using (4.3) and the derivatives of this function with respect to γ_1 and γ_2, it is easy to compute the integrals. This delivers the following linear system of equations in the two unknowns γ_1 and γ_2:

$$-4 = 8\frac{2}{3}\,\gamma_1 + 16\gamma_2,$$
$$-6\frac{2}{3} = 16\gamma_1 + 33\frac{1}{15}\,\gamma_2.$$

Figure 4.1 shows that the approximate solution is not too far from the true function e^{-t}. Of course, we can get a better approximation if we use a higher degree polynomial.

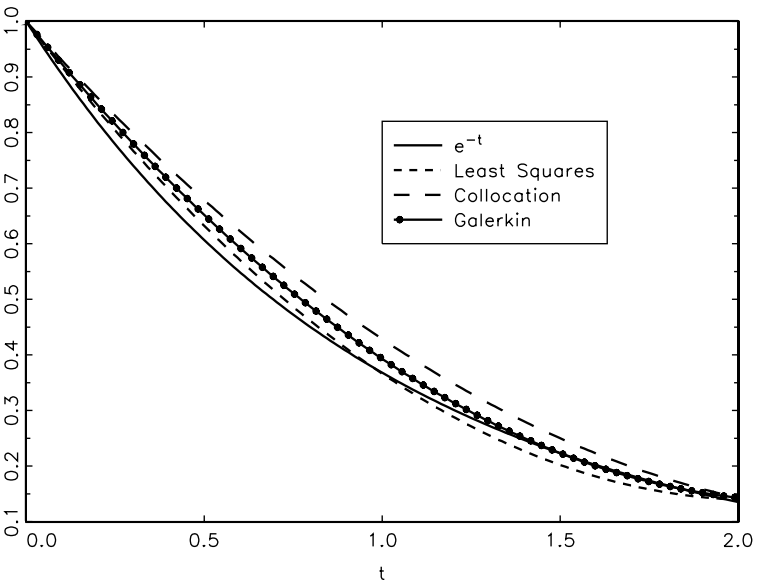

Figure 4.1: Polynomial Approximation of e^{-t}

Using a well known property of the least squares estimator delivers another solution concept, the Galerkin method. Remember, the least squares residuals are orthogonal to the space spanned by the vectors that represent the observations of the independent variables. Here, the functions t and t^2 play the role of these vectors. Thus, we demand

$$\begin{aligned} 0 &= \int_0^2 R(\boldsymbol{\gamma}, t) t dt, \\ 0 &= \int_0^2 R(\boldsymbol{\gamma}, t) t^2 dt. \end{aligned} \tag{4.5}$$

Computing the integrals on the rhs of (4.5) gives a second set of linear equations in the unknown parameters γ_1 and γ_2:

$$\begin{aligned} -2 &= 4\frac{2}{3}\gamma_1 + 9\frac{1}{3}\gamma_2, \\ -2\frac{2}{3} &= 6\frac{2}{3}\gamma_1 + 14\frac{2}{5}\gamma_2. \end{aligned}$$

The dotted line in Figure 4.1 represents the Galerkin approximate solution of the differential equation (4.1).

Finally, we may want that the residual function is equal to zero at a given set of points. Suppose we choose $t_1 = 1$ and $t_2 = 2$. This gives the linear system

$$\begin{aligned} -1 &= 2\gamma_1 + 3\gamma_2, \\ -1 &= 3\gamma_1 + 8\gamma_2. \end{aligned}$$

The solution based on this principle is known as collocation method. Figure 4.1 reveals that this approximation is about as close to the true curve as are the other solutions.

4.1.2 The General Framework

The three different solutions that we have just considered may be obtained from the following setting. We want to approximate an unknown function $f : X \to Y$, where X and Y are subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. This function is implicitly defined by the functional equation $F(f) = 0$, where $F : C_1 \to C_2$. C_1 and C_2 are given spaces of functions, e.g., the set of all continuously differentiable functions on $[a, b]$. Examples of functional equations are the Bellman equation (1.13) of the deterministic growth model considered in Chapter 1 and the Euler equation of the stochastic growth model (1.48c) also presented in Chapter 1. Given a family of polynomials $\mathcal{P} := \{\varphi_i\}_{i=0}^{\infty}$, we approximate f by a finite linear combination of the first $p + 1$ members of this family:

$$\hat{f}(\mathbf{x}) = \sum_{i=0}^{p} \gamma_i \varphi_i(\mathbf{x}), \ \mathbf{x} \in X \subset \mathbb{R}^n. \tag{4.6}$$

The residual function is obtained by substituting $\hat{f}$ into the functional equation:

$$R(\boldsymbol{\gamma}, \mathbf{x}) := F(\hat{f}(\boldsymbol{\gamma}, \mathbf{x})), \quad \boldsymbol{\gamma} := (\gamma_0, \ldots, \gamma_p).$$

Suppose there is a set of test functions $\{g_i(\mathbf{x})\}_{i=0}^{p}$ and a weight function $w(\mathbf{x})$. Together with R they define an inner product given by

$$\int_X w(\mathbf{x}) R(\boldsymbol{\gamma}, \mathbf{x}) g_i(\mathbf{x}) d\mathbf{x}.$$

On a function space, this inner product induces a norm (i.e., a measure of distance) on this space and we choose the vector of parameters $\boldsymbol{\gamma}$ so that

$$\int_X w(\mathbf{x}) R(\boldsymbol{\gamma}, \mathbf{x}) g_i(\mathbf{x}) d\mathbf{x} = 0, \quad \forall i = 0, 1, \ldots, n. \tag{4.7}$$

Now, it is easy to see that the three different solutions considered above are derived from (4.7) for special choices of g_i and w.

- The least squares solution puts $g_i \equiv \partial R / \partial \gamma_i$ and $w \equiv 1$.
- The Galerkin solution chooses $g_i \equiv \varphi_i$ and $w \equiv 1$.
- The collocation method uses the Dirac delta function as weight function,

$$w(\mathbf{x}) = \begin{cases} 0 \text{ if } \mathbf{x} \neq \mathbf{x}_i, \\ 1 \text{ if } \mathbf{x} = \mathbf{x}_i, \end{cases}$$

  and puts $g_i \equiv 1$.

In the following, we restrict ourselves to these three definitions of a solution being close to the true function. In the next section we consider the different steps to implement a specific solution in more detail. Before that, we summarize the general procedure that underlies projection methods in the following algorithm.

Algorithm 4.1.1 (Projection Method)

Purpose: *Approximate the solution $f : X \to Y$ of a functional equation $F(f) = 0$.*

Steps:

Step 1: Choose a bounded state space $X \subset \mathbb{R}^n$ and a family of functions $\varphi_i(\mathbf{x}) : X \to Y$, $i = 0, 1, \ldots$.

Step 2: Choose a degree of approximation p and let

$$\hat{f}(\boldsymbol{\gamma}, \mathbf{x}) = \sum_{i=0}^{p} \gamma_i \varphi_i(\mathbf{x}).$$

Step 3: Define the residual function:

$$R(\boldsymbol{\gamma}, \mathbf{x}) := F(\hat{f}(\boldsymbol{\gamma}, \mathbf{x})).$$

Step 4: Choose a projection function g_i, a weight function w and compute the inner product:

$$G_i := \int_X w(\mathbf{x}) R(\boldsymbol{\gamma}, \mathbf{x}) g_i(\mathbf{x}) d\mathbf{x}, \quad i = 0, \ldots, n.$$

Find the value of $\boldsymbol{\gamma}$ that solves $G_i = 0$, or, in the case of least squares projection ($g_i = \partial R / \partial \gamma_i$ and $w \equiv 1$), minimize

$$\int_X R(\boldsymbol{\gamma}, \mathbf{x})^2 d\mathbf{x}$$

with respect to $\boldsymbol{\gamma}$.

Step 5: Verify the quality of the candidate solution $\boldsymbol{\gamma}$. If necessary, return to step 2 and increase the degree of approximation p or even return to step 1 and choose a different family of basis functions.

4.1.3 Parameterized Expectations and Projection Methods

Lawrence Christiano and Jonas Fisher (2000) point out that the conventional parameterized expectations approach (PEA) presented in Chapter 3 is a particular projection method. Consider again the non-linear regression step 4 of algorithm 3.1.1. In this step, we solve the nonlinear-equation system (3.9), which we restate for the readers' convenience:

$$0 = \frac{-2}{T}\sum_{t=0}^{T-1}[\phi_j(\mathbf{u}_{t+1}(\boldsymbol{\gamma})) - \psi_j(\boldsymbol{\gamma}_j, \mathbf{x}_t(\boldsymbol{\gamma}))]\frac{\partial \psi_j}{\partial \gamma_{ij}}(\boldsymbol{\gamma}_j, \mathbf{x}_t(\boldsymbol{\gamma})),$$
$$\text{for all } i = 1, 2, \ldots, p, \text{ and } j = 1, 2, \ldots, k.$$

In this equation, the n-vector of states $\mathbf{x}_t \in X \subset \mathbb{R}^n$ contains all information relevant to predict the conditional expectation $\boldsymbol{\phi}(\mathbf{u}_{t+1})$. The function that approximates $\phi_j(\cdot)$ is given by $\psi_j(\cdot)$ and parameterized by $\boldsymbol{\gamma}_j$.

In order to generate the time series for $\mathbf{x}_t$ and for $\mathbf{u}_{t+1}$ in the stochastic growth model, for example, an exogenous process of the technology shock is generated. If the time horizon T is getting large, the probability distribution of the state variable $\mathbf{x}_t \in X \subset \mathbb{R}^n$ approaches the ergodic distribution, say $\pi(\mathbf{x})$, induced by the choice of $\boldsymbol{\phi}(\cdot)$ and the behavior of the exogenous stochastic expectations approach (PEA) can now readily be identified with a specific projection method. The residual is given by

$$R_j(\boldsymbol{\gamma}, \mathbf{x}) = \phi_j(\mathbf{u}_{t+1}(\boldsymbol{\gamma})) - \psi_j(\boldsymbol{\gamma}_j, \mathbf{x}_t(\boldsymbol{\gamma}))$$

and the weight for the i-th component of $\boldsymbol{\gamma}$ is simply

$$w^i(\mathbf{x}_t, \boldsymbol{\gamma}) = \pi(\mathbf{x}_t)\frac{\partial \psi_j}{\partial \gamma_{ij}}(\boldsymbol{\gamma}_j, \mathbf{x}_t(\boldsymbol{\gamma})).$$

In particular, the residual at point $\mathbf{x}_t \in X$ is also weighted by its probability. The Monte-Carlo simulation used in the conventional PEA more likely generates data points near the steady state for two reasons. First, since the innovations in the AR(1) process are drawn from a normal distribution, realizations far from the mean

of 0 are less likely. Second, the economy tends to return to the steady state after a shock.

This property of the Monte-Carlo simulation, however, also constitutes a major weakness of the conventional PEA. If we numerically approximate a function, we use nodes that are not concentrated in a certain area of the interval over which we want to approximate this function.[5] For example, with Chebyshev regression a relatively large proportion of the nodes is distributed close to the limits of the interval $[-1, 1]$. Therefore, we could get a much better quality of fit with much less fitting points if we modify the PEA accordingly. We will show this in 4.3.2, where we solve the stochastic growth model.

4.2 The Building Blocks of Projection Methods

In this subsection we consider the separate steps of Algorithm 4.1.1 in more detail. We begin with the choice of the family of functions.

4.2.1 Approximating Function

In the applications of Chapter 3 we use the family of monomials $\{1, x, x^2, ...\}$ to approximate the conditional expectations function. There, we already encountered the problem that the fitting step may fail due to the fact that higher degree monomials may be nearly indistinguishable from each other numerically. We can circumvent this problem by using a family of orthogonal polynomials, as, e.g., the Chebyshev polynomials described in Section 8.2.4. There are further considerations that make Chebyshev polynomials a prime candidate for projection methods. Some of them are mentioned in Section 8.2.4 and others will become obvious in the next paragraphs.

One further issue must be resolved at this step. Polynomials are single valued functions. So, how are we going to interpret the term $\varphi_i(\mathbf{x})$ in equation (4.6)? As we explain in Section 8.2.5, $\varphi_i(\mathbf{x})$, is

[5] See Section 8.2.

the i–th member of a so called product base, which consists of products of members of a family of polynomials. If $\mathbf{x} \in X \subset \mathbb{R}^n$, let $(k_1, k_2, \ldots k_n)$ denote the n-tuple of integers from the set $\mathcal{I} := \{0, 1, 2, \ldots, p\}$. Then:

$$\varphi_i(\mathbf{x}) := \prod_{j=1}^{n} \varphi_{k_j}(x_j).$$

Note that there are $(1+p)^n$ different n-tuples that can be built from the set $\mathcal{I}$. The respective product space is called the n-fold tensor product base. A smaller set, the complete set of polynomials, derives from the condition $\sum_{j=1}^{n} k_j = p$. For instance, if $n = 2$, this set consists of $(p+1)(p+2)/2$ members, whereas the tensor product base has $(1+p)^2$ members.

4.2.2 Residual Function

In many economic applications there are several ways to solve the model. For instance, in Chapter 3 we approximate the agent's conditional expectations function. Yet, we can solve the models considered there also by computing the agent's policy function. In some cases it is not always obvious, which way is best, and some experimentation with different solution concepts may be warranted. In other cases a particular solution may suggest itself on a priori reasons. In the stochastic growth model with a binding constraint on investment the agent's policy function will have a kink at the point where the constraint becomes binding. As we demonstrate in Section 8.2.4, it is difficult to approximate a kinked function with a linear combination of differentiable functions as the Chebyshev polynomials. Thus, in this case it is better to solve the model by computing the agent's conditional expectations function.

Even if we have decided on the function that we wish to approximate it is not always obvious how to define the residual function in step 3 of Algorithm 4.1.1. Consider the Euler equation of the deterministic growth model from (1.11):

$$0 = \frac{u'(C_t)}{u'(C_{t+1})} - \beta f'(K_{t+1}),$$
$$K_{t+1} = f(K_t) - C_t,$$

where C_t is consumption in period t, K_t the agent's stock of capital, $u'(\cdot)$ the marginal utility of consumption, and $f'(\cdot)$ the marginal productivity of capital. Assume we want to solve this model in terms of the policy function $C(K_t)$. Letting $\hat{C}(\boldsymbol{\gamma}, K)$ denote the approximate solution, the residual function may be computed from

$$R(\boldsymbol{\gamma}, K) := \frac{u'[\hat{C}(\boldsymbol{\gamma}, K)]}{u'[\hat{C}(\boldsymbol{\gamma}, f(K) - \hat{C}(\boldsymbol{\gamma}, K))]} - \beta f'[f(K) - \hat{C}(\boldsymbol{\gamma}, K)]. \tag{4.8}$$

Notice that by this formulation we do not put more weight on low asset values K (and, hence, low consumption C) with a corresponding high value of marginal utility because we form the fraction of current and next-period marginal utilities. However, if we chose the alternative residual function

$$R(\boldsymbol{\gamma}, K) := u'(\hat{C}(\boldsymbol{\gamma}, K)) - \beta u'[\hat{C}(\boldsymbol{\gamma}, f(K) - \hat{C}(\boldsymbol{\gamma}, K))] f'[f(K) - \hat{C}(\boldsymbol{\gamma}, K)]$$

small errors in the approximation of the true consumption function $C(K)$ would result in large residuals at low values of the capital stock K, while relatively larger deviations of the approximated function from the true solution for high values of K would result in a much smaller residual. As we aim to find a good uniform approximation of the policy function over the complete state space, we should be careful with respect to the choice of the residual function and rather use (4.8).

4.2.3 Projection and Solution

Depending on the choice of the projection function and the weight function this step may become more or less involved. Note that for $\mathbf{x} \in X \subset \mathbb{R}^n$ the shorthand $\int_X$ denotes the n-fold integral:

$$\int_X w(\mathbf{x})R(\boldsymbol{\gamma},\mathbf{x})g_i(\mathbf{x})d\mathbf{x} := \int_{\underline{x}_1}^{\bar{x}_1}\int_{\underline{x}_2}^{\bar{x}_2}\dots\int_{\underline{x}_n}^{\bar{x}_n} w(\mathbf{x})R(\boldsymbol{\gamma},\mathbf{x})g_i(\mathbf{x})dx_1dx_2\dots dx_n. \tag{4.9}$$

If the dimension of the state space is small, we can use one of several methods to compute numeric approximations to these integrals. For instance, in Section 8.3.2 we consider the Gauss-Chebyshev quadrature that replaces the integral by a weighted sum of m function values, computed at the zeros of the m-th degree Chebyshev polynomial. Suppose $\mathbf{x} = (x_1, x_2)$ so that the double integral is approximated by a double sum over the pairs (x_i, x_j), $i, j = 1, 2, \dots, m$. If we use $m = 100$ nodes to compute the integral, this amounts to adding up 10,000 elements. In general, using Gauss-Chebyshev quadrature to evaluate (4.9), requires $m^n - 1$ summations. In higher dimensional problems, the integration step can become a binding constraint. For instance, Heer and Maussner (2004b) use the Galerkin method to solve a multi-country, representative agent model. For eight countries with idiosyncratic productivity shocks the state space of this model has dimension $n = 16$. Even with only 3 nodes Gauss-Chebyshev quadrature requires 43,046,721 evaluations of the integrand. In this paper we employ an integration formula that uses $2^n + 2n + 1$ points. On a personal computer with Pentium III, 846 MHz processor it takes 14 days, 16 hours and 32 minutes to find the solution.

If we project the residual against the Dirac delta function, we circumvent the computation of integrals. Of course, this will save a lot of computer time if the state space is large. Instead, the task is to solve the non-linear equation system

$$R(\boldsymbol{\gamma}, \mathbf{x}_j) = 0, \quad j = 0, 1, \dots, p.$$

But at which set of points $\mathbf{x}_j$ should the residual function equal zero? It is well known from the so called Chebyshev interpolation theorem[6] that the Chebyshev zeros minimize the maximal interpolation error. For this reason, one should use the Chebyshev

[6] See, e.g., Judd (1998), Theorem 6.7.2, p. 221.

nodes of the Chebyshev polynomial of order $p+1$. This particular projection method is called Chebyshev collocation.

We have seen that the least squares projection derives from minimizing $\int_X R(\boldsymbol{\gamma}, \mathbf{x})^2 d\mathbf{x}$. Thus, instead of solving the set of $p+1$ non-linear equations

$$\int_X R(\boldsymbol{\gamma}, \mathbf{x}) \frac{\partial R(\boldsymbol{\gamma}, \mathbf{x})}{\partial \gamma_j} d\mathbf{x} = 0, \quad \forall j = 0, 1, \ldots, p,$$

we can also employ numerical optimization techniques to find the minimizer of $\int_X R(\boldsymbol{\gamma}, \mathbf{x})^2 d\mathbf{x}$. Otherwise we must compute the partial derivatives of the residual function either analytically or numerically. Depending on the structure of the problem, the latter approach – though not as accurate as the former – may be preferable since it requires less programming. What is required is just passing the function $R(\cdot)$ to a subroutine that returns the gradient of a user supplied function.

4.2.4 Accuracy of Solution

A first and simple to perform check of the accuracy of the solution is to compute the residuals $R(\boldsymbol{\gamma}, \mathbf{x})$ over a grid of points in X. To get an idea of how good your solution is, you must compare it to a second solution. This second solution could use different projection functions $\tilde{g}_i$ or a more accurate, but perhaps more time-consuming integration routine.

A second accuracy check is to simulate the model. From this simulation the DM-statistic (see Section 9.3) can be computed. Furthermore, the moments of the simulated time series may be compared to those obtained from a different solution method (see Heer and Maussner (2004a) on this approach).

In the case of Chebyshev polynomials, there is a third, easy to use criterium. From Theorem 8.2.3 we know that the coefficients γ_j drop off rapidly and that γ_p is small. If your solution does not displays this pattern, you should return to step 2 or even step 1 of Algorithm 4.1.1.

4.3 Applications

In this section, we present several applications. First, we compute the policy function of the deterministic growth model that we considered extensively in Section 1.2. The state space of this model is one-dimensional and consists of the capital stock K_t only. In the second example, we extend the analysis to a two-dimensional state space considering the stochastic growth model of example 1.3.3. Finally, we model asset pricing . Towards this end, we introduce habit persistence and adjustment costs of capital into the stochastic growth model. The state space of this model consists of the productivity shock, the stock of capital and past consumption.

4.3.1 The Deterministic Growth Model

The Model. In Section 1.2 we introduce the deterministic growth model. For your convenience, we restate the farmer's decision problem given in (1.9):

$$\begin{aligned}
&\max_{C_0,C_1,\ldots} \quad U_0 = \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\eta}-1}{1-\eta}, \quad \beta \in (0,1), \eta > 0, \\
&\text{s.t.} \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (4.10) \\
&\left.\begin{array}{rcll} K_{t+1}+C_t & \leq & K_t^{\alpha}+(1-\delta)K_t, & \alpha \in (0,1), \\ 0 & \leq & C_t, & \\ 0 & \leq & K_{t+1}, & \end{array}\right\} t = 0,1,\ldots, \\
&K_0 \quad \text{given},
\end{aligned}$$

where C_t is consumption in period t and K_t the farmer's stock of capital. Here, we assume that the current period utility function $u(C_t)$ has a constant elasticity of marginal utility with respect to consumption of $-\eta$. The production function K_t^{α} is of the Cobb-Douglas type and capital depreciates at the rate $\delta \in (0,1]$. As we know from Example 1.2.1 this model has an analytic solution in the case of $\eta = 1 \Leftrightarrow u(C_t) = \ln C_t$ and $\delta = 1$ given by $C_t = (1-\alpha\beta)K_t^{\alpha}$.

The Euler equation of this problem is given by:

$$\left[\frac{C_{t+1}}{C_t}\right]^{-\eta} \beta\left(1-\delta+\alpha K_{t+1}^{\alpha-1}\right)-1=0. \tag{4.11}$$

From this equation we derive the steady state value of the capital stock:

$$K^* = \left[\frac{\alpha\beta}{1-\beta(1-\delta)}\right]^{1/(1-\alpha)}.$$

Implementation. The state space of the problem is one-dimensional and consists of the capital stock K_t. In order to perform the computation, we need to specify an upper and lower bound for the state space. We proceed as in Section 1.2.5 and choose a state space $[\underline{K}, \overline{K}]$ that embeds the steady state value K^*. For illustrative purpose, we choose the steady state to be the middle point of the interval and consider $X := [0.5K^*, 1.5K^*]$. Depending on the nature of our problem, we might want to specify a smaller or larger interval. For example, if we consider the transition dynamics from an initial capital stock K_0, we may choose an interval $[\underline{K}, \overline{K}]$ that contains K_0 and K^*, and choose the borders $\underline{K}$ and $\overline{K}$ to be close to these values.

In the first and second step of Algorithm 4.1.1, we choose a family of functions for the approximation. In particular, we approximate the consumption function $C(K)$ with the help of a Chebyshev polynomial of order p,

$$\hat{C}(\boldsymbol{\gamma}, K) := \sum_{j=0}^{p} \gamma_j T_j(\tilde{K}(K)),$$

where $\tilde{K}(K)$ maps the capital stock K into the interval $[-1, 1]$ according to equation (8.36).

In step 3, we define the residual function $R(\boldsymbol{\gamma}, K)$. We argued in the previous subsection that it is best to use a version of the Euler equation, where the fraction of current and future marginal utility appears. For this reason we use (4.11) as our starting point. We compute the residual $R(\boldsymbol{\gamma}, K)$ in the following steps:

1) Given a parameter vector $\boldsymbol{\gamma}$ and $K_0 \in [\underline{K}, \overline{K}]$ we use Algorithm 8.2.1 to compute $\hat{C}_0 := \hat{C}(\boldsymbol{\gamma}, K_0)$. At this step we will

terminate the algorithm if the returned value of consumption is non-positive. This may occur if $\boldsymbol{\gamma}$ is far from the solution.

2) From $\hat{C}_0$ we compute the future stock of capital K_1 from the resource constraint:

$$K_1 = K^\alpha + (1-\delta)K - \hat{C}_0.$$

Here we must check whether K_1 is in $[\underline{K}, \overline{K}]$. If this condition is not satisfied, we must stop the algorithm: for values of K_1 outside the interval $[\underline{K}, \overline{K}]$ the transformed variable $\tilde{K}(K)$ is outside the interval $[-1, 1]$ where the Chebyshev polynomial is not defined. We know from Section 1.2 that the true solution converges to K^*. Therefore, the true policy function $C(K)$ always satisfies

$$\underline{K} < \underline{K}^\alpha + (1-\delta)\underline{K} - C(\underline{K}),$$
$$\overline{K} > \overline{K}^\alpha + (1-\delta)\overline{K} - C(\overline{K}).$$

Of course, a bad approximate solution needs not to satisfy this requirement. We invite you to discover what strange things can happen if you ignore this condition in the computation of a candidate solution. Just out-comment the respective line in the file `Ramsey2b.g`.

3) Given K_1 we use Algorithm 8.2.1 again to get $\hat{C}_1 := \hat{C}(\boldsymbol{\gamma}, K_1)$.

4) In this final step we compute the residual from

$$R(\boldsymbol{\gamma}, K_0) := \left[\frac{\hat{C}_1}{\hat{C}_0}\right]^{-\eta} \left(1 - \delta + \alpha K_1^{\alpha-1}\right) - 1.$$

The fourth step of Algorithm 4.1.1 concerns the projection method. The least squares method requires the minimization of

$$\int_{\underline{K}}^{\overline{K}} R(\boldsymbol{\gamma}, K)^2 dK$$

with respect to the parameter vector $\boldsymbol{\gamma}$. We use Gauss-Chebyshev quadrature (see equation (8.64)) and approximate this integral by the sum

$$S(\boldsymbol{\gamma}) := \frac{\pi(\overline{K} - \underline{K})}{2L} \sum_{l=1}^{L} R(\boldsymbol{\gamma}, K(\tilde{K}_l))^2 \sqrt{1 - \tilde{K}_l^2},$$

where $\tilde{K}_l$ are the zeros of the L-th degree Chebyshev polynomial and $K(\tilde{K}_l)$ is the transformation of these zeros to the interval $[\underline{K}, \overline{K}]$ given by equation (8.37). The minimization of $S(\boldsymbol{\gamma})$ via a quasi Newton algorithm requires good starting values. With bad initial values it is not possible to evaluate S at all. It turns out that it is no trivial task to pick admissible points from which Algorithm 8.6.3 converges. In the case where an analytic solution exists we get starting values from a regression of the analytic solution on a Chebyshev polynomial. For this purpose we employ Algorithm 8.2.2. In all other cases we use a genetic search routine [7] that provides an initial point for Algorithm 8.6.3.

For the Galerkin projection method we use again Gauss-Chebyshev quadrature. With this, we must solve the system of $p + 1$ non-linear equations:

$$\frac{\pi(\overline{K} - \underline{K})}{2L} \sum_{l=1}^{L} R(\boldsymbol{\gamma}, K(\tilde{K}_l)) T_i(\tilde{K}_l) \sqrt{1 - \tilde{K}_l^2}, \qquad (4.12)$$
$$i = 0, 1, \ldots, p.$$

The simplest method in terms of computer code required to specify the respective system of non-linear equations is the Chebyshev collocation method. Here, we determine the coefficients $\gamma_0, \ldots, \gamma_p$ from the conditions:

$$R(\boldsymbol{\gamma}, K(\tilde{K}_i)) = 0, \quad \forall i = 0, 2, \ldots, p, \qquad (4.13)$$

where, again, $\tilde{K}_i$ is the i-th zero of the Chebyshev polynomial of order $p + 1$.

To solve both the non-linear system (4.12) and (4.13), we use the modified Newton-Raphson algorithm with line search explained in Section 8.5. Again, it is difficult to find good initial values. Our short cut to solve this problem was to regress the analytic solution on a Chebyshev polynomial using Algorithm 8.2.2

[7] See the GAUSS procedure `GSearch1` in the file `Toolbox.src`.

if an analytic solution is available. Otherwise we use the solution returned by the search algorithm employed to minimize $S(\boldsymbol{\gamma})$.

Programs. The program `Ramsey2b.g` implements the different solutions. Table 4.1 displays the solutions for $p = 4$ in the case of logarithmic preferences and full depreciation. The parameter values that we used are $\alpha = 0.28$ and $\beta = 0.994$. The coefficients differ only in the fifth digit to the right of the decimal point. They drop off nicely so that we can be confident to have found a good solution. As can be seen from the last row of Table 4.1 the maximum absolute distance to the analytic solution is negligible. Interestingly, the collocation solution is closest to the true solution.

Table 4.1

Coefficient	Least Squares	Galerkin	Collocation
γ_0	2.39944	2.39944	2.39944
γ_1	0.51614	0.51614	0.51607
γ_2	−0.03455	−0.03459	−0.03472
γ_3	0.00509	0.00509	0.00520
γ_4	−0.00093	−0.00090	−0.00084
Distance	0.00096	0.00090	0.00111

With respect to computation time collocation is the fastest method, requiring just 0.04 seconds (on the same machine used to solve Example 1.2.1 in Section 1.2.5). The least squares solution needs about 1.17 seconds. The Galerkin solution is found in 1.62 seconds.

The yardstick that we use to determine what is a good solution in the case where no analytic solution exists is the maximum absolute distance to the solution from value function iteration with linear interpolation between grid points. For $\alpha = 0.28$, $\beta = 0.994$, $\eta = 2$, and $\delta = 0.011$ we used Algorithm 1.2.3 to compute a solution on a grid of thousand points. The last line in Table 4.2 shows that the Galerkin method delivers the best solution, followed by the least squares method.

4.3.2 The Stochastic Growth Model with Non-negative Investment

The Model. The stochastic Ramsey model with a non-negative investment constraint, Example 3.3.1, is the following planning problem: the representative agent solves

$$\max_{C_0} \quad E_0\left[\sum_{t=0}^{\infty}\beta^t\frac{C_t^{1-\eta}}{1-\eta}\right], \quad \eta>0,\ \beta\in(0,1),$$

s.t.

$$\left.\begin{array}{rcl} K_{t+1}+C_t &\leq& Z_tK_t^{\alpha}+(1-\delta)K_t,\ \alpha\in(0,1),\\ Z_t &=& Z_{t-1}^{\varrho}e^{\epsilon_t},\ \varrho\in(0,1),\ \epsilon_t\sim N(0,\sigma^2),\\ 0 &\leq& C_t,\\ 0 &\leq& K_{t+1}-(1-\delta)K_t. \end{array}\right\} \begin{array}{l} t=0,\\ 1,\ldots,\end{array}$$

$K_0,\ Z_0$ given,

The symbols have the usual interpretation: t is the time index, C, K, and Z denote consumption, the stock of capital and the level of productivity, respectively. The first-order conditions are stated in (3.13). They may be condensed into:

$$C_t^{-\eta}-\mu_t=\beta E_t\left[C_{t+1}^{-\eta}(1-\delta+\alpha Z_{t+1}K_{t+1}^{\alpha-1})-(1-\delta)\mu_{t+1}\right], \tag{4.14a}$$

$$K_{t+1}=\max\{Z_tK_t^{\alpha}+(1-\delta)K_t-C_t,(1-\delta)K_t\}. \tag{4.14b}$$

Implementation. The state space of this model consists of all pairs $(K_t,Z_t)\in\mathbb{R}^2$. To apply any of our projection methods, we must restrict this space to a bounded subset $X\subset\mathbb{R}^2$. Since the unconditional variance of the AR(1) process governing the log of the productivity shock, $z_t:=\ln(Z_t)$, is given by[8]

$$\operatorname{var}(z)=\frac{\sigma^2}{1-\varrho^2},$$

we choose a multiple of this variance, say λ, being large enough so that virtually all realizations of this process are in $[\underline{z},\overline{z}]$, where $\underline{z}=$

[8] See Section 9.2.

$-\lambda\sigma/\sqrt{1-\varrho^2}$ and $\overline{z} = -\underline{z}$. Alternatively, we may approximate the AR(1) process by a finite Markov chain as explained in Section 9.2. By this, we force the realizations of z_t to be taken from a finite grid of points. Yet another approach is to use a transformation that maps $(-\infty, \infty)$ into $[-1, 1]$ as, e.g., the hyperbolic tangens, $\tanh(x) := (e^x - e^{-x})/(e^x + e^{-x})$. The boundaries for the stock of capital are a fraction and a multiple, respectively, of the stationary stock of capital from the deterministic growth model. The latter is given by

$$K^* = \left[\frac{\alpha\beta}{1-\beta(1-\delta)}\right]^{\frac{1}{1-\alpha}}.$$

We must choose this interval large enough for K_t to stay in $[\underline{K}, \overline{K}]$. In our program we verify this condition during simulations so that we are able to adjust this interval upwards if necessary.

We have already noted in Section 4.2 that the consumption function in this model has a kink at the points where the non-negativity constraint binds. For this reason we approximate the conditional expectation on the rhs of (4.14a) by a Chebyshev polynomial. As in Section 3.3.1, we use the exponential of a polynomial in $\ln(Z)$ and $\ln(K)$, respectively:

$$\psi(\boldsymbol{\gamma}, Z, K) := \beta \exp\left(\sum_{i=0}^{p_1}\sum_{j=0}^{p_2} \gamma_{ij} T_i[X(\ln(Z))] T_j[X(\ln(K))]\right),$$

where X denotes the transformation of $\ln(z) \in [\underline{z}, \overline{z}]$ and $\ln(K) \in [\ln(\underline{K}), \ln(\overline{K})]$ to the interval $[-1, 1]$ as given by (8.36).

We derive the residual function from (4.14a) in the following steps.

1) Given a pair (Z, K), we compute

$$C = \psi(\boldsymbol{\gamma}, Z, K)^{-1/\eta}$$

and check the non-negativity constraint: if

$$K_1 := ZK^\alpha + (1-\delta)K - C \geq (1-\delta)K,$$

we accept this choice, else we put

$$C = ZK^{\alpha},$$

$$K_1 = (1-\delta)K.$$

If $K_1 \in [\underline{K}, \overline{K}]$, we proceed to the next step; otherwise we stop and try a different vector $\boldsymbol{\gamma}$.

2) Let $K_1(Z,K)$ denote the solution from the first step and let Z_1 be the size of next period's shock. We solve for

$$C_1 = \psi(\boldsymbol{\gamma}, Z_1, K_1(Z,K))^{-1/\eta}.$$

Again, we must check whether this solution violates the non-negativity constraint on investment. Thus, if

$$K_2 = Z_1 K_1^{\alpha} + (1-\delta)K_1 - C_1 \geq (1-\delta)K_1$$

we accept and put $\mu_1 = 0$, else we set

$$C_1 = Z_1 K_1^{\alpha},$$
$$\mu_1 = C_1^{-\eta} - \psi(\boldsymbol{\gamma}, Z_1, K_1).$$

Given this, we can compute the expression

$$g(Z_1, Z, K) := C_1^{-\eta}(1 - \delta + \alpha Z_1 K_1^{\alpha-1}) - (1-\delta)\mu_1.$$

3) Now, observe that the conditional expectation on the rhs of (4.14a) is given by

$$\beta \int_{\underline{Z}}^{\overline{Z}} g(Z_1, Z, K)\text{prob}(Z_1|Z)dZ_1, \quad \underline{Z} = e^{\underline{z}},\ \overline{Z} = e^{\overline{z}},$$

where $\text{prob}(Z_1|Z)$ denotes the probability of Z_1 given Z. Since $Z_1 = e^{\varrho z + \epsilon}$ and since the probability density function of ϵ is given by the normal distribution, we may also write

$$\phi(Z,K) := \beta \int_{\underline{z}}^{\overline{z}} g(e^{\varrho z+\epsilon}, Z, K)(2\pi\sigma)^{-1/2} e^{-(1/2\sigma)\epsilon^2} d\epsilon.$$

In our program we compute this integral via Gauss-Chebyshev quadrature of the function to the right of the integral sign.
As an alternative, we may approximate the AR(1) process by a Markov chain with m states. Thus, if $Z = e^z$ is the i-th element of the grid $\mathscr{Z} = \{z_1, z_2, \ldots, z_m\}$ and if p_{ij}, $i, j = 1, 2, \ldots, m$ denotes the probability of moving from i to j, the conditional probability on the rhs of (4.14a) is approximated by

$$\phi(Z, K) := \beta \sum_{j=1}^{m} g(Z_j, Z_i, K) p_{ij}.$$

In any way, we are now finished, and the residual function is defined by

$$R(\boldsymbol{\gamma}, Z, K) = \phi(Z, K) - \psi(\boldsymbol{\gamma}, Z, K).$$

The Program. The GAUSS program `Ramsey3d.g` (and the various procedures in the file `Toolbox.src`) as well as the Fortran program `Ramsey3d.for` implement the solution. The user has several options. First, you may choose between two different projection methods: least squares and Chebyshev collocation. Second, you can decide between a complete polynomial or a product base polynomial. Third, you can either use a finite state Markov chain or a normal distribution tailored to $[\underline{z}, \overline{z}]$. Fourth, you can provide an initial choice of the parameter vector $\boldsymbol{\gamma}_0$ to start the algorithms or you can apply the genetic algorithm programmed in `GSearch1` to find an appropriate starting point. There are many other details of the various algorithms that the user may change, as, e.g., the boundaries of the intervals for the stock of capital and the log of the productivity shock, the number of elements of the Markov chain, or the number of nodes used in the Gauss-Chebyshev quadrature to compute the least squares objective function.

The program computes the DM-statistic defined in (9.15) to check the accuracy of the found solution. Towards this end it generates a number of artificial time series and counts the number of instances where the DM-statistic is below (above) the 2.5 (97.5) critical value of the respective χ^2 distribution.

Results. Table 4.2 presents the results from a few experiments that we conducted. For reasons of comparison with Table 3.1 we report the results for a second degree complete polynomial in $(\ln(Z), \ln(K))$ as well as for a few product base polynomials.

Table 4.2

Simulation	σ_C	DM-Stat	
$\sigma = 0.0072$			
d=2	0.02240	2.8	2.8
$d_1 = 3, d_2 = 3$	0.02195	2.4	1.6
$d_1 = 4, d_2 = 3$	0.02194	2.2	3.0
$\sigma = 0.05$			
d=2	0.15114	1.6	3.8
$d_1 = 3, d_2 = 3$	0.15488	0.6	6.4
$d_1 = 4, d_2 = 3$	0.15683	2.2	4.6

Notes: σ is the standard deviation of the innovations of the productivity shock, d denotes the degree of the complete polynomial used to approximate $\psi(\cdot)$, d_1 (d_2) is the degree of the product base polynomial with respect to z (K), σ_C is the standard deviation of consumption, DM-Stat: the percentage of simulations (out of 500) whose DM-statistic is below (above) the 2.5 (97.5) percentile of the $\chi^2(11)$ distribution.

The results are from the collocation solution. We restricted the log of the productivity shock to an interval that is six times the size of the unconditional standard deviation of the AR(1) process for $\ln(z)$. This guaranteed that in all our simulations the realized values of $\ln(z)$ stayed in this interval. Very often we used homotopy to initialize our non-linear equations solver `FixVMN`. A genetic search routine provided the first set of acceptable initial values for a first degree complete polynomial. To find the solutions for higher order polynomials we used the solution for the lower degree polynomial with zeros in the places of yet unknown elements. This worked just fine.

Compared to the PEA reported in Table 3.1 there is no noteworthy difference in accuracy in terms of the DM-statistic. For our benchmark values of $\alpha = 0.27$, $\beta = 0.994$, $\eta = 2$, $\delta = 0.011$,

$\varrho = 0.9$, and $\sigma = 0.0072$ about 2.8 percent out of 500 simulations have a DM-statistic below the 2.5 critical value of the $\chi^2(11)$ distribution and 2.8 percent are above the 97.5 critical value. For the PEA solution figures are 2.0 and 2.6. For the larger shocks, $\sigma = 0.05$, the PEA solution delivered 4.8 percent of the computed DM-statistics either in the lower or in the upper tail of the $\chi^2(11)$ distribution, as compared to 5.4 percent for the collocation method. As can be seen from the fourth and eighth line of Table 4.2, higher degree product base polynomials do not always provide a better solution than a smaller complete polynomial.

Our experiments revealed no clear answer as to what type of projection method provides the best solution. In some cases the least squares solution performed better in others the collocation solution delivered better results. However, the least squares solution requires much more computer time.

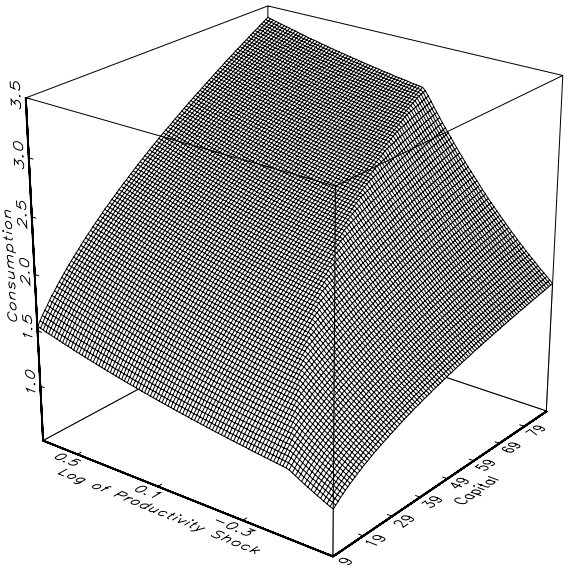

Figure 4.2: Polynomial Approximation of the Consumption Function $C(K)$ in the Deterministic Growth Model

As we already know from the PEA solution of the stochastic growth model, the non-negativity constraint does not bind for the low value of $\sigma = 0.0072$. In the case of $\sigma = 0.05$, however, the kink in the policy function for consumption is clearly visible in Figure 4.2, which displays the solution in the case of Chebyshev collocation for a 4×4 product base polynomial. Out of 500 simulations 4.8 percent where either in the 2.5-lower or 97.5-upper tail of the $\chi^2(11)$ distribution, indicating a very accurate solution.

4.3.3 The Equity Premium Puzzle

The Puzzle. One of the most regarded puzzles in the theory of financial economics is the equity premium puzzle: Why has the average real return on stocks in the US been six percentage points higher than the return on US Treasury Bills over the last century?[9] In this chapter, we present a model of asset pricing in a production economy based on the work of Jermann (1998).

The model is an extension of the stochastic growth model that you are, by now, most familiar with. In the latter model the expression

$$R_t := \alpha Z_t K_t^{\alpha-1} - \delta$$

(i.e., the marginal product of capital less the rate of depreciation) is the net return on one unit of output invested in the capital stock of a representative firm. We also know from this model that the household's lifetime utility does not change, if she trades one unit of consumption today against

$$\frac{u'(C_t)}{\beta E_t u'(C_{t+1})} = \frac{\Lambda_t}{\beta E_t \Lambda_{t+1}}$$

units of consumption tomorrow. Thus, the household is willing to pay

[9] An excellent overview of this issue is provided by Kocherlakota (1996). In their recent work, Jagannathan et al. (2001) argue that the equity premium has declined significantly in the last two decades and is likely to remain at a lower level on average as the transaction costs for trading stocks have been reduced substantially.

$$p_t := \beta E_t \frac{\Lambda_{t+1}}{\Lambda_t}$$

for a bond that promises one unit of consumption tomorrow for certain. For this reason we can use

$$r_t := \frac{1}{p_t} - 1 \equiv \frac{\Lambda_t}{\beta E_t \Lambda_{t+1}} - 1$$

as a measure of the risk free rate of return. Note that the time subscript t in this definition refers to the date on which the return becomes known. The return materializes in period $t+1$ when the bond pays one unit of consumption. The mean equity premium in the simple stochastic growth model is $E(R_{t+1} - r_t)$, where, as usual, $E(\cdot)$ denotes the unconditional mathematical expectation taken over the probability distribution of (Z, K, C).

In the simple stochastic growth model with less than full depreciation there is not much variation in the marginal product of capital, since investment is only a small portion of the stock of capital. One way to raise the variability of the stock of capital is to provide further incentives for investment. For instance, if the household's current period utility depends not only on current but also on past consumption, its desire to smooth consumption increases. This is usually referred to as habit persistence. A second way to obtain more variation in the return on equity is to allow for a variable price of shares. In the simple stochastic growth model the price of capital in terms of consumption goods is constant and equal to one, since it is possible to consume the stock of capital. The most common way to allow for a variable price of capital goods is to introduce adjustment costs.

In the following we extend the simple stochastic growth model along these two lines. We consider a decentralized economy inhabited by a continuum of identical households of mass one and a continuum of identical firms of the same size.

Households. The representative household provides one unit of labor to firms and earns the competitive real wage w_t. As a shareholder she is entitled to receive dividends d_t per unit of stocks of the representative firm S_t. The current price of stocks in terms

of the consumption good is v_t. Thus, total income is given by $w_t + d_t S_t$. Since stocks are the single asset of this economy, the current period budged constraint is:

$$v_t(S_{t+1} - S_t) \leq w_t + d_t S_t - C_t. \tag{4.15}$$

The household's current period utility function is specified in the following way:

$$u(C_t, C_{t-1}) := \frac{(C_t - bC_{t-1})^{1-\eta} - 1}{1-\eta}, \quad b \in [0,1),\ \eta > 0. \tag{4.16}$$

Habit persistence occurs if $b > 0$, otherwise we get the standard iso-elastic current period utility function. The household maximizes expected lifetime utility

$$E_0 \sum_{t=0}^{\infty} \beta^t \frac{(C_t - bC_{t-1})^{1-\eta} - 1}{1-\eta}, \quad \beta \in (0,1),$$

subject to (4.15) and given the initial number of shares S_0. Employing the techniques presented in Section 1.3 we derive the following first-order conditions:

$$\Lambda_t = (C_t - bC_{t-1})^{-\eta} - \beta b E_t (C_{t+1} - bC_t)^{-\eta}, \tag{4.17a}$$

$$\Lambda_t = \beta E_t \Lambda_{t+1} R_{t+1}, \tag{4.17b}$$

$$R_t := \frac{d_t + v_t}{v_{t-1}}. \tag{4.17c}$$

The term R_t gives the current period (ex post) gross rate of return on equities. As usual, Λ_t is the Lagrange multiplier of the budget constraint, and E_t denotes expectations conditional on information available at the beginning of the current period t.

Firms. The representative firm uses labor services N_t and capital services K_t to produce output according to

$$Y_t = Z_t N_t^{1-\alpha} K_t^{\alpha}, \tag{4.18}$$

where Z_t measures the level of total factor productivity. There are two ways to finance investment expenditures I_t: by issuing equities $v_t(S_{t+1} - S_t)$ or out of retained earnings RE_t:

$$I_t = v_t(S_{t+1} - S_t) + RE_t. \tag{4.19}$$

We assume that investment expenditures of size I_t result in net additions to the firm's current capital stock K_t as given by

$$K_{t+1} - K_t = \phi(I_t/K_t)K_t - \delta K_t, \quad \delta \in (0, 1], \tag{4.20}$$

where $\phi(I_t/K_t)$ is an increasing, concave function of its argument. The case $\phi(I_t/K_t) \equiv I_t/K_t$ specifies the capital accumulation equation in the standard model. Here we use the following parameterization of this function:

$$\phi(I_t/K_t) := \frac{a_1}{1-\zeta}\left(\frac{I_t}{K_t}\right)^{1-\zeta} + a_2, \quad \zeta > 0, \tag{4.21}$$

so that further increases of the capital stock become more and more costly in terms of the required investment expenditures. This is what we mean when we speak of adjustment costs of capital. The dividends payed to the firm's shareholders equal profits less any retained earnings:

$$d_t S_t = \Pi_t - RE_t. \tag{4.22}$$

Profits equal revenues Y_t less labor costs $w_t N_t$:

$$\Pi_t = Y_t - w_t N_t. \tag{4.23}$$

To motivate the firm's objective function we consider the deterministic case first. The value of the firm at the beginning of time $t + 1$ is given by

$$V_{t+1} = v_t S_{t+1}.$$

Using (4.19), (4.22), and (4.23) this may be expressed as

$$\begin{aligned} V_{t+1} = v_t S_{t+1} &= v_t S_t + I_t - RE_t, \\ &= v_t S_t + I_t + d_t S_t - \Pi_t, \\ &= \left(\frac{d_t + v_t}{v_{t-1}}\right) v_{t-1} S_t - (Y_t - w_t N_t - I_t), \end{aligned}$$

or, using the definition of R_t in (4.17c), as

$$V_{t+1} + (Y_t - w_t N_t - I_t) = R_t V_t.$$

Iterating this equation forward beginning with $t = 1$ and ending with $t = T$ yields:

$$V_1 = \frac{V_T}{R_1 R_2 \dots R_{T-1}} + \sum_{t=1}^{T-1} \frac{Y_t - w_t N_t - I_t}{R_1 R_2 \dots R_t}.$$

For the present value of the firm V_1 to be finite if T tends to infinity requires:

$$\lim_{T \to \infty} \frac{V_T}{R_1 R_2 \dots R_{T-1}} = 0.$$

In effect, this condition rules out speculative bubbles. Thus, we end up with the following formula for the present value of the firm:

$$V_1 = \sum_{t=1}^{\infty} \varrho_t [Y_t - w_t N_t - I_t], \quad \varrho_t := \frac{1}{R_1 R_2 \cdots R_t}. \tag{4.24}$$

Note that the firm is not able to choose its discount factor. In equilibrium the household sector requires a return on equities given by

$$R_t = \frac{\Lambda_t}{\beta \Lambda_{t+1}}, \tag{4.25}$$

which follows from (4.17b) in the case of no aggregate uncertainty. As a consequence, the firm's value depends upon the sequence of cash flows and the sequence of shadow prices Λ_t but not upon the firm's dividend policy.[10]

The firm aims at maximizing its present value (4.24) subject to (4.20). The respective Lagrangean for this problem is

[10] This is not generally true. Here it follows since we neglect income and corporate taxes. See Turnovsky (2000), 292ff.

$$\mathscr{L} = \sum_{t=1}^{\infty} \varrho_t \Big\{ Z_t N_t^{1-\alpha} K_t^{\alpha} - w_t N_t - I_t + q_t \big[\phi(I_t/K_t)K_t + (1-\delta)K_t - K_{t+1}\big] \Big\},$$

where q_t is the period t value of the Lagrange multiplier attached to the constraint (4.20).

The first-order conditions for the optimal choice of N_t, I_t, and K_{t+1} are:

$$w_t = (1-\alpha) Z_t N_t^{-\alpha} K_t^{\alpha}, \tag{4.26a}$$

$$q_t = \frac{1}{\phi'(I_t/K_t)}, \tag{4.26b}$$

$$q_t = \frac{1}{R_{t+1}} \Big\{ \alpha Z_{t+1} N_{t+1}^{1-\alpha} K_{t+1}^{\alpha-1} - (I_{t+1}/K_{t+1}) + q_{t+1} \big[\phi(I_{t+1}/K_{t+1}) + 1 - \delta\big] \Big\}. \tag{4.26c}$$

In addition, the transversality condition

$$\lim_{t\to\infty} \varrho_t q_t K_{t+1} = 0 \tag{4.26d}$$

must hold.

The first condition determines labor input in the usual way and deserves no further comment, except that it implies

$$\Pi_t := Y_t - w_t N_t = \alpha Z_t N_t^{1-\alpha} K_t^{\alpha} = \alpha Y_t. \tag{4.27}$$

Given q_t, the shadow value of an additional unit of new capital in terms of the firm's output, the second equation can be solved for the optimal amount of investment expenditures I_t. We want adjustment costs of capital to play no role in the deterministic stationary state of the model. This has two consequences: q^* must equal one and I^* must equal δK^*. Using (4.21) and (4.26b) the first condition requires us to choose

$$a_1 = \delta^{\zeta}.$$

Via the second condition this in turn implies:

$$a_2 = \frac{-\zeta}{1-\zeta}\delta.$$

It is easy to see that for $\zeta > 1$ condition (4.26b) implies $q_t \to 0 \Rightarrow I_t \to 0$. Thus, in this case, there is always a solution featuring $q_t, I_t > 0$. Using (4.27) condition (4.26c) may be rewritten as

$$q_t = \frac{1}{R_{t+1}}\Big\{Y_{t+1} - w_{t+1}N_{t+1} - I_{t+1} + q_{t+1}K_{t+2}\Big\}\frac{1}{K_{t+1}}.$$

Iterating on this equation delivers

$$q_0 K_1 = \sum_{t=1}^{T} \varrho_t (Y_t - w_t N_t - I_t) + \varrho_T q_T K_{T+1}.$$

Taking the limit for $T \to \infty$, invoking the transversality condition (4.26d) and comparing the result to the definition of the period $t = 1$ value of the firm, establishes

$$q_0 K_1 = V_1.$$

Since the choice of the current period is arbitrary, we have just shown that $V_{t+1} = q_t K_{t+1}$, i.e., q_t is the ratio of the firm's stock market value to the replacement costs of its capital stock.

This result carries over to the stochastic case to which we turn next. Since we have already seen that the management of the firm has to use the household's marginal valuation of wealth Λ_t, we define the expected present value of the firm in the following way:

$$V_1 = E_0 \sum_{t=1}^{\infty} \beta^t \frac{\Lambda_t}{\Lambda_0}\left(Z_t N_t^{1-\alpha} K_t^{\alpha} - w_t N_t - I_t\right).$$

Proceeding in a way analogous to Section 1.3.2 we can derive first-order conditions. With regard to optimal labor input and the optimal amount of investment these conditions are equal to (4.26a) and (4.26b), respectively. The condition with respect to K_{t+1} is the obvious modification of (4.26c):

$$\begin{aligned} q_t = \beta E_t \frac{\Lambda_{t+1}}{\Lambda_t}\Big(&\alpha Z_{t+1}(K_{t+1}/N_{t+1})^{\alpha-1} - \frac{I_{t+1}}{K_{t+1}} \\ &+ q_{t+1}\left[\phi(I_{t+1}/K_{t+1}) + 1 - \delta\right]\Big). \end{aligned} \tag{4.26c'}$$

Market Equilibrium. Since the size of the household sector is one and since leisure is not an argument of the household's utility function, total labor supply always equals unity: $N_t \equiv 1$. The household's budget constraint (4.15) together with definitions (4.19), (4.22), and (4.23) implies the economy's resource constraint:

$$Z_t K_t^{\alpha} = C_t + I_t. \tag{4.28}$$

Thus, the model's dynamics is governed by the stochastic Euler equations (4.17a) and (4.26c'), the capital accumulation equation (4.20), the resource constraint (4.28), and the pricing condition (4.26b).

Deterministic Stationary State. Remember that we have assumed that adjustment costs of capital play no role in the deterministic stationary state of the model, i.e., $q^* = 1$ and $\phi(I^*/K^*) = \delta$. Using this and $\Lambda_{t+1} = \Lambda_t = \Lambda^*$ in the Euler equation (4.26c') implies:

$$K^* = \left[\frac{\alpha\beta}{1-\beta(1-\delta)}\right]^{1/1-\alpha}.$$

Hence, the deterministic stationary stock of capital coincides with the solution for the same variable in the simple deterministic growth model. From the resource constraint we get

$$C^* = (K^*)^{\alpha} - \delta K^*,$$

where we assume $Z \equiv 1$. Finally, the Euler equation (4.17a) delivers the stationary value of Λ :

$$\Lambda^* = (1-\beta b)[(1-b)C^*]^{-\eta}.$$

Solution. We can solve for all period t variables, if we know q_t and Λ_t. These variables in turn depend upon the predetermined variables K_t and C_{t-1} and the level of total factor productivity Z_t. Therefore, the state space of our model is a subspace of $\mathbb{R}^3$, given by

$$X := [\underline{Z}, \overline{Z}] \times [\underline{K}, \overline{K}] \times [\underline{C}, \overline{C}],$$

for suitable lower and upper bounds on Z_t, K_t, and C_{t-1}. In the following, we use a Markov chain approximation of the stochastic process for $\ln Z_t = \rho \ln Z_{t-1} + \epsilon_t$ that allows us to compute the transition probability π_j from any $Z \in [\underline{Z}, \overline{Z}]$ to $Z_j \in Z := \{Z_1, Z_2, \ldots, Z_m\}$ using the formula given under Step 2 in Algorithm 9.2.1. As in the previous subsection we choose to solve the model by looking for polynomials in (Z, K, C) for the conditional expectations that appear on the rhs of the Euler equations (4.17a) and (4.26c'). In the case of equation (4.26c') this is tantamount to approximating q. Since both q and Λ are non-negative, we use exponentials of complete polynomials denoted by $\psi^1(\boldsymbol{\gamma}^1, Z, K, C)$ and $\psi^2(\boldsymbol{\gamma}^2, Z, K, C)$. The parameters of these functions are collected in the yet to be determined vectors $\boldsymbol{\gamma}^1$ and $\boldsymbol{\gamma}^2$. We compute the residual functions R^1 and R^2 in the following steps.

1) Given a triple (Z, K, C) we compute

$$q_1 = \psi^1(\boldsymbol{\gamma}^1, Z, K, C).$$

Using equation (4.26b) and our parametrization of ϕ given in (4.21), we find

$$I_1 = K(a_1 q_1)^{1/\zeta}.$$

The resource constraint (4.28) delivers

$$C_1 = ZK^\alpha - I_1,$$

and from the capital accumulation equation (4.20) we get

$$K_2 = \phi(I_1/K)K + (1-\delta)K.$$

Finally, we compute Λ_1 from

$$\Lambda_1 = (C_1 - bC)^{-\eta} - \beta b \psi^2(\boldsymbol{\gamma}^2, Z, K, C).$$

Before we proceed, we must check whether C_1 and K_2 are in $[\underline{C}, \overline{C}]$ and $[\underline{K}, \overline{K}]$.

2) Let π_j denote the transition probability from state Z to state Z_j. For all $j = 1, 2, \ldots, m$ we repeat the previous sequence of computations using the solutions found in the first step:

$$
\begin{aligned}
q_2 &= \psi^1(\boldsymbol{\gamma}^1, Z_j, K_2, C_1), \\
I_2 &= K_2(a_1 q_2)^{1/\zeta}, \\
C_2 &= Z_j K_2^{\alpha} - I_2, \\
\Lambda_2 &= (C_2 - bC_1)^{-\eta} - \beta b \psi^2(\boldsymbol{\gamma}^2, Z_j, K_2, C_1), \\
x_{1j} &:= \beta(\Lambda_2/\Lambda_1)\left[\alpha Z_j K_2^{\alpha-1} - (I_2/K_2) + q_2(\phi(I_2/K_2) + 1 - \delta)\right], \\
x_{2j} &:= (C_2 - bC_1)^{-\eta}.
\end{aligned}
$$

3) Finally, the residual from the Euler equation (4.26c') is computed from

$$R^1(\boldsymbol{\gamma}, Z, K, C) := q_1 - \sum_{j=1}^{m} \pi_j x_{1j},$$

and the residual from the Euler equation (4.17a) equals

$$R^2(\boldsymbol{\gamma}, Z, K, C) := \psi^2(\boldsymbol{\gamma}^2, Z, K, C) - \sum_{j=1}^{m} \pi_j x_{2j},$$

where $\boldsymbol{\gamma} := [\boldsymbol{\gamma}^1, \boldsymbol{\gamma}^2]'$.

We use the collocation method that determines $\boldsymbol{\gamma}^1$ and $\boldsymbol{\gamma}^2$ so that the residuals R^1 and R^2 are equal to zero at the zeros of the respective Chebyshev polynomials.

Notice that we do not need to solve for the share price v_t appearing in the definition of the return on equities $R_{t+1} := (d_{t+1}+v_{t+1})/v_t$. We have shown above that the firm's total value at $t+1$ is given by $V_{t+1} = q_t K_{t+1}$ and that it does not depend on the firm's dividend policy. The share price $v_t := V_{t+1}/S_{t+1}$, however, and thus, the return on equity, depends on the dividend policy. In the following we assume that the firm finances its investment expenditures entirely out of retained earnings. Therefore, $S_{t+1} = S_t$ for all t, and we are free to normalize the number of outstanding shares to one. Consequently, we can compute R_{t+1} from

$$R_{t+1} = \frac{\alpha Y_{t+1} - I_{t+1} + q_{t+1} K_{t+2}}{q_t K_{t+1}}.$$

The conditional expectation $E_t \Lambda_{t+1}$ that appears in the definition of the risk free rate $r_t := \Lambda_t / (\beta E_t \Lambda_{t+1})$ is computed in a manner analogous to steps 2) and 3) above.

Implementation. The Fortran program `Epp.for` as well as the GAUSS program `Epp.g` implement the solution. It was not an easy task to find a good solution. We had to use a combination of stochastic search and homotopy. For our benchmark set of parameters ($\alpha = 0.27$, $\beta = 0.994$, $\delta = 0.011$, $\eta = 2.0$, $\rho = 0.90$, and $\sigma = 0.0072$) we started with a first or second degree polynomial and used stochastic search to find acceptable starting values for our non-linear equations solver. Again, we used our own routine, since it keeps track of the various error flags that signal when C or K are outside their respective bounds or when other error conditions (e.g., $\Lambda < 0$) occur. Having found a solution, we increased the degree of both polynomials by one and used the previously found solution with zeros in the places of the yet unknown coefficients as starting values. We proceeded in this way until the solution was sufficiently accurate. A simple way to determine this is to see whether the polynomial for q evaluated at the deterministic stationary values of C and K is very close to one. To find the solution for different parameter values, we started with the final solution for the benchmark set and increased the selected parameter in small steps.

Two more points are worth mentioning. The first concerns the choice of nodes, the second the choice of the state space.

1) Consider, for example, a complete polynomial of degree 2 in three variables. This polynomial has 10 parameters. Thus, we must choose 10 points in the three dimensional space $X := [\underline{Z}, \overline{Z}] \times [\underline{K}, \overline{K}] \times [\underline{C}, \overline{C}]$. Chebyshev collocation uses the zeros of the Chebyshev polynomial of degree three adjusted to the respective subinterval, i.e., to $[\underline{Z}, \overline{Z}]$, $[\underline{K}, \overline{K}]$, or $[\underline{C}, \overline{C}]$. We can combine these zeros to $3^3 = 27$ different three tuples $(\tilde{Z}, \tilde{K}, \tilde{C})$. Yet, we need only 10 points from this set to determine the parameters. Our program chooses the points implicitly via the ordering of the variables. The

parameters in the vector $\boldsymbol{\gamma}^j$, $j = 1, 2$, are determined by the requirement $R(\boldsymbol{\gamma}^j, Z(i_1), K(i_2), C(i_3)) = 0$, where the indices i_1, i_2, and i_3 run over $1, 2, \ldots, p+1$ and obey $i_1+i_2+i_3 \leq 1, 2, \ldots, p+1$.

2) Since we use only a few points to determine the parameters of both polynomials, it seems natural to use the state space as small as possible. Yet, then it may easily happen that one of the conditions $C \in [\underline{C}, \overline{C}]$ or $K \in [\underline{K}, \overline{K}]$ becomes violated. Our solution to this problem rests on the following observation. Remember, if C or K are outside their boundaries, we are not able to evaluate the respective Chebyshev polynomial. Yet, this does not require to approximate the solution on the same state space. Usually the model economy stays in a small neighborhood of the deterministic stationary state. What we need is a good approximation on this subspace. Therefore, if we want to approximate the solution on the compact space $X_1 := [\underline{Z}, \overline{Z}] \times [\underline{K}_1, \overline{K}_1] \times [\underline{C}_1, \overline{C}_1]$, we are nevertheless free to choose a sufficiently larger domain $X_2 := [\underline{Z}, \overline{Z}] \times [\underline{K}_2, \overline{K}_2] \times [\underline{C}_2, \overline{C}_2]$ for the Chebyshev polynomials.

A guess for appropriate boundaries of X_1 can rely on the following: in the deterministic model and for an arbitrary level of total factor productivity Z the capital stock converges to

$$K(Z) := \left[\frac{\alpha\beta Z}{1-\beta(1-\delta)}\right]^{1/(1-\alpha)},$$

and consumption approaches

$$C(Z) := ZK(Z)^\alpha - \delta K(Z).$$

Thus, one can try $\underline{K} = K(\underline{Z})$ and $\overline{K} = K(\overline{Z})$ (and similarly for C) where the boundaries for Z are chosen in the usual way (i.e., as a multiple of the unconditional variance of the AR(1) process for the log of Z). Of course, in simulations of the model, we must check whether the state variables leave X_1.

Results. Table 4.3 displays the results from a few simulations of the model. The equity return $R - 1$ and the equity premium $E(R - r)$ are averages of 500 simulations. In each simulation the

Table 4.3

	$\sigma = 0.0072$			$\sigma = 0.01$		
	$b = 0.1$	$b = 0.5$	$b = 0.8$	$b = 0.1$	$b = 0.5$	$b = 0.8$
$R - 1$	0.62	0.71	1.27	0.64	0.81	1.85
$E(R - r)$	0.02	0.11	0.70	0.03	0.22	1.28

$\sigma = 0.01$, $b = 0.8$

	$\eta = 3$	$\eta = 4$	$\eta = 5.0$
$R - 1$	2.02	2.52	2.61
$E(R - r)$	1.50	1.90	2.12

Notes: The entries are averages over 500 simulations. Each simulation computes the return on equity R and the equity premium $E(R - r)$ as average over 120 periods.
If not noted otherwise in the table, the parameter values are $\alpha = 0.27$, $\beta = 0.994$, $\delta = 0.011$, $\eta = 2.0$, $\rho = 0.90$.

respective magnitudes are obtained from averaging over 120 periods. The results clearly indicate the importance of habit persistence for the model's ability to generate a noticeable equity premium. If habit persistence plays only a minor role, $b = 0.1$, the difference between the risk free rate and the return on equity is negligible. For the benchmark parameter set we need $b = 0.8$ to find an equity premium of 0.7 percentage points. Since our model is calibrated to a period length of one quarter, this equals an annual equity premium of about 2.8 percentage points. The premium increases with the variance of the productivity shock. The value of $\sigma = 0.01$ implies that the variance of output is slightly larger than its empirical counterpart. In this case, the model delivers an equity premium of about 1.3 percentage points per quarter. The last row indicates that this rate increases with the risk aversion parameter η. For $\eta = 5$, $b = 0.8$, and $\sigma = 0.01$ we find an equity premium of around 2.0 percentage points. Thus, the model is in principle able to predict an equity premium of about 6 percentage points per annum.

The ability of the model to predict a sizeable equity premium is sensitive to the modelling of labor supply. If labor supply is endogenous, agents can smooth consumption over time quite effectively with the help of their labor supply and, in this case,

the equity premium is substantially smaller again. Boldrin et al. (2001) consider a two-sector model where labor is immobile between the two sectors within periods. In this model, they are able to replicate the average equity premium.

Problems

4.1 Consider the following discrete time version of Lucas' (1988) model of growth through human capital accumulation. In this deterministic model the social planner solves the following problem:

$$\max \sum_{t=0}^{\infty} \beta^t \frac{C_t^{1-\eta} - 1}{1-\eta}, \quad \beta \in (0,1),\ \eta > 0,$$

subject to

$$K_{t+1} = (u_t H_t)^{\alpha} K_t^{1-\alpha} + (1-\delta)K_t - C_t,$$
$$H_{t+1} = A(1-u_t)H_t + (1-\delta)H_t, \quad A > 0,$$
$$K_0, H_0 \text{ given.}$$

Here C_t is consumption in period t, K_t the stock of capital, H_t the stock of human capital. The size of the working population N is normalized to 1 so that u_t is the fraction of human capital adjusted labor $H_t N$ devoted to the production of output. The state variables of this model are physical capital K_t and human capital H_t. The control variables are consumption C_t and the fraction of hours spent in the production of output u_t.
In the steady state of this model all variables grow at the rate $g_H = A(1-u^*) + (1-\delta)$, where u^* is the steady state value of u_t. Therefore, variables that are stationary (and, thus, remain within a compact space) are, for instance, $k_t := K_t/H_t$, $c_t := C_t/H_t$, and $h_{t+1} := H_{t+1}/H_t$.
Use projection methods to approximate the functions $c(k_t, h_t)$ and $u(k_t, h_t)$. We propose the following values of the model's parameters: $\alpha = 0.27$, $\beta = 0.994$, $\eta = 2$, $\delta = 0.011$. Choose A so that the steady-state growth rate is 0.005 per quarter. Compute the transitional dynamics of the model for both an economy with a relative shortage of physical and a relative shortage of human capital. Is there any difference?

4.2 Use projection methods and the parameters from Table 1.1 to solve the benchmark business cycle model from Example 1.4.1. Simulate the model and compare the second moments to those found in Tables 1.2, 2.1, and 3.2. (Hint: Heer and Maussner (2004) use the Galerkin method to solve this model. See this paper for details of the computations.)

4.3 Consider the model of Section 4.3.3.
a) Use the Galerkin method to compute an approximate solution to the model of Section 4.3.3. To speed up computations you may also try to use the Gauss-Hermite integration formula (8.65) for the computation of the conditional expectations in step 3 (see page 230).
b) In our computation of the equity premium, we assume that the firm finances its investment expenditures entirely out of retained earnings. Consider the opposite dividend policy: investment expenditures are

solely financed by issuing new equities. Compute the equity premium for this alternative specification.

c) Modify the instantaneous utility function of the household to include leisure:

$$u(C_t, C_{t-1}, 1-N_t) := \frac{(C_t - bC_{t-1})^{(1-\eta)}(1-N_t)^{\theta(1-\eta)} - 1}{1-\eta}$$

and solve this model. Are you still able to produce a sizeable equity premium?

4.4 Consider the following model with a variable utilization rate of capital u_t and a second shock that represents exogenous variations in the price of imported oil p_t (this is adapted from Finn (1995)). The representative agent solves

$$\max E_0 \sum_{t=0}^{\infty} \beta^t \left[\ln C_t + \theta \ln(1-N_t)\right], \quad \beta \in (0,1),\ \theta > 0,$$

subject to

$$K_{t+1} = (Z_t N_t)^{\alpha}(u_t K_t)^{1-\alpha} + (1-\delta(u_t))K_t - C_t - p_t Q_t,$$

$$\delta(u_t) := \frac{u_t^{\gamma}}{\gamma},$$

$$\frac{Q_t}{K_t} = \frac{u_t^{\zeta}}{\zeta},$$

$$\ln Z_t = \ln Z + \ln Z_{t-1} + \epsilon_t^Z, \quad \epsilon_t^Z \sim n(0, \sigma^Z),$$

$$\ln p_t = \rho^p \ln p_{t-1} + \epsilon_t^p, \quad \epsilon_t^p \sim n(0, \sigma^p),$$

$$K_0 \text{ given.}$$

As usual, C_t denotes consumption in period t, N_t are working hours, K_t is the stock of capital, and Q_t it the quantity of oil imported at the price of p_t. A more intense utilization of capital increases the amount of energy required per unit of capital. Thus, if the price of oil rises, capital utilization will decrease. Verify this claim as follows.

In this model, labor augmenting technical progress follows a random walk with drift rate $\ln Z$. Define the following stationary variables $c_t := C_t/Z_t$, $k_t := K_t/Z_{t-1}$, and $z_t := Z_t/Z_{t-1}$. The state variables of the model are k_t, z_t, and p_t. Solve the model for the consumption function $(C_t/Z_t) = c(k_t, z_t, p_t)$. Given this solution, compute the time path of the utilization rate of capital for a one-time oil price shock of the size of one standard deviation of ϵ^p. Use the following parameter values taken from Finn (1995): $\beta = 0.9542$, $\theta = 2.1874$, $\alpha = 0.7$, $\gamma = 1.4435$, $\zeta = 1.7260$, $\rho^p = 0.9039$, $\sigma^p = 0.0966$, $Z = 1.0162$, $\sigma^Z = 0.021$.

Part II

Heterogeneous Agent Models

Chapter 5

Computation of Stationary Distributions

Overview. This chapter introduces you to the modelling and computation of heterogeneous-agent economies. In this kind of problem, we have to compute the distribution of the individual state variable(s). While we focus on the computation of the stationary equilibrium in this chapter, you will learn how to compute the dynamics of such an economy in the next chapter.

The representative agent framework has become the standard tool for modern macroeconomics. It is based on the intertemporal calculus of the household that maximizes life-time utility. Furthermore, the household behaves rationally. As a consequence, it is a natural framework for the welfare analysis of policy actions. In the following chapters, we will consider an important extension of the deterministic Ramsey model. In the remaining part of the book, agents are no longer homogeneous and cannot be represented by a single agent.

There are many different characteristics by which agents may differ. Agents are different with regard to their abilities, their education, their age, their marital status, their number of children, their wealth holdings, to name but a few. In the introductory section of this chapter, we model the real life feature that some agents are employed, while others are unemployed. For simplicity, we assume that the agent cannot influence his employment probability, e.g. by searching harder for a new job or asking for a lower wage. In addition, agents cannot insure against the idiosyncratic risk of being unemployed. Accordingly, agents in our economy differ with regard to their employment status and their employment history. Those agents who were lucky and have been employed for many years are able to save more and build up more wealth

than their unlucky contemporaries who have been unemployed for longer periods of time. As a consequence, agents also differ with regard to their wealth. Besides, all agents are equal. In the second part of this chapter, we will compute the stationary distribution of the individual state variables. In the final section, we present two prominent applications from macroeconomic theory, the risk-free rate of interest puzzle and the distributional effects of a switch from an income tax to a consumption tax. In addition, we give you a short survey of the modern literature on the theory of income distribution.

5.1 A Simple Heterogeneous-Agent Model with Aggregate Certainty

In Section 1.2, we present the deterministic infinite horizon Ramsey problem. The model considers a representative household that, at the same time, is also the producer. In Section 1.4.3, we show that the equilibrium of this economy is equivalent to the one of a decentralized economy and that the fundamental theorems of welfare economics hold. In this section, we consider heterogeneity at the household level, but keep the simplifying assumption that all firms are equal and, hence, can act as a representative firm. As a consequence, we most conveniently formulate our model in terms of a decentralized economy and study the behavior of the households and the firm separately.

As a second important characteristic of our model, we only consider idiosyncratic risk. In our economy, households can become unemployed and cannot insure against this risk. However, there is no aggregate uncertainty. For example, the technology is deterministic. As you will find out, the economy will display a long-run behavior that is easily amenable to computational analysis. In the *stationary equilibrium* of the economy, the distribution of the state variable, the aggregate wage and the aggregate interest rate are all constant, while the employment status and the wealth level of the individual households vary.[1]

[1] Aggregate uncertainty will be introduced into the heterogeneous-agent extension of the Ramsey model in Chapter 6.

In our simple model, three sectors can be distinguished: households, production, and the government. Households maximize their intertemporal utility subject to their budget constraint. In order to insure against the risk of unemployment, they build up precautionary savings during good times. Firms maximize profits. The government pays unemployment compensation to the unemployed agents that is financed by an income tax. We will describe the behavior of the three sectors in turn.

Households. The economy consists of many infinitely lived individuals. In particular, we consider a continuum of agents of total mass equal to one.[2] Each household consists of one agent and we will speak of households and agents interchangeably. Households differ only with regard to their employment status and their asset holdings. Households maximize their intertemporal utility

$$E_0 \sum_{t=0}^{\infty} \beta^t u\left(c_t\right), \tag{5.1}$$

where $\beta < 1$ is the subjective discount factor and expectations are conditioned on the information set at time 0. At time zero, the agent knows his beginning-of-period wealth a_0 and his employment status $\epsilon_0 \in \{e, u\}$. If $\epsilon = e$ ($\epsilon = u$), the agent is employed (unemployed). The agent's instantaneous utility function is twice continuously differentiable, increasing and concave in his consumption c_t and has the following form:

$$u(c_t) = \frac{c_t^{1-\eta}}{1-\eta}, \quad \eta > 0, \tag{5.2}$$

where η, again, denotes the coefficient of relative risk aversion. In the following, lowercase letters denote individual variables and uppercase letters denote aggregate variables. For example, c_t is individual consumption, while C_t is aggregate consumption in the economy. We, however, keep the notation that real prices are denoted by lower case letters, while nominal prices are denoted by upper case letters.

[2] This amounts to assume that the number of individual households is infinite and, if we index the household with $i \in [0, 1]$, the probability that $i \in [i_0, i_1]$ is simply $i_1 - i_0$.

Agents are endowed with one indivisible unit of time in each period. If the agent is employed ($\epsilon = e$) in period t, he earns gross wage w_t. If the agent is unemployed ($\epsilon = u$) in period t, he receives unemployment compensation b_t. We will assume that $(1-\tau)w_t > b_t$, where τ denotes the income tax rate. The individual-specific employment state is assumed to follow a first-order Markov chain. The conditional transition matrix is given by:

$$\pi(\epsilon'|\epsilon) = Prob\left\{\epsilon_{t+1} = \epsilon'|\epsilon_t = \epsilon\right\} = \begin{pmatrix} p_{uu} & p_{ue} \\ p_{eu} & p_{ee} \end{pmatrix}. \tag{5.3}$$

where, for example, $Prob\left\{\epsilon_{t+1} = e|\epsilon_t = u\right\} = p_{ue}$ is the probability that an agent will be employed in period $t+1$ given that the agent is unemployed in period t. Households know the law of motion of the employment status ϵ_t.

In our economy, unemployment is exogenous. We have not modelled any frictions which might be able to explain this feature. In this regard, we follow Hansen and İmrohoroğlu (1992) in order to simplify the exposition and the computation. Of course, it would be straightforward to introduce endogenous unemployment into this model. For example, various authors have used search frictions in the labor market in order to explain unemployment with the help of either endogenous search effort as in Costain (1997) or Heer (2003) or endogenous separation from the firms as in Den Haan et al. (2000). In addition, we assume that there are no private insurance markets against unemployment and unemployed agents only receive unemployment compensation from the government.[3]

The household faces the following budget constraint

$$a_{t+1} = \begin{cases} (1 + (1-\tau)r_t)\, a_t + (1-\tau)w_t - c_t, & \text{if } \epsilon = e \\ (1 + (1-\tau)r_t)\, a_t + b_t - c_t, & \text{if } \epsilon = u, \end{cases} \tag{5.4}$$

[3] One possible reason why there are no private insurance markets against the risk of unemployment is moral hazard. Agents may be reluctant to accept a job if they may receive generous unemployment compensation instead. Chiu and Karni (1998) show that the presence of private information about the individual's work effort helps to explain the failure of the private sector to provide unemployment insurance.

where r_t denotes the interest rate in period t. Interest income and wage income are taxed at rate τ. Each agent smoothes his consumption $\{c_t\}_{t=0}^{\infty}$ by holding the asset a. An agent accumulates wealth in good times ($\epsilon = e$) and runs it down in bad times ($\epsilon = u$). As a consequence, agents are also heterogeneous with regard to their assets a. We impose the asset constraint $a \geq a_{min}$, so that households cannot run down their assets below $a_{min} \leq 0$.

The first-order condition of the household that is not wealth-constrained can be solved by introducing the Lagrange multiplier λ and setting to zero the derivatives of the Lagrangean expression

$$\mathcal{L} = E_0 \sum_{t=0}^{\infty} \Big\{ \beta^t \left[u(c_t) + \lambda_t \left(1_{\epsilon_t = u} b_t + (1 + (1-\tau) r_t) \right. \right. \\ \left. \left. + 1_{\epsilon_t = e} (1-\tau) w_t + a_t - a_{t+1} - c_t \right) \right] \Big\}$$

with respect to c_t and a_{t+1}. $1_{\epsilon_t = e}$ ($1_{\epsilon_t = u}$) denotes an indicator function that takes the value one if the agent is employed (unemployed) in period t and zero otherwise. The first-order condition for the employed and unemployed agent in period t is

$$\frac{u'(c_t)}{\beta} = E_t \left[u'(c_{t+1}) (1 + (1-\tau) r_t) \right]. \tag{5.5}$$

The solution is given by the policy function $c(\epsilon_t, a_t)$ that is a function of the employment status ϵ_t and the asset holdings a_t in period t. In particular, the policy function is independent of calendar time t. Together with (5.4), the policy function $c(\epsilon_t, a_t)$ also gives next-period asset holdings $a_{t+1} = a'(\epsilon_t, a_t)$.

Production. Firms are owned by the households and maximize profits with respect to their labor and capital demand. Production $F(K_t, N_t)$ is characterized by constant returns to scale using capital K_t and labor N_t as inputs:

$$F(K_t, N_t) = K_t^{\alpha} N_t^{1-\alpha}, \quad \alpha \in (0,1). \tag{5.6}$$

In a market equilibrium, factors are compensated according to their marginal products and profits are zero:

$$r_t = \alpha \left(\frac{N_t}{K_t} \right)^{1-\alpha} - \delta, \tag{5.7a}$$

$$w_t = (1-\alpha) \left(\frac{K_t}{N_t} \right)^{\alpha}, \tag{5.7b}$$

where δ denotes the depreciation rate of capital.

Government. Government expenditures consists of unemployment compensation B_t which are financed by a tax on income. The government budget is assumed to balance in every period:

$$B_t = T_t, \tag{5.8}$$

where T_t denotes government revenues.

Stationary Equilibrium. First, we will analyze a stationary equilibrium. We may want to concentrate on the stationary equilibrium, for example, if we want to analyze the long-run effects of a permanent change in the government policy, e.g. a once-and-for-all change in the unemployment compensation b. In a stationary equilibrium, the aggregate variables and the factor prices are constant and we will drop the time indices if appropriate, e.g. for the aggregate capital stock K or the interest rate r and the wage w. Furthermore, the distribution of assets is constant for both the employed and unemployed agents, and the numbers of employed and unemployed agents are constant, too. The individual agents, of course, are not characterized by constant wealth and employment status over time. While we focus on a stationary distribution in this chapter, we will also analyze 1) the transition dynamics for given initial distribution of the assets to the stationary distribution and 2) the movement of the wealth and income distribution over the business cycle in the next chapter.

For the description of the stationary equilibrium, we need to describe the heterogeneity in our economy. In this book, we use a very pragmatic and simple way to define the stationary equilibrium. In particular, we only use basic concepts from probability theory and statistics which all readers should be familiar

with, namely the concept of a distribution function.[4] In the stationary equilibrium, the distribution of assets is constant and we will refer to it as either the stationary, invariant or constant distribution. In our particular model, we are aiming to compute the two distribution functions of the assets for the employed and unemployed agents, $F(e,a)$ and $F(u,a)$, respectively. The corresponding density functions are denoted by $f(e,a)$ and $f(u,a)$, respectively. The individual state space consists of the sets $(\epsilon,a) \in \mathcal{X} = \{e,u\} \times [a_{min},\infty)$.

The concept of a stationary equilibrium uses a recursive representation of the consumer's problem. Let $V(\epsilon,a)$ be the value of the objective function of a household characterized by productivity ϵ and wealth a. $V(\epsilon,a)$ for the benchmark government policy is defined as the solution to the dynamic program:

$$V(\epsilon,a) = \max_{c,a'} \left[u(c) + \beta E\left\{V(\epsilon',a')\right\}\right], \tag{5.9}$$

subject to the budget constraint (5.4), the government policy $\{b,\tau\}$, and the stochastic process of the employment status ϵ as given by (5.3).[5]

A stationary equilibrium for a given set of government policy parameter b is a value function $V(\epsilon,a)$, individual policy rules $c(\epsilon,a)$ and $a'(\epsilon,a)$ for consumption and next-period capital, respectively, a time-invariant density of the state variable $x = (\epsilon,a) \in \mathcal{X}$, $f(e,a)$ and $f(u,a)$, respectively, time-invariant

[4] A description of more general heterogeneous-agent economies might necessitate the use of more advanced concepts from measure theory. Since the algorithms and solution methods developed in this chapter do not require a thorough understanding of measure theory and should already be comprehensible with some prior knowledge of basic statistics, we dispense of an introduction into measure and probability theory. For a more detailed description of the use of measure theory in recursive dynamic models please see Stokey and Lucas (1989).

[5] The solution obtained by maximizing (5.1) s.t. (5.4) and (5.3) corresponds to the solution obtained by solving (5.9) s.t. (5.4) and (5.3) under certain conditions on the boundedness of the value function $V(\cdot)$ (Theorems 4.2, 4.3, 9.2, and 9.4 in Stokey and Lucas (1989)). This correspondence has been called the *Principle of Optimality* by Richard Bellman.

relative prices of labor and capital $\{w, r\}$, and a vector of aggregates K, N, C, T, and B such that:

a) Factor inputs, consumption, tax revenues, and unemployment compensation are obtained aggregating over households:

$$K = \sum_{\epsilon \in \{e,u\}} \int_{a_{min}}^{\infty} a \; f(\epsilon, a) \;\; da, \tag{5.10a}$$

$$N = \int_{a_{min}}^{\infty} f(e, a) \;\; da, \tag{5.10b}$$

$$C = \sum_{\epsilon \in \{e,u\}} \int_{a_{min}}^{\infty} c(\epsilon, a) \; f(\epsilon, a) \;\; da, \tag{5.10c}$$

$$T = \tau(wN + rK), \tag{5.10d}$$

$$B = (1 - N)b. \tag{5.10e}$$

b) $c(\epsilon, a)$ and $a'(\epsilon, a)$ are optimal decision rules and solve the household decision problem described in (5.9).

c) Factor prices (5.7a) and (5.7b) are equal to the factors' marginal productivities, respectively.

d) The goods market clears:

$$F(K, L) + (1 - \delta)K = C + K' = C + K. \tag{5.11}$$

e) The government budget (5.8) is balanced: $T = B$.

f) The distribution of the individual state variable (ϵ, a) is stationary:

$$F(a', \epsilon') = \sum_{\epsilon \in \{e,u\}} \pi(\epsilon'|\epsilon) \; F\left(a'^{-1}(a', \epsilon), \epsilon\right) \tag{5.12}$$

for all $(a', \epsilon') \in \mathcal{X}$. Here, $a'^{-1}(a', \epsilon)$ denotes the inverse of the function $a'(a, \epsilon)$ with respect to its first argument a.[6] Accordingly, the distribution over states $(\epsilon, a) \in \mathcal{X}$ is unchanging.

[6] In particular, we assume that $a'(a, \epsilon)$ is invertible. As it turns out, $a'(a, \epsilon)$ is invertible in our example economy in this chapter. In Section 5.2., we will also discuss the changes in the computation of the model that are necessary if $a'(a, \epsilon)$ is not invertible. This will be the case if the non-negativity constraint on assets is binding.

Calibration. As we will often use the model as an example in the subsequent sections, we will already assign numerical values to its parameters in this introductory part. Following İMROHOROĞLU (1989), periods are set equal to six weeks ($\approx 1/8$ of a year). Preferences and production parameters are calibrated as commonly in the dynamic general equilibrium models. In particular, we pick the values $\alpha = 0.36$ and $\eta = 2.0$. Our choice of $\beta = 0.995$ implies a real annual interest rate of approximately 4% before taxes. The employment probabilities are set such that the average duration of unemployment is 2 periods (=12 weeks) and average unemployment is 8%.[7] The employment transition matrix is given by:

$$\begin{pmatrix} p_{uu} & p_{ue} \\ p_{eu} & p_{ee} \end{pmatrix} = \begin{pmatrix} 0.500 & 0.500 \\ 0.0435 & 0.9565 \end{pmatrix}. \tag{5.13}$$

The non-capital income of the unemployed household b amounts to 1.199 and is set equal to one fourth of the steady-state gross wage rate in the corresponding representative agent model,[8] where the gross interest rate is equal to the inverse of the discount factor β and, therefore, the capital stock amounts to $K = (\alpha/(1/\beta - 1+\delta))^{1/(1-\alpha)}N$. In the literature, the ratio of unemployment compensation to net wage income is also called the replacement ratio which will be approximately equal to 25.6% in our model. In addition, the income tax rate is determined endogenously in the computation with the help of the balanced budget rule. Finally, the annual depreciation rate is set equal to 4% implying a six-week depreciation rate of approximately 0.5%.

5.2 Computation of the Stationary Equilibrium of a Heterogeneous Agent Economy

With only very few exceptions, dynamic heterogeneous-agent general equilibrium models do not have any analytical solution

[7] Notice that unemployed agents stay unemployed with probability 0.5. As a consequence, the average duration of unemployment is simply 1/0.5=2 periods. In Section 9.2, you will learn how to compute the stationary unemployment rate from the employment transition matrix.

[8] In such a model, the 'representative' household consists of $(1 - N)$ unemployed workers and N employed workers.

or allow for the derivation of analytical results. Algorithms to solve heterogeneous-agent models with an endogenous distribution have only recently been introduced into the economic literature. Notable studies in this area are AIYAGARI (1994, 1995), DEN HAAN (1997), HUGGETT (1993), İMROHOROĞLU ET AL. (1995), KRUSELL and SMITH (1998) or RÍOS-RULL (1999). We will use Example 5.2.1 as an illustration for the computation of the stationary equilibrium of such an economy.

Example 5.2.1

Consider the following stationary distribution:

a) Households are allocated uniformly on the unit interval $[0,1]$ and are of measure one. The individual household maximizes

$$V(\epsilon,a) = \max_{c,a'} \left[\frac{c^{1-\eta}}{1-\eta} + \beta E\left\{V(\epsilon',a')\right\} \right],$$

s.t.

$$a' = \begin{cases} (1+(1-\tau)r)\,a + (1-\tau)w - c & \epsilon = e, \\ (1+(1-\tau)r)\,a + b - c & \epsilon = u, \end{cases}$$

$$a \geq a_{min},$$

$$\pi(\epsilon'|\epsilon) = Prob\left\{\epsilon_{t+1} = \epsilon' | \epsilon_t = \epsilon\right\} = \begin{pmatrix} p_{uu} & p_{ue} \\ p_{eu} & p_{ee} \end{pmatrix}.$$

b) The distribution of (ϵ, a) is stationary and aggregate capital K, aggregate consumption C, and aggregate employment N are constant.

c) Factors prices are equal to their respective marginal products:

$$\begin{aligned} r &= \alpha \left(\frac{N}{K}\right)^{1-\alpha} - \delta, \\ w &= (1-\alpha)\left(\frac{K}{N}\right)^{\alpha}. \end{aligned}$$

d) The government budget balances: $B = T$.

e) The aggregate consistency conditions hold:

$$K = \sum_{\epsilon \in \{e,u\}} \int_{a_{min}}^{\infty} a \; f(\epsilon, a) \; da,$$
$$N = \int_{a_{min}}^{\infty} f(e, a) \; da,$$
$$C = \sum_{\epsilon \in \{e,u\}} \int_{a_{min}}^{\infty} c(\epsilon, a) \; f(\epsilon, a) \; da,$$
$$T = \tau(wN + rK),$$
$$B = (1 - N)b.$$

The computation of the solution of Example 5.2.1 consists of two basic steps, the computation of the policy function and the computation of the invariant distribution. For this reason, we will apply several elements of numerical analysis that we introduced in the first part of this book. In order to solve the individual's optimization problem, we need to know the stationary factor prices and the tax rate. For a given triplet $\{K, N, \tau\}$, we can use the methods of Chapters 2-4 in order to compute the individual policy functions $c(\epsilon, a)$ and $a'(\epsilon, a)$. The next step is the basic new element that you have not encountered in the computation of representative agent economies. We need to compute the distribution of the individual state variables, aggregate the individual state variables, and impose the aggregate consistency conditions. The complete solution algorithm for Example 5.2.1 is described by the following steps:

Algorithm 5.2.1 (Computation of Example 5.2.1)

Purpose: *Computation of the stationary equilibrium.*

Steps:

Step 1: Compute the stationary employment N.
Step 2: Make initial guesses of the aggregate capital stock K and the tax rate τ.
Step 3: Compute the wage rate w and the interest rate r.
Step 4: Compute the household's decision functions.

Step 5: Compute the stationary distribution of assets for the employed and unemployed agents.
Step 6: Compute the capital stock K and taxes T that solve the aggregate consistency conditions.
Step 7: Compute the tax rate τ that solves the government budget.
Step 8: Update K and τ and return to step 2 if necessary.

In Step 1, we compute the stationary employment N. In our simple Example 5.2.1, employment N_t does not depend on the endogenous variables w_t, r_t, or the distribution of assets a_t in period t. N_t only depends on the number of employed in the previous period N_{t-1}. Given employment N_{t-1} in period $t-1$, we know that next-period employment is simply the sum of the lucky unemployed agents who find a job and the lucky employed agents that keep their job

$$N_t = p_{ue}(1 - N_{t-1}) + p_{ee}N_{t-1}. \tag{5.15}$$

Given any employment level N_0 in period 0, we can iterate over (5.15) for $t = 1, 2, \ldots$. In fact, if we use the probabilities $p_{ue} = 0.50$ and $p_{ee} = 0.9565$ from (5.13) and iterate some ten to twenty times for any given employment level $N_0 \in (0, 1)$, the percentage of employed people in the economy, or equally, the number of employed, N_t, converges to the so-called stationary employment $N = 0.92$. In essence, we are computing the invariant distribution of a simple 2-state Markov Chain. There are, however, more efficient methods in order to compute the stationary values of a Markov chain process and we describe them in more detail in Section 9.2.

In Step 5, we compute the stationary distribution of assets for the employed and unemployed workers. The wealth distribution is continuous and, hence, is an infinite-dimensional object that can only be computed approximately. Therefore, in general, we apply other methods for its computation than in the case of a finite-state Markov chain. Three different kinds of methods are presented in order to compute the invariant distribution $f(\epsilon, a)$ of the heterogeneous-agent model. First, we will compute the distribution function on a discrete number of grid points over the

assets. Second, we will use Monte-Carlo simulations by constructing a sample of households and tracking them over time. And third, a specific functional form of the distribution function will be assumed and we will use iterative methods to compute the approximation.

Discretization of the Distribution Function. We first consider a method which relies upon the discretization of the state space. Our individual state space consists of two dimensions, the employment status ϵ and the wealth level a. However, the first state variable ϵ can only take two different values, $\epsilon \in \{e, u\}$, so that we only need to discretize the second state variable, the asset level a. Assume that we choose a grid over the state space with m points. If the policy function has been computed with the help of methods that rely upon the discretization of the state space, for example discrete value function approximation, we want to choose a finer grid for the computation of the state space following Ríos-Rull (1999). Denote the distribution function by $F(\epsilon, a)$ and the density function by $f(\epsilon, a)$.

If we discretize the distribution function, the state variable (ϵ, a) can only take a discrete number of values $2 \cdot m$. In this case, we are in essence trying to compute the Markov transition matrix between these states (ϵ, a). For the computation of the transition matrix between employment state ϵ, we presented several methods in the previous section. These methods are not all applicable for the computation of the transition matrix between the states (ϵ, a). In particular, with current computer technology, we will run into problems using the procedure `equivec1.g` to compute the ergodic distribution due to the curse of dimensionality because the Markov transition matrix has $(2m)^2$ entries. For reasonable values of grid points $2m$, we have a storage capacity problem and GAUSS, for example, will be unable to compute the ergodic matrix.[9] In the following we will present two iterative methods that rely upon the discretization of the state space

[9] The transition matrix between the $2m$ states mainly consists of zero entries, i.e. the matrix is sparse. As a consequence, we may still be able to apply the procedure `equivec1.g`; however, we have to change the computer code applying sparse matrix methods. In essence, we only store the non-zero

in order to compute the discretized invariant distribution function. Algorithm 5.2.2 computes the invariant distribution function based on the equilibrium condition (5.12), while Algorithm 5.2.3 computes the invariant density function.

Algorithm 5.2.2 (Computation of the Invariant Distribution Function $F(\epsilon, a)$)

Purpose: *Computation of the stationary equilibrium.*

Steps:

Step 1: *Place a grid on the asset space $\mathcal{A} = \{a_1 = a_{min}, a_2, \ldots, a_m = a_{max}\}$ such that the grid is finer than the one used to compute the optimal decision rules.*

Step 2: *Choose an initial piecewise distribution function $F_0(\epsilon = e, a)$ and $F_0(\epsilon = u, a)$ over the grid. The vectors have m rows each.*

Step 3: *Compute the inverse of the decision rule $a'(\epsilon, a)$.*

Step 4: *Iterate on*

$$F_{i+1}(\epsilon', a') = \sum_{\epsilon=e,u} \pi(\epsilon', \epsilon) F_i \left(a'^{-1}(\epsilon, a'), \epsilon \right) \tag{5.16}$$

on grid points (ϵ', a').

Step 5: *Iterate until F converges.*

The Algorithm 5.2.1 that computes the stationary equilibrium of the heterogeneous-agent economy 5.2.1 and the Algorithm 5.2.2 that computes the invariant distribution function are implemented in the GAUSS program `ch5_disf.g`. The individual policy functions are computed with the help of value function iteration with linear interpolation as described in Chapter 1. We compute the value function at $n = 200$ equidistant grid points a_j in the interval $[-2, 3000]$. The interval is found by some trial and error. The interval, of course, should contain the steady state capital stock of the corresponding representative agent economy,

entries. GAUSS, for example, provides sparse matrix commands in a source file named `sparse.src`.

$K = (\alpha/(1/\beta - 1 + \delta))^{1/(1-\alpha)} N = 247.6$. We would also love to choose an ergodic set so that once the individual's capital stock is inside the set, it stays inside the interval. As it turns out, this interval is rather large and we choose the smaller interval $[-2, 3000]$ instead. In the stationary equilibrium, all employed agents have strictly positive net savings over the complete interval $[-2, 3000]$. However, the number of agents that will have assets exceeding 1500 is extremely small. In fact, fewer than 0.01% of the agents have assets in the range of $[1500, 3000]$ so that we can be very confident that our choice of the interval is not too restrictive. The reason for the low number of very rich people is the law of large numbers. We simulate the economy over 25,000 periods or more and sooner or later, the employed agents will loose their job and start decumulating their wealth again.

After we have computed the individual policy function $a'(\epsilon, a)$ for given capital stock K, unemployment compensation b, and income tax τ, we compute the invariant distribution function according to Algorithm 5.2.2. In step 1, we choose an equidistant grid with $m = 3n = 600$ points on $[-2, 3000]$ for the computation of the distribution function.[10] In step 2, we initialize the distribution function with the equal distribution so that each agent has the steady-state capital stock of the corresponding representative agent economy.

In step 3, we compute the inverse of the policy function $a'(\epsilon, a)$. $a = a'^{-1}(\epsilon, a_j)$, $j = 1, \ldots, m$, over the chosen grid. Since the unemployed agent with low wealth may want to spend all his wealth and accumulate debt equal or exceeding $-a_{min}$, a' may not be invertible when $a' = a_{min}$. For this reason, we define $a'^{-1}(\epsilon, a_{min})$ as the maximum a such that $a'(\epsilon, a) = a_{min}$.[11] Furthermore, the computation of $a'(\epsilon, a)$ involves some type of interpolation, as $a'(\epsilon, a)$ is stored for only a finite number of values $n < m$. We use linear interpolation for the computation of $a'(\epsilon, a)$ for $a_j < a < a_{j+1}$.

In step 4, the invariant distribution is computed. F is computed for every wealth level $a' = a_j$, $j = 1, \ldots, m$, and $\epsilon = e, u$. In the

[10] The grid over the asset space for the value function and the distribution function do not need to be equally spaced.

[11] Huggett (1993) establishes that a' is strictly nondecreasing in a.

computation, we impose two conditions: 1) If $a'^{-1}(\epsilon, a_j) < a_{min}$, $F(\epsilon, a_j) = 0$, and 2) if $a'^{-1}(\epsilon, a_j) \geq a_m$, $F(\epsilon, a_j) = g(\epsilon)$, where $g(\epsilon)$ denotes the ergodic distribution of the employment transition matrix. The first condition states that the number of employed (unemployed) agents with a current-period wealth below a_{min} is equal to zero. The second condition states that the number of the employed (unemployed) agents with a current-period wealth equal to or below a_{max} is equal to the number of all employed (unemployed) agents. In addition, as there may be some round-off errors in the computation of the next-period distribution $F_{i+1}(\epsilon', a')$, we normalize the number of all agents equal to one and multiply $F_{i+1}(e, a')$ and $F_{i+1}(u, a')$ by $0.92/F_{i+1}(e, a_{max})$ and $0.08/F_{i+1}(u, a_{max})$, respectively. Again, we need to use an interpolation rule, this time for the computation of $F_i(\epsilon, a)$. In (5.16), $a_0 = a'^{-1}(\epsilon, a_j)$, $j = 1, \ldots, m$, need not be a grid point. As we have only stored the values of $F_i(\epsilon, a_0)$ for grid points $a = a_j$, $j = 1, \ldots, m$, we need to interpolate the value of F_i at the point a_0. We use linear interpolation for the computation of $F_i(\epsilon, a)$ for $a_j < a < a_{j+1}$.

Once we have computed the distribution function, we are also able to compute the aggregate capital stock in step 6 of the Algorithm 5.2.1:

$$\begin{aligned} K &= \sum_{\epsilon \in \{e,u\}} \int_{a_{min}}^{\infty} a \; f(\epsilon, a) \;\; da \\ &\approx \sum_{\epsilon} \left(\sum_{j=2}^{m} \left(F(\epsilon, a_j) - F(\epsilon, a_{j-1}) \right) \frac{a_j + a_{j-1}}{2} + F(\epsilon, a_1) a_1 \right). \end{aligned}$$

In this computation, we assume that the distribution of the individual asset holdings is uniform in the interval $[a_{j-1}, a_j]$ for $j = 2, \ldots, m$. Of course, the accuracy of our computation will increase with a finer grid and increasing number of grid points m. If the capital stock K is close to the capital stock in the previous iteration, we are done. We stop the computation if two successive values of the capital stock diverge by less than 0.1%.

In the program `ch5_disf.g`, we also increase the number of iterations over the invariant distribution as the algorithm slowly

converges to the invariant aggregate capital stock K. We start with an initial number of 500 iterations i over $F_i(\cdot)$ which we increase by 500 in each iteration to 25,000 iterations in the iteration $q = 50$ over the capital stock. In the first iterations over the capital stock, we do not need a high accuracy in the computation of the invariant distribution. It saves computational time to increase the accuracy as we get closer to the solution for the aggregate capital stock. Similarly, the value function is getting more accurate as the algorithm converges to the aggregate capital stock. The reason is that we use a better initialization of the value function in each iteration, namely the solution of the last iteration.

The divergence between the capital stocks in iteration 50 and 51 is less than 0.1% so that we stop the computation. The computational time is very long and amounts to 28 hours and 25 minutes using GAUSS and a Pentium III with 846 MHz. For our calibration, the invariant aggregate capital stock is $K = 243.7$. The implied values for the wage rate, the interest rate, and the tax rate are $w = 4.770$, $r = 0.513\%$, and $\tau = 1.724\%$. Notice that $\beta = 0.99500 \approx 0.99499 = 1/(1 + r(1 - \tau))$, where the deviation is due to numerical round-off errors. As in the representative agent deterministic Ramsey model, the inverse of β is equal to the gross interest rate (after taxes). In the heterogeneous-agent economies of Example 5.2.1, this equation need not always hold. For our calibration, the wealth constraint $a \geq a_{min}$ is found to be non-binding. Huggett and Ospina (2001) show that the stationary interest rate is always larger in any equilibrium with idiosyncratic shocks as long as the consumers are risk averse ($\eta > 0$) and if the liquidity constraint binds for some agents. We will also demonstrate this result to hold in the application of Section 5.3.1.

At this point, we need to direct your attention to an important point. For our Example 5.2.1, it is rather the exception than the rule that the Algorithm 5.2.2 converges. E.g., if you increase the number of simulations over the distribution from $\{500, 1000, 1500, \ldots, 25000\}$ to $\{2500, 5000, \ldots, 125000\}$ while you iterate over the capital stock $q = 1, \ldots, 50$, the algorithm will not converge. Similarly, if we choose the uniform distribution over the interval $[-2, 3000]$:

$$F(\epsilon, a) = \frac{a - a_{min}}{a_{max} - a_{min}}, \quad a \in [a_{min}, a_{max}]$$

for the initial distribution rather than the equal distribution:

$$F(\epsilon, a) = \begin{cases} 1 & \text{if } a \geq K \\ 0 & \text{else,} \end{cases}$$

where all agents hold the representative-agent economy steady-state capital stock, the algorithm does not converge either. Therefore, computing the stationary solution to Example 5.2.1 involves a lot of trial and error. Furthermore, as the computation time amounts to several hours, the solution might be very time-consuming.

Why is convergence so hard to achieve with the help of Algorithm 5.2.2? Consider what happens if we are not close to the stationary solution and, for example, our choice of the stationary capital stock is too low. As a consequence, the interest rate is too high and agents save a higher proportion of their income than in the stationary equilibrium. Consequently, if we choose rather too many time periods for the simulation of the distribution when we start the algorithm (and are far away from the true solution), the distribution of wealth among the employed agents becomes increasingly concentrated in the upper end of the wealth interval $[-2, 3000]$. As a result, we have a new average capital stock that is much higher than the stationary capital stock. In the next iteration over the capital stock, we might, therefore, also choose a capital stock that is much higher than the stationary capital stock and an interest rate that is lower than the stationary rate. As a consequence, agents may now save a much lower proportion of their wealth than in the stationary equilibrium. For this reason, as we simulate the distribution over many periods, the distribution may now become increasingly centered in the lower part of the interval $[-2, 3000]$. If we are unlucky, the distribution might alternate between one that is concentrated in the lower part of the interval for individual wealth and one that is concentrated close to the upper end of the interval.

The algorithm, furthermore, fails to converge at all if we do not fix the unemployment compensation b,[12] but, for example, calibrate it endogenously to amount to 25% of the net wage rate in each iteration over the capital stock. In this case, you will not be able to generate convergence even with the choice of the equal distribution for the initial distribution. Our choice of $b = 1.299$ serves as an anchor. If we do not fix it, b starts to alternate between high and low values and, as a consequence, precautionary savings of the employed agents also switch between low and high values, respectively. The convergence of the algorithm improves considerably if we could also fix the wage income of the agents. In fact, you will get to know two prominent applications from the literature in Sections 5.3.1 and 6.3.1, where we will exactly do this. By this device, we will be able to compute the stationary equilibrium in the models of Huggett (1993) and İmrohoroğlu (1989) without any problems and convergence can be achieved for any initial distribution. In Section 5.3.2., you will encounter another example where convergence is not a problem. Different from Example 5.2.1, we will then introduce endogenous labor supply. In this case, richer agents supply less labor ceteris paribus and, as a consequence, the wage income decreases with higher wealth and so do savings. This mechanism, of course, improves convergence. In Chapter 6, where we compute the dynamics of the distribution endogenously, this problem does also not occur. In these models, as we will argue, an increase in the average capital stock over the number of simulation periods is then accompanied by a decrease in the endogenous interest rate and, hence, an endogenous reduction of the savings rate.

The convergence of the mean of the distribution during the final iteration over the capital stock is displayed in Figure 5.1. Notice that the rate of convergence is extremely slow. We also made this observation in all our other applications: Convergence of the distributions' moments[13] only occurs after a substantial

[12] We encourage you to recompute Example 5.2.1 with the help of `ch5_disf.g` for the cases discussed.

[13] The same result holds for the second and third moments of the distributions.

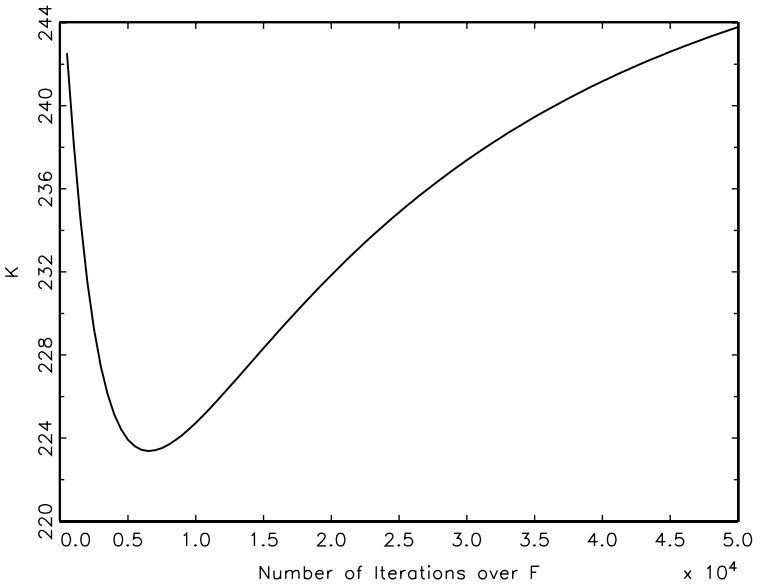

Figure 5.1: Convergence of the distribution mean

number of iterations well in excess of several thousands. It is for this reason that the computation of the stationary equilibrium of a heterogeneous-agent economy is extremely time-consuming.

Figure 5.1 also suggests that we should increase the number of iterations over the distribution function further to perhaps $n = 50,000$ or more.[14] In order to judge if our results are already accurate it is instructive to look at Figure 5.2 which displays the convergence of the aggregate capital stock K. At the first iteration over the capital stock, $q = 1$, we only use 500 iterations over the distribution functions and our value functions are highly inaccurate. For higher values of $q > 30$, our aggregate capital stock remains rather constant no matter if we iterate 15,000, 20,000 or 25,000 times over the distribution function (corresponding to q=30, 40, and 50, respectively). This result indicates that we have indeed found the stationary solution.

With the help of the stationary distribution function that we computed with the help of `ch5_disf.g`, we can also compute the

[14] We encourage the reader to change the program `ch5_disf.g` accordingly.

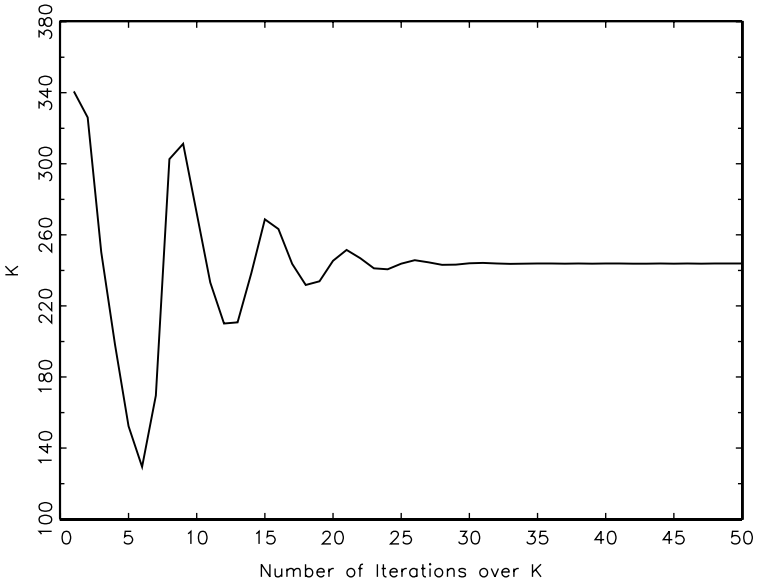

Figure 5.2: Convergence of K

invariant density function such that $f(\epsilon, a) = F(\epsilon, a_j) - F(\epsilon, a_{j-1})$ for $a = (a_j + a_{j-1})/2$, $a_j, a_{j-1} \in \mathcal{A}$. The invariant density function of the employed (unemployed) worker that is computed with the help of Algorithm 5.2.2 is displayed by the solid (broken) line in the Figure 5.3. Notice that the wealth constraint $a \geq a_{min}$ is non-binding and that the number of agents with wealth above $a = 1000$ is almost zero. Therefore, our choice of the wealth interval $[a_{min}, a_{max}] = [-2, 3000]$ is sensible. Notice further, that, as observed empirically, the distribution is skewed to the left.

Discretization of the Density Function. Alternatively, we may approximate the continuous density function $f(\epsilon, a)$ by a discrete density function, which, for notational convenience, we also refer to as $f(\epsilon, a)$. Again, we discretize the asset space by the grid $\mathcal{A} = \{a_1 = a_{min}, a_2, \ldots, a_m = a_{max}\}$. We assume that the agent can only choose a next-period asset a' from the set $\mathcal{A}$. Of course, the optimal next-period capital stock $a'(\epsilon, a)$ will be on the grid with probability zero. For this reason, we introduce a simple lottery: If the optimal next-period capital stock

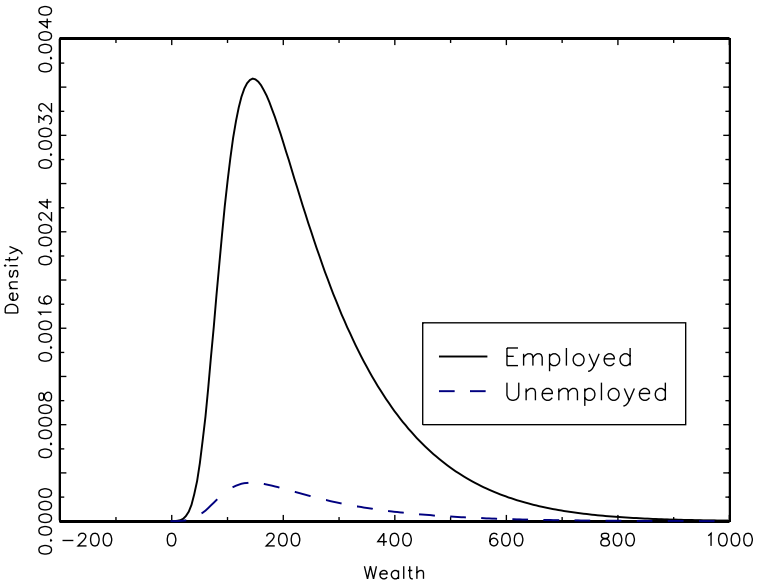

Figure 5.3: Invariant density function of wealth

happens to lie between a_{j-1} and a_j, $a_{j-1} < a' < a_j$, we simply assume that the next-period capital stock will be a_j with probability $(a' - a_{j-1})/(a_j - a_{j-1})$ and a_{j-1} with the complementary probability $(a_j - a')/(a_j - a_{j-1})$. With these simplifying assumptions, we can compute the invariant discrete density function with the help of the following algorithm:

Algorithm 5.2.3 (Computation of the Invariant Density Function $f(\epsilon, a)$)

Purpose: *Computation of the stationary equilibrium.*

Steps:

Step 1: Place a grid on the asset space $\mathcal{A} = \{a_1 = a_{min}, a_2, \ldots, a_m = a_{max}\}$ *such that the grid is finer than the one used to compute the optimal decision rules.*

Step 2: Set $i = 0$. *Choose initial discrete density functions* $f_0(\epsilon = e, a)$ *and* $f_0(\epsilon = u, a)$ *over that grid. The two vectors have* m *rows each.*

Step 3: Set $f_{i+1}(\epsilon, a) = 0$ *for all* ϵ *and* a. *i) For every* $a \in \mathcal{A}$, $\epsilon \in \{e, u\}$, *compute the optimal next-period wealth* $a_{j-1} \leq a' = a'(\epsilon, a) < a_j$ *and ii) for all* $a' \in \mathcal{A}$ *and* $\epsilon' \in \{e, u\}$ *the following sums:*

$$f_{i+1}(\epsilon', a_{j-1}) = \sum_{\epsilon=e,u} \sum_{\substack{a \in \mathcal{A} \\ a_{j-1} \leq a'(\epsilon,a) < a_j}} \pi(\epsilon'|\epsilon) \frac{a_j - a'}{a_j - a_{j-1}} f_i(\epsilon, a),$$

$$f_{i+1}(\epsilon', a_j) = \sum_{\epsilon=e,u} \sum_{\substack{a \in \mathcal{A} \\ a_{j-1} < a'(\epsilon,a) < a_j}} \pi(\epsilon'|\epsilon) \frac{a' - a_{j-1}}{a_j - a_{j-1}} f_i(\epsilon, a).$$

Step 4: Iterate until f *converges.*

The Algorithm 5.2.3 is implemented in the GAUSS program `ch5_denf.g`. The invariant discrete density function is computed with the same policy functions and parameterization as in the case of the approximation of the invariant distribution function. In particular, we use the equal distribution as initial distribution and increase the number of iterations over the density function from 500 to 25,000 by 500 in each iteration over the capital stock K. Again, we stop the computation as soon as two successive values of the capital stock diverge by less than 0.1% and the number of iterations over the density function is equal to 25,000.

Table 5.1

	Invariant Distribution	Invariant Density	Monte Carlo	Exponential Function n=2
Mean	243.7	243.7	243.4	246.6
Runtime*	28 h 25 m	20 h 16 m	75 h 51 m	19 h 5 m
Iterations**	51	51	54	63

* in hours and minutes; ** over the variable K.

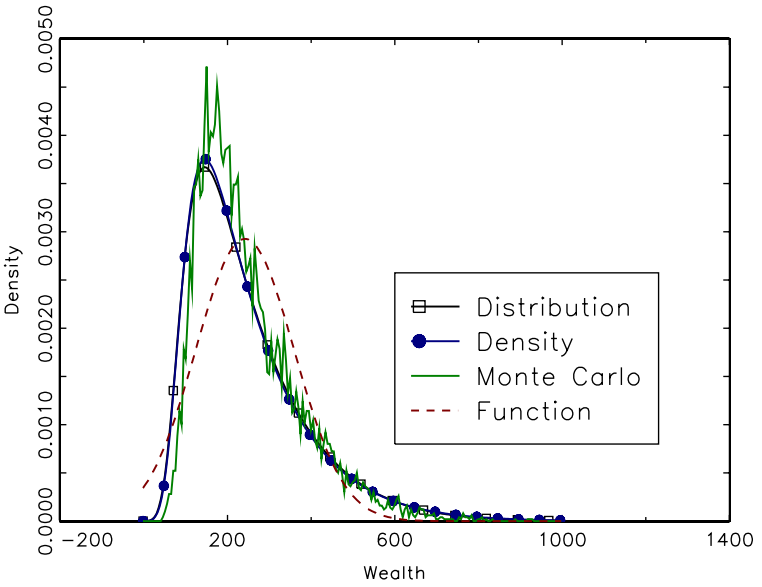

Figure 5.4: Invariant density function of wealth for the employed worker

The density function of the employed worker that is computed with the help of Algorithm 5.2.3 is displayed by the dotted line in Figure 5.3. The two density functions for the wealth of the employed worker computed with the help of the Algorithms 5.2.2 and 5.2.3 almost coincide and cannot be discerned. The two means $K = 243.7$ are identical. However, the computational time is much longer in the case of the discretized distribution function. The computation with the help of Algorithm 5.2.2 takes 40% longer than the one with Algorithm 5.2.3 due to the Step 3 where we compute the inverse of the policy function $a'(\epsilon, a)$. Table 5.1 summarizes the computational time and the first moment of the distribution for the various methods introduced in this section.

Monte Carlo-Simulation. The second method to compute the invariant distribution is by means of Monte Carlo simulation. In this method, we choose a large sample of households, typically in excess of thousand, and track their behavior over time. The household is subject to an employment shock which follows the

Markov process (5.13). We simulate this individual employment shock with the help of a random number generator. The algorithm is as follows:

Algorithm 5.2.4 (Computation of the Invariant Distribution Function $F(\epsilon, a)$ by Monte Carlo Simulation)

Purpose: *Computation of the stationary equilibrium.*

Steps:

Step 1: Choose a sample size N equal to some tens of thousands.
Step 2: Initialize the sample. Each household $i = 1, \ldots, N$ is assigned an initial wealth level a_0^i and employment status ϵ_0^i.
Step 3: Compute the next-period wealth level $a'(\epsilon^i, a^i)$ for all $i = 1, \ldots, N$.
Step 4: Use a random number generator to obtain ϵ'^i for all $i = 1, \ldots, N$.
Step 5: Compute a set of statistics from this sample. We choose the mean and the standard deviation of a and ϵ.
Step 6: Iterate until the distributional statistics converge.

The algorithm is implemented in the program `ch5_mont.g`. As an initial asset level, the agent is assigned the wealth level $a = 247.6$, which is equal to the steady-state capital stock in the corresponding representative-agent model. Similarly, the agent is employed with the ergodic employment probability $g(e)$ and unemployed with the complementary probability. The statistics of the sample distribution are the mean of the wealth and the standard deviations of the wealth and the employment status. We choose $N = 1000$ individuals in order to keep the computational time to a reasonable amount and in order to demonstrate the necessity to choose a high number of individuals. Furthermore, we need to adjust the share of employed and unemployed agents in each iteration. As we use a random number generator in step 4, the number of employed agents need not be equal to the number of employed agents in the ergodic distribution, $Ng(e)$. If the number of employed agents in any iteration of the simulation is higher

than the ergodic number, we select a random set of employed agents and change their employment status to 'unemployed' until the number of employed agents, again, is equal to the respective number in the ergodic distribution. If the number of unemployed agents is higher than the respective number in the ergodic distribution, we change the employment status of the unemployed agents in an analogous way.

The distribution function that is computed with the help of Monte Carlo Simulation is displayed by the solid line in Figure 5.4. Notice that the distribution function has a lower variance than those computed with the help of Algorithms 5.2.2 and 5.2.3. It is also less smooth than these two distributions and we take this as a hint that we should use a higher number of individuals N. The algorithm converges after 54 iterations and the mean capital stock computed with the help of Algorithm 5.2.4 amounts to $K = 243.4$. Also, the number of households $N = 1000$ is chosen too small as the convergence of the mean capital stock is not smooth (not displayed) and fluctuates between $K = 242$ and $K = 244$. If we increased the number to $N = 5000$ or $N = 10,000$, the convergence of the mean would be much smoother. However, in our application, the computational time becomes exorbitant and is rather a matter of days than of hours. In fact, for $N = 1000$, the computational time already amounts to more than 75 hours! It is for this very reason that we discourage the use of Monte Carlo simulations in many applications.

Function Approximation. In this section, we introduce a third method to compute the invariant distribution function. In particular, we approximate the distribution function by a flexible functional form with a finite number of coefficients. In Chapter 4, we approximated the policy function with a linear combination of Chebyshev polynomials. Chebyshev polynomials, however, can take a value below zero. For this reason, it is advisable to use another class of functions. We follow DEN HAAN (1997) and use the class of exponential functions for the $n-th$ order approximation of the wealth holdings of the agents with employment status $\epsilon \in \{e, u\}$:

$$F(\epsilon, a) = 0 \qquad a < a_{min}, \tag{5.18a}$$

$$F(\epsilon, a) = \rho_0^\epsilon \int_{-\infty}^{a} e^{\rho_1^\epsilon x + \dots \rho_n^\epsilon x^n} \, dx \qquad a \geq a_{min}. \tag{5.18b}$$

This approximation allows for having a positive number of agents at the borrowing constraint $a = a_{min}$. Of course, this is a very desirable feature of the distribution function in the present case and might be very useful for models with binding constraints and heterogeneity.

For the exponential family, the first n moments capture the same information as the $n + 1$ coefficients ρ_i. Suppose that we have found the first n moments of a distribution. In particular, we will use the first two moments, the mean μ^ϵ and the variance $(\sigma^\epsilon)^2$ for the wealth distribution of the employed and the unemployed, respectively. To find the values $\rho^\epsilon = (\rho_0^\epsilon, \rho_1^\epsilon, \rho_2^\epsilon)$, $\epsilon \in \{e, u\}$ that correspond to μ^ϵ and $(\sigma^\epsilon)^2$, we have to solve the following set of non-linear equations:

$$g(\epsilon) = \rho_0^\epsilon \int_{-\infty}^{a_{max}} e^{\rho_1^\epsilon a + \rho_2^\epsilon a^2} da, \tag{5.19a}$$

$$\mu^\epsilon = \rho_0^\epsilon \int_{-\infty}^{a_{max}} \max(a, a_{min}) e^{\rho_1^\epsilon a + \rho_2^\epsilon a^2} da, \tag{5.19b}$$

$$(\sigma^\epsilon)^2 = \rho_0^\epsilon \int_{-\infty}^{a_{max}} \left(\max(a, a_{min}) - \mu^\epsilon\right)^2 e^{\rho_1^\epsilon a + \rho_2^\epsilon a^2} da, \tag{5.19c}$$

where $g(\cdot)$, again, denotes the ergodic distribution of ϵ. The solution of this non-linear equation problem is not trivial, especially for a higher-order approximation n. As the problem is highly non-linear, a good first-order approximation is needed. In fact, some experimentation with a good starting value for ρ^ϵ might be necessary. Often, one might want to try to start with a low-order approximation of the exponential function, e.g. $n = 1$, and increase n subsequently. In our application, we will use the uniform distribution as an initial guess for $g(\cdot)$.

Algorithm 5.2.5 (Approximation of $F(\epsilon, a)$ by an Exponential Function of Order 2)

Purpose: *Computation of the stationary equilibrium.*

Steps:

Step 1: Choose initial moments μ^ϵ and $(\sigma^\epsilon)^2$ for the wealth distribution for $\epsilon \in \{e, u\}$ and compute the corresponding parameters ρ^ϵ of the exponential distribution by solving the non-linear equation problem (5.19a)-(5.19c).

Step 2: Compute the moments of the next-period wealth distribution for the employed and unemployed agents, respectively, e.g. for the employed agent ($\epsilon = e$):

$$\begin{aligned}
\mu^{e\prime} &= \pi(e|e)\rho_0^e \int_{-\infty}^{a_{max}} \max(a'(a,e), a_{min}) e^{\rho_1^e a + \rho_2^e a^2} da \\
&\quad + \pi(e|u)\rho_0^u \int_{-\infty}^{a_{max}} \max(a'(a,u), a_{min}) e^{\rho_1^u a + \rho_2^u a^2} da, \\
\left(\sigma^{e\prime}\right)^2 &= \pi(e|e)\rho_0^e \int_{-\infty}^{a_{max}} \left(\max(a'(a,e), a_{min}) - \mu^e\right)^2 e^{\rho_1^e a + \rho_2^e a^2} da \\
&\quad + \pi(e|u)\rho_0^u \int_{-\infty}^{a_{max}} \left(\max(a'(a,u), a_{min}) - \mu^u\right)^2 e^{\rho_1^u a + \rho_2^u a^2} da,
\end{aligned}$$

and compute the parameters of the distribution function ρ^ϵ, $\epsilon \in \{e, u\}$, corresponding to the computed next-period moments μ' and σ'^2.

Step 3: Iterate until the moments μ^ϵ and σ^ϵ converge.

The GAUSS program `ch5_func.g` implements the Algorithm 5.2.5. We parameterize the model and compute the value function in exactly the same way as in the other methods in this section. However, we do not use the equal distribution for the initialization of the distribution function as we would like to start with a continuous function. Therefore, in the first step of Algorithm 5.2.5, the function is approximated by the uniform distribution with $\rho_1^\epsilon = \rho_2^\epsilon = 0$ for $\epsilon \in \{e, u\}$ and $\rho_0^e = \rho_0^u = 1/(a_{max} - a_{min})$. We choose a smaller interval for individual wealth,$[a_{min}, a_{max}] =$

$[-2, 1000]$. The approximation, of course, is likely to be more accurate on a smaller interval and our computations in this section indicate that the number of agents with wealth exceeding 1000 is practically zero. In the second step, we need to compute an integral. We will apply Gauss-Chebyshev quadrature as described in Section 8.3.2 using 20 nodes.

The Algorithm 5.2.5 needs more iterations over the capital stock in order to converge than the other algorithms and stops after 63 iterations. The computational time is close to the one for the computation of the density function and amounts to 19 hours. The average wealth of the distribution is higher than the one found with the help of the discretization methods presented in Algorithm 5.2.2 and 5.2.3. The density function approximated with the help of the Algorithm 5.2.5 is displayed by the broken line in Figure 5.4. Obviously, the density function is much more symmetrical and described by smaller variance than in the case of the discretization methods and we are sceptical if the approximation of the density function with the exponential function is accurate.

In conclusion, we like to emphasize that our experience with the computation of the stationary distribution points to the following suggestions: Probably the first best try to compute the stationary distribution is by means of Algorithm 5.2.3 as implemented in the program `ch5_denf.g`. If the functional form of the density function is similar to the one of a parameterized function, you may also want to try to approximate the density function using Algorithm 5.2.5. This involves some experience with functional approximation. Approximation methods, however, may better work locally. In our example, approximation over the complete state space is poor. Monte Carlo simulations have the advantage that they are easy to implement. For our simple Example 5.2.1 with only one dimension for the (continuous) state space, discretization methods are much faster. For state spaces with higher dimension, however, Monte Carlo simulation may become an important alternative. In Algorithms 5.2.2 and 5.2.3, and different from Algorithm 5.2.4, the computational time increases exponentially with the number of dimensions.

5.3 Applications

5.3.1 The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies

Two different phenomena have been observed in financial markets during the last hundred years: 1) the low risk-free rate and 2) the large equity premium. During the last 100 years, the average real return on US Treasury Bills has been about one percent. The average real return on US stocks has been six percent higher. The representative agent model has difficulties to resolve this problem as we discussed in Section 4.3.3. Mehra and Prescott (1985) show that the representative agent model can only explain the large equity premium and the low risk-free rate if the typical investor is implausibly risk averse. Consequently, the representative agent model in an Arrow-Debreu economy can be regarded as largely unsuccessful to explain the two observations from financial markets. Kocherlakota (1996) argues that one of the three assumptions of the representative-agent model need to be abandoned in order to explain the two puzzles: 1) the standard utility function, 2) complete markets, and 3) costless trading. In this section, we try to explain the first phenomenon by abandoning the assumption of complete markets. We further consider a heterogeneous-agent economy in order to have both individuals who supply credit and individuals who demand credit. Following Huggett (1993), we compute the equilibrium interest rate which balances credit supply and credit demand and show that the consideration of incomplete asset markets implies a lower risk-free rate.

The Exchange Economy. Huggett (1993) considers a simple exchange economy without production. Agents receive an endowment of the only good in the economy. The endowment set $\mathcal{E}$ consists of only two values, $\mathcal{E} = \{e_h, e_l\}$ which we interpret again as in the previous section as the earnings during employment (e_h) and unemployment (e_l). The endowment (or employment) process follows a first-order Markov process with transition probability $\pi(e'|e) = \text{Prob}(e_{t+1} = e'|e_t = e) > 0$ for $e', e \in \mathcal{E}$. The agent maximizes expected discounted utility:

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t u(c_t) \right], \tag{5.20}$$

where $\beta < 1$ denotes the discount factor and instantaneous utility $u(\cdot)$ is a CES function of consumption:

$$u(c) = \frac{c^{1-\eta}}{1-\eta}. \tag{5.21}$$

As in previous chapters, $1/\eta$ denotes the intertemporal elasticity of substitution.

Agents may hold a single asset. A credit balance of a units entitles the agent to a units of consumption goods this period. To obtain a credit balance of a' goods next period, the agent has to pay $a'q$ goods this period. q is the price of the next-period credit balances and we can interpret $r = 1/q - 1$ as the interest rate in the economy. Furthermore, there is a credit constraint so that agents cannot run a credit balance below $\bar{a} < 0$. The asset space is denoted by $\mathcal{A}$. Furthermore, assume that the central credit-authority who administers the credit balances has no transaction costs.

The budget constraint of the household is

$$c + a'q = a + e, \quad \text{where } a' \geq \bar{a}. \tag{5.22}$$

We can also formulate the individuals problem recursively with the help of the value function $v(\cdot)$:

$$v(e, a; q) = \max_{c, a'} u(c) + \beta \sum_{e'} \pi(e'|e) v(e', a'; q') \tag{5.23}$$

subject to the budget constraint (5.22).

We consider a stationary equilibrium where the price q of the next-period credit balance a' is constant and the distribution of assets, $f(e, a)$, is invariant. In a stationary equilibrium, furthermore, markets clear so that the average credit balance is equal to zero.

A stationary equilibrium for the exchange economy are policy functions for consumption and next period assets, $c(e, a)$ and $a'(e, a)$, respectively, a price of next period credit balances q and a distribution function $f(e, a)$ satisfying:

a) $c(e,a)$ and $a'(e,a)$ are optimal decision rules given q.
b) Markets clear:

$$\sum_e \int_a c(e,a)\ f(e,a)\ da = \sum_e \int_a e f(e,a)\ da, \tag{5.24}$$

$$\sum_e \int_a a'(e,a)\ f(e,a)\ da = 0. \tag{5.25}$$

c) $f(e,a)$ is a stationary distribution:

$$f(e',a') = \pi(e'|e_h) f(e_h,a_h) + \pi(e'|e_l) f(e_l,a_l), \tag{5.26}$$

for all $a' \in \mathcal{A}$ and $e' \in \{e_l, e_h\}$ and with $a' = a'(e_h,a_h) = a'(e_l,a_l)$.

A detailed discussion of the equilibrium concept and the uniqueness and existence of the solution can be found in Huggett (1993). The model is also calibrated as in this paper. The endowments are set equal to $e_h = 1.0$ and $e_l = 0.1$. One model period corresponds to 8.5 weeks so that 6 periods are equal to one year. The transition probabilities are calibrated such that the average duration of the low endowment shock (unemployment) is two model periods and the standard deviation of annual earnings is 20%:

$$\pi(e'|e) = \begin{pmatrix} 0.925 & 0.075 \\ 0.5 & 0.5 \end{pmatrix}.$$

The equilibrium unemployment rate is 13%.

Huggett sets the discount factor equal to $\beta = 0.99322$ implying an annual discount rate of 0.96. In his benchmark case, the risk aversion coefficient η is set equal to 1.5. For the credit limit $\bar{a}$, he considers different values, $\bar{a} \in \{-2,-4,-6,-8\}$. A credit limit of -5.3 corresponds to one year's average endowment.

Programs. The model is computed with the help of the GAUSS routine `ch5_hug.g`. The algorithm for this problem is analogous to the Algorithm 5.2.1. First, we make an initial guess of the interest rate r and compute the policy functions. Second, we compute the stationary equilibrium and the equilibrium average asset holdings.

Finally, we update the interest rate and return to the first step, if necessary.

For the computation of the policy functions, we applied the techniques developed in Chapter 1. In particular, the model is solved with value function iteration and linear interpolation between grid points. We compute the value function over a discrete grid $\mathcal{A}$. We choose an equispaced grid over $[\bar{a}, a_{max}]$. The upper limit of $\mathcal{A}$ can only be found by experimentation. We set $a_{max} = 4$ and find that agent do not hold assets in excess of a_{max} in the stationary equilibrium of our exchange economy. We further use 100 evenly spaced grid points on $\mathcal{A}$. The invariant distribution is computed over a finer grid. In particular, we use 300 grid points over the same interval.

The value function is initialized assuming that each agent consumes his endowment infinitely and does not change his endowment type. The value function and optimal policy rules are computed with value function iteration. The value of the value function is stored at grid points. Between grid points, we interpolate linearly. The maximum value of the rhs of the Bellman equation is found by Golden Section search as described in Section 8.7.1. The optimal next-period assets of the employed and unemployed agents are displayed in Figures 5.5 and 5.6, respectively. Notice that only the employed agent has a higher next-period asset a' than the current period asset a and $a'(e, a) > a$ only for $a \leq 1$ (a' crosses the 45^o degree line). In other words, the ergodic set for the asset is approximately $[-2, 1]$ and the chosen upper limit a_{max} is not binding. Even if the initial distribution has agents with $a > 1$, after the transition to the stationary equilibrium, no agent has a credit balance exceeding a_{max}. The change in the asset level $a' - a$ is illustrated in Figure 5.7

Once we have computed the decision functions $c(\cdot)$ and $a'(\cdot)$, we are able to compute the invariant distribution. We apply the methods of Section 5.2 and iterate over the density function until convergence using Algorithm 5.2.3. We use 20,000 iterations over the distribution in each iteration over the interest rate r. The stationary distribution is displayed in Figure 5.8 for the employed

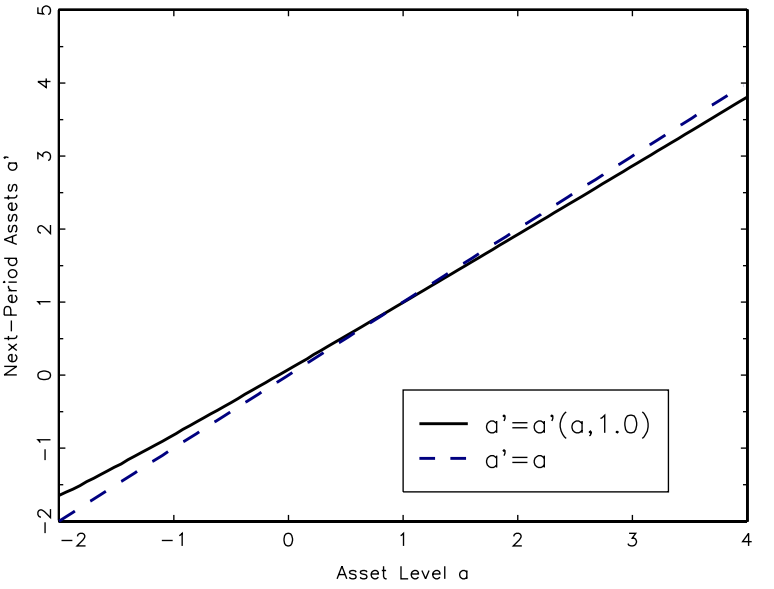

Figure 5.5: Next-period assets $a'(a, e_h)$ of the employed agent

agent (solid line) and unemployed agent (broken line), respectively. The mean of this distribution is equal to zero.

The computation of the stationary equilibrium is almost identical to the one in the production economy of Example 5.2.1 with only one exception. In Section 5.2, we analyzed a production economy where the equilibrium interest rate can be computed from the marginal product of capital. In the present exchange economy, we can only guess the equilibrium price of next-period capital q which clears the credit market. We need to modify our computation as follows: First, make two initial guesses of the interest rate $r = 1/q - 1$. We choose the values $r_1 = 0\%$ and $r_2 = 1\%$, respectively. Next, compute the average asset holding of the economy for the two cases, a_1 and a_2. We compute the following guesses for the equilibrium interest rate with the help of the secant method which is described in more detail in Section 8.5. Given two points (a_s, r_s) and (a_{s+1}, r_{s+1}) we compute r_{s+2} from:

$$r_{s+2} = r_{s+1} - \frac{r_{s+1} - r_s}{a_{s+1} - a_s} a_{s+1}. \tag{5.27}$$

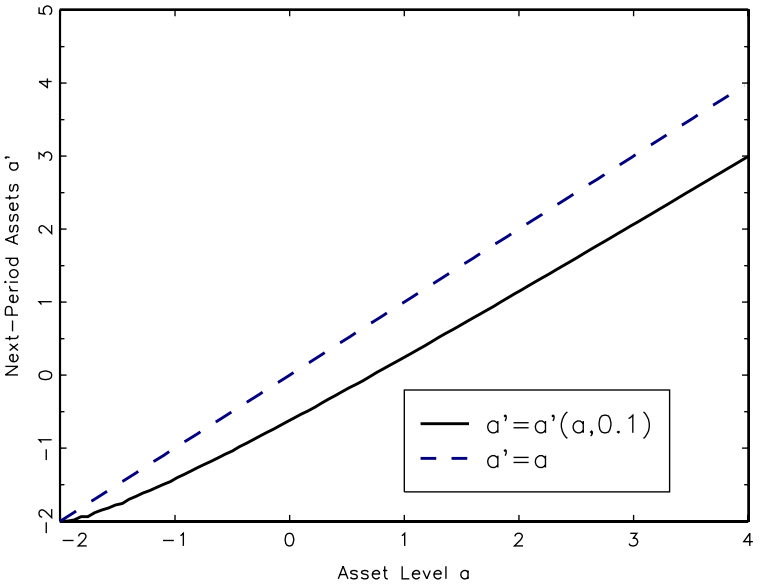

Figure 5.6: Next-period assets $a'(a, e_l)$ of the unemployed agent

In order to improve convergence, we use extrapolation and use the interest rate $r = \phi r_{s+1} + (1 - \phi) r_s$ in the next iteration. We choose a value $\phi = 0.5$ in our computation.

We stopped the computation as soon as the absolute average asset level is below 10^{-5}. We need approximately 20 iterations over the interest rate with a computational time of approximately 20 minutes. Table 5.2 presents the results from our computation.[15]

Clearly, the interest rate is lower for a stricter credit limit. For a credit limit $\bar{a} = -2$ approximately equal to one half of the average income, the interest rate is even below zero. With a lower credit limit, the interest rate increases as agents can borrow more. For $\bar{a} = -8$, the interest rate is already equal to 0.63%. In the corresponding representative-agent economy, the risk-free rate is equal to the time preference rate $(1-\beta)/\beta = 0.682\%$. Notice that

[15] Our results deviate less than 1% from Huggett's result. Notice that our interest rates are computed for one period equal to 1/6 of a year, while the corresponding numbers in Huggett's table 1 are computed for one year instead.

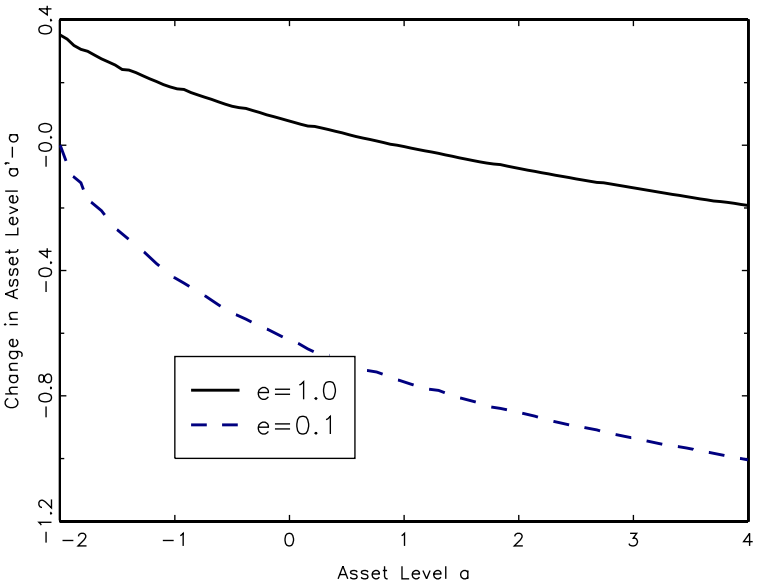

Figure 5.7: Change in assets $a' - a$

Table 5.2

Credit limit $\bar{a}$	Interest rate r	price q
-2	-1.27%	1.0129
-4	0.196%	0.9983
-6	0.507%	0.9949
-8	0.627%	0.9938

for a less binding credit constraint, the interest rate approaches the value of the representative agent economy. As noted above, the risk-free rate is strictly less than the time preference rate in a heterogeneous-agent economy with incomplete insurance markets and binding liquidity constraints. In conclusion, we find that incomplete insurance (against the risk of a negative endowment shock) helps to explain that the empirically observed risk-free rate of return is lower than the one found in standard representative-

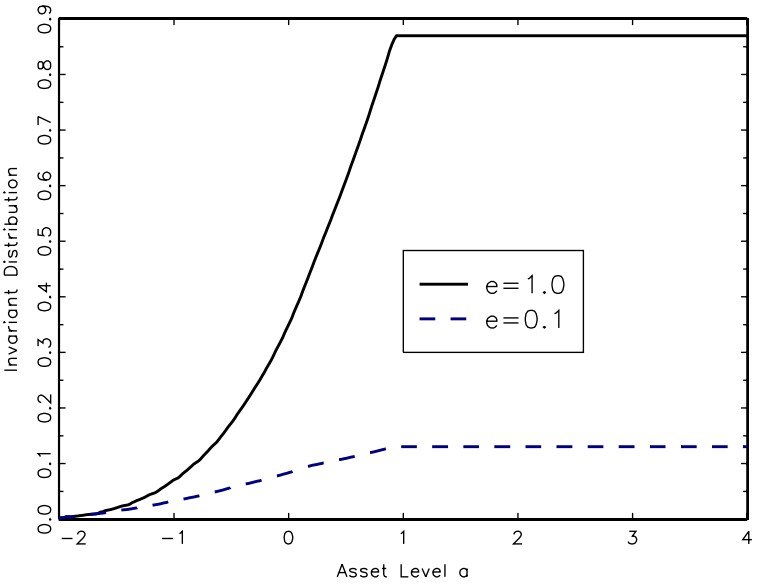

Figure 5.8: Stationary distribution function

agent models. As a consequence, the representative agent model might not be appropriate for the analysis of some problems in finance, but rather the application of heterogeneous-agent models is warranted.

5.3.2 Heterogeneous Productivity and Income Distribution

Naturally, we are unable to study redistributive problems in the representative-agent model. The representative-agent model cannot answer the question how, for example, different fiscal policies affect the distribution of income and wealth. Furthermore, it does not provide an answer to the question how the dispersion of income and wealth arises in the first place.

The explanation of the income and the wealth distribution has been a central objective of the early literature on heterogeneous-agent models. In this section, we analyze how we can model the income heterogeneity of the economy. Like in most heterogeneous-agent models, the source of income heterogeneity like e.g. differ-

ent levels of individual productivity or education is assumed to be exogenous.[16] Agents with different income build up different savings so that the wealth distribution can be computed endogenously and compared to the empirical distribution. We will find that our simple model is unable to replicate the empirical wealth distribution successfully and we will discuss possible solutions to this problem in the next chapters.

This section is organized as follows. First, empirical facts from the US and the German economy with regard to the distribution of wealth and income are reviewed.[17] Second, we discuss the standard way of introducing income heterogeneity into heterogeneous-agent models. Finally, we present a model with income heterogeneity and compute the endogenous invariant wealth distribution. We also analyze the steady-state effects of a fiscal policy reform that consists of a switch from a flat-rate income tax to a consumption tax.

Empirical Facts on the Income and Wealth Distribution. US households hold different levels of wealth and income. To be precise, we define *earnings* to be wages and salaries plus a fraction of business income, *income* as all kinds of revenue before taxes, and wealth as the net worth of households.[18] One striking feature of the US (and most industrialized and developed countries) is that wealth is much more unequally distributed than earnings and income. Using data from the 1992 Survey of Consumer Finances, Díaz-Giménez et al. (1997) compute Gini coefficients of income, earnings, and wealth equal to 0.57, 0.63, and 0.78, respectively. The Lorenz curves of US earnings, income, and wealth in 1992 are displayed in Figure 5.9.[19]

The distribution of income in many countries is a little less concentrated than the one in the US. For example, in Germany,

[16] As one of the few exceptions, Heckman et al. (1998) model the decision to attend college or not endogenously.

[17] This is only a very concise presentation of facts and the interested reader is encouraged to consult any of the references cited in this section.

[18] For a more detailed definition, see Díaz-Giménez et al. (1997).

[19] The data on the US economy from the 1992 Survey of Consumer Finance is provided in Díaz-Giménez et al. (1997).

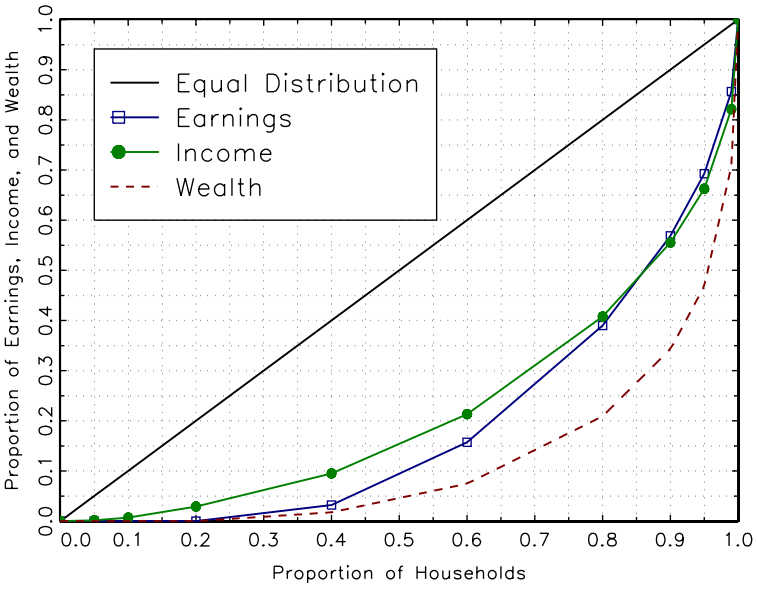

Figure 5.9: Lorenz Curve of US Wealth, Income, and Earnings in 1992

the Gini coefficient of labor income amounts to 0.317.[20] The distribution of wages is even less concentrated with a Gini coefficient equal to 0.275. Again, wealth is much more unequally distributed in Germany than earnings and the wealth distribution is characterized by a Gini coefficient in the range 0.59-0.89 depending on the assets included in the computation of wealth.[21]

[20] We computed the empirical Gini coefficient of gross wage income using the German Socio-Economic Panel (SOEP) data on annual individual labor income. The SOEP is a wide-ranging representative longitudinal study of private households. It provides information on all household members, consisting of Germans living in the Old and New German States, Foreigners, and recent Immigrants to Germany. The Panel was started in 1984. In 2002, there were more than 12,000 households, and nearly 24,000 persons sampled. For the computation, we deleted individuals with implausibly low or high implied hourly wage rates. We chose 7 DM as the lower limit and 200 DM as the upper limit. The number of deletions is small (about 0.17% at the top and about 6.5% at the bottom of the distribution).

[21] Bomsdorf (1989) analyzes Gini coefficients of the wealth distribution for different kinds of assets in the periods 1973, 1978, and 1983 for West Germany. Within each asset group, Gini coefficients are remarkably stable. The

One crucial aspect for the analysis of redistributive effects of economic policy is the consideration of mobility. Households move up and down the different income, earnings, and wealth groups. Some people fulfil their and the American dream and become rich. Others have simply bad luck (such as an accident or a divorce) and become poor. A redistribution of income, therefore, may have multiple effects. For example, an increase in income taxes may help to finance a rise in unemployment benefits and redistributes income from the income-rich to the income-poor. This may increase welfare as utility is a concave function of consumption. On the other hand, higher income taxes reduce incentives both to supply labor and to accumulate savings. As a consequence, total income decreases and welfare is reduced because of the increased distortions in the economy. Redistribution comes at the expense of efficiency. If we also consider income mobility, the welfare effect of such a policy is reduced further. The reason is simple: income-poor agents may move up the income hierarchy and will also be harmed by higher taxes and a reduction in the efficiency of the economy in the future. Therefore, if we consider the redistributive effects of an economic policy in a heterogeneous-agent model, mobility is a crucial ingredient.

The US earnings mobility is presented in the following matrix which is taken from Díaz-Giménez et al. (1997):[22]

$$\begin{array}{c} \\ \text{1984 Quintile} \end{array} \begin{array}{c} \text{1989 Quintile} \\ \begin{bmatrix} 0.858 & 0.116 & 0.014 & 0.006 & 0.005 \\ 0.186 & 0.409 & 0.300 & 0.071 & 0.034 \\ 0.071 & 0.120 & 0.470 & 0.262 & 0.076 \\ 0.075 & 0.068 & 0.175 & 0.465 & 0.217 \\ 0.058 & 0.041 & 0.055 & 0.183 & 0.663 \end{bmatrix} \end{array} \tag{5.28}$$

The matrix can be interpreted as follows: The entry in the first row, second column is equal to 0.116 and signifies that 11.6% of the households in the lowest earnings quintile in 1984 were in the

distribution of savings, securities, and real estate in 1983 are characterized by Gini coefficients equal to 0.59, 0.89, and 0.74, respectively.

[22] Díaz-Giménez et al. (1997) use data from the 1984, 1985, 1989 and 1990 Panel Study of Income Dynamics in order to compute the transition matrix.

second lowest earnings quintile in 1985. Notice that the entries in the diagonal are the maximums of each row so that there is a tendency to remain in the same earnings group. These values range between 40.9% and 85.8% and the low-income group is the least mobile group in the US. The income mobility is almost the same in Germany. For the 1980s, BURKHAUSER ET AL. (1997) find that, even though earnings are more unequally distributed in the US than in Germany, the patterns of the quintile to quintile mobility, surprisingly, are similar in the two countries.

Modelling Income Heterogeneity. In models of income heterogeneity, you have to introduce an exogenous source of such heterogeneity. Agents have either different abilities, inherit different levels of wealth or just happen to be unemployed after experiencing bad luck. Somewhere you have to start and assume heterogeneity. People are not the same. In Example 5.2.1, agents face idiosyncratic risk of unemployment which they cannot insure against. In this section, we consider income heterogeneity. One can either assume that the individual's earnings y_t^i are stochastic or that labor productivity ϵ_t^i is stochastic. In the first case, labor income is an exogenous variable, where in the latter case, agents may still be able to vary their labor supply so that labor income $y_t^i = \epsilon_t^i w_t n_t^i$, which is the product of individual productivity ϵ, wage w_t, and labor time n_t^i, is endogenous.

There have been many empirical studies on the time-series behavior of earnings and wages. For example, LILLARD and WILLIS (1978) estimate an AR(1) process for log earnings, while MACURDY (1982) considers an ARMA(1,2) equation for log earnings. They find substantial persistence in the shocks to earnings (the autoregressive coefficients equal 0.406 and 0.974 for annual data, respectively). More recently, some computable general equilibrium models with income heterogeneity and exogenous labor supply have used a regression to the mean process for log-labor earnings. Examples include AIYAGARI (1994), HUBBARD ET AL. (1995), HUGGETT (1996), or HUGGETT and VENTURA (2000). In these models, individual earnings y_t follow the process:

$$\ln y_t - \overline{\ln y} = \rho \left(\ln y_{t-1} - \overline{\ln y} \right) + \eta_t, \tag{5.29}$$

where $\eta_t \sim N(0, \sigma_\eta^2)$. Atkinson et al. (1992) report that estimates of the regression towards the mean parameter ρ vary from 0.65 to 0.95 in annual data. In Huggett and Ventura (2000) who study a life-cycle economy, the income is also age-dependant and follows the process:

$$\ln y_j - \overline{\ln y_j} = \rho \left(\ln y_{j-1} - \overline{\ln y_{j-1}} \right) + \eta_j, \tag{5.30}$$

where y_j is the income of the j-year-old household and $\eta_j \sim N(0, \sigma_\eta)$ and $\ln y_1 \sim N(\overline{\ln y_1}, \sigma_{y_1}^2)$. The parameters ρ, σ_{y_1}, and σ_η are calibrated in order to reproduce the Gini coefficient of US earnings of different cohorts and the overall economy on the one hand and the estimated variance of the persistence of the shocks to log earnings on the other hand.

Having specified the log earnings process as an AR(1) process, we need to discretize the process for computational purpose. The earnings process can easily be approximated with a finite-state Markov chain using Tauchen's method as presented in Section 9.2.

In the model of the next section, labor supply n is endogenous. Accordingly, earnings are endogenous and we cannot specify an exogenous earnings process in such a model. Rather, the exogenous variable is productivity ϵ or, similarly, the wage per unit labor, $w\epsilon$, as all agents face the same wage rate w per efficiency unity ϵn. Similar to related studies, e.g. Ventura (1999) or Castañeda et al. (2004), we assume productivity ϵ to follow a first-order Markov chain with conditional transition probabilities given by:

$$\pi(\epsilon'|\epsilon) = Prob\{\epsilon_{t+1} = \epsilon' | \epsilon_t = \epsilon\}, \tag{5.31}$$

where $\epsilon, \epsilon' \in \mathcal{E} = \{\epsilon^1, \ldots, \epsilon^{n\epsilon}\}$. Empirical evidence provided by Shorrocks (1976) suggests that the dynamics of productivity (and income) may be modelled slightly better by a second-order Markov chain, but the improvement in accuracy is rather small and does not justify the considerable increase in the model's complexity. The productivities $\epsilon \in \mathcal{E} = \{\epsilon^1, \ldots, \epsilon^{n\epsilon}\}$ are chosen to replicate the discretized distribution of hourly wage rates which,

in our model, are proportional to productivity. The number of productivities is set equal to $n\epsilon = 5$. We also consider unemployment and let ϵ^1 characterize the state of unemployment by setting ϵ^1 equal to zero. The productivities $\{\epsilon^2, \epsilon^3, \epsilon^4, \epsilon^5\}$ are estimated from the empirical distribution of hourly wages in Germany in 1995. The productivity ϵ^i corresponds to the average hourly wage rate of earners in the $(i-1)$-th quartile. Normalizing the average of the four nonzero productivities to unity we arrive at

$$\{\epsilon^2, \epsilon^3, \epsilon^4, \epsilon^5\} = \{0.4476, 0.7851, 1.0544, 1.7129\}. \tag{5.32}$$

The transition probability into and out of unemployment, $\pi(\epsilon' = 0|\epsilon > 0)$ and $\pi(\epsilon' > 0|\epsilon = 0)$ where ϵ' represents next period's productivity, are chosen in order to imply an average unemployment rate of 10.95% and an average duration of unemployment equal to slightly more than one year (we assume that the average transition takes place in the middle of the year). Periods correspond to one year. Further, we assume that the probability to loose one's job does not depend on the individual productivity. During unemployment, the worker's human capital depreciates or, equivalently, his productivity decreases. We assume that the worker can only reach productivity ϵ^2 after unemployment and set $\pi(\epsilon' = \epsilon^2|\epsilon = 0) = 1 - \pi(\epsilon' = 0|\epsilon = 0)$ and $\pi(\epsilon' > \epsilon^2|\epsilon = 0) = 0$.[23] The remaining $(n\epsilon - 1)^2 = 16$ transition probabilities are calibrated such that (i) each row in the Markov transition matrix sums to one, (ii) the model economy matches the observed quartile transition probabilities of the hourly wage rate from 1995 to 1996 as given by the German Socio-economic panel data.[24] Our transition matrix is given by:

[23] Alternatively, we could have assumed that the worker's productivity does not decrease during unemployment. In this case, however, we had to introduce an additional state variable into the model which makes the computation and calibration even more cumbersome.

[24] A different approach is followed by Castañeda et al. (2004) who calibrate the transition matrix in order to replicate the U.S. earnings and wealth distribution as closely as possible. As a consequence, the diagonal elements of the transition matrix calibrated by Castañeda et al. (1998a) are far larger than the empirical counterparts.

$$\pi(\epsilon'|\epsilon) = \begin{pmatrix} 0.3500 & 0.6500 & 0.0000 & 0.0000 & 0.0000 \\ 0.0800 & 0.6751 & 0.1702 & 0.0364 & 0.0383 \\ 0.0800 & 0.1651 & 0.5162 & 0.2003 & 0.0384 \\ 0.0800 & 0.0422 & 0.1995 & 0.5224 & 0.1559 \\ 0.0800 & 0.0371 & 0.0345 & 0.1606 & 0.6879 \end{pmatrix}. \tag{5.33}$$

You may want to compare the German wage mobility of the employed agents (the lower 4x4-matrix of (5.33) divided by $1-10.95\%$ in order to imply a mass equal to unity for the employed agents) with the US earnings mobility as described by (5.28). Notice, however, that (5.33) considers a 1-year transition period while (5.28) considers a time horizon of 5 years. If you assume that earnings follow an AR(1) process, you may derive the 5-year transition matrix for Germany by multiplying (5.33) 4 times with itself.[25] If you compare these two matrices, you cannot help noticing that German workers are much more mobile than the US workers. While the diagonal elements in (5.28) are in the range 0.409-0.885, the corresponding elements in the 5-year transition matrix in Germany amount to values between 0.27-0.37. This result, however, is an artefact of our approximation. As pointed out above, the earnings process might be better modelled with the help of an AR(2) process as suggested by Shorrocks (1976).

Modelling Income Tax Reforms and Their Effect on the Distribution. In the following, we consider a heterogeneous-agent economy where agents differ with regard to their productivity and employment status. Agents are also mobile and, between periods, the productivity and employment status may change. As a consequence, individual labor income also changes. The model is able to account for both the observed heterogeneity in wage rates and the observed labor income mobility in Germany. In addition to the economy studied in Example 5.2.1, we model the household's labor supply decision. As a consequence, the labor income distribution is endogenous. As one major implication of our modelling framework, we are able to replicate the German labor income distribution quite closely. The model follows Heer

[25] See also Section 9.2.

and TREDE (2003).[26] Three sectors can be depicted: households, firms, and the government.

Households. Households are of measure one and infinitely-lived. Households are heterogeneous with regard to their employment status, their productivity ϵ^j, and their wealth k^j, $j \in [0,1]$.[27] Individual productivity $\epsilon^j \in \mathcal{E} = \{0, 0.4476, 0.7851, 1.0544, 1.7129\}$ follows the first-order finite-state Markov chain with conditional transition probabilities given by (5.31).

Agents are not allowed to borrow, $k^j \geq 0$. In addition, the household faces a budget constraint. He receives income from labor n_t and capital k_t which he spends on consumption c_t and next-period wealth k_{t+1}:

$$k_{t+1}^j = (1+r)k_t^j + w_t n_t^j \epsilon_t^j - (1+\tau_c)c_t^j - \tau_y y_t^j + 1_{\epsilon=\epsilon^1} b_t, \quad (5.34)$$

where r_t, w_t, τ_c, and τ_y denote the interest rate, the wage rate, the consumption tax rate, and the tax rate on income y, respectively. $1_{\epsilon=\epsilon^1}$ is an indicator function which takes the value one if the household is unemployed ($\epsilon = \epsilon^1$) and zero otherwise. If the agent is unemployed, he receives unemployment compensation b_t. Taxable income is composed of interest income and labor income:

$$y_t^j = y_t^j(\epsilon_t^j, k_t^j) = r k_t^j + w_t n_t^j \epsilon_t^j. \quad (5.35)$$

Household j, which is characterized by productivity ϵ_t^j and wealth k_t^j in period t, maximizes his intertemporal utility with regard to consumption c_t^j and labor supply n_t^j:

[26] HEER and TREDE also study the more complicated case of a progressive income tax. In this case, the policy function for labor supply does not have a continuous derivative and the computation is a little bit more complicated. The interested reader is referred to the original article. For the US economy, we know of various other studies who consider the effects of a flat-rate tax versus a progressive income tax. VENTURA (1999) considers a life-cycle model, CASTAÑEDA ET AL. (2004) use a model similar to ours, but with a different calibration procedure for the Markov process (5.31), and CAUCUTT ET AL. (2003) also model endogenous human capital formation.

[27] As we only consider one type of asset, we will refer to k as capital, wealth, and asset interchangeably.

$$E_0 \sum_{t=0}^{\infty} \beta^t u(c_t^j, 1 - n_t^j), \tag{5.36}$$

where $\beta < 1$ is a discount factor and expectations are conditioned on the information set of the household at time 0. Instantaneous utility $u(c_t, 1 - n_t)$ is assumed to be additively separable in the utility from consumption and the utility from leisure as given by:

$$u(c_t, 1 - n_t) = \frac{c_t^{1-\eta}}{1 - \eta} + \gamma_0 \frac{(1 - n_t)^{1-\gamma_1}}{1 - \gamma_1}. \tag{5.37}$$

Our choice of the functional form for utility follows Castañeda et al. (2004). Most quantitative studies of general equilibrium model specify a Cobb-Douglas functional form of utility. In this case, however, the elasticity of individual labor supply with regard to wealth is larger than for the utility function (5.37) and, consequently, the distribution of working hours varies more (and is less in accordance with empirical observations) than for our choice of the utility function (5.37). Notice that this utility function is only applicable to an economy which is not growing over time. To see this point assume that we analyze a perfect-foresight economy with exogenous productivity growth at a rate $g > 0$ and no uncertainty. In steady state, capital, wages and consumption grow at rate $g > 0$, while labor supply is constant. The first-order condition of the household is given by

$$\gamma_0 \frac{(1 - n_t)^{\gamma_1 - 1}}{c_t^{-\eta}} = (1 - \tau_y) \epsilon w_t. \tag{5.38}$$

Consequently, for a steady state growth $c_{t+1}/c_t = w_{t+1}/w_t = 1 + g$ with constant labor supply $n_t = n$, either $g \neq 0$ and $\sigma = 1$ or $g = 0$.

Production. Firms are owned by the households and maximize profits with respect to their labor and capital demand. Production $F(K_t, N_t)$ is characterized by constant returns to scale using capital K_t and labor N_t as inputs:

$$F(K_t, N_t) = K_t^{\alpha} N_t^{1-\alpha}. \tag{5.39}$$

In a market equilibrium, factors are compensated according to their marginal products and profits are zero:

$$r_t = \alpha \left(\frac{N_t}{K_t} \right)^{1-\alpha} - \delta, \tag{5.40}$$

$$w_t = (1-\alpha) \left(\frac{K_t}{N_t} \right)^{\alpha}, \tag{5.41}$$

where δ denotes the depreciation rate of capital.

Government. Government expenditures consists of government consumption G_t and unemployment compensation B_t. In our benchmark case, government expenditures are financed by an income tax and a consumption tax. We will compare the employment and distribution effects of two tax systems with equal tax revenues: (i) a flat-rate income tax structure and (ii) only a consumption tax ($\tau_y = 0$).

The government budget is assumed to balance in every period so that government expenditures are financed by tax revenues T_t in every period t:

$$G_t + B_t = T_t. \tag{5.42}$$

Stationary Equilibrium. We will define a stationary equilibrium for given government tax policy and a constant distribution $f(e,k)$ over the individual state space $(e,k) \in \mathcal{E} \times [0,\infty)$.

A stationary equilibrium for a given set of government policy parameters is a value function $V(\epsilon,k)$, individual policy rules $c(\epsilon,k)$, $n(\epsilon,k)$, and $k'(\epsilon,k)$ for consumption, labor supply, and next-period capital, respectively, a time-invariant distribution $f(\epsilon,k)$ of the state variable $(\epsilon,k) \in \mathcal{E} \times [0,\infty)$, time-invariant relative prices of labor and capital $\{w,r\}$, and a vector of aggregates K, N, B, T, and C such that:

a) Factor inputs, consumption, tax revenues, and unemployment compensation are obtained aggregating over households:

$$K = \sum_{\epsilon \in \mathcal{E}} \int_0^{\infty} k \; f(\epsilon,k) \;\; dk, \tag{5.43a}$$

$$N = \sum_{\epsilon \in \mathcal{E}} \int_0^\infty \epsilon \; n(\epsilon, k) \; f(\epsilon, k) \;\; dk, \tag{5.43b}$$

$$C = \sum_{\epsilon \in \mathcal{E}} \int_0^\infty c(\epsilon, k) \; f(\epsilon, k) \;\; dk, \tag{5.43c}$$

$$T = \tau_y \left(K^\alpha N^{1-\alpha} - \delta K \right) + \tau_c C \tag{5.43d}$$

$$B = \int_0^\infty b \; f(\epsilon_1, k) \;\; dk, \tag{5.43e}$$

b) $c(\epsilon, k)$, $n(\epsilon, k)$, and $k'(\epsilon, k)$ are optimal decision rules and solve the household decision problem

$$V(\epsilon, k) = \max_{c,n,k'} \left[u(c, 1-n) + \beta E \left\{ V(\epsilon', k') \right\} \right], \tag{5.44}$$

where ϵ' and k' denote next-period productivity and wealth, subject to the budget constraint (5.34), the tax policy, and the stochastic mechanism determining the productivity level (5.31).

c) Factor prices (5.40) and (5.41) are equal to the factors' marginal productivities, respectively.

d) The goods market clears:

$$F(K, L) + (1 - \delta)K = C + K' + G = C + K + G. \tag{5.45}$$

e) The government budget (5.42) is balanced.

f) The distribution of the individual state variables is constant:

$$f(\epsilon', k') = \sum_{\epsilon \in \mathcal{E}} \pi(\epsilon' | \epsilon) \; f(\epsilon, k), \tag{5.46}$$

for all $k' \in [0, \infty)$ and $\epsilon' \in \mathcal{E}$ and with $k' = k'(\epsilon, k)$.[28]

[28] Our definition of the stationary equilibrium, again, does not use advanced concepts of the measure theory. In particular, our formulation of the characteristics of the stationary distribution is assuming that the number of households with zero capital is zero. This will be the case for our calibration.

Calibration. The model is calibrated as in Heer and Trede (2003). The preference parameters are set equal to $\eta = 2$, $\gamma_0 = 0.13$, and $\gamma_1 = 10$. The latter two parameters are selected in order to imply an average working time of $\bar{n} = 32\%$ and a coefficient of variation for hours worked equal to $\sigma_n/\bar{n} = 0.367$. The empirical value for Germany for the coefficient of variation is equal to 0.385. The discount factor β amounts to 0.96. The productivities $\epsilon \in \{0, 0.4476, 0.7851, 1.0544, 1.7129\}$ imply a Gini coefficient of wages equal to 0.254, which compares favorably with the empirical counterpart (0.275). The Markov transition matrix is given by (5.33). The income tax rate is set equal to 17.4%, while the consumption tax rate is computed endogenously in order to imply a government consumption share in GDP equal to 19.6%. The replacement ratio of unemployment compensation b relative to the gross wage of the lowest wage quartile is equal to 52%, $b = 0.52\epsilon^2 w\bar{n}^2$, where $\bar{n}^2$ denotes the average working time of the lowest productivity workers. The production elasticity α is set equal to 0.36 and the annual depreciation rate is estimated at $\delta = 4\%$.

Computation. The solution algorithm for the benchmark case with a flat-rate income tax is described by the following steps:

1) Make initial guesses of the aggregate capital stock K, aggregate employment N, the consumption tax τ_c, and the value function $V(\epsilon, k)$.
2) Compute the wage rate w, the interest rate r, and unemployment compensation b.
3) Compute the household's decision functions $k'(\epsilon, k)$, $c(\epsilon, k)$, and $n(\epsilon, k)$.
4) Compute the steady-state distribution of assets.
5) Compute K, N, and taxes T that solve the aggregate consistency conditions.
6) Compute the consumption tax τ_c that solves the government budget.
7) Update K, N, and τ_c, and return to step 2 if necessary.

In step 3, the optimization problem of the household is solved with value function iteration. For this reason, the value function

is discretized using an equispaced grid $\mathcal{K}$ of 1000 points on the interval $[0, k^{max}]$. The upper bound on capital $k^{max} = 12$ is found to never be binding. The value function is initialized assuming that working agents supply 0.2 units of time as labor and that each agent consumes his current-period income infinitely. The matrix that stores the values of the value function has 1000×5 entries. We also assume that the agent can only choose discrete values from the interval $[0, 1]$ for his labor supply. We choose an equispaced grid $\mathcal{N}$ of 100 points. The algorithm is implemented in the program `ch5_tax.g`.

In order to find the maximum of the rhs of the Bellman equation (5.44), we need to iterate over the next-period capital stock $k' \in \mathcal{K}$ and the optimal labor supply $n' \in \mathcal{N}$ for every $k \in \mathcal{K}$ and $\epsilon^i, i = 1, \ldots, n\epsilon$. This amounts to $1000 \times 100 \times 1000 \times 4 + 1000 \times 1000$ iterations (the labor supply of the unemployed is equal to 0). In order to reduce the number of iterations, we can exploit the fact that the value function is a monotone increasing function of assets k, that consumption is strictly positive and monotone increasing in k, and that the labor supply is a monotone decreasing function of assets k. Therefore, given an optimal next-period capital stock $k'(\epsilon, k_i)$ and labor supply $n(\epsilon, k_i)$, we start the iteration over the next-period capital stock for the optimal next-period capital stock $k'(\epsilon, k_{i+1})$ at $k'(\epsilon, k_i)$ with $k_{i+1} > k_i$. Similarly, we start the iteration over the labor supply n at $n(\epsilon, k_i)$ and decrease the labor supply at each iteration in order to find $n(\epsilon, k_{i+1}) \leq n(\epsilon, k_i)$. We also stop the iteration as soon as $c \leq 0$. The number of iterations is reduced substantially by the exploitation of the monotonicity conditions.

During the first iterations over the aggregate capital stock, we do not need a high accuracy of the value function and the policy functions. Therefore, we iterate only 10 times over the value function and increase the number of iterations to 20 as the algorithm converges to the true solution. By this device, we save a lot of computational time. The computer program is already very time-consuming and runs approximately one day. As a much faster alternative, we may compute the optimal labor supply functions with the help of the first-order condition (5.38) and you

will be asked to perform this computation in the exercises. Using the time-consuming value function iteration over both the capital stock and the labor supply, however, might be a good starting point i) if you would like to compute a rough approximation of the final solution as an initial guess for more sophisticated methods or ii) if your policy function is not well-behaved. The latter case might arise in the presence of a progressive income tax where the optimal labor supply does not have a continuous first derivative.[29]

As soon as we have computed the optimal policy function, we might want to check the accuracy of our computation. For this reason, we compute the residual function for the two first-order conditions:

$$R^1(\epsilon,k) \equiv \frac{u_l\left(c(\epsilon,k),1-n(\epsilon,k)\right)\ (1+\tau_c)}{u_c\left(c(\epsilon,k),1-n(\epsilon,k)\right)\ (1-\tau_y)w\epsilon} - 1,$$

$$R^2(\epsilon,k) \equiv E\left[\beta\frac{u_c\left(c(\epsilon',k'),1-n(\epsilon',k')\right)}{u_c\left(c(\epsilon,k),1-n(\epsilon,k)\right)}\left(1+r(1-\tau_y)\right)\right] - 1.$$

The mean absolute deviations are about 1.07% and 3.71% for the two residual functions R^1 and R^2, respectively. The maximum deviations even amount to 11% and 47% for R^1 and R^2, respectively. For a closer fit, we either need to increase the number of grid points or to compute the optimal policy functions at points off the grid (see the exercises).

The remaining steps of the algorithm are straightforward to implement using the methods presented in the previous chapters. For the computation of the invariant distribution, in particular, we discretize the wealth density and compute it as described in the Algorithm 5.2.3.

Results. In Table 5.3, the effects of the two different tax policies on the aggregate capital stock K, the effective labor N, average working hours $\bar{n}$, the real interest rate r, the Gini coefficients of the labor income and the wealth distribution, and the variational coefficient of working time and effective labor are presented. In the

[29] Again, the interested reader is referred to either Ventura (1999) or Heer and Trede (2003) for further reference.

stationary equilibrium, the unemployment rate is equal to 10.95%. Aggregate effective labor supply amounts to $N = 0.251$ with an average working time approximately equal to $\bar{n} = 0.324$. Working hours vary less than effective labor. The variational coefficient of working hours (effective labor) is equal to 0.367 (0.691) (see the last two columns of Table 5.3). The two variational coefficients are in very good accordance with the empirical estimates 0.385 (0.638) which we computed using data from the German Socio-Economic Panel during 1995-96. The higher variation of effective labor relative to working hours reflects the optimizing behavior of the working agents who work longer if they are more productive as the substitution effect of a rise in the wage predominates the income effect. The labor supply elasticity with regard to the wage rate, η_{nw}, is moderate, amounting to 0.213 for the average worker. Again, this compares favorably with the data. Sieg (2000), for example, estimates that elasticities for male labor supply are small and in the range between 0.02 and 0.2.

Table 5.3

Tax Policy	K	N	$\bar{n}$	r	Gini $w\epsilon n$	Gini k	$\sigma_n/\bar{n}$	$\sigma_{\epsilon n}/N$
τ_y	2.70	0.251	0.324	3.88%	0.317	0.406	0.367	0.691
τ_c	3.24	0.249	0.323	3.01%	0.316	0.410	0.366	0.685

Notes: τ_y refers to the case of a flat-rate income tax and τ_c to the case where the income tax rate is zero and the consumption tax rate τ_c is increased such that to guarantee a balanced budget.

The aggregate capital stock amounts to $K = 2.70$ which is associated with a capital-output coefficient equal to $K/Y = 4.57$. During 1991-97, the empirical value of K/Y was equal to 5.0 (2.6) in Germany for the total economy (producing sector). The distribution of wealth, however, is not modelled in a satisfactory manner. In our model, the concentration of wealth is too low with a Gini coefficient equal to $\text{GINI}_{wealth} = 0.406$ and compares unfavorably with empirical estimates of the wealth Gini coefficient reported

above (which were well in excess of 0.6). We will discuss the reasons why the simple heterogeneous-agent model of this section is unable to replicate the empirical wealth distribution in the next chapters.

In our second tax experiment, we set the income tax rate to zero and increase the consumption tax rate in order to generate the same tax revenues as in the benchmark case. The new steady-state consumption tax amounts to $\tau_c = 39.5\%$ (compared to 20.5% under tax policy τ_y). As interest income is not taxed any more, households increase their savings. Accordingly, the aggregate capital stock K rises from 2.70 to 3.24. As labor is not taxed either any more, the incentives to supply labor increases on the one hand. On the other hand, average wealth of the agents is higher and, for this reason, labor supply decreases. The net effect is rather small so that employment approximately remains constant. Associated with these changes of the input factors is a strong decline of the interest rate r by 0.8 percentage points. The distribution effect of the tax reform is rather modest. The Gini coefficient of gross labor income almost remains constant and wealth is only a little more concentrated. Similarly, the coefficients of variation for labor supply and effective labor are hardly affected. In summary, the most marked effect of a switch to a consumption tax consists of a pronounced rise of savings.

Problems

5.1 **Sparse Linear Systems**

The program `equivec1.g` computes the stationary distribution $\pi = [\pi_1, \pi_2, \ldots, \pi_n]'$ of an n-state Markov chain by solving the system

$$\boldsymbol{\pi}' \underbrace{\begin{bmatrix} p_{11}-1 & p_{12} & \ldots & p_{1,n-1} & 1 \\ p_{21} & p_{22}-1 & \ldots & p_{2,n-1} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{n-1,1} & p_{n-1,2} & \ldots & p_{n-1,n-1}-1 & 1 \\ p_{n1} & p_{n2} & \ldots & p_{n,n-1} & 1 \end{bmatrix}}_{P} = (0, \ldots, 0, 1). \tag{Sparse}$$

Recode this program to allow for a sparse transition matrix P. Recompute the Example 5.2.1 solving the sparse linear system (Sparse). How do you find the transition matrix?

(Hint: Choose $n = 50$ points for the equispaced grid over wealth a. The transition matrix has dimension $2n \times 2n$ where the upper left $n \times n$ matrix, for example, describes the transition probability of the employed agent with wealth a_j, $j = 1, \ldots, n$, to the state of employment with wealth a_i, $i = 1, \ldots, n$. Similarly, the upper right $n \times n$ square matrix of this transition matrix describes the transition from employment with wealth a_j to unemployment with wealth a_i. The rows $(n+1)$-$(2n)$ describe the transition from unemployment. Most of the entries are zero. For example, if the optimal next-period asset of the employed agent with wealth a_j happens to be a' with $a_{i-1} < a' < a_i$, and p_{ee} denotes the transition probability from employment to employment, the entry in row j and column $i-1$ of the transition matrix is equal to $(a_i - a')/(a_i - a_{i-1})p_{ee}$. Similarly, the entry in row j and column i is equal to $(a' - a_{i-1})/(a_i - a_{i-1})p_{ee}$. Likewise, the entry in columns $i-1+n$ and $i+n$ are equal to $(a_i - a')/(a_i - a_{i-1})p_{eu}$ and $(a' - a_{i-1})/(a_i - a_{i-1})p_{eu}$. The remaining entries of the row j in the transition matrix are equal to zero and we can apply sparse matrix methods.)

5.2 **Function Approximation**

Compute the invariant distribution of example 5.2.1 with the help of functional approximation as described in algorithm 5.2.5. However, choose an exponential function of order $n = 3$ for the approximation of the density function.

5.3 **The Risk-Free Rate of Return**

a) Compute the model with production in Example 5.2.1 with $\beta = 0.96$ and for different levels of minimum asset levels, $a_{min} \in \{0, -2, -4, -8\}$,

and show that the equilibrium interest rate decreases with a more binding credit constraint.

b) Compute the equilibrium prices in the exchange economy of HUGGETT (1993) for a higher coefficient of risk aversion $\eta = 3$ and compare your results with table 2 in HUGGETT (1993).

5.4 **Unemployment Insurance and Moral Hazard** (adapted from HANSEN and İMROHOROĞLU, 1992)
Consider the following extension of Example 5.2.1. The agents' utility function is now a function of both consumption and leisure,

$$u(c_t, l_t) = \frac{\left(c_t^{1-\sigma} l_t^{\sigma}\right)^{1-\eta}}{1-\eta}.$$

All agents are either offered an employment opportunity ($\epsilon = e$) or not ($\epsilon = u$). The Markov transition matrix is again described by (5.3). Agents that receive an employment offer may either accept the offer and work full-time, $n = 1 - l = 0.3$, or reject the offer and receive unemployment insurance b_t with probability $q(\epsilon_{t-1})$. In particular, the probability of unemployment benefits may be different for a *searcher*, $\epsilon_{t-1} = u$, and a quitter, $\epsilon_{t-1} = e$, $q(e) \neq q(u)$. Agents that turn down employment offers in order to extend unemployment spells may have different chances to receive unemployment benefits than quitters. Compute the stationary equilibrium of the model for the parameters of Example 5.2.1. In addition, set $\sigma = 0.67$. Compute the model for different replacement ratios $b_t/(1-\tau)w_t \in \{0.25, 0.5, 0.75\}$ and different probabilities to receive unemployment benefits $g(e) = g(u) = 0.9$, $g(e) = g(u) = 0.8$, $g(e) = 0.9, g(u) = 0.8$. How does the optimal unemployment insurance (as measured by the average value of the households) look like?

5.5 **Income Tax Reform**
Recompute the model of Section 5.3.2 implementing the following changes:

a) Compute the optimal labor supply with the help of the first-order condition (5.38) (do not forget to check if the constraint $0 \leq n \leq 1$ is binding). Therefore, you need to solve a non-linear equation.

b) Compute the optimal next-period capital k' where k' need not be a gridpoint. Use linear interpolation to evaluate the value function between gridpoints. Apply the Golden Section Search algorithm presented in the appendix in Chapter 8 in order to compute the maximum rhs of the Bellman equation.

5.6 **Superneutrality in the Sidrauski Model** (follows HEER, 2004)
As is well-known, money is superneutral in the SIDRAUSKI (1967) model. A change of the money growth rate does not affect the real variables of the Ramsey model that is augmented by a monetary sector if 1) money

demand is introduced with the help of money-in-the-utility and 2) labor supply is inelastic. Consider the following heterogeneous-agent extension of the standard Sidrauski model that consists of the three sectors households, firms, and the monetary authority:

Households. The household $j \in [0,1]$ lives infinitely and is characterized by her productivity ϵ_t^j and her wealth a_t^j in period t. Wealth a_t^j is composed of capital k_t^j and real money $m_t^j \equiv M_t^j/P_t$, where M_t^j and P_t denote the nominal money holdings of agent j and the aggregate price level, respectively. Individual productivity ϵ_t^j is assumed to follow a first-order Markov chain with conditional probabilities given by:

$$\Gamma(\epsilon'|\epsilon) = Prob\left\{\epsilon_{t+1} = \epsilon'|\epsilon_t = \epsilon\right\},$$

where $\epsilon, \epsilon' \in \mathcal{E} = \{\epsilon_1, \ldots, \epsilon_{n\epsilon}\}$.
The household faces a budget constraint. She receives income from labor n_t^j, capital k_t^j, and lump-sum transfers tr_t which she either consumes at the amount of c_t^j or accumulates in the form of capital or money:

$$k_{t+1}^j + (1+\pi_{t+1})m_{t+1} = (1+r)k_t^j + m_t + w_t\epsilon_t^j n_t^j + tr_t - c_t^j,$$

where $\pi_t \equiv \frac{P_t - P_{t-1}}{P_{t-1}}$, r_t, and w_t denote the inflation rate, the real interest rate, and the wage rate in period t.
The household j maximizes life-time utility:

$$W = \sum_{t=0}^{\infty} \beta^t u(c_t^j, m_t^j)$$

subject to the budget constraint. The functional form of instantaneous utility $u(.)$ is chosen as follows:

$$u(c,m) = \gamma \ln c + (1-\gamma)\ln m.$$

Labor supply is exogenous, $n = \bar{n} = 0.3$.

Production. Firms are also allocated uniformly along the unit interval and produce output with effective labor N and capital K. Let $f_t(k,m,\epsilon)$ denote the period-t distribution of the household with wealth $a = k + m$ and idiosyncratic productivity ϵ, respectively. Effective labor N_t is given by:

$$N_t = \sum_{\epsilon \in \mathcal{E}} \int_k \int_m \bar{n} \cdot \epsilon \cdot f_t(k,m,\epsilon)\, dm\, dk.$$

Effective labor N is paid the wage w. Capital K is hired at rate r and depreciates at rate δ. Production Y is characterized by constant returns to scale and assumed to be Cobb-Douglas:

$$Y_t = F(K_t, N_t) = K_t^{\alpha} N_t^{1-\alpha}.$$

In a factor market equilibrium, factors are rewarded with their marginal product:

$$w_t = (1-\alpha) K_t^{\alpha} N_t^{-\alpha},$$
$$r_t = \alpha K_t^{\alpha-1} N_t^{1-\alpha} - \delta.$$

Monetary Authority. Nominal money grows at the exogenous rate θ_t:

$$\frac{M_t - M_{t-1}}{M_{t-1}} = \theta_t.$$

The seignorage is transferred lump-sum to the households:

$$tr_t = \frac{M_t - M_{t-1}}{P_t}.$$

a) Define a recursive stationary equilibrium which is characterized by a constant money growth rate θ and constant distribution $f(\epsilon, k, m)$.
b) Show that in the homogeneous-agent case, $\epsilon^j = \bar{\epsilon}$, money is superneutral in the stationary equilibrium, i.e. the steady-state growth rate of money θ has no effect on the real variables of the model.
c) Compute the heterogeneous-agent model for the following calibration: Periods correspond to years. The number of productivities is set to $n\epsilon = 5$ with

$$\mathcal{E} = \{0.2327, 0.4476, 0.7851, 1.0544, 1.7129\}.$$

Further, $\gamma = 0.990$, $\beta = 0.96$, $\alpha = 0.36$, and $\delta = 0.04$. The transition matrix is given by:

$$\pi(\epsilon'|\epsilon) = \begin{pmatrix} 0.3500 & 0.6500 & 0.0000 & 0.0000 & 0.0000 \\ 0.0800 & 0.6751 & 0.1702 & 0.0364 & 0.0383 \\ 0.0800 & 0.1651 & 0.5162 & 0.2003 & 0.0384 \\ 0.0800 & 0.0422 & 0.1995 & 0.5224 & 0.1559 \\ 0.0800 & 0.0371 & 0.0345 & 0.1606 & 0.6879 \end{pmatrix}.$$

Show that money is not superneutral (consider $\theta \in \{0, 5\%, 10\%\}$). Can you think of any reason for this result?

Chapter 6

Dynamics of the Distribution Function

Overview. This chapter presents methods in order to compute the dynamics of an economy that is populated by heterogenous agents. In the first section, we show that this amounts to compute the law of motion of the density function $f(\epsilon, a)$ that describes the distribution of wealth among agents. In the second section, we concentrate on an economy without aggregate uncertainty. The initial distribution is not stationary. For example, this might be the case after a change in policy, e.g. after a change in the income tax schedule, or during a demographic transition, as many modern industrialized countries experience it right now. Given this initial distribution, we compute the transition to the new stationary equilibrium. With the methods developed in this section we are able to answer questions as to how the concentration of wealth evolves following a change in capital taxation or how the income distribution evolves following a change in the unemployment compensation system. In the third section, we consider a model with aggregate risk. There are many ways to introduce aggregate risk, but we will focus on a simple case. We distinguish good and bad times which we identify with the boom and recession during the business cycle. In good times, employment probabilities increase and productivity rises. The opposite holds during a recession. As one application, we study the income and wealth distribution dynamics over the business cycle in the final section of this chapter. We will need to find an approximation to the law of motion $f' = G(f)$ and introduce you to the method developed by Krusell and Smith (1998).

6.1 The Dynamics of Heterogeneous-Agent Economies

In the previous chapter, we have focused on the case of a stationary equilibrium where the distribution of wealth is invariant. If we want to compute the non-stationary state of an economy, we face severe problems, which we will explore in the following. Towards this end we consider Example 5.2.1 with one important difference: the economy is not in the stationary equilibrium.

In our illustrative example, households maximize their intertemporal utility (5.1)

$$E_0 \sum_{t=0}^{\infty} \beta^t u\left(c_t\right),$$

subject to the budget constraint (5.4)

$$a_{t+1} = \begin{cases} \left(1 + (1-\tau_t) r_t\right) a_t + (1-\tau_t) w_t - c_t, & \text{if } \epsilon_t = e, \\ \left(1 + (1-\tau_t) r_t\right) a_t + b_t - c_t, & \text{if } \epsilon_t = u, \end{cases}$$

the employment transition probability (5.3)

$$\pi(\epsilon'|\epsilon) = Prob\left\{\epsilon_{t+1} = \epsilon' | \epsilon_t = \epsilon\right\} = \begin{pmatrix} p_{uu} & p_{ue} \\ p_{eu} & p_{ee} \end{pmatrix},$$

the aggregate consistency condition (5.10a)

$$K_t = \sum_{\epsilon_t \in \{e,u\}} \int_{a_{min}}^{\infty} a_t f_t(\epsilon_t, a_t)\, da_t, \tag{6.1}$$

and the dynamics of the distribution

$$f_{t+1}(\epsilon_{t+1}, a_{t+1}) = \sum_{\epsilon_t \in \{e,u\}} \pi(\epsilon_{t+1}|\epsilon_t) f_t\left(\epsilon_t, a_{t+1}^{-1}(\epsilon_t, a_{t+1})\right) \equiv G(f_t), \tag{6.2}$$

where $a_{t+1}^{-1}(\epsilon_t, a_{t+1})$ is the inverse of the optimal policy $a_{t+1} = a_{t+1}(\epsilon_t, a_t)$ with regard to current period wealth a_t. Again, we are assuming that $a_{t+1}(\cdot)$ is invertible which will be the case in our example economy. Furthermore, (6.2) constitutes a functional

equation as it describes a map of the function f from the function space.

The factor prices depend on the aggregate capital stock K and the aggregate employment N (which is assumed to be constant in our economy in the case of aggregate certainty):

$$w_t = w(K_t, N_t), \tag{6.3}$$

$$r_t = r(K_t, N_t). \tag{6.4}$$

The household's first-order condition with respect to its intertemporal consumption allocation depends on consumption in this and the next period, c_t and c_{t+1}:

$$u'(c_t) = \beta E_t \left[u'(c_{t+1})(1 + r_{t+1}) \right]. \tag{6.5}$$

Consumption in this and the next period is given by the budget constraint, $c_t = (1 + (1 - \tau_t)r_t)a_t + (1 - \tau_t)w_t - a_{t+1}$ and $c_{t+1} = (1 + (1 - \tau_{t+1})r_{t+1})a_{t+1} + (1 - \tau_{t+1})w_{t+1} - a_{t+2}$, respectively. What do we need in order to compute the solution to (6.5)? The household observes the following current-period aggregate variables: the aggregate capital stock K_t, aggregate employment N_t, the wage rate w_t, the interest rate r_t, the tax rate τ_t, and the distribution of the assets $f_t(\epsilon_t, a_t)$. His individual state space are his employment status ϵ_t and his individual assets a_t. The solution of (6.5), as we will argue in the following, consists of a time-invariant function $a'(\epsilon, a, f)$ which gives him the optimal next-period capital stock $a_{t+1} = a'(\epsilon_t, a_t, f_t)$ and also the optimal capital stock in two periods, $a_{t+2} = a'(\epsilon_{t+1}, a_{t+1}, f_{t+1})$. Different from the optimal policy function in the computation of the stationary state in Chapter 5, we also include the distribution f as an additional argument.[1] Why? In Chapter 5, the aggregate capital stock $K_t = K$ is constant. Therefore, the interest rate and the wage rate are constant, too. In the present model, K_t and, hence, r_t are not constant. As the solution of (6.5) clearly depends on r_t (via c_t and the budget constraint), K_t needs to be an argument of the policy function

[1] In our special model, aggregate employment N is constant so we are able to drop it from the list of arguments. In other models with endogenous labor supply, N is also an additional argument of the policy functions, as we will argue below.

$a_{t+1} = a'(\cdot)$ as well.[2] K_t, however, can be computed with the help of the distribution $f_t(\epsilon_t, a_t)$ using (6.1). Now we only need to explain why we also have to include the distribution of the individual states, $f_t(\epsilon_t, a_t)$, as an additional argument and not only use the capital stock K_t instead. Consider again (6.5). The next-period interest rate r_{t+1} appears on the rhs of the equation. Therefore, the households needs to predict r_{t+1}. In the stationary economy, this is not a problem: $r_{t+1} = r$. In the representative-agent economy, this is not a problem either: If the agent chooses the next-period capital stock a_{t+1}, he also knows that all other agents (with mass one) choose a_{t+1} such that the aggregate capital stock is given by $K_{t+1} = a_{t+1}$. In the heterogeneous-agent economy, however, the individual household is unable to infer the value of the next-period aggregate capital stock K_{t+1} from his own decision. He needs to know how all the other (infinitely-numbered) households in the economy decide and how much they save. As households with different wealth may also have different savings rates and incomes, he needs to know the distribution of the individual states and to sum the next-periods assets over all households in order to compute K_{t+1} and, hence, r_{t+1}.[3] As a consequence, the distribution $f_t(\epsilon_t, a_t)$ is also an argument of the policy function. Put differently, if we consider different distributions $f_t(\epsilon_t, a_t)$ that are characterized by the same mean $\bar{a}_t = K_t$, we will have different next-period distributions $f_{t+1}(\epsilon_{t+1}, a_{t+1})$ which only by chance will all have the same mean $\overline{a}_{t+1} = K_{t+1}$. For this reason, $f_t(\cdot)$ needs to be included as an additional argument of the policy function.

We are now ready to formulate the recursive problem. We will omit the time index t from the variables in the definition of the recursive equilibrium in order to keep the notation as simple as possible. The households maximizes his or her value function:

[2] Alternatively, we could have used the variable r_t rather than K_t as an argument of $a'(\cdot)$.

[3] In order to compute the savings of the other households, he also needs to know the policy functions $a'(\epsilon_t, a_t, f_t)$ of the other agents. As all agents, however, have the same optimization problem, the policy functions of all agents are identical. Consequently, as the individual households knows his own policy function, he also knows the policy functions of the other agents.

$$V(\epsilon, a, f) = \max_{c,a'} \left[u(c) + \beta E \left\{ V(\epsilon', a', f') \right\} \right], \tag{6.6}$$

subject to the budget constraint (5.4), the government policy $\{b, \tau\}$, the stochastic process of the employment status ϵ as given by (5.3), and the distribution dynamics (6.2). Again, the value function is a function of individual states ϵ and a, and the distribution $f(\cdot)$. The distribution of the assets, f, however, is an infinite-dimensional object and we cannot track it. Furthermore, finding the law of motion for the distribution, $G(f)$, is not trivial as G is a map from the set of functions (an infinite dimensional space) into itself.

6.2 Transition Dynamics

In this section, we consider the transition dynamics for given initial state in the economy with aggregate certainty as described by the following Example:[4]

Example 6.2.1

Households are allocated uniformly along the unit interval and are of measure one. The individual household maximizes

$$V(\epsilon, a, f) = \max_{c,a'} \left[\frac{c_t^{1-\eta}}{1-\eta} + \beta E \left\{ V(\epsilon', a', f') \right\} \right],$$

s.t.

$$a' = \begin{cases} (1 + (1-\tau)r)\, a + (1-\tau)w - c, & \text{if } \epsilon = e, \\ (1 + (1-\tau)r)\, a + b - c, & \text{if } \epsilon = u, \end{cases}$$

$$a \geq a_{min},$$

$$\pi(\epsilon'|\epsilon) = Prob\left\{ \epsilon_{t+1} = \epsilon' | \epsilon_t = \epsilon \right\} = \begin{pmatrix} p_{uu} & p_{ue} \\ p_{eu} & p_{ee} \end{pmatrix}.$$

The distribution f of (ϵ, a) is described by the following dynamics:

$$f'(\epsilon', a') = \sum_{\epsilon \in \{e,u\}} \pi(\epsilon'|\epsilon)\; f(\epsilon, a'^{-1}(\epsilon, a', f)).$$

[4] Please keep in mind that aggregate variables like r, w, or τ vary over time and are not constant out of the steady state, even though we do not index them.

Factor prices are equal to their respective marginal products:

$$\begin{aligned} r &= \alpha \left(\frac{N}{K}\right)^{1-\alpha} - \delta, \\ w &= (1-\alpha)\left(\frac{K}{N}\right)^{\alpha}. \end{aligned}$$

The aggregate consistency conditions hold:

$$\begin{aligned} K &= \sum_{\epsilon \in \{e,u\}} \int_{a_{min}}^{\infty} a f(\epsilon, a)\, da, \\ C &= \sum_{\epsilon \in \{e,u\}} \int_{a_{min}}^{\infty} c(\epsilon, a, f)\, da, \\ T &= \tau(wN + rK), \\ B &= \int_{0}^{\infty} b f(u, a)\, da, \end{aligned}$$

The government policy is characterized by a constant replacement ratio $\zeta = b/(1-\tau)w$ and a balanced budget: $T = B$.

We will introduce two ways in order to approximate the dynamics of the distribution (6.2). The first way is to use partial information and has been applied by Den Haan (1997) and Krusell and Smith (1998). The basic idea is that households do not use all the information at hand, i.e., the distribution f, but only use a little bit of information about f, for example the first moment. By this device, we reduce the infinite-dimensional problem of finding a law of motion for f to a finite-dimensional one. The second method is a shooting method which is only applicable to models with aggregate certainty. In this case, one assumes that one reaches the new stationary equilibrium after T periods and projects a transition path for the prices $\{w, r\}_{t=0}^{T}$ over the next T periods.[5] Given the dynamics of the prices, the optimal policy functions and the dynamics of the distribution can be computed. From this, we can update the time path for the factor prices until the algorithm

[5] If we considered aggregate uncertainty, we would have to project a distribution over the factor prices which would make the problem much more complicated, again.

converges. We will present the two approaches in turn. Both approaches assume that the stationary equilibrium is stable and that the distribution function converges to the invariant distribution function.

6.2.1 Partial Information

In the subsection, we assume that agents only use partial information in order to predict the law of motion for the state variable(s) or, equivalently, are boundedly rational. Agents perceive the dynamics of the distribution $f' = G(f)$ in a simplified way. In particular, they characterize the distribution f by I statistics $m = (m_1, \ldots, m_I)$. In Chapter 5, we approximated the invariant distribution function with an exponential function. One might use the parameters ρ_i of the approximated exponential distribution function as statistics m_i, for example. In this section, we follow Krusell and Smith (1998) and use the moments of the distribution function, instead. In particular, we only consider the simple case that agents only use the first moment m_1, ie the aggregate capital stock K. Krusell and Smith find that the forecast error due to the omission of higher moments is extremely small.[6] The economic intuition for this result is straightforward. Higher moments of the wealth distribution only have an effect on aggregate next-period capital stock if agents of different wealth level have different propensities to save out of wealth. However, most agents (except for the very poor ones, who, of course, do not contribute much to total savings) have approximately the same savings rate.[7] Therefore, the omission of higher moments is justified for the present case.

Accordingly, we assume that agents perceive the law of motion for m as follows:

[6] In exercise 2, you will be asked to verify this hypothesis.

[7] Empirically, high income households save a larger fraction than low income households in the US. Huggett and Ventura (2000), however, show that age and relative permanent earnings differences across households together with the social security system are sufficient to replicate this fact. All these factors are absent from the model in Example 6.2.1.

$$m' = H_I(m). \tag{6.7}$$

Given the law of motion for m and the initial value of m, each agent optimizes his intertemporal consumption allocation by solving the following problem:

$$V(\epsilon, a, m) = \max_{c,a'} \left[u(c) + \beta E\left\{V(\epsilon', a', m')\right\}\right], \tag{6.8}$$

subject to the budget constraint (5.4), the government policy $\{b, \tau\}$, the stochastic process of the employment status ϵ as given by (5.3), and the distribution dynamics (6.7). Again, the factor prices are computed as functions of the aggregate capital stock and employment, $w = w(K, N)$ and $r = r(K, N)$, where the aggregate capital stock is given by the first moment of the distribution $K = m_1$.[8] Similarly, we can compute the income tax rate τ and the unemployment compensation b from the balanced budget and for given replacement ratio ζ, aggregate capital K and employment N for every period t:

$$\begin{aligned} T &= \tau K^{\alpha} N^{1-\alpha} = B = (1-N)b, \\ b &= \zeta(1-\tau)w = \zeta(1-\tau)(1-\alpha)\left(\frac{K}{N}\right)^{-\alpha}. \end{aligned}$$

The remaining problem is to approximate the law of motion for the moments m of the distribution, $m = m_1 = K$. We will choose a simple parameterized functional form for $H_I(m)$ following Krusell and Smith (1998):

$$\ln K' = \gamma_0 + \gamma_1 \ln K. \tag{6.9}$$

Given the function H_I, we can solve the consumer's problem and compute optimal decision functions. For given initial distribution f_0 with mean K_0, we can simulate the behavior of the economy over time, and in particular, can compute the law of motion for K and can compare it to our projection (6.9). If the goodness of fit is not satisfactory, we might want to try a different functional

[8] Again, we drop the time index from the variables in order to keep the notation simple.

form for H_I or try a higher order I. As it turns out, one moment, $I = 1$, and the functional form (6.9) are quite satisfactory. The algorithm can be described by the following steps:

Algorithm 6.2.1 (Transition Dynamics with Bounded Rationality)

Purpose: *Computation of the transition dynamics of the distribution function for Example 6.2.1 with given initial distribution f_0 and the dynamics as given by* (6.7)*.*

Steps:

Step 1: Choose the initial distribution of assets f_0 with mean K_0.
Step 2: Choose the order I of moments m.
Step 3: Guess a parameterized functional form for H_I and choose initial parameters of H_I.
Step 4: Solve the consumer's optimization problem and compute $v(\epsilon, a, m)$.
Step 5: Simulate the dynamics of the distribution.
Step 6: Use the time path for the distribution to estimate the law of motion for the moments m.
Step 7: Iterate until the parameters of H_I converge.
Step 8: Test the goodness of fit for H_I. If the fit is satisfactory, stop, otherwise increase I or choose a different functional form for H_I.

Calibration. The model of Example 6.2.1 is calibrated as the model of Example 5.2.1. In particular, the parameter values are given by $\alpha = 0.36$, $\beta = 0.995$, $\sigma = 2$, $\delta = 0.005$, $\zeta = 0.25$,[9] and the employment transition matrix:

$$\begin{pmatrix} p_{uu} & p_{ue} \\ p_{eu} & p_{ee} \end{pmatrix} = \begin{pmatrix} 0.500 & 0.500 \\ 0.0435 & 0.9565 \end{pmatrix}.$$

[9] Different from Example 5.2.1, we, however, do not need to fix unemployment benefits b in order to ensure convergence of the algorithm, but can rather use the replacement ratio ζ instead. This facilitates the calibration because the replacement ratio ζ is readily observable from empirical data, contrary to the absolute amount of unemployment benefits b.

The minimum wealth a_{min}, again, is set equal to -2.

Implementation. The algorithm is implemented in the program `ch6_part.g`. We choose an equispaced grid $\mathcal{A} = \{a_1, \ldots, a_n\} = \{-2, \ldots, 600\}$ for wealth a with $n = 101$ nodes. For the aggregate capital stock K, we also choose an equispaced grid $\mathcal{K} = \{K_1, \ldots, K_{n_K}\} = \{100, \ldots, 500\}$.

In the first step, we have to initialize the distribution function. We assume that at time period $t = 0$, the distribution is uniform over the interval $[-2, 300]$ implying an aggregate capital stock K equal to the mean $\bar{K} = 149$. If we considered a policy change, e.g. an increase of unemployment benefits, we would have computed the invariant distribution of wealth prior to the policy change with the help of the methods developed in the previous chapter and would have used this distribution for our initialization of f. In the second step, we set the order I equal to one, i.e. consumers only use the first moment of the wealth distribution as information about the distribution f. In step 3, we choose the log-linear law of motion (6.9) for the capital stock and initialize the parameters $\gamma_0 = 0.09$ and $\gamma_1 = 0.95$.

For the solution of the consumer's optimization problem in step 4, we resort to the methods presented in the first part of this book. In the computer program `ch6_part.g`, we use value function iteration with linear interpolation. Our objective is to compute the value function $v(\epsilon, a, K)$ for $\epsilon = e, u$. As we only consider the first moment of the distribution f, the value function is only a function of the employment status ϵ, the individual wealth a, and the aggregate capital stock K. We choose a coarse grid for K and again use linear interpolation to approximate the value function at points off the grid points. The grid for the aggregate capital stock K consists of $n_K = 6$ nodes. The initial value function for the employed agent, $v^e = v(e, a, K)$, and the unemployed agent, $v^u = v(u, a, K)$, are computed at the grid points (a_i, K_j), $i = 1, \ldots, n$ and $j = 1, \ldots, n_K$, assuming that agents consume their current income permanently:

$$v_0^e(a_i, K_j) = \sum_{t=0}^{\infty} \beta^t u\left((1-\tau)r(K_j)a_i + (1-\tau)w(K_j)\right)$$

$$= \frac{1}{1-\beta} u\left((1-\tau) r(K_j) a_i + (1-\tau) w(K_j)\right),$$

$$v_0^u(a_i, K_j) = \frac{1}{1-\beta} u\left((1-\tau) r(K_j) a_i + b(K_j)\right).$$

The interest rate and the wage rate, of course, are functions of the aggregate capital stock K_j (and so are the unemployment benefits $b = \zeta(1-\tau)w(K_j)$).

For given value function in iteration l, we can compute the value function of the employed agent, for example, in the next iteration $l+1$ from:

$$v_{l+1}^e(a_i, K_j) = \max_{a' \in \mathcal{A}} \; u(c) + \beta \left\{ p_{ee} v_l^e\left(a', e^{\gamma_0 + \gamma_1 \ln K_j}\right) + p_{eu} v_l^u\left(a', e^{\gamma_0 + \gamma_1 \ln K_j}\right) \right\}$$

with

$$c = \left(1 + (1-\tau) r(K_j)\right) a_i + (1-\tau) w(K_j) - a'.$$

The value function is computed for every aggregate capital stock $K_j \in \mathcal{K}$ and $a_i \in \mathcal{A}$. The outer loop of the iteration is over the capital stock K_j. Given $K = K_j$, we can compute the factor prices w and r, unemployment compensation b, income taxes τ, and the next-period capital stock K'. For given w, r, b, τ, and K', we can compute the value function $v^e(a, K_j)$ and $v^u(a, K_j)$ at every grid point $a = a_i$. Notice that we do not have to compute the function $v^e(a', K')$ and $v^u(a', K')$ on the rhs of the Bellman equation for $a' \in \mathcal{A}$ at each iteration over the individual wealth $a_i \in \mathcal{A}$ but only once before we start the iteration over a because we know K' in advance. In the program `ch6_part`, we store the value functions $v^e(a', K')$ and $v^u(a', K')$ for $K' = K'(K_j)$ and $a' = a_1, \ldots, a_n$ in the vectors `ve1` and `vu1`, respectively, before we start the iteration over $a \in \mathcal{A}$. In order to find the optimum a', we only need to use the values from these one-dimensional vectors (or interpolate linearly between two values $a_m < a' < a_{m+1}$ of these vectors, respectively). The maximization of the rhs of Bellman's equation is performed using the Golden Section Search procedure introduced in Section 8.6.1.

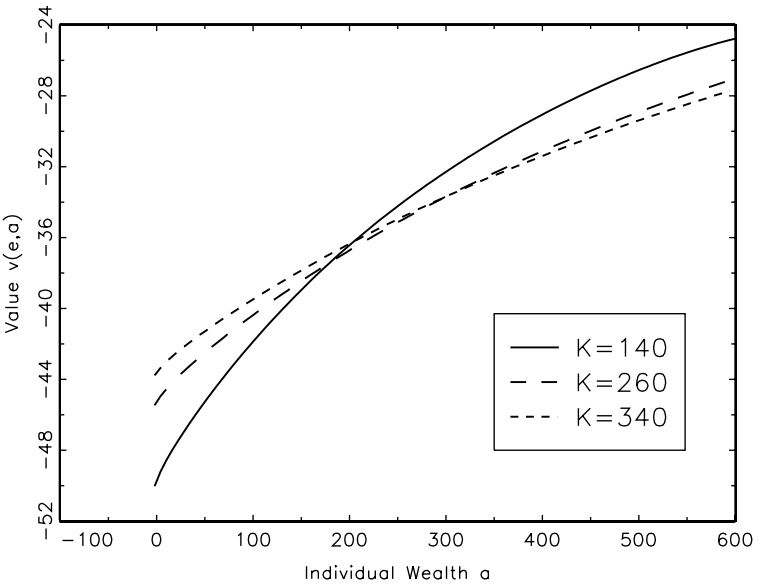

Figure 6.1: Value function of the employed worker

The computed value function of the employed consumer is displayed in Figure 6.1.[10] The value function is a concave increasing function of individual wealth a. For the low aggregate capital stock $K = 140$ (solid line), the interest rate r is high and the wage rate w is low. For low wealth a, therefore, the value is lower than the value for a larger aggregate capital stock $K = 260$ or $K = 340$ (broken and dotted line, respectively). At an intermediate wealth level a, approximately equal to 200, the value function of the agents is again higher for low capital stock $K = 140$ compared to the other cases with a higher capital stock K because the interest income ra becomes a more important component of total income.

The savings behavior of the households is displayed in Figure 6.2. Savings increase with higher interest rate r or, equally, lower capital stock K. The savings functions $a' - a$ of the employed worker is presented by the solid and broken line for aggregate

[10] The value function of the unemployed worker behaves exactly as the one of the employed worker.

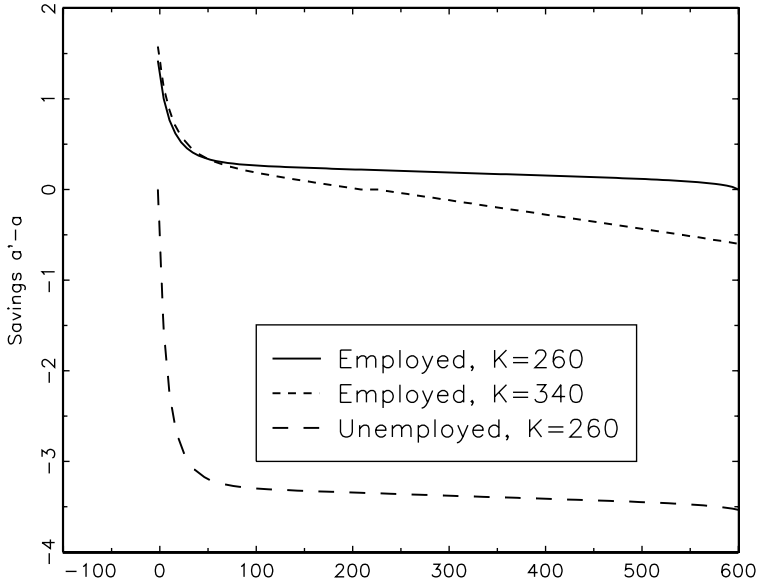

Figure 6.2: Savings of the employed and unemployed worker

capital stocks $K = 260$ and $K = 340$, respectively. Savings of the unemployed workers is smaller than those of the employed workers as they have lower wage income. The dotted line in Figure 6.2 displays the savings function of the unemployed worker if the aggregate capital stock is equal to $K = 260$.

In step 5, we compute the dynamics of the distribution function. Given f_0, we can compute f_t from f_{t-1} with the help of the agent's savings function. As in the previous chapter, we choose a finer grid for the wealth level a for the computation of the distribution function. In the program `ch6_part.g`, we choose $2n$ equispaced nodes. The optimal policy functions off grid points (a, K) are computed with the help of linear interpolation. The dynamics of the distribution function are displayed in Figure 6.3, which presents the distribution function of the employed workers at period $t = 1$, $t = 10$, $t = 20$, and $t = 1000$. After 1000 iterations, the distribution function is stationary and the transition is complete. The mean of the distribution is also constant after 1000 iterations

and amounts to $\bar{K} = 247.5$. The convergence of the distribution's mean is displayed in Figure 6.4.

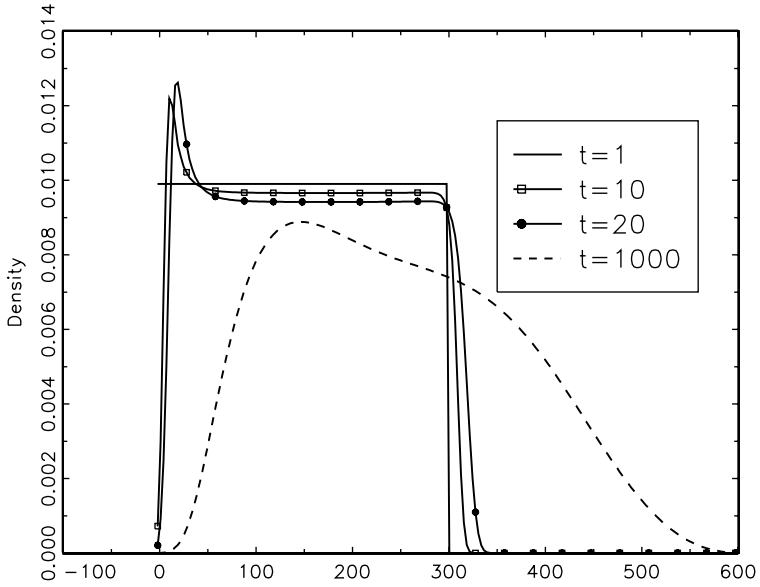

Figure 6.3: Dynamics of the distribution function over time

If you compare Figures 6.4 and 5.1, you cannot help noticing that convergence of the distribution's mean (and also of the higher moments) occurs much faster in Figure 6.4 than in Figure 5.1, i.e. the distribution function approaches the stationary distribution much faster in the case where we model the transition of the aggregate capital stock. What is the reason for this observation? Assume we start with a distribution that has an initial mean $\bar{a}_0$ below the stationary mean $\bar{K}$. In the computation of the stationary equilibrium in Chapter 5, we assumed that $K_t = \bar{K}$ in every period, while we use the aggregate capital stock $K_0 = \bar{a}_0$ in the Algorithm 6.2.1. Accordingly, the interest rate is lower in the economy with a constant interest rate than in the economy where we model the transition dynamics. Therefore, agents have lower savings in each period in the constant price economy and the mean of the distribution adjusts at a slower pace.

Consequently, the step 5 in Algorithm 5.2.1 where we compute the stationary distribution is much more time-consuming than the corresponding step 5 in Algorithm 6.2.1. However, it would be wrong to conclude that Algorithm 6.2.1 is faster than Algorithm 5.2.1. In fact, Algorithm 6.2.1 is about 3-4 times slower in the computation of the Example 5.2.1 than Algorithm 5.2.1. Importantly, using Algorithm 6.2.1, we i) need to iterate over the law of motion for the aggregate capital stock and ii) have to compute the policy functions for a state space that is characterized by a higher dimension. As a consequence, the computational time is much higher in the latter case and, therefore, if we only aim at computing stationary equilibria, we would rather apply the methods presented in Chapter 5.

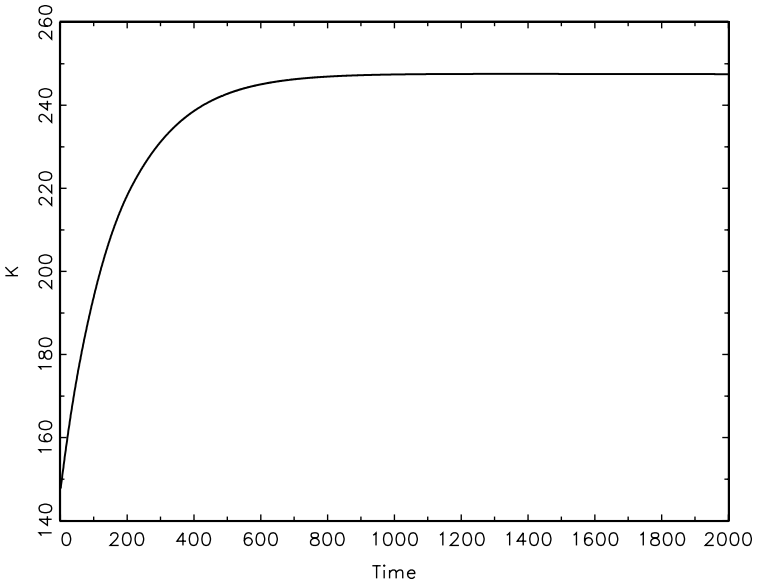

Figure 6.4: Convergence of the aggregate capital stock

We can use the time path of the capital stock displayed in Figure 6.4 to update the coefficients γ_0 and γ_1 (step 6 and 7). We use ordinary least squares regression (OLS) to compute the two coefficients (step 6). However, we only use the first 500 values for

the capital stock $\{K_t\}_{t=0}^{T}$. Close to the stationary value of K, we only have observation points (K, K') where K and K' are almost equal and we have only little variation in K and K'. In this regression, we get a very high R^2, almost equal to one, $R^2 = 0.995$. The computed dynamics (K, K') (simulated) and the regression line (as predicted by the households with the help of (6.9)) are almost identical and the mean absolute deviation amounts to 0.053%. Obviously, the fit is extremely well. In step 7, we update the parameters γ_0 and γ_1 until they converge. The final solution for the law of motion for the capital stock is given by

$$\ln K' = 0.041910 + 0.992388 \ln K. \tag{6.11}$$

This equation implies a stationary capital stock equal to $\bar{K} = e^{\frac{\gamma_0}{1-\gamma_1}} = 246.1$, which is a little lower than the one computed from the simulation ($\bar{K} = 247.5$). For γ_1 close to one, small errors in the estimation of γ_i, $i = 0, 1$, imply high errors in the computation of $\bar{K}$. For $\gamma_0 = 0.041953$ and $\gamma_1 = 0.992388$, the stationary capital stock is already equal to $\bar{K} = 247.5$.

In the present model, K is a sufficient statistic for factor prices and taxes. We can compute the wage rate w, the interest rate r, the tax rate τ that balances the government budget, and the unemployment compensation if we only know K. In many economic applications, however, the distribution of wealth and its mean are not a sufficient statistic of factor prices. Consider the case of an elastic labor supply. Households maximize their utility by their choice of leisure. For example, if we assume instantaneous utility to be of the form

$$u(c, 1-n) = \frac{\left(c(1-n)^{\theta}\right)^{1-\eta}}{1-\eta}, \tag{6.12}$$

where n denotes labor supply and $1-n$ is leisure (the time endowment of the household is normalized to one). Labor income of the employed worker is simply the net wage rate times the working hours, $(1-\tau)wn$, and aggregate labor N in period t is given by

$$N = \int_{a_{min}}^{\infty} n(a; K, N)\ f(e, a)\ da, \tag{6.13}$$

where the labor supply of the unemployed is equal to zero. In this case, individual labor supply depends on individual wealth a and, consequently, aggregate labor supply N depends on the distribution of wealth.[11] In this case, we also need to estimate a prediction function for aggregate labor

$$N' = J(N, K), \tag{6.14}$$

that, for example, might take the log-linear form $\ln N' = \psi_0 + \psi_1 \ln N + \psi_2 \ln K$. The household maximizes intertemporal utility subject to the additional constraint (6.14) and the value function $v(\epsilon, a, K, N)$ has the aggregate labor N as an additional argument. Of course, the method developed in this chapter is still applicable to such more complicated problems and you will be asked to solve the growth model with leisure in the exercises.

6.2.2 Guessing a Finite Time Path for the Factor Prices

In the previous section, we computed the value function as a function of the aggregate capital stock. If the model is getting more complex, e.g. if we consider endogenous labor supply, endogenous technology or multiple financial assets, the number of arguments in the value function rises and the computation becomes more cumbersome. In this section, we introduce another method for the computation of the transition path that only considers the individual variables as arguments of the value function (or policy functions). The only additional variable of both the value function and the policy functions is time t. The method presented in this chapter, however, is only applicable to deterministic economies.

Again, we consider the transition to a stationary equilibrium. For the computation, we assume that the stationary equilibrium is reached in finite time, after T periods. Typically, we choose T large enough, say $T = 1000$ or higher. Furthermore, we can compute the stationary equilibrium at period $t \geq T$ with the help of the methods developed in the previous chapter. We also

[11] Individual labor supply also depends on the wage rate and, hence, on K and N.

know the distribution of wealth in the initial period $t = 1$ and, therefore, the aggregate capital stock and factor prices in period $t = 1$ and $t = T$. In order to compute the policy functions during the transition, we need to know the time path of the factor prices, or, equally, the time path of the aggregate capital stock. We start with an initial guess for the time path of the factor prices, compute the decision functions, and with the help of the initial distribution and the computed decision functions, we are able to compute the time path of the factor prices. If the guess is different from the values implied by our simulation, we update the guess accordingly. The algorithm can be described by the following steps:[12]

Algorithm 6.2.2 (Computation of Example 6.2.1)

Purpose: *Computation of the transition dynamics by guessing a finite-time path for the factor prices*

Steps:

Step 1: Choose the number of transition periods T.

Step 2: Compute the stationary distribution $\tilde{f}$ of the new stationary equilibrium. Initialize the first-period distribution function f^1.

Step 3: Guess a time path for the factor prices r and w, unemployment compensation b, and the income tax rate τ that balances the budget. The values of these variables in both periods $t = 1$ and $t = T$ are implied by the initial and stationary distribution, respectively.

Step 4: Compute the optimal decision functions using the guess for the interest rate r, the wage income w, the tax rate τ and the unemployment compensation b. Iterate backwards in time, $t = T - 1, \ldots, 1$.

Step 5: Simulate the dynamics of the distribution with the help of the optimal policy functions and the initial distribution for the transition from $t = 1$ to $t = T$.

[12] The algorithm follows Ríos-Rull (1999) with some minor modifications.

Step 6: Compute the time path for the interest rate r, the wage w,unemployment compensation b, and the income tax rate τ, and return to step 3, if necessary.

Step 7: Compare the simulated distribution f^T with the stationary distribution function $\tilde{f}$. If the goodness of fit is poor, increase the number of transition periods T.

In step 4, we compute the optimal policy functions by backward iteration. In period T, we know the new stationary distribution, optimal policy functions, and the factor prices. For periods $t = T-1, \ldots, 1$, we may compute the policy functions $c_t(\epsilon_t, a_t)$ and $a_{t+1}(\epsilon_t, a_t)$ for consumption and next-period assets with the methods developed in Part 1 of this book recursively.[13] For example, we may compute $c_t(\epsilon_t, a_t)$ and $a_{t+1}(\epsilon_t, a_t)$ for given policy functions $c_{t+1}(\epsilon_{t+1}, a_{t+1})$ and $a_{t+2}(\epsilon_{t+1}, a_{t+1})$ from the Euler equation (5.5) with the help of projection methods:[14]

$$\frac{u'(c_t(\epsilon_t, a_t))}{\beta} = E_t\left[u'(c_{t+1}(\epsilon_{t+1}, a_{t+1}))(1 + (1-\tau_{t+1})r_{t+1})\right], \quad \epsilon_t = e, u, \tag{6.15}$$

where $c_t(e, a_t) = (1 + r_t(1-\tau_t))a_t + (1-\tau_t)w_t - a_{t+1}(e, a_t)$ and $c_t(u, a_t) = (1 + r_t(1-\tau_t))a_t + b_t - a_{t+1}(u, a_t)$ for the employed and unemployed worker, respectively. Alternatively, we may compute the optimal policy functions with value function iteration from the Bellman equation (5.9):

$$V_t(\epsilon_t, a_t) = \max_{c_t, a_{t+1}} \left[u(c_t) + \beta E_t\left\{V_{t+1}(\epsilon_{t+1}, a_{t+1})\right\}\right]. \tag{6.16}$$

In period $t = T-1$, again, we know the optimal next-period consumption policy c_T and the value function V_T which are equal

[13] Notice that with the present method, the policy functions are no longer time-invariant. Optimal consumption $c_t(\cdot)$ depends on the period t via the factor prices w_t and r_t which are not arguments of the policy function. Therefore, we have to compute the optimal policy function in every period $t = 1, \ldots, T$.

[14] You will be asked to compute the transition dynamics using projection methods for the computation of the policy functions in the exercises.

to the stationary optimal consumption policy and value function, respectively. Notice that, in order to compute the policy functions from (6.16), we need to store V_t, $t = 1, \ldots, T$, but the value function is only a function of individual state variables ϵ and a. We iterate backwards in time and compute c_t given V_{t+1}.

As soon as we have computed the optimal consumption policy for $t = 1, \ldots, T$, with the help of the value function or the Euler equation, it is straightforward to simulate the behavior of the economy with the help of the first-period distribution function and compute the time path for the capital stock $\{K_t\}_{t=1}^T$.

Implementation. The algorithm is implemented in the program `ch6_gues.g`. The model of Example 6.2.1 is calibrated in exactly the same way as in the previous section. We also choose the same grid over the asset space $\mathcal{A}$ for the value function and the distribution function as in the program `ch6_part.g` of the previous section. The transition time is set equal to $T = 2000$ periods. The program loads the stationary policy function and distribution function as an input into the computation.[15] In step 2, the initial distribution is chosen to be the uniform distribution over the interval of $[a_{min}, a_{max}/2]$ as in the previous section.

There are various ways to update the time path for $\{K_t\}_{t=0}^T$ in step 6. We may either resort to a parameterized recursive equation (6.9) as in the previous section and adjust the parameters γ_0 and γ_1 as in the program `ch6_part.g` . Alternatively, we may use a tatonnement algorithm guessing an initial sequence $\{K_t\}_{t=0}^T$ and updating it after each iteration i, $K_t^i = K_t^{i-1} + \phi\left(K_t^{i-1} - K_t^i\right)$, $t = 2, \ldots, T-1$. This approach is used in the program `ch6_gues.g`. A third approach is to use any standard nonlinear equation solution method, e.g. Newton's method to find the sequence $\{K_t\}_{t=0}^T$ that implies the same sequence for the simulated model. In the present case, the algorithm converges after approximately 20 iterations over the sequence $\{K_t\}_{t=0}^{t=2000}$. The simulated time path and the projected time path of the capital stock are almost identical

[15] The function and distribution are computed with the help of the program `ch5_disf.g` as described in Chapter 5.

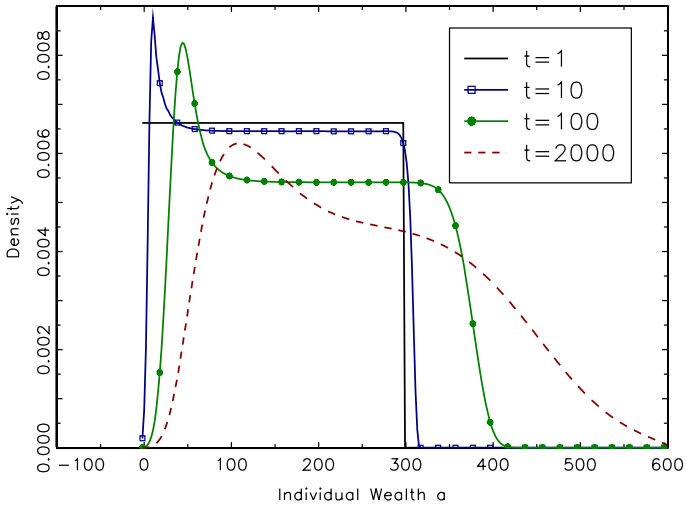

Figure 6.5: The dynamics of the distribution function

and the deviation only amounts to 0.098% on average during the transition.

The dynamics of the distribution over time is displayed in Figure 6.5. From the initial uniform distribution (solid line), the distribution slowly converges to the final distribution in period $T = 2000$. The distribution at the end of the transition (solid line) and the new stationary distribution of wealth a (broken line) are shown in Figure 6.6. The two distributions have the almost the same mean, $K = 244.29$ and $K = 244.38$, for the simulated distribution after 2000 periods and the new stationary distribution, respectively. However, the second moments vary as the tail of the distribution function after 2000 periods is thinner than the one of the new stationary distribution. A longer transition period may even improve the fit.

6.3 Aggregate Uncertainty

So far, we have only considered individual risk in this chapter. Agents faced the idiosyncratic risk of getting unemployed, while

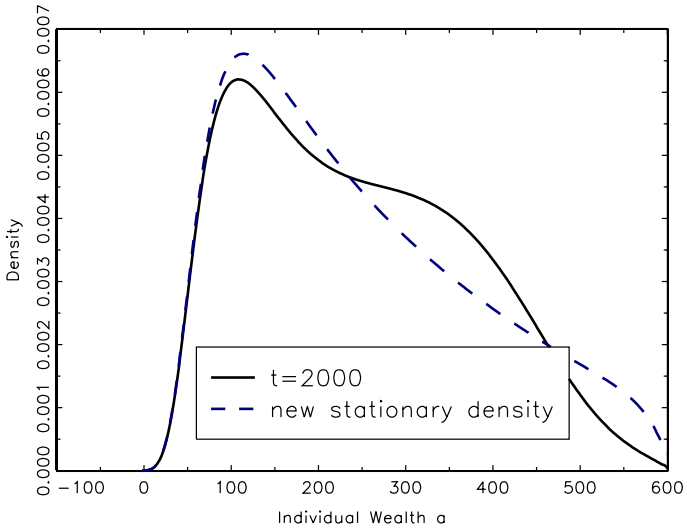

Figure 6.6: Goodness of fit for stationary distribution

the real interest rate and the factor prices were constant in the stationary state. Only during transition to the new steady state did factor prices vary. In this section, we also take a look at aggregate risk. As in Chapter 1, aggregate risk is introduced by a stochastic technology level Z_t in period t. In particular, the productivity shock follows a Markov process with transition matrix $\Gamma_Z(Z'|Z)$, where Z' denotes next-period technology level and $\pi_{ZZ'}$ denotes the transition probability from state Z to Z'. This assumption is not very restrictive. Given empirical evidence, we assumed in Chapter 1 that productivity Z_t followed an AR(1)-process. As you also learn in Section 9.2, an AR(1) process can easily be approximated by a finite Markov-chain.[16]

With stochastic technology level Z_t, aggregate production is given by:

$$F(K_t, N_t) = Z_t K_t^{\alpha} N_t^{1-\alpha}. \tag{6.17}$$

[16] In the exercises, you will be asked to compute the solution for a heterogenous-agent economy with aggregate uncertainty where the productivity shocks follows an AR(1)-process using Algorithm 9.2.1.

We assume competitive factor and product markets implying the factor prices:

$$w_t = Z_t(1-\alpha)K_t^{\alpha}N_t^{-\alpha}, \tag{6.18}$$

$$r_t = Z_t\alpha K_t^{\alpha-1}N_t^{1-\alpha} - \delta. \tag{6.19}$$

The individual employment probabilities, of course, depend on the aggregate productivity Z_t. In good times (high productivity Z_t), agents have higher employment probabilities than in bad times. The joint process of the two shocks, Z_t and ϵ_t, can be written as a Markov process with transition matrix $\Gamma(Z', \epsilon'|Z, \epsilon)$. We use $p_{Z\epsilon Z'\epsilon'}$ to denote the probability of transition from state (Z, ϵ) to state (Z', ϵ'). In the following, we restrict our attention to a very simple example. The economy only experiences good and bad times with technology levels Z_g and Z_b, respectively, $Z_g > Z_b$. As before, agents are either employed ($\epsilon = e$) or unemployed ($\epsilon = u$). Consequently, the joint processes on (Z, ϵ) are Markov chains with 4 states.

Households are assumed to know the law of motion of both $\{\epsilon_t\}$ and $\{Z_t\}$ and they observe the realization of both stochastic processes at the beginning of each period. Besides, the model is identical to the one in Example 6.2.1 and is summarized in the following:

Example 6.3.1

Households are of measure one. The individual household maximizes

$$V(\epsilon, a, Z, f) = \max_{c,a'}\left[\frac{c_t^{1-\eta}}{1-\eta} + \beta E\left\{V(\epsilon', a', Z', f')\right\}\right],$$

s.t.

$$a' = \begin{cases} (1+(1-\tau)r)\,a + (1-\tau)w - c & \text{if } \epsilon = e, \\ (1+(1-\tau)r)\,a + b - c & \text{if } \epsilon = u, \end{cases}$$

$$a \geq a_{min},$$

$$\Gamma(Z', \epsilon'|Z, \epsilon) = Prob\left\{Z_{t+1} = Z', \epsilon_{t+1} = \epsilon'|Z_t = Z, \epsilon_t = \epsilon\right\}$$

$$= \begin{pmatrix} p_{Z_geZ_ge} & p_{Z_geZ_gu} & p_{Z_geZ_be} & p_{Z_geZ_bu} \\ p_{Z_guZ_ge} & p_{Z_guZ_gu} & p_{Z_guZ_be} & p_{Z_guZ_bu} \\ p_{Z_beZ_ge} & p_{Z_beZ_gu} & p_{Z_beZ_be} & p_{Z_beZ_bu} \\ p_{Z_buZ_ge} & p_{Z_buZ_gu} & p_{Z_buZ_be} & p_{Z_buZ_bu} \end{pmatrix}.$$

The distribution of the individual states (ϵ, a) for given aggregate state variables (Z, K) in period t is denoted by $f(\epsilon, a; Z, K)$. The dynamics of the distribution of the individual states are described by the following equations:

$$f'(\epsilon', a'; Z', K') = \sum_{\epsilon} \Gamma(Z', \epsilon' | Z, \epsilon) f(\epsilon, a; Z, K),$$

where $a' = a'(\epsilon, a; Z, K)$ and

$$K' = \sum_{\epsilon} \int_a a' \; f(\epsilon, a; Z, K) \; da.$$

Factors prices are equal to their respective marginal products:

$$\begin{aligned} r &= \alpha Z_t \left(\frac{N}{K}\right)^{1-\alpha} - \delta, \\ w &= (1-\alpha) Z \left(\frac{K}{N}\right)^{\alpha}. \end{aligned}$$

The aggregate consistency conditions hold:

$$\begin{aligned} K &= \sum_{\epsilon} \int_a a \; f(\epsilon, a; Z, K) \; da, \\ N &= \int_a f(e, a; Z, K) \; da, \\ C &= \sum_{\epsilon} \int_a c(\epsilon, a; Z, K) \; f(\epsilon, a; Z, K) \; da, \\ T &= \tau(wN + rK), \\ B &= \int_a b f(u, a; Z, K) \; da. \end{aligned}$$

The government policy is characterized by a constant replacement ratio $\zeta = b/(1-\tau)w$ and a balanced budget: $T = B$. ———————

Due to the presence of aggregate uncertainty, there are three major changes in the computation of the model compared to the one in Section 6.2:

1) we have fluctuating employment levels.

2) When we approximate the distribution function of wealth by its first I moments, for example, the value function is a function of the employment status ϵ, individual wealth a, the first I moments of wealth, and, in addition, aggregate technology Z.
3) The distribution of wealth is not stationary. We will discuss these three points in turn.

1) Individual employment probabilities depend on both the current employment status ϵ and current and next-period productivity, Z and Z'. Given an employment distribution in period t, the next-period employment distribution depends on the technology level Z' because agents have a higher job finding probability in good times, $Z' = Z_g$, than in bad times, $Z' = Z_b$. As a consequence, we have an additional state variable in the model, namely aggregate employment. As the measure of households is normalized to one, aggregate employment is equal to $N_t = 1 - u_t$, where u_t is the unemployment rate of the economy in period t. As the factor prices are functions of both aggregate capital K_t and employment N_t, the households need to predict the law of motion for both state variables. Aggregate employment next period N', however, only depends on aggregate employment in the current period N and the technology level in this and the next-period, Z and Z', because we assume inelastic labor supply. As a consequence, the agents know that next-period aggregate employment is either N'_g if $Z' = Z_g$ or N'_b if $Z' = Z_b$ because they know the transition matrix Γ, the current period employment N, and the technology level Z.[17]

We will simplify the analysis further following Krusell and Smith (1998). In particular, we assume that the unemployment rate takes only two values u_g and u_b in good times and in bad times, respectively, with $u_g < u_b$. In order to simplify the dynamics of aggregate employment accordingly, the following restrictions have to be imposed on the transition matrix Γ:

$$u_Z \frac{p_{ZuZ'u}}{p_{ZZ'}} + (1 - u_Z) \frac{p_{ZeZ'u}}{p_{ZZ'}} = u_{Z'}, \tag{6.21}$$

[17] Furthermore, the law of large numbers holds.

for $Z, Z' \in \{Z_g, Z_b\}$. Condition (6.21) implies that unemployment is u_g and u_b if $Z' = Z_g$ and $Z' = Z_b$, respectively. Consequently, we do not need to consider employment as an additional state variable in the special case (6.21) as the technology level Z' is a sufficient statistic for N'. Example 6.3.1 has already been formulated accordingly and the state variable is given by $\{\epsilon, a, Z, f\}$ rather than $\{\epsilon, a, Z, N, f\}$.[18]

2) In comparison with Example 6.2.1, the households' value function has an additional argument, the technology level Z. The Bellman equation can be formulated as follows:

$$V(\epsilon, a, Z, f) = \max_{c,a'} \left[u(c) + \beta E \left\{V(\epsilon', a', Z', f')\right\}\right].$$

The additional state variable Z has a finite number of values and the computation of the value function is analogous to the one in Section 6.2.1. In particular, the household is assumed to be boundedly rational and use only the first I moments m in order to predict the law of motion for the distribution $f(\cdot)$ with $m_1 = K$:

$$V(\epsilon, a, Z, m) = \max_{c,a'} \left[u(c) + \beta E \left\{V(\epsilon', a', Z', m')\right\}\right].$$

3) In the economy with aggregate uncertainty, the distribution of capital is not stationary. The household's income and savings depend on the aggregate productivity level and, for this reason, the distribution of capital changes over time. Similarly, the law of motion of the aggregate capital stock depends on the productivity level Z and (6.7) needs to be modified:[19]

[18] To be more precise, if the household knows the distribution $f(\epsilon, a)$, the argument N is redundant (even if Z' is not a sufficient statistic for N'), as he can compute aggregate employment N' with the help of $f'(\epsilon, a)$ and the aggregate consistency condition for N'. In the numerical computation, however, we assume that the household is boundedly rational and only uses the first I moments of the wealth distribution $f(\cdot)$ as information. In this case, he is unable to compute N from the aggregate consistency conditions and we have to introduce N as an additional state variable into the value and policy functions.

[19] In an economy, where (6.21) does not hold, employment N is an additional state variable and enters the function $H_I(m, Z, N)$ in (6.22) as an additional variable.

$$m = H_I(m, Z). \tag{6.22}$$

In our economy with $Z \in \{Z_g, Z_b\}$, we will again analyze the simple case where the agents only use the first moment $\bar{a} = K$ to predict the law of motion for the aggregate capital stock in good and bad times, respectively, according to:

$$\ln K' = \begin{cases} \gamma_{0g} + \gamma_{1g} \ln K & \text{if } Z = Z_g, \\ \gamma_{0b} + \gamma_{1b} \ln K & \text{if } Z = Z_b. \end{cases} \tag{6.23}$$

As the aggregate productivity is a stochastic variable, we can only simulate the dynamics of the economy. We follow Krusell and Smith (1998) and use 5000 agents in order to approximate the population. We choose an initial distribution of assets a and employment status ϵ over the 5000 households in period $t = 1$. In particular, we assume that every household is endowed with the initial asset holdings a_1 equal to the average capital stock of the economy and that the number of unemployed is equal to $u_1 \in \{u_g, u_b\}$. In the first iteration, the average capital stock is computed from the stationary Euler equation $1/\beta - \delta = \alpha(N/K)^{1-\alpha}$ with $N = 0.95$. We simulate the dynamics of the economy over 3,000 periods and discard the first 500 periods. As a consequence, the initialization of the distribution of (a, ϵ) in period $t = 1$ does not have any effect on our results for the statistics of the distribution in period 501-3,000. As in the previous section, we use the dynamics of the capital stock $\{K_t\}_{t=501}^{t=3,000}$ in order to estimate the law of motion for K_t in good and bad times, $Z_t = Z_g$ and $Z_t = Z_b$, respectively. For this reason, we separate the observation points (K_t, K_{t+1}) into two samples with either $Z_t = Z_g$ or $Z_t = Z_b$ and estimate the parameters $\{\gamma_0, \gamma_1\}$ for each subsample separately.

In order to simulate the dynamics of the households' wealth distribution, we use the optimal policy functions of the households. The optimal next-period asset level a' is a function of the employment status ϵ, current period wealth a, the aggregate productivity level Z, and the aggregate capital stock K, $a' = a'(\epsilon, a, Z, K)$. We use value function iteration in order to compute the decision functions so that the individual asset level a and the aggregate capital stock K need not be a grid point a_i or K_j, respectively. Therefore, we have to use bilinear interpolation in order to compute

the optimal next-period asset level a' off grid points in our simulation. Assume that we would like to compute $a'(\epsilon, a, Z, K)$ and that $a_i < a < a_{i+1}$ and $K_j < K < K_{j+1}$, where a_i, a_{i+1} and K_j, K_{j+1} denote the neighboring grid points of a and K on the grids $\mathcal{A}$ and $\mathcal{K}$, respectively. In order to compute $a'(\epsilon, a, Z, K)$, we interpolate bilinearly. Define h_a and h_K as follows:

$$h_a \equiv \frac{a - a_i}{a_{i+1} - a_i}, \qquad h_K \equiv \frac{K - K_j}{K_{j+1} - K_j}.$$

The optimal next-period asset is given by:

$$\begin{aligned} a'(\epsilon, a, Z, K) &= (1 - h_a)(1 - h_K)a'(\epsilon, a_i, Z, K_j) \\ &\quad + h_a(1 - h_K)a'(\epsilon, a_{i+1}, Z, K_j) \\ &\quad + h_a h_K a'(\epsilon, a_{i+1}, Z, K_{j+1}) \\ &\quad + (1 - h_a)h_K a'(\epsilon, a_i, Z, K_{j+1}). \end{aligned}$$

Finally, we impose the law of large numbers on our simulation results. While we track the behavior of 5,000 agents, the fraction of unemployed agents need not be equal to u_g in good times and u_b in bad times. We use a random number generator in order to simulate the motion of the individuals' employment status according to their appropriate conditional probabilities. In each period t, we check if the fraction of unemployed is equal to either u_g or u_b. If not, we choose a random number of agents and change their employment status accordingly. For example, if the number of unemployed agents is above $u_g \times 5,000$ in period t with $Z_t = Z_g$, we choose an unemployed agent at random and switch his employment status to employed and continue this process until $u_t = u_g$. The complete algorithm can be described by the following steps:[20]

Algorithm 6.3.1 (Computation of Example 6.3.1)

Purpose: *Computation of the dynamics in the heterogenous-agent economy with aggregate uncertainty assuming bounded rationality of the consumers.*

[20] The algorithm follows Krusell and Smith (1998) with some modifications.

Steps:

Step 1: Compute aggregate next-period employment N as a function of current productivity Z: $N = N(Z)$.
Step 2: Choose the order I of moments m.
Step 3: Guess a parameterized functional form for H_I in (6.22) and choose initial parameters of H_I.
Step 4: Solve the consumer's optimization problem and compute $V(\epsilon, a, Z, m)$.
Step 5: Simulate the dynamics of the distribution function.
Step 6: Use the time path for the distribution to estimate the law of motion for the moments m.
Step 7: Iterate until the parameters of H_I converge.
Step 8: Test the goodness of fit for H_I using, for example, R^2. If the fit is satisfactory, stop, otherwise increase I or choose a different functional form for H_I.

Implementation. The algorithm is implemented in the program `ch6_unc.g`. The parameterization is chosen for a model period equal to one year. We set the technology level equal to $Z_g = 1.03$ in good times and $Z_b = 0.97$ in bad times. The average duration of a boom or a recession is 5 years. Booms and recessions are of equal length so that the transition matrix Γ_Z is equal to:

$$\Gamma_Z = \begin{pmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{pmatrix}. \tag{6.24}$$

The following conditional employment probabilities are taken from Castañeda et al. (1998b) who consider the annual employment mobility for the US economy:

$$\Gamma(\epsilon'|Z' = Z_g, Z = Z_g, \epsilon) = \begin{pmatrix} 0.9615 & 0.0385 \\ 0.9581 & 0.04919 \end{pmatrix},$$
$$\Gamma(\epsilon'|Z' = Z_b, Z = Z_b, \epsilon) = \begin{pmatrix} 0.9525 & 0.0475 \\ 0.0.3952 & 0.6048 \end{pmatrix}.$$

These employment probabilities imply ergodic distributions with unemployment rates $u_g = 3.86\%$ and $u_b = 10.73\%$, respectively. The conditional employment probabilities for the transition from

good times to bad times, $Z = Z_g$ and $Z' = Z_b$ are calibrated such that all unemployed agents stay unemployed and that the unemployment rate is u_b in the next period using (6.21). The transition matrix from $Z = Z_b$ to $Z' = Z_g$ is calibrated so that all employed agents remain employed and the unemployment rate is equal to u_g in the next period, again making use of (6.21). The asset grids over individual wealth a, $\mathcal{A} = \{a_1 = 0, \ldots, a_{na} = 12\}$ and aggregate capital K, $\mathcal{K} = \{K_1 = 2.5, \ldots, K_{nk} = 5.5\}$, are chosen to be equispaced with $na = 101$ and $nk = 10$ nodes, respectively. The upper and lower bounds of these two intervals are found to be non-binding. The remaining parameters are also taken from Castañeda et al. (1998b): $\alpha = 0.36$, $\beta = 0.96$, $\delta = 0.1$, and $\sigma = 1.5$.

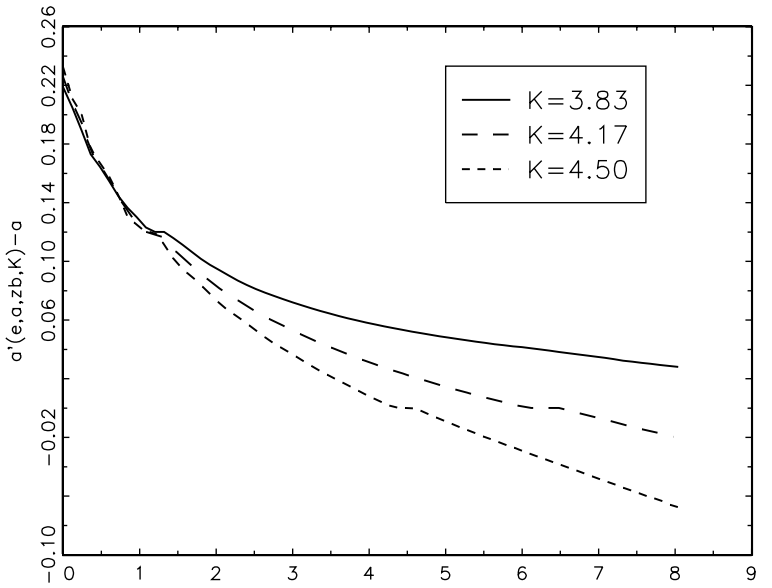

Figure 6.7: Net savings of the employed worker during bad times in Example 6.3.1

The optimal policy functions and the value functions behave as expected. Savings $a'(\epsilon, a)$ and consumption $c(\epsilon, a)$ increase with higher individual wealth a, while net savings $a' - a$ decline. In

addition, households save a higher proportion of their income for higher interest rates r or, equally, lower aggregate capital K. The net savings behavior of the employed and unemployed workers in bad times is displayed in Figures 6.7 and 6.8, respectively. The savings functions are graphed for different levels of aggregate wealth. Notice that the first derivative of the savings functions of the unemployed agents is not very smooth in the range $1.0 - 3.0$ of the individual wealth a. In order to get a more accurate solution, we need to increase the number of grid points over both assets, $\mathcal{A}$ and $\mathcal{K}$. Alternatively, we might want to try different methods in order to compute the policy functions, e.g. projection methods or PEA methods presented in the Chapters 3 and 4, respectively.

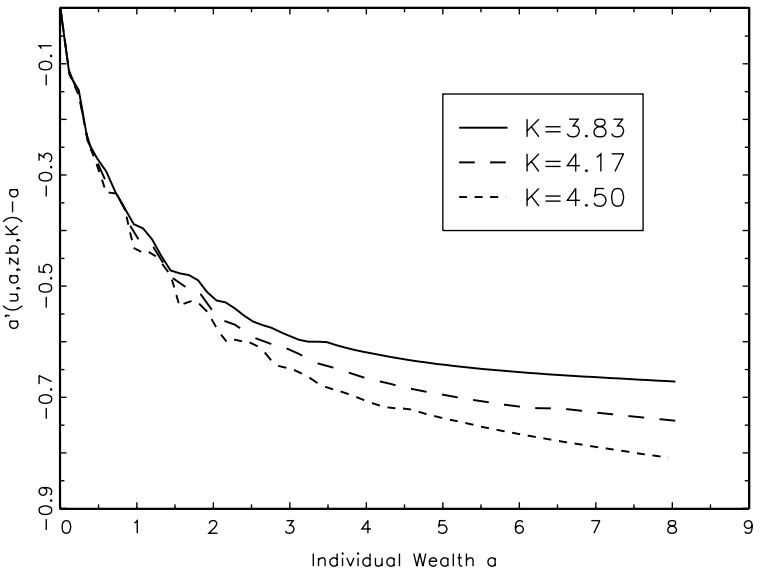

Figure 6.8: Net savings of the unemployed worker during bad times in Example 6.3.1

The mean capital stock in the economy with uncertainty is equal to $\bar{K} = 4.35$. The distribution of individual wealth in period $T = 3000$ is graphed in Figure 6.9. In our simulation, the aggregate capital stock in the last period $t = 3,000$ has been

equal to $K_{3000} = 4.18$ and the economy has been in a recession, $Z_b = 0.97$. Notice, in particular, that the upper grid point of $\mathcal{A}$, $a_{na} = 12$, is not binding and the maximum wealth of the households is approximately equal to $a = 7.5$.

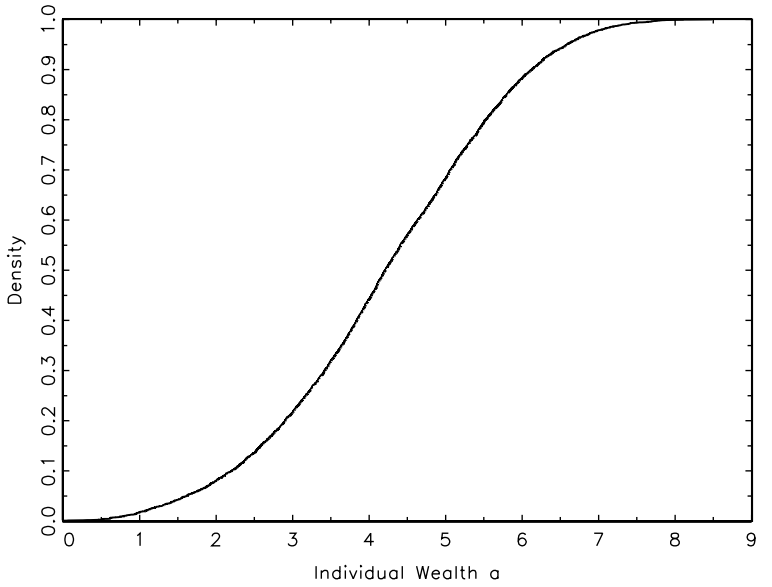

Figure 6.9: Distribution function in period $T = 3,000$

The law of motion for capital (6.23) is estimated at:

$$\ln K' = \begin{cases} 0.178 + 0.886 \ln K, & \text{if } Z = Z_g, \\ 0.135 + 0.900 \ln K, & \text{if } Z = Z_b. \end{cases} \tag{6.26}$$

Using (6.26), the mean prediction error of the capital stock amounts to 3.2%. The dynamics of the capital stock in our simulation are displayed in Figure 6.10. The standard deviation of capital[21] is equal to $\sigma_K = 0.195$. Our simple model falls short of replicating important business cycle characteristics. For example, the standard deviation of output ($\sigma_y = 0.458\%$) is much smaller

[21] The log of the time series of the aggregate capital stock K_t, output y_t, and aggregate consumption c_t have been HP-filtered with $\mu = 100$.

than the one of consumption ($\sigma_c = 2.97\%$). In the next section, you will get to know two more elaborate models of the business cycle dynamics.

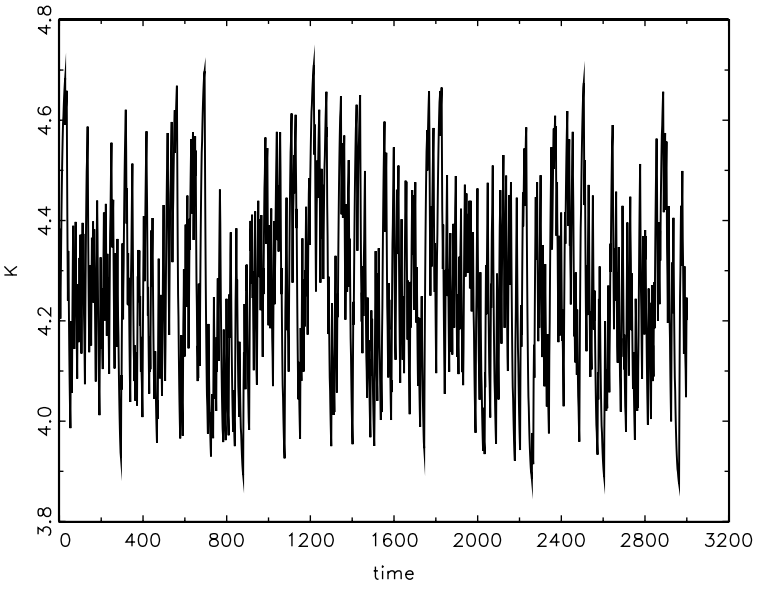

Figure 6.10: Time path of the aggregate capital stock K_t

6.4 Applications

In this section, we will look at two prominent applications of computational methods for heterogenous-agent economies with uncertainty that consider business cycles dynamics. One of the first papers in the area of computable general equilibrium models of heterogenous-agent economies is the article by Ayşe İmrohoroğlu published in the *Journal of Political Economy* in 1989.[22] Her model is similar to the economy described in Example 6.2.1, and we will recompute her benchmark equilibrium. She

[22] Her pioneering work, even though the model is only partial equilibrium, can be considered as the very first milestone in the literature on computable heterogeneous-agent economies that our second part of this book is

shows that the costs of business cycles depend on the assumption whether agents can borrow or not. Her equilibrium is partial in the sense that the factor prices are exogenous. As a consequence, agents do not need to project the cyclical behavior of the interest rate and the labor income. They only need to consider the probability of being employed in the next period. Therefore, the computation is straightforward applying the methods presented in Chapter 5.[23] In the second application, we consider the business cycle dynamics of the income distribution. Our model follows closely Castañeda et al. (1998b) and we need to apply the methods developed in the previous section.

6.4.1 Costs of Business Cycles with Indivisibilities and Liquidity Constraints

The model. There are many infinitely lived households of mass one who differ with regard to the assets a_t and their employment status ϵ_t. Households maximize their intertemporal utility

$$E_0 \sum_{t=0}^{\infty} \beta^t u\left(c_t\right), \tag{6.27}$$

where $\beta < 1$ is the subjective discount factor and expectations are conditioned on the information set at time 0. At time zero, the agent knows his beginning-of-period wealth a_0 and his employment status $\epsilon_0 \in \{e, u\}$. The agent's instantaneous utility function is a CES function of his consumption:

$$u(c_t) = \frac{c_t^{1-\eta}}{1-\eta}, \quad \eta > 0, \tag{6.28}$$

concerned with. Moreover, Ayşe İmrohoroğlu was already publishing other important contributions in the field of computable heterogeneous-agent economies at this very early time, where the computer technology started to allow for such computations. Among others, she also made an important contribution to the study of the welfare costs of inflation that was published in the *Journal of Economic Dynamics and Control* in 1992.

[23] We, however, included this model in the present section because it also studies the effects of business cycle fluctuations.

where η, again, denotes the coefficient of relative risk aversion.

If $\epsilon = e$ ($\epsilon = u$), the agent is employed (unemployed). If the agent is employed he produces $y(e) = 1$ units of income. If he is unemployed, he engages in home production and produces $y(u) = \theta$ units of consumption goods, where $0 < \theta < 1$. Furthermore, the agents cannot insure against unemployment.

İmrohoroğlu (1989) considers two different economies: In the first economy, agents cannot borrow, $a \geq 0$. They can insure against fluctuations in their income by storing the asset. The budget constraint is given by:

$$a_{t+1} = a_t - c_t + y(\epsilon_t). \tag{6.29}$$

In the second economy, the agents can borrow at rate r_b. Agents can save assets by either lending at rate $r_l = 0$ or storing them. There is an intermediation sector between borrowing and lending households. The borrowing rate r_b exceeds the lending rate $r_b > r_l$. The intermediation costs that are equal to the difference of the borrowing rate and the lending rate times the borrowed assets are private costs and reduce total consumption.

In the case without business cycle fluctuations, the individual-specific employment state is assumed to follow a first-order Markov chain. The conditional transition matrix is given by:

$$\pi(\epsilon'|\epsilon) = Prob\left\{\epsilon_{t+1} = \epsilon'|\epsilon_t = \epsilon\right\} = \begin{pmatrix} p_{uu} & p_{ue} \\ p_{eu} & p_{ee} \end{pmatrix}, \tag{6.30}$$

where, for example, $Prob\left\{\epsilon_{t+1} = e|\epsilon_t = u\right\} = p_{ue}$ is the probability that an agent will be employed in period $t+1$ given that the agent is unemployed in period t.

In the case with business cycle fluctuations, the economy experiences good and bad times. In good times, employment is higher and both employed and unemployed agents have a higher probability to find a job. We can distinguish four states $s \in \{s_1, s_2, s_3, s_4\}$:

$s = s_1$: the agent is employed in good times,
$s = s_2$: the agent is unemployed in good times,
$s = s_3$: the agent is employed in bad times,

$s = s_4$: the agent is unemployed in bad times.

The transition between the four states is described by a first-order Markov chain with conditional transition matrix $\pi(s'|s)$. The economies with and without business cycles have the same average unemployment rate.

Calibration and Computation. The model is calibrated for a model period of 6 weeks or approximately 1/8 of a year. The discount factor $\beta = 0.995$ implies an annual subjective time discount rate of approximately 4%. The relative coefficient of risk aversion η is set equal to 1.5. The annual borrowing rate is set equal to 8% corresponding to a rate of $r_b = 1\%$ in the model period.

The conditional transition matrices $\pi(\epsilon'|\epsilon)$ and $\pi(s'|s)$ are calibrated so that average unemployment is 8%, unemployment in good times and bad times is 4% and 12%, respectively. In the economy with business cycles, the average duration of unemployment is 1.66 and 2.33 periods (10 and 14 weeks) in good and bad times, respectively. Furthermore, the probability that good or bad times continue for another period is set equal to 0.9375 so that the average duration of good and bad times is equal to 24 months implying an average duration of the business cycle equal to 4 years. The transition matrices are then given by:

$$\pi(\epsilon'|\epsilon) = \begin{pmatrix} 0.5000 & 0.5000 \\ 0.9565 & 0.0435 \end{pmatrix}, \tag{6.31}$$

and

$$\pi(s'|s) = \begin{pmatrix} 0.9141 & 0.0234 & 0.0587 & 0.0038 \\ 0.5625 & 0.3750 & 0.0269 & 0.0356 \\ 0.0608 & 0.0016 & 0.8813 & 0.0563 \\ 0.0375 & 0.0250 & 0.4031 & 0.5344 \end{pmatrix}. \tag{6.32}$$

The Markov process described by matrix (6.32) implies average unemployment rates of 4.28% and 11.78% during good and bad times, respectively. Finally, the households' home production is equal to $\theta = 0.25$.

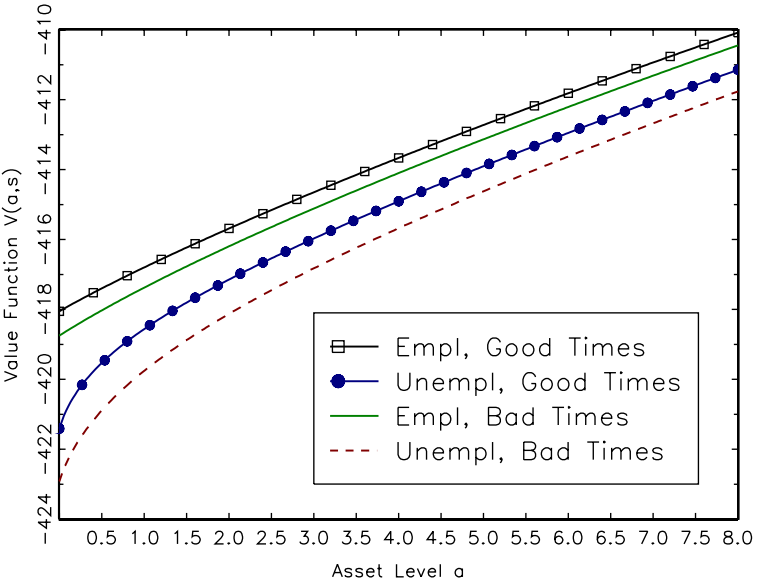

Figure 6.11: Value function $V(a, s)$ in the storage economy

In the following, we describe the computation of the economy with business cycle fluctuations. The computation of the model is simpler than the one for the economy considered in section 6.1. In Example 6.2.1 with the endogenous interest rate r, we had to pick an initial value of the interest rate, compute the decision functions and the invariant distribution and update the interest rate subsequently until it converged. In the present economy, the interest rate is given. We first compute the decision functions by value function iteration. The value function of the individual is a function of his assets a and the state s:

$$\begin{aligned} V(a,s) &= \max_{c,a'} \left[u(c) + \beta E\left\{V(a',s')\right\}\right] \qquad (6.33) \\ &= \max_{c,a'} \left[u(c) + \beta \sum_{s'} \pi(s'|s) V(a',s')\right]. \end{aligned}$$

From Chapter 1, we know how to solve this simple dynamic programming problem. In the program `ch63_imo.g`, we use value function iteration with linear interpolation between grid points.

The maximum of the rhs of the Bellman equation (6.33) is computed with Golden Section Search. We use $n_a = 301$ grid points for the asset space so that we have to store a matrix with $n_a \times 4 = 1204$ entries. The grid is chosen to be equispaced on the interval $[0, 8]$ and $[-8, 8]$ for the economy with only a storage technology and the economy with intermediation, respectively. The value function for the economy with a storage technology only is displayed in Figure 6.11.

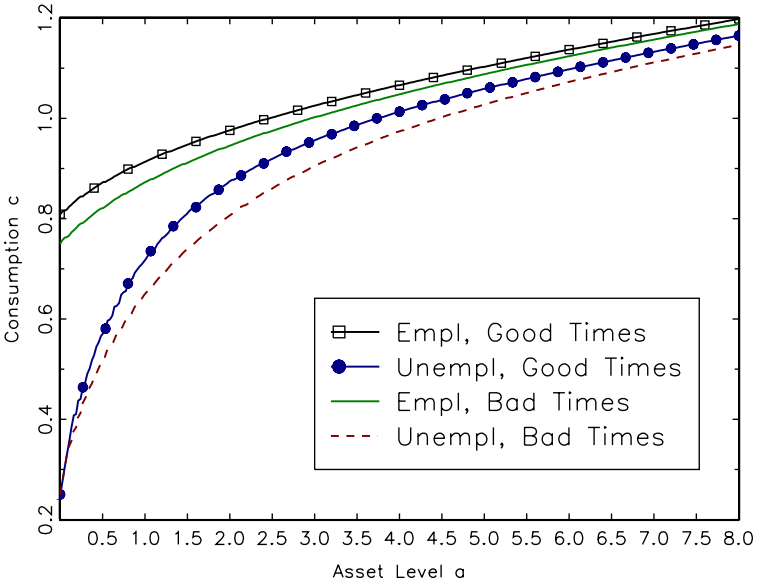

Figure 6.12: Consumption $c(a, s)$ in storage economy

Consumption is an increasing function of income and is also higher in good times as agents have a higher expected next-period income (compare Figure 6.12). The optimal next-period asset $a'(a, s)$ is a monotone increasing function of assets a. Figure 6.13 displays the net savings $a' - a$ which are always negative for the unemployed agent and become negative for the employed agents at a wealth level approximately equal to 4 so that the ergodic set is contained in the interval $[0, 8]$.

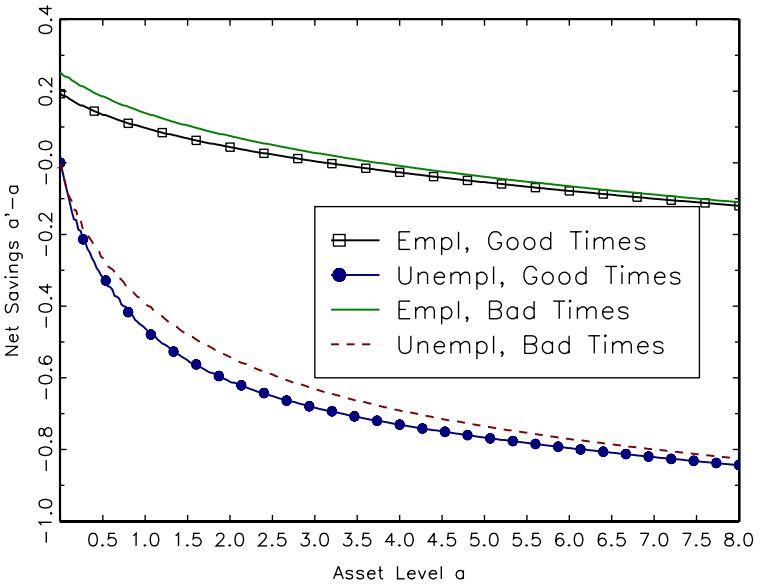

Figure 6.13: Net savings $a' - a$ in the storage economy

Next, we compute the invariant density function $f(a, s)$. The invariant distribution can be interpreted as the fraction of time a particular individual spends in the different states (a, s). For an economy with business cycle fluctuations, the invariant distribution is the limit of the predictive probability distribution of an individual in n periods where n goes to infinity. We compute the invariant density function as described in Chapter 5. We use a finer grid over the asset space for the computation of the distribution than for the computation of the policy function. In particular, we compute the density function at $n_{ag} = 903$ equispaced points over the interval $[0, 8]$ and $[-8, 8]$, respectively. The invariant density function $f(a, s)$ can be computed from the following dynamics:

$$f'(a', s') = \sum_{a'=a'(a,s)} \sum_{s'} \pi(s'|s) f(a, s). \tag{6.34}$$

As the optimal next-period asset level a' may not be a grid point, we simply assume that it will be on the lower or higher

neighboring point with a probability that corresponds to the distance from the higher or lower neighboring point, respectively (see Chapter 5). The invariant density function $f(a, s)$ is displayed in Figure 6.14. The ergodic set is approximately $[0, 3.8]$ and the density is zero for $a > 3.8$.[24]

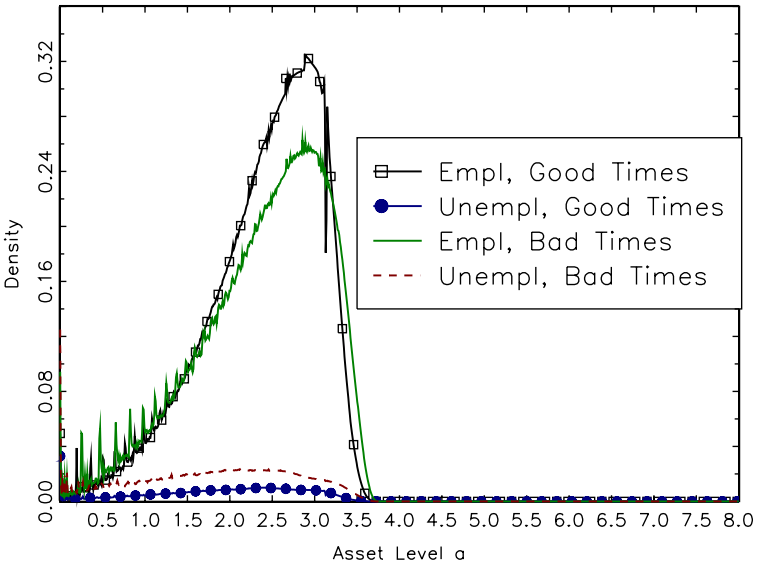

Figure 6.14: Invariant density function $g(a, s)$ in the storage economy

On average, assets at the amount of $\bar{a} = 2.35$ are stored in this economy. As the average employment rate is 8%, average income (which is equal to average consumption) is equal to $\bar{y} = 0.92 + 0.08 \cdot 0.25 = 0.94$.

[24] Different from our density function, the density function computed by İmrohoroğlu (1989) (Fig. 1 in her article) displays two spikes in the range 1.5-2.0 of individual wealth and maximum values of approximately 0.05 in good times. This, however, is an artefact of her computational methods. She only computes the policy functions at grid points and does not interpolate between grid points. As a consequence, our results differ to a slight degree and our policy functions and distribution functions are much smoother.

The consumption and savings behavior in the economy with intermediation is different from the one in the economy with a storage technology only. In particular, the consumption behavior changes around $a = 0$ as the interest rate on assets changes from the low lending rate $r^l = 0\%$ to the high borrowing rate $r^b = 8\%$ (see Figures 6.15 and 6.16). It is for this reason that we have used value function iteration. If we had used a computational method like projection methods that does not rely on the discretization of the individual asset grid, we may have had problems in capturing this non-monotonocity of the first derivative of $a'(a, \epsilon)$ at $a = 0$.

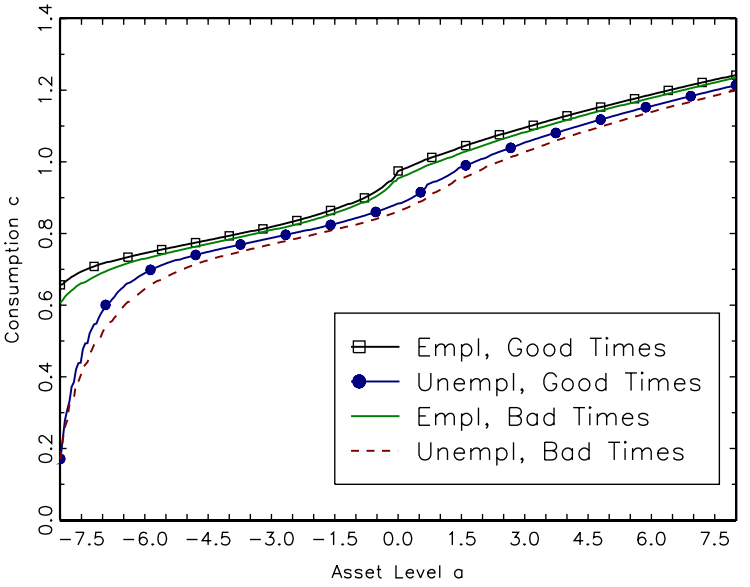

Figure 6.15: Consumption $c(a, s)$ in an economy with intermediation

In our model, we only study a partial equilibrium. The average of assets borrowed amounts to 0.510 and is not equal to the amount of assets saved (=0.287). In a general equilibrium, the interest rate r^b and r^l would adjust in order to clear the capital market. The average income $\bar{y}$ is equal to the one in the economy with a storage technology only and amounts to 0.940. As, however, intermediation costs are private costs, average consumption

is smaller than average income, $\bar{c} = 0.935 < 0.940 = \bar{y}$. The difference between average income and average consumption is simply the borrowing rate times the assets borrowed.

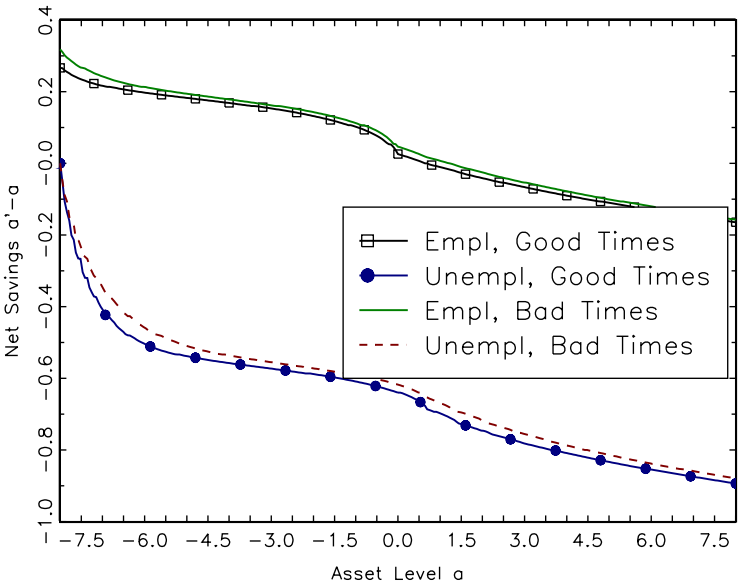

Figure 6.16: Net savings $a' - a$ in an economy with intermediation

As one central problem of her work, Ayşe İmrohoroğlu (1989) computes the welfare gain from eliminating business cycle fluctuations. For this reason, she computes average utility in the economy with and without business cycle fluctuations, either using (6.31) or (6.32) for the state transition matrix of the economy. For the benchmark calibration, the elimination of business cycles is equivalent to a utility gain corresponding to 0.3% of consumption in the economy with a storage technology. If the relative risk aversion η is increased to 6.2, the welfare gain rises to 1.5% of consumption.[25] An intermediation technology significantly re-

[25] Notice that this is a huge welfare effect. Lucas (1987) estimates the costs of business cycle model to be very small and only equivalent to 0.1% of total US consumption. Different from the present model, agents can insure against the idiosyncratic risk in his model.

duces the business cycle costs. For $\eta = 1.5$, the fluctuations only cause a utility loss equivalent to 0.05% of consumption.

The computation of the welfare effects from business cycle fluctuations in İmrohoroğlu (1989) is only sensible if the average asset holdings for the economies with and without business cycles do not change significantly. This is the case in İmrohoroğlu (1989). Heer (2001) considers an economy with endogenous prices where asset holdings are different in the economies with and without business cycles. Agents may hold much higher assets for precautionary reasons in a fluctuating economy. As a consequence, average asset holdings may change and, in a general equilibrium, average consumption may also change significantly. In his model, welfare changes that result from business cycle fluctuations are even more pronounced than in the present model.

6.4.2 Business Cycle Dynamics of the Income Distribution

Ana Castañeda, Javier Díaz-Giménez and José-Victor Ríos-Rull (1998b) explore the business cycle dynamics of the income distribution both empirically and in a theoretical computable general equilibrium model. They find that, in the US, the income share earned by the lowest quintile is more procyclical and more volatile than the other income shares. In particular, the income shares earned by the 60%-95% group are even countercyclical, while the share earned by the top 5% is still acyclical.

To address these issues, they construct a heterogenous-agent economy with aggregate uncertainty based on the stochastic neoclassical growth model. The aggregate uncertainty is modelled as in Example 6.3.1 and follows a Markov process. Again, unemployment is higher during bad times and contributes to the explanation of the business cycle dynamics of the lowest income share. Contrary to the model by İmrohoroğlu (1989) presented in the previous section, the aggregate capital stock and the interest rate are endogenous variables of the model. As one of their major results, cyclical unemployment helps to reconcile the model's behavior with the data on the dynamics of the income distribution.

In the following, we will present a slightly modified version of the model by Castañeda et al. (1998b). In particular, we will introduce income mobility in addition to Castañeda et al. and consider its effect on the cyclical behavior of the income shares of each income quintile. First, we will describe the model. Second, we present the calibration and the computational method. We conclude with the presentations of our results.

The Model. There are many infinitely lived households of mass one who differ with regard to the assets a_t, their employment status ϵ_t, and their efficiency type $i \in \{1, \ldots, 5\}$. The mass of type i household is equal to $\mu_i = 20\%$ for $i = 1, \ldots, 5$.

Households maximize their intertemporal utility[26]

$$E_0 \sum_{t=0}^{\infty} \beta^t u\left(c_t\right), \tag{6.35}$$

where $\beta < 1$ is the subjective discount factor and expectations are conditioned on the information set at time 0. At time zero, the agent knows his beginning-of-period wealth a_0, his employment status $\epsilon_0 \in \{e, u\}$ and his efficiency type i. The agent's instantaneous utility function is a CES function of his consumption:

$$u(c_t) = \frac{c_t^{1-\eta}}{1-\eta}, \quad \eta > 0, \tag{6.36}$$

where η, again, denotes the coefficient of relative risk aversion.

The model is characterized by both idiosyncratic and aggregate risk. At the individual level, the household may either be employed ($\epsilon = e$) or unemployed ($\epsilon = u$). Aggregate risk is introduced by a stochastic technology level Z_t in period t. In particular, the productivity shock follows a Markov process with transition matrix $\Gamma_Z(Z'|Z)$, where Z' denotes next-period technology level and $\pi_{ZZ'}$ denotes the transition probability from state Z to Z'. The individual employment probabilities, of course, depend on the

[26] For the sake of notational simplicity, we refrain from indexing individual variables like consumption, wealth, the employment status, and the efficiency index with a subscript $j \in [0, 1]$.

aggregate productivity Z_t. In good times (high productivity Z_t), agents have higher employment probabilities than in bad times. The joint process of the two shocks, Z_t and ϵ_t, can be written as a Markov process with transition matrix $\Gamma_i(Z', \epsilon'|Z, \epsilon)$ and depends on the efficiency type of the agent. We use $\pi_i(Z', \epsilon'|Z, \epsilon)$ to denote the probability of transition from state (Z, ϵ) to state (Z', ϵ') for an individual with efficiency type i. In the following, we restrict our attention to the simple case presented in Example 6.3.1. The economy only experiences good and bad times with technology levels Z_g and Z_b, respectively, $Z_g > Z_b$. Consequently, the joint processes on (Z, ϵ) are Markov chains with 4 states for each efficiency type $i = 1, \ldots, 5$.

We study a simple extension of the model by Castañeda et al. (1998b). In particular, we assume that agents may change their efficiency type i. Castañeda et al. assume that agents keep their efficiency type i forever. Given the empirical evidence on income mobility, however, we examine the effect of this assumption on the cyclical behavior of the income distribution.[27] We assume that the efficiency type i follows a Markov chain $\pi(i'|i)$ that is independent of the aggregate productivity. Furthermore, the employment probability next period does only depend on the efficiency type of this period. In other words, $\pi_i(Z', \epsilon'|Z, \epsilon)$ is not a function of next-period type i'. The probability of an employed type i agent in the good state $Z = Z_g$ to be employed next period as a type i' agent in the bad state, for example, is given by the product $\pi(i'|i)\, \pi_i(e, Z_b|e, Z_g)$.

Households are assumed to know the law of motion of both $\pi_i(Z', \epsilon'|Z, \epsilon)$ and $\pi(i'|i)$ and they observe the realization of both stochastic processes at the beginning of each period. In good times, agents work $h(Z_g)$ hours, and, in bad times, agents work $h(Z_b)$ hours. Let ζ_i denote the efficiency factor of a type i agent. If employed, the agent receives the labor income $h(Z)\zeta_i w$; otherwise, he produces home production $\bar{w}$.

[27] In addition, the assumption of income mobility may change the wealth inequality in the model as we will argue below. Still, wealth heterogeneity is too small in both our model and the model by Castañeda et al.

Let $N_i(Z)$ denote the number of employed households of type i for current productivity Z. We will calibrate these values below so that $N_i(Z)$ is constant for $Z \in \{Z_g, Z_b\}$ and does not depend on the history of the productivity level Z, $\{Z_\tau\}_{\tau=-\infty}^{\tau=t}$. The assumption that employment only depends on current productivity greatly simplifies the computation. Agents do not have to form expectations about the dynamics of the aggregate employment level but only need to consider aggregate productivity and the distribution of wealth. The aggregate labor input measured in efficiency units is given by $N(Z) = \sum_i \zeta_i h(Z) N_i(Z)$. For the technology level Z_t, aggregate production is given by:

$$F(K_t, N_t) = Z_t K_t^\alpha N_t^{1-\alpha}. \tag{6.37}$$

We assume competitive factor and product markets implying the factor prices:

$$w_t = Z_t(1-\alpha)K_t^\alpha N_t^{-\alpha}, \tag{6.38}$$

$$r_t = Z_t \alpha K_t^{\alpha-1} N_t^{1-\alpha} - \delta, \tag{6.39}$$

where δ denotes the rate of depreciation. Notice that the agents only need to forecast the aggregate capital stock K' (and, therefore, the dynamics of the distribution of capital) and the aggregate technology level Z' in order to form a prediction of the next-period factor prices w' and r', respectively, as we assume N' to be a function of Z' only.

In the following, we describe the household decision problem in a recursive form. Let f denote the distribution of the individual state variables $\{i, \epsilon, a\}$. For each household, the state variable consists of his efficiency type i, his employment status ϵ, his individual asset level a, the aggregate technology level Z, the aggregate capital stock K (which is implied by the distribution $f(\cdot)$), and the distribution of efficiency types, employment, and individual wealth, $f(i, \epsilon, a)$.

The recursive problem can be formulated as follows:

$$V(i, \epsilon, a; Z, f) = \max_{c, a'} \left[u(c) + \beta E\left\{V(i', \epsilon', a'; Z', f')\right\}\right] \tag{6.40}$$

subject to the budget constraint:

$$a' = \begin{cases} (1+r)\,a + w\zeta_i h(Z) - c, & \text{if } \epsilon = e, \\ (1+r)\,a + \bar{w} - c, & \text{if } \epsilon = u, \end{cases} \tag{6.41}$$

and subject to (6.38) and (6.39), the stochastic process of the employment status ϵ and the aggregate technology Z, $\pi_i(Z', \epsilon'|Z, \epsilon)$, the agent's efficiency mobility as given by $\pi(i'|i)$, and the distribution dynamics $f' = G(f, Z, Z')$, where G describes the law of motion for the distribution f.

The definition of the equilibrium is analogous to the one in Example 6.3.1 and we will omit it for this reason. The interested reader is referred to Section 3.5.2 of Castañeda et al. (1998b).

Calibration. The parameters, if not mentioned otherwise, are taken from Castañeda et al. (1998b).[28] The model's periods correspond to 1/8 of a year (=6.5 weeks). The coefficient of relative risk aversion is set equal to $\eta = 1.5$ and the discount factor is set equal to $0.96^{1/8}$ implying an annual discount rate of 4%.

Castañeda et al. assume that agents are immobile implying $\pi(i'|i) = 1$ if $i' = i$ and zero otherwise. Budría et al. (2001) provide an estimate of the US earnings transition matrix between the different earnings quintile from 1984 to 1989:[29]

$$P = \begin{pmatrix} 0.58 & 0.28 & 0.09 & 0.03 & 0.02 \\ 0.22 & 0.44 & 0.22 & 0.08 & 0.03 \\ 0.10 & 0.15 & 0.43 & 0.23 & 0.09 \\ 0.06 & 0.09 & 0.18 & 0.46 & 0.21 \\ 0.06 & 0.02 & 0.06 & 0.21 & 0.65 \end{pmatrix}. \tag{6.42}$$

We still have to transform this 5-year transition matrix into a 1/8-year transition matrix. For this reason, we diagonalize the transition matrix P, $P = U \cdot D \cdot U^{-1}$, where D is a diagonal matrix, U is the matrix of the eigenvectors, and U^{-1} is the inverse of U. For expositional reasons, we only illustrate how we find the square root of P such that $P^{1/2} \cdot P^{1/2} = P$. Therefore, we simply need to compute the square root of the diagonal elements which we denote by $D^{1/2}$. Notice that due to its diagonality $D^{1/2} \cdot D^{1/2} = D$.

[28] We would like to thank Victor Ríos-Rull for providing us with the calibration data on the transition matrices.

[29] See table 24 in their appendix.

Furthermore, $P^{1/2} = U \cdot D^{1/2} \cdot U^{-1}$ is the solution as $P^{1/2} \cdot P^{1/2} = U \cdot D^{1/2} \cdot U^{-1} \cdot U \cdot D^{1/2} \cdot U^{-1} = U \cdot D \cdot U^{-1} = P$. Similarly, we can compute $P^{1/40}$ which we set equal to $\pi(i'|i)$:[30]

$$\pi(i'|i) = \begin{pmatrix} 0.983 & 0.015 & 0.001 & 0.000 & 0.000 \\ 0.011 & 0.974 & 0.013 & 0.001 & 0.000 \\ 0.003 & 0.007 & 0.974 & 0.013 & 0.002 \\ 0.001 & 0.004 & 0.010 & 0.976 & 0.010 \\ 0.002 & 0.000 & 0.001 & 0.010 & 0.987 \end{pmatrix}. \tag{6.43}$$

The 6 weekly earnings mobility is rather small as the entries on the diagonal of (6.43) are close to unity.[31] Therefore, we would expect little influence from the neglect of the income mobility on the results by Castañeda et al.(1998b).

Castañeda et al. use five types of households with efficiency $\zeta^i \in \{0.509, 0.787, 1.000, 1.290, 2.081\}$. The efficiency factors are chosen to be the relative earnings of the different income groups. The variation of hours worked of these 5 income groups between good and bad times are treated as if they were variations in employment rates. With the help of the coefficient of variation of average hours, the employment rates are calibrated as follows:[32]

[30] The computation is performed by the procedure `matroot` in the program `ch63cas1.g`. Furthermore, we set all negative entries of the matrix root equal to zero and normalize the sum of each row equal to one in the routine `matroot`. The error is rather small for our case (you may check this by computing $\pi(i'|i) \cdot \pi(i'|i) \cdot \ldots$ and compare it to the matrix P from (6.42)).

[31] In order to obtain (6.43) from (6.42), we have assumed that earnings follow an AR(1) process. As we argued in Chapter 5, the behavior of earnings is much better described by an AR(2) process. Due to the lack of high-frequency data on the earnings mobility, however, we do not have another choice.

[32] Please see Castañeda et al. (1998b) for a detailed description of the calibration procedure.

i	$N(Z_g)$	$N(Z_b)$
1	0.8612	0.8232
2	0.9246	0.8854
3	0.9376	0.9024
4	0.9399	0.9081
5	0.9375	0.9125

Given the employment of type i household in good and bad times, $N_i(Z_g)$ and $N_i(Z_b)$, respectively, and the average duration of unemployment in good times (10 weeks) and in bad times (14 weeks), we are able to compute the matrices $\pi_i(\epsilon', Z'|\epsilon, Z)$ with $Z = Z'$. For this reason, we have to solve a system of 4 equations (including non-linear equations) which is carried out in the routine `transp` in the program `ch63cas1.g`. Two equations are given by the conditions that agents are either employed or unemployed in the next period. Taking $\pi_i(\epsilon', Z_g|\epsilon, Z_g)$ as an example, we impose the following two conditions on the transition matrix:

$$\pi_i(e, Z_g|e, Z_g) + \pi_i(u, Z_g|e, Z_g) = \pi(Z_g|Z_g),$$
$$\pi_i(e, Z_g|u, Z_g) + \pi_i(u, Z_g|u, Z_g) = \pi(Z_g|Z_g).$$

Furthermore, the average duration of unemployment is 10 weeks or 10/6 periods in good times implying $\pi_i(u, Z_g|u, Z_g) = 4/10 \cdot \pi(Z_g|Z_g)$. The fourth condition is given by the equilibrium employment in good times, $N_i(Z_g)$. We impose as our fourth non-linear equation that the ergodic distribution of the employed agents is equal to $N_i(Z_g)$. The matrix $\pi_1(\epsilon', Z_g|\epsilon, Z_g)$, for example, is given by:

$$\pi_1(\epsilon', Z_g|\epsilon, Z_g) = \begin{pmatrix} 0.9033 & 0.09067 \\ 0.60000 & 0.40000 \end{pmatrix} \pi(Z_g|Z_g).$$

Similarly, we are able to compute $\pi_i(\epsilon', Z_b|\epsilon, Z_b)$ for all $i = 1, \ldots, 5$.

It remains to compute the transition matrix $\pi(\epsilon', Z_b|\epsilon, Z_g)$ between good and bad times on the one hand and the transition matrix $\pi(\epsilon', Z_g|\epsilon, Z_b)$ between bad and good times on the other hand.

First, we assume that all unemployed agents remain unemployed if the economy transits from good to bad times, $\pi_i(u, Z_b|u, Z_g) = \pi(Z_b|Z_g)$ and $\pi_i(e, Z_b|u, Z_g) = 0$ for all $i = 1, \ldots, 5$. Second, we assume that employments $N_i(Z_g)$ and $N_i(Z_b)$ are constant, respectively. For this reason, $N_i(Z_g)\pi_i(e, Z_b|e, Z_g) = N_i(Z_b)\pi(Z_b|Z_g)$ must hold. Together with the condition that $\pi_i(e, Z_b|e, Z_g) + \pi_i(u, Z_b|e, Z_g) = \pi(Z_b|Z_g)$, we have four conditions that help us to determine the matrix $\pi_i(\epsilon', Z_b|\epsilon, Z_g)$. For the computation of the matrix $\pi_i(\epsilon', Z_b|\epsilon, Z_g)$, we assume that all employed agents remain employed if the economy transits from the bad to the good state. Furthermore, we assume $N_i(Z_g)$ to be constant so that we also impose the restriction that $(1 - N_i(Z_g))\pi(Z_g|Z_b) = \pi_i(u, Z_g|u, Z_b)(1 - N_i(Z_b))$. Together with the two conditions that the sum of each row must be unity we can determine the matrix $\pi_i(\epsilon', Z_g|\epsilon, Z_b)$ for all $i = 1, \ldots, 5$.

The transition matrix between good and bad states is set equal to:

$$\pi(Z'|Z) = \begin{pmatrix} 0.9722 & 0.0278 \\ 0.0278 & 0.9722 \end{pmatrix},$$

implying equal length of booms and recession averaging 4.5 years. Furthermore, employment is constant both in good times and bad times, respectively, and the factor $Zh(Z)^{1-\alpha}$ is set equal to 1 and 0.9130 for $Z = Z_g$ and $Z = Z_b$, respectively. We assume that average working hours amount to $h(Z_g) = 32\%$ and $h(Z_b) = 30\%$ of the available time during good and bad times, respectively.[33]

Finally, the household production $\bar{w}$ is set equal to 25% of average earnings in the economy. In particular, the earnings during unemployment $\bar{w}$ are constant over the business cycle.

Computation. The solution is computed with the help of program `ch63cas1.g` using the methods described in Section 6.2. In particular, we apply Algorithm 6.3.1 with the following steps 1-8:

Step 1) We choose computational parameters and compute the aggregate employment levels in good and bad times, $N(Z_g) =$

[33] Together with the income mobility transition matrix, these are the only parameters that differ from the calibration of Castañeda et al. (1998b).

$\sum_i \mu_i \zeta_i h(Z_g) N_i(Z_g)$ and $N(Z_b) = \sum_i \mu_i \zeta_i h(Z_b) N_i(Z_b)$. The agents form very simple expectations about the next-period employment. Employment next period only depends on productivity in the next period: $N' = N'(Z')$. The policy functions are computed on the interval $\mathcal{A} \times \mathcal{K} = [a_{min}, a_{max}] \times [K_{min}, K_{max}] = [0, 800] \times [80, 400]$. The interval limits are found with some trial and error and do not bind. The policy and value functions are computed on an equispaced grid of the state space using $na = 50$ and $nk = 5$ grid points on the intervals $\mathcal{A}$ and $\mathcal{K}$, respectively.

As an initial guess for the interest rate, we use the steady state capital stock for the corresponding representative agent model as implied by $1/\beta = 1 + r - \delta$. For the computation of the distribution function $f(\cdot)$, we, again, need to discretize the continuous variables of the individual state space. We use $na = 100$ equispaced points over the individual asset space $\mathcal{A}$. Furthermore, we have $ni = 5$ types of agents, $nz = 2$ states of the productivity, and $ne = 2$ states of employment.

Step 2) Agents need to predict next-period factor prices w' and r', which are functions of both aggregate capital K' and aggregate employment N' as well as the exogenous technology level Z'. In order to predict the capital stock K', agents need to know the dynamics of the distribution. They only use partial information about the distribution, namely its first m moments. In accordance with Castañeda et al. (1998b), we choose $m = 1$. Agents only consider the aggregate capital stock as a statistic for the distribution. As argued above, this assumption is warranted if agents of different wealth have approximately equal savings rates. Therefore, the value function of the agents, $V(i, \epsilon, a, Z, K)$, and the consumption function, $c(i, \epsilon, a, Z, K)$, are functions of the individual efficiency type i, the employment status ϵ, the asset holdings a, the aggregate productivity Z, and the aggregate capital stock K.

The value function and the policy functions are both five-dimensional objects. This may impose some computational problems. For example, in older versions of GAUSS, only two-dimensional objects can be stored. There are two ways to solve this problem. First, in our model, there is only a small number of effi-

ciency types $i = 1, \ldots, 5$, two states of technology, $Z \in \{Z_g, Z_b\}$, and two employment status, $\epsilon \in \{e, u\}$. Consequently, we can store the two-dimensional value matrices $V(a, K; i, \epsilon, Z)$ for the $5 \times 2 \times 2 = 20$ different values of i, ϵ, and Z, separately. That's how we proceed. If the number of states is getting larger, of course, this procedure becomes cumbersome. In the latter case, you may want to store the value function in one matrix, reserving the first na rows for $i = 1$, $Z = Z_g$, $\epsilon = e$, the next na rows for $i = 2$, $Z = Z_g$, $\epsilon = e$, and so forth. In the second case, of course, it is very convenient for the computation to write a subroutine that returns you the value function $V(a, K; i, \epsilon, Z)$ for a state vector (i, ϵ, Z).

For the initialization of the consumption function for each (i, ϵ, Z), we assume that the agents consume their respective income. We further initialize the distribution of assets assuming that every agent holds equal wealth. The initial state of the economy is chosen by random choice. With probability 0.5, $Z = Z_g$. Otherwise, the bad state $Z = Z_b$ prevails. As we dispense of the first 100 simulated time periods, the initial choice of the distribution and the productivity does not matter.

Step 3) We impose again a very simple law of motion for the capital stock. As in (6.23), we assume that the aggregate capital stock follows a loglinear law of motion in good and bad times, respectively:

$$\ln K' = \begin{cases} \gamma_{0g} + \gamma_{1g} \ln K, & \text{if } Z = Z_g, \\ \gamma_{0b} + \gamma_{1b} \ln K, & \text{if } Z = Z_b. \end{cases} \tag{6.45}$$

We initialize the parameters as follows: $\gamma_{0g} = \gamma_{0b} = 0.09$ and $\gamma_{1g} = \gamma_{1b} = 0.95$.

Step 4) In this step, we compute the optimal next-period asset level $a'(i, \epsilon, a, Z, K)$ by value function iteration. Between grid points, we interpolate linearly. The maximization of the right-hand side of the Bellman equation is performed using the Golden Section Search Algorithm 8.6.1. We need to find the optimum for $50 \times 5 \times 5 \times 2 \times 2 = 5,000$ grid points. The computation is much faster i) if we compute and store the next-period value

$V(i', \epsilon', a', Z', K')$ for all nk values $K'(K)$ where K' is computed from the dynamics (6.45), before we start iterating over i, ϵ, a and Z. ii) We make use of both the monotonicity of next-period asset level $a'(a, .)$ and the value function $V(a, .)$ with respect to a and the concavity of the value function $V(a, .)$ with respect to a'. In particular, we stop searching over the next-period asset grid a' if the rhs of the Bellman equation decreases and we do not search for the optimal next-period asset level for values of $a'(a_i)$ below $a'(a_{i-1})$ for $a_i > a_{i-1}$.

Step 5) In order to simulate the dynamics of the wealth distribution, we choose a sample of $nh = 5,000$ households. We divide the households in 10 subsamples (i, ϵ), $i = 1, \ldots, 5$, $\epsilon \in \{e, u\}$. We know that the relative numbers of these subsamples are equal to $N_i(Z)$ and $1 - N_i(Z)$, respectively, for $Z = Z_g$ and $Z = Z_b$. We initialize the distribution so that each agent has equal wealth in period 1. In particular, the average wealth in period 1 is equal to the aggregate capital stock in the economy.[34] The assets of the next period are computed with the help of the optimal decision rule $a'(a, K; i, \epsilon, Z)$ for each household. The aggregate capital stock of the economy is equal to average wealth in the economy. We further use a random number generator in order to find i) the productivity level of next period Z' using the transition matrix $\pi(Z'|Z)$, ii) the employment status of the next period ϵ' using the transition matrix $\pi_i(\epsilon', Z'|\epsilon, Z)$ and iii) the efficiency type of the individual using the income mobility matrix $\pi(i'|i)$.

In the period t, we have a sample of 5,000 households with wealth holdings a_t and a distribution with mean K_t. The productivity level is equal to Z_t. The number of the employed households of type i, for example, may not be equal to $N_i(Z_t)$. For this reason, we choose a random number of agents accordingly. We also may have to switch the productivity type i. For this reason, we start looking at the households with efficiency $i = 1$ and $\epsilon = e$. If their number is smaller than $N_1(Z_t)$ we switch the missing number of

[34] In the very first simulation, we use the aggregate capital stock as an initial guess that is computed from the steady-state of the corresponding representative-agent model.

the households with $i = 2$ and $\epsilon = e$ to $i = 1$ and $\epsilon = e$ at random. Otherwise, we switch the surplus number of households with type $i = 1$ to type $i = 2$. We continue this process for $i = 1, \ldots, 5$, $\epsilon = e, u$. By this procedure, agents of type $i = 1$ may not be switched to agents of type $i = 4$, for example. We judge this to be a reasonable imposition of the law of large numbers.

Step 6) We divide the simulated time series of the aggregate capital stock $\{K_t\}_{t=101}^{t=2,000}$ into two subsamples, with $Z_t = Z_g$ or $Z_t = Z_b$, respectively. For the two subsamples, we estimate the coefficients γ_0 and γ_1 of the equation (6.23) with the help of an OLS-regression.

Step 7) We continue this iteration until the estimated OLS regressors of the loglinear law of motion for the capital stock converge.

Step 8) As it turns out,the fit of the regression is very accurate with an R^2 close to one.

Results. The economy with efficiency mobility behaves very similarly to the one without efficiency mobility. For this reason, we concentrate on displaying the results for the former economy if not mentioned otherwise. The law of motion (6.45) is given by:

$$\ln K' = \begin{cases} 0.0754 + 0.986 \ln K, & \text{if } Z = Z_g, \\ 0.0620 + 0.988 \ln K, & \text{if } Z = Z_b. \end{cases} \tag{6.46}$$

The stationary average aggregate capital stock amounts to $K = 219$.[35]

The distribution of earnings among the employed agents, i.e., $wh(Z)\zeta_i$, is more or less exogenous in our model and is proportional to the efficiency type ζ_i. Of course, the wage w is endogenous in our model. As home production is assumed to be constant over the business cycle, while the earnings of the employed agents increases during booms and decreases during recessions, the distribution of earnings is not constant over the business cycle. During

[35] For example, aggregate capital amounts to $K = 229$ in the economy without income mobility, and the law of motion for the capital stocks are given by $\ln K' = 0.0755 + 0.986 \ln K$ and $\ln K' = 0.0622 + 0.988 \ln K$ in good and bad times, respectively.

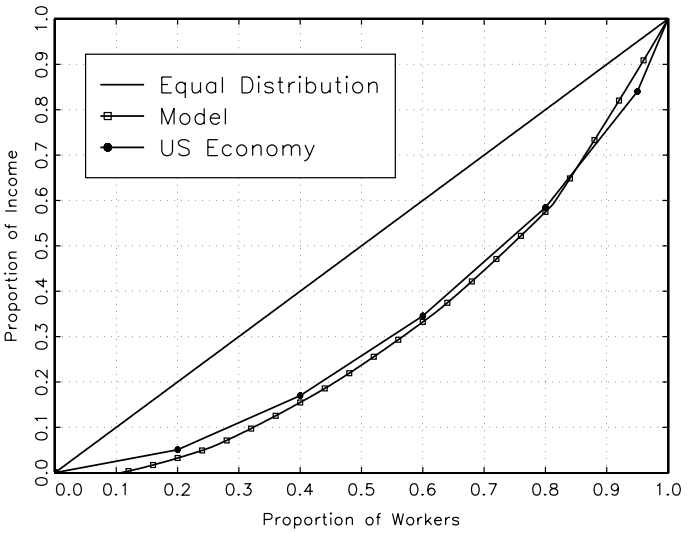

Figure 6.17: Lorenz curve of income

booms, $Z = Z_g$, earnings are more concentrated and characterized by a Gini coefficient equal to 0.305. During a recession, the Gini coefficient of earnings drops to 0.291. The Gini coefficient of income (earnings plus interest income) is more volatile than the Gini coefficient of earnings because the concentration of wealth and the interest rate are procyclical. Consequently, the concentration of interest income and total income increases during booms. The Gini coefficients of income varies between 0.285 and 0.325 over the cycle. The Lorenz curve of income is displayed in Figure 6.17. The income shares are computed as averages over 2000 periods. Notice that we are able to replicate the empirical distribution of income (line with solid circles) very closely.[36] The income distribution implied by the model (dotted and broken line) is almost identical to the one in the US during 1946-84.

Table 6.1 reports the cyclical behavior of income shares for the US and for the model economy with varying efficiency types. The empirical correlations of US output and income shares are taken

[36] The empirical values for the US income and wealth distribution during 1948-86 are provided in Table 1 and 6 of Castañeda et al. (1998b), respectively.

Table 6.1

Income Quintile	Correlation between output and income: US	Correlation between output and income: model
lowest quintile (0-20%)	0.53	0.79
second quintile (20-40%)	0.49	0.79
third quintile (40-60%)	0.31	-0.74
fourth quintile (60-80%)	-0.29	-0.80
next 15% (80-95%)	-0.64	-0.80
top 5% (95-100%)	0.00	-0.78

from Table 2 in CASTAÑEDA ET AL. (1998b). The sample period, again, is 1948-86. The yearly output data is logged and detrended using a Hodrick-Prescott filter with a smoothing parameter $\mu = 100$. The income share of the lower quintiles (0-60%) is procyclical, the income share of the fourth quintile and next 15% (60-95%) is anticyclical, while the top 5% of the income are acyclical.

In the third column of Table 6.1, we report the statistics computed from our simulation over 2000 periods for the economy with time-varying efficiency types. The cyclical behavior of income shares in the economy without mobility is found to be almost identical to the one in the economy with mobility. Again, output is logged and detrended using the Hodrick-Prescott filter with $\mu = 100$ in order to compare it to the empirical numbers. Therefore, we need to compute annual averages of output and income and earnings shares for 2000/8=250 years. The simulated correlation of income is only in good accordance with the empirical observations for the first and second income quintiles as well as for the 80-95% income percentile class. As one possible explanation for the rather poor modelling of the other percentiles, we do not allow for endogenous labor supply (which may result in more procyclical behavior of the 3rd and 4th income quintiles) and we are not very successful in replicating the wealth distribution (which may result in more procyclical interest and profit income for the top 5% of the income distribution).

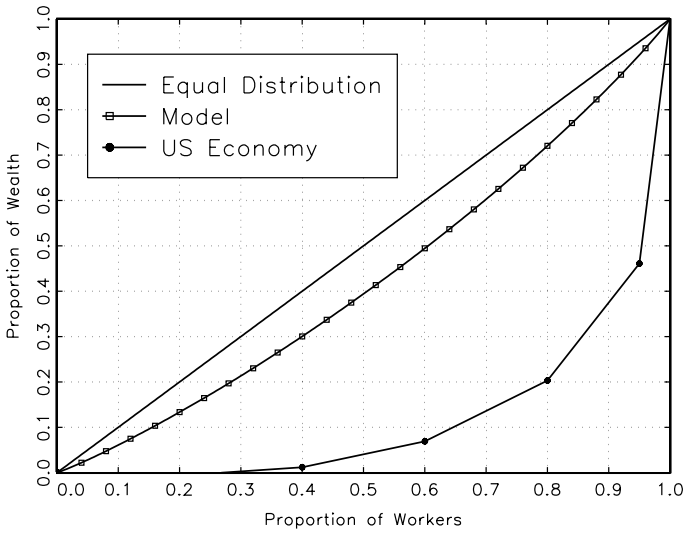

Figure 6.18: Lorenz curve of wealth

The most pronounced effect of income mobility on the distribution of the individual variables earnings, income, and wealth is on the concentration of wealth. There are two opposing effects of income mobility on wealth heterogeneity: In the economy with time-varying efficiency type, wealth-rich and income-rich agents of type $i = 2, 3, 4, 5$ accumulate higher savings for precautionary reasons in case they move down the income ladder. This effect, of course, increases wealth concentration in our economy and we would expect the Gini coefficient of wealth to be higher in the economy with efficiency mobility for this reason. On the other hand, agents of type $i = 5$, for example, might have had efficiency type $i = 4$ or even lower in previous periods so that they have accumulated less wealth than agents who have had the efficiency type $i = 5$ forever. For this reason, wealth heterogeneity is less in an economy with time-varying efficiency types. As it turns out, the former effect dominates and wealth heterogeneity is lower in the case of no mobility. In both economies, the endogenous wealth concentration is much lower than observed empirically. This can be seen in Figure 6.18, which displays the Lorenz curve for the case of time varying efficienty types. The respective Gini coeffi-

cient of wealth only amounts to 0.298. In the economy without efficiency mobility the distribution of wealth is slightly more unequal with a Gini coefficient of 0.347. distribution (broken line). In the next chapter, you will find out how we can improve the modelling of the wealth distribution.

Problems

6.1 **Example 6.2.1**

a) Assume that agents use the first two moments to forecast future factor prices in example 6.2.1. Show that the consideration of an additional moment does not result in much higher accuracy in the prediction of the factor prices as compared to the case of one moment.

b) Assume that the initial distribution of the economy described in example 6.2.1 is given by the stationary distribution and consider a policy that increases the replacement ratio of unemployment insurance to 40%. Compute the transition dynamics and assume that the income tax rate always adjusts in order to balance the budget. How does the wealth distribution change? Compute the Gini coefficients of the income and wealth distribution during the transition and in the new stationary state.

c) Compute the stationary state and the transition dynamics for the growth model 6.2.1 with leisure. Use the utility function (6.12) with $\eta = 0.5$. Use the prediction function $\ln N' = \eta_0 + \eta_1 \ln N$ for aggregate employment N.

d) Implement the Algorithm 6.2.2 using projection methods for the computation of the policy functions.

6.2 **Aggregate Uncertainty:**

a) Assume that agents use the first two moments to forecast future factor prices in example 6.3.1. Show that the consideration of an additional moment does not result in much higher accuracy in the prediction of the factor prices as compared to the case of one moment.

b) Assume that, in example 6.3.1, technology z follows the first-order autoregressive process $z_t = \rho z_{t-1} + \eta_t$ with $\rho = 0.9$ and $\eta_t \sim N(0, 0.01)$. Compute a 5-state Markov-chain approximation of the AR(1)-process using Algorithm 9.2.1. Compute the model of example 6.3.1. How do results change if you use 9 states instead of 5 states for the Markov chain approximation?

c) Assume that the unemployment rate is not constant during booms or recessions. Assume that leisure is an argument of the utility function. How does the program `ch63cas1.g` need to be adjusted?

6.3 **Costs of Business Cycles**

a) Compute the gain in average utility from eliminating the business cycle fluctuations in the model of İmrohoroğlu (1989) presented in Section 6.3.1.

b) Assume that there is perfect insurance in the economy described in chapter 6.3.1. Each agent receives the average income of the economy. By how much is the utility gain from the elimination of cyclical fluctuations reduced?

6.4 **Dynamics of the Income Distribution** Compute the model of Section 6.3.2 with the same calibration as presented in Secton 6.3.2 except that $\pi(\epsilon', z_g|\epsilon, z_b) = \pi(\epsilon', z_b|\epsilon, z_b)$ and $\pi(\epsilon', z_b|\epsilon, z_g) = \pi(\epsilon', z_g|\epsilon, z_g)$. Notice that for this calibration, next-period aggregate employment N is not only function of next-period aggregate productivity z', but also of current-period productivity and employment, $N' = N'(N, z, z')$. Recompute the mean Gini coefficients of income and earnings and the correlation of income and earnings with output.

Chapter 7

Overlapping Generations Models

Overview. In this chapter, we introduce an additional source of heterogeneity. Agents do not only differ with regard to their individual productivity or their wealth, but also with regard to age. First, you will learn how to compute a simple overlapping generations model (OLG model) where each generation can be represented by a homogeneous household. Subsequently, we study the dynamics displayed by the typical Auerbach-Kotlikoff model. Finally, we present methods in order to solve more elaborate heterogeneous-agent OLG models. First, we introduce stochastic individual productivity into the OLG model. Second, we study aggregate uncertainty.

The previous chapter concentrated on the computation of models that were based on the Ramsey model. In this chapter, we will analyze overlapping generations (OLG) models. The central difference between the OLG model and the Ramsey model is that there is a continuous turnover of the population. The life-time is finite and every period, a new generation is born and the oldest generation is dying. In such models, many cohorts coexist at any time. In the pioneering work on OLG models by Samuelson (1958) and Diamond (1965), the number of coexisting cohorts only amounts to two, the young and working generation on the one hand and the old and retired generation on the other hand. In these early studies of simple OLG models, Samuelson and Diamond focused on the analysis of theoretical problems, i.e., if there is a role for money and what are the effects of national debt, respectively.

The OLG model is a natural framework to analyze life-cycle problems such as the provision of public pensions, endogenous fer-

tility, or the accumulation of human capital and wealth. In order to study the quantitative effects of economic policy, subsequent work has been directed towards the study of large scale numerical OLG models. Typically, cohorts were identified with the members of the population of the same age. One seminal work in this area is the study of dynamic fiscal policy by Auerbach and Kotlikoff (1987).[1] In their work, the first cohort was identified with the 20-year-old cohort, who were the new entrant in the labor market. 55 different generations were distinguished so that at age 75, all agents were dying. In their 55-period overlapping generations model of a representative household, they showed, among others, that a 60% benefit level of unfunded social security decreases welfare by approximately 5-6% of total wealth (depending on the financing of the social security expenditures).

In recent years, there has been a wide range of problems studied with the help of OLG models. In addition to the early work by Auerbach and Kotlikoff, subsequent authors have introduced various new elements in the study of overlapping generations, like, for example, stochastic survival probabilities, bequests, or individual income mobility, to name but a few. In this vein, Huggett and Ventura (2000) look at the determinants of savings and use a calibrated life-cycle model in order to evaluate why high income households save as a group a much higher fraction of income than do low income households in US cross-section data. Relatedly, Huggett (1996) shows that the life-cycle model is able to reproduce the US wealth Gini coefficient and a significant fraction of the wealth inequality within age groups. Heer (2001b) studies the role of bequests in the explanation of observed wealth inequality. The US tax system and the US social security system have also attracted substantial attention: İmrohoroğlu (1998) analyzes the effects of capital income taxation, İmrohoroğlu et al. (1998) evaluate the benefits of tax favored retirement accounts, and Ventura (1999) considers the effects of a flat-rate versus a progressive income tax. The effects of social security and

[1] Other early studies of life-cycle economies include Summers (1981), Auerbach et al. (1983), Evans (1983), or Hubbard and Judd (1987).

unemployment compensation are studied by İmrohoroğlu et al. (1995), Hubbard et al. (1995), and Heer (2003), among others. İmrohoroğlu et al. (1995), for example, examine the effects of a change in the public pensions on economic welfare in a 60-period OLG model with liquidity constraints and income uncertainty. In their model, welfare may even increase following the introduction of unfunded social security. The OLG framework is also the natural framework in order to study questions related to the demographic transition. As the population is ageing, the pension system gets under pressure. De Nardi et al. (1999) look at different policy plans in order to cope with the transition. Heckman et al. (1998) explain the rising wage inequality since the 1960s with the enlarged cohorts of the Baby Boom. As one of the very few studies, Heckman et al. endogenizes the schooling choice of the young cohort. The OLG model framework has also been successfully applied to the study of business cycle fluctuations or the pricing of assets and equities.[2] The list of recent applications is only selective and by no means exhaustive.

This chapter is organized differently from the previous chapters. In the Chapters 1-6, the sections have been devoted to the study of numerical problems and methods. In this chapter, you will not get to know new methods, but you will rather adapt previously introduced methods to the study of OLG models. For this reason, the sections are rather organized with regard to the particular model types. In the first section, you will be introduced to the basic life-cycle model with age-dependent cohorts. In the following section, the transition between two steady states is examined. In the final section, we also introduce both stochastic individual and stochastic aggregate productivity.

7.1 Life-cycle Model with Perfect Foresight

In this section, we solve an Overlapping Generations model without uncertainty. All agents of one cohort are identical and their

[2] Please see Ríos-Rull (1996), Storesletten et al. (2001), and Brooks (2002), among others.

behavior is analyzed by means of the behavior of a representative agent.

7.1.1 An Illustrative Example

We will use a 60-period Overlapping Generations model as an illustration. The periods correspond to years. The model is a much simplified version of the economy studied by AUERBACH and KOTLIKOFF (1987).[3] Three sectors can be depicted: households, production, and the government.

Households. Every year, a generation of equal measure is born. The total measure of all generations is normalized to one. Their first period of life is period 1. A superscript s of a variable denotes the age of the generation, a subscript t denotes time. For example, c_t^s is the consumption of the s-year old generation at time t.

Households live $T + T^R = 40 + 20$ years. Consequently, the measure of each generation is 1/60. During their first $T = 40$ years, agents supply labor n_t^s at age s in period t enjoying leisure $l_t^s = 1 - n_t^s$. After T years, retirement is mandatory ($n_t^s = 0$ for $s > T$). Agents maximize life-time utility at age 1 in period t:

$$\sum_{s=1}^{T+T^R} \beta^{s-1} u(c_{t+s-1}^s, l_{t+s-1}^s), \tag{7.1}$$

where β denotes the discount factor. Notice that, different from the discount factor β in the Ramsey model, β does not necessarily need to be below one in an OLG model so that life-time utility is finite.[4]

Instantaneous utility is a function of both consumption and leisure:

$$u(c, l) = \frac{\left((c + \psi) l^\gamma\right)^{1-\eta} - 1}{1 - \eta}. \tag{7.2}$$

[3] For example, we do not consider different types of agents among one cohort and model the tax and pension system in a very stylized way.

[4] For restrictions on the size of β in economies with infinitely-lived agents, see DEATON (1991).

The small constant $\psi = 0.001$ is added in order to ensure that utility is finite even for zero consumption in the case of no income. This choice will turn out to be very convenient in the subsequent computations as we will be able to use a grid over the individual capital stock with a lower bound equal to zero.

Agents are born without wealth, $k_t^1 = 0$, and do not leave bequests, $k^{61} = 0$. Since capital k is the only asset held by individuals, the terms capital and wealth will henceforth be used interchangeably. Agents receive income from capital k_t^s and labor n_t^s. The real budget constraint of the working agent is given by

$$k_{t+1}^{s+1} = (1 + r_t)k_t^s + (1 - \tau_t)w_t n_t^s - c_t^s, \quad s = 1, \ldots, T, \tag{7.3}$$

where r_t and w_t denote the interest rate and the wage rate in period t, respectively. Wage income in period t is taxed at rate τ_t. We can also interpret $\tau_t w_t n_t^s$ as the worker's social security contributions.

The first-order conditions of the working household are given by:

$$\begin{aligned} \frac{u_l(c_t^s, l_t^s)}{u_c(c_t^s, l_t^s)} &= \gamma \frac{c_t^s + \psi}{l_t^s} = (1 - \tau_t)w_t, & (7.4) \\ \frac{1}{\beta} &= \frac{u_c(c_{t+1}^{s+1}, l_{t+1}^{s+1})}{u_c(c_t^s, l_t^s)} \left[1 + r_{t+1}\right] & (7.5) \\ &= \frac{\left(c_{t+1}^{s+1} + \psi\right)^{-\eta} \left(l_{t+1}^{s+1}\right)^{\gamma(1-\eta)}}{\left(c_t^s + \psi\right)^{-\eta} \left(l_t^s\right)^{\gamma(1-\eta)}} \left[1 + r_{t+1}\right]. \end{aligned}$$

During retirement, agents receive public pensions b irrespective of their employment history and the budget constraint of the retired worker is given by

$$k_{t+1}^{s+1} = (1 + r_t)k_t^s + b - c_t^s, \quad s = T + 1, \ldots, T + T^R. \tag{7.6}$$

The first-order condition of the retired worker is given by (7.5) with $l_t^s = 1$.

Production. The production sector is identical to the one used in previous chapters. Firms are of measure one and produce output Y_t in period t with labor N_t and capital K_t. Labor N_t is paid the

wage w_t. Capital K_t is hired at rate r_t and depreciates at rate δ. Production Y_t is characterized by constant returns to scale and assumed to be Cobb-Douglas:

$$Y_t = F(K_t, N_t) = K_t^{\alpha} N_t^{1-\alpha}. \tag{7.7}$$

In a factor market equilibrium, factors are rewarded with their marginal product:

$$w_t = (1-\alpha) K_t^{\alpha} N_t^{-\alpha}, \tag{7.8}$$

$$r_t = \alpha K_t^{\alpha-1} N_t^{1-\alpha} - \delta. \tag{7.9}$$

Government. The government uses the revenues from taxing labor in order to finance its expenditures on social security:

$$\tau_t w_t N_t = \frac{T^R}{T + T^R} b. \tag{7.10}$$

Following a change in the provision of public pensions b or in gross labor income $w_t N_t$, the labor income tax rate τ_t adjusts in order to keep the government budget balanced.

Equilibrium. The concept of equilibrium applied in this section uses a recursive representation of the consumer's problem following Stokey and Lucas (1989). This specification turns out to be very amenable to one of the two solution methods described in this section. For this reason, let $V^s(k_t^s, K_t, N_t)$ be the value of the objective function of the s-year old agent with wealth k_t^s. K_t and N_t denote the aggregate capital stock and employment. $V^s(k_t, K_t, N_t)$ is defined as the solution to the dynamic program:

$$V^s(k_t^s, K_t, N_t) = \begin{cases} \max\limits_{k_{t+1}^{s+1}, c_t^s, n_t^s} \left[u\left(c_t^s, l_t^s\right) + \beta V^{s+1}(k_{t+1}^{s+1}, K_{t+1}, N_{t+1}) \right], \\ \qquad s = 1, \ldots, T \\ \\ \max\limits_{k_{t+1}^{s+1}, c_t^s} \left[u\left(c_t^s, 1\right) + \beta V^{s+1}(k_{t+1}^{s+1}, K_{t+1}, N_{t+1}) \right], \\ \qquad s = T+1, \ldots, T+T^{R-1}, \end{cases} \tag{7.11}$$

subject to (7.3) and (7.6), respectively, and

$$V^{T+T^R}(k_t^{T+T^R}, K_t^{T+T^R}, N^{T+T_t^R}) = u(c_t^{T+T^R}, 1). \tag{7.12}$$

The value function $V^s(\cdot)$, in particular, depends on the aggregate state variables K_t and N_t that determine the wage rate w_t and the interest rate r_t in period t via (7.8) and (7.9), and, in addition, τ_t with the help of the balanced budget (7.10). Furthermore, $V^s(\cdot)$ depends on the age s of the household, but not on calendar time t.

An equilibrium for a given government policy b and initial distribution of capital $\{k_0^s\}_{s=1}^{T+T^R}$ is a collection of value functions $V^s(k_t^s, K_t, N_t)$, individual policy rules $c^s(k_t^s, K_t, N_t)$, $n^s(k_t^s, K_t, N_t)$, and $k^{s+1}(k_t^s, K_t, N_t)$, relative prices of labor and capital $\{w_t, r_t\}$, such that:

a) Individual and aggregate behavior are consistent:

$$N_t = \sum_{s=1}^{T} \frac{n_t^s}{T + T^R}, \tag{7.13}$$

$$K_t = \sum_{s=1}^{T+T^R} \frac{k_t^s}{T + T^R}. \tag{7.14}$$

The aggregate labor supply N_t is equal to the sum of the labor supplies of each cohort, weighted by its mass $1/(T + T^R) = 1/60$. Similarly, the aggregate capital supply is equal to the sum of the capital supplies of all cohorts.

b) Relative prices $\{w_t, r_t\}$ solve the firm's optimization problem by satisfying (7.8) and (7.9).

c) Given relative prices $\{w_t, r_t\}$ and the government policy b, the individual policy rules $c^s(\cdot)$, $n_t^s(\cdot)$, and $k_{t+1}^s(\cdot)$ and solve the consumer's dynamic program (7.11)-(7.12).

d) The goods market clears:

$$K_t^{\alpha} N_t^{1-\alpha} = \sum_{s=1}^{T+T^R} \frac{c_t^s}{T + T^R} + K_{t+1} - (1 - \delta)K_t. \tag{7.15}$$

e) The government budget (7.10) is balanced.

Calibration. Our model just serves as an illustration. Therefore, we calibrate our model with the functional forms and parameters as commonly applied in DGE life-cycle models. Our benchmark case is characterized by the following calibration: $\eta = 2$, $\beta = 0.99$, $\alpha = 0.3$, $\delta = 0.1$, replacement ratio $\zeta = \frac{b}{(1-\tau)w} = 0.3$, $T = 40$, $T^R = 20$. γ is chosen in order to imply a steady state labor supply of the working agents approximately equal to 35% of available time and amounts to $\gamma = 2.0$. The small constant ψ is set equal to 0.001.

7.1.2 Computation of the Steady State

In this section, we compute the steady state that is characterized by a constant distribution of the capital stock over the generations, $\{k_t^s\}_{s=1}^{60} = \{k_{t+1}^s\}_{s=1}^{60} = \{\bar{k}^s\}_{s=1}^{60}$. In the steady-state economy, the aggregate capital stock and aggregate employment are constant, $K_t = \bar{K}$ and $N_t = \bar{N}$, respectively. As a consequence, prices w and r are constant, too, and so are taxes τ. Therefore, in the steady state, the computation of the equilibrium is simplified, as for given aggregate capital stock $\bar{K}$ and employment $\bar{N}$, the value function and the individual policy function are only functions of the age s and individual wealth k^s. For notational convenience, we drop the time index t in this section and will only reintroduce it in the next section.

The general solution algorithm is described by the following steps:

Algorithm 7.1.1 (Computation of the Stationary Equilibrium of the OLG Model in Section 7.1)

Purpose: *Computation of the stationary equilibrium.*

Steps:

Step 1: Make initial guesses of the steady state values of the aggregate capital stock $\bar{K}$ and employment $\bar{N}$.

Step 2: Compute the values w, r, and τ, which solve the firm's Euler equations and the government budget.

Step 3: Compute the optimal path for consumption, savings, and employment for the new-born generation by backward induction given the initial capital stock $k^1 = 0$.
Step 4: Compute the aggregate capital stock $\bar{K}$ *and employment* $\bar{N}$.
Step 5: Update $\bar{K}$ *and* $\bar{N}$ *and return to step 2 until convergence.*

In step 3, the household's optimization problem needs to be solved. Our aim is to compute the steady-state distribution of capital $\left\{\bar{k}^s\right\}_{s=1}^{60}$. There are basically two different numerical techniques in order to solve this problem in an economy with perfect foresight. Assume that we would like to compute the optimal next-period capital stock $k^{s+1}(k^s, \bar{K}, \bar{N})$, current consumption $c^s(k^s, \bar{K}, \bar{N})$, and current labor supply $n^s(k^s, \bar{K}, \bar{N})$. Then, we may either compute the policy functions only for $k^s = \bar{k}^s$ or we may compute the policy function over an interval $[k_{min}, k_{max}]$. In the first case, we have simultaneously computed the distribution of individual capital. In the second case, we compute the time path of savings, employment, and consumption using the optimal decision functions and the initial condition $k^1 = 0 = \bar{k}^1$. With the help of $k^2(k^1, K, N,)$, we can compute $\bar{k}^2 = k^2(\bar{k}^1, K, N)$ and similarly $\bar{k}^s$, $s = 3, \ldots, 60$. The methods are described in turn.

Direct Computation of the Steady State Distribution. In order to illustrate the direct computation of the steady-state distribution, consider the first-order conditions of the working household with regard to labor supply and next-period capital stock, (7.4) and (7.5), respectively. Inserting the working households budget (7.3) in these two equations, we derive the following two steady-state equations for $s = 1, \ldots, T-1$:

$$(1-\tau)w = \gamma \frac{(1+r)k^s + (1-\tau)wn^s - k^{s+1} + \psi}{1-n^s}, \tag{7.16}$$

$$\begin{aligned} \frac{1}{\beta} =& \frac{\left((1+r)k^{s+1} + (1-\tau)wn^{s+1} - k^{s+2} + \psi\right)^{-\eta}}{\left((1+r)k^s + (1-\tau)wn^s - k^{s+1} + \psi\right)^{-\eta}} \\ & \times \frac{\left(1-n^{s+1}\right)^{\gamma(1-\eta)}}{\left(1-n^s\right)^{\gamma(1-\eta)}} \left[1+r\right]. \end{aligned} \tag{7.17}$$

Similarly, (7.16) also holds for $s = T$, while (7.17) needs to be adjusted:

$$\begin{aligned}\frac{1}{\beta} =& \frac{\left((1+r)k^{T+1} + b - k^{T+2} + \psi\right)^{-\eta}}{\left((1+r)k^{T} + (1-\tau)wn^{T} - k^{T+1} + \psi\right)^{-\eta}} \\ & \times \frac{1}{(1-n^{T})^{\gamma(1-\eta)}} \left[1+r\right].\end{aligned} \tag{7.18}$$

For the retired agent, the labor supply is zero, $n^s = 0$, and the Euler equation is given by

$$\frac{1}{\beta} = \frac{\left((1+r)k^{s+1} + b - k^{s+2} + \psi\right)^{-\eta}}{\left((1+r)k^{s} + b - k^{s+1} + \psi\right)^{-\eta}} \left[1+r\right] \tag{7.19}$$

for $s = T+1, \ldots, T+T^R-1 = 41, \ldots, 59$. Remember that the optimal capital stock after death is also set equal to zero, $k^{61} \equiv 0$. The equations (7.16)-(7.19) for $s = 1, \ldots, 59$ constitute a system of 59+40=99 equations in the 59+40=99 unknowns $\{k^s\}_{s=2}^{60}$ and $\{n^s\}_{s=1}^{40}$. Therefore, we have the same type of problem that we already encountered in Example 1.1.1 in Chapter 1. We, again, need to compute a non-linear equations system in n unknowns, in our case with $n = 99$. However, as you may have learned by now, the computation of such large-scale non-linear problems may become cumbersome. Therefore, we better make further use of the recursive structure of the problem.

In the program `ch71d.g`, we compute the solution of this problem. We know that agents are born without wealth at age 1, $k^1 = 0$ and do not leave bequests. Therefore, $k^{61} = 0$. Let us start by providing an initial guess of the wealth in the last period of life, k^{60}. With the help of this initial guess and the retired worker's first-order condition (7.19) at age $s = 59$, we are able to compute k^{59}. In this case, we only have to solve a non-linear equation problem with one unknown. Having computed k^{s+1} and k^{s+2} for the retired agent, we simply iterate backwards and compute k^s for $s = 59, 58, \ldots, 41$. From (7.16) for $s = 40$ and (7.18), we are able to compute n^{40} and k^{40}. We continue to compute k^s and n^s with the help of the values n^{s+1}, k^{s+1}, and k^{s+2} found in the previous two iterations and with the help of equations (7.16)

and (7.17) until we have computed k^1 and n^1. If $k^1 = 0$, we are finished. Otherwise, we need to update our guess for k^{60} and recompute the distribution of individual capital and labor supply, $\{k^s\}_{s=1}^{60}$ and $\{n^s\}_{s=1}^{40}$. Notice, in particular, that we need to iterate backwards. We cannot start with a guess of k^2 given $k^1 = 0$ and iterate forwards in the presence of endogenous labor supply. With exogenous labor supply, we would also be able to find the optimal capital distribution with forward iteration (why?).

Finally, we need to mention how we update successive values for k^{60}. In `ch71d.g`, we apply the *Secant Method* that we present in Section 8.5. Successive values of k^{60} are found by:

$$k_{i+2}^{60} = k_{i+1}^{60} - \frac{k_{i+1}^{60} - k_i^{60}}{k_{i+1}^1 - k_i^1} k_{i+1}^1,$$

where the subscript i denotes the number of the iteration. As the first two guesses for k^{60}, we choose the values $k_1^{60} = 0.4$ and $k_2^{60} = 0.5$. After 5 iterations, we find the absolute value of k_5^1 to be below 10^{-8}.

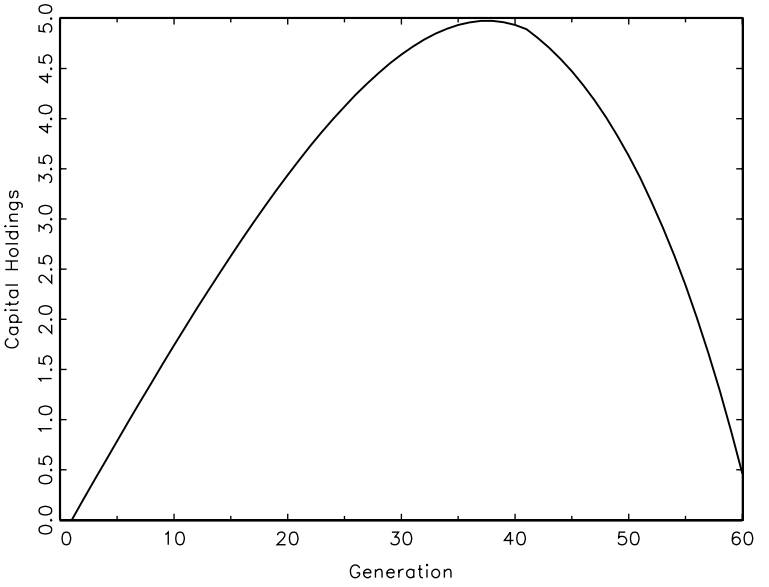

Figure 7.1: Age-wealth profile

The solution for $\{\bar{k}^s\}_{s=1}^{60}$ is displayed in Figure 7.1. Typically for the life-cycle model, savings k^s increase until retirement at age $s = T$ and decline monotonically thereafter. The aggregate capital stock amounts to $\bar{K} = 3.166$. Optimal labor supply is graphed in Figure 7.2. The labor supply declines with increasing age s because older agents hold higher stocks of capital. As a consequence, marginal utility of income declines for older age. Average employment amounts to 0.357 so that aggregate employment is equal to $\bar{N} = 0.2138$. The steady state values of pensions, the interest rate, and taxes are given by $b = 0.1313$, $r = 4.546\%$, and $\tau = 13.04\%$, respectively.

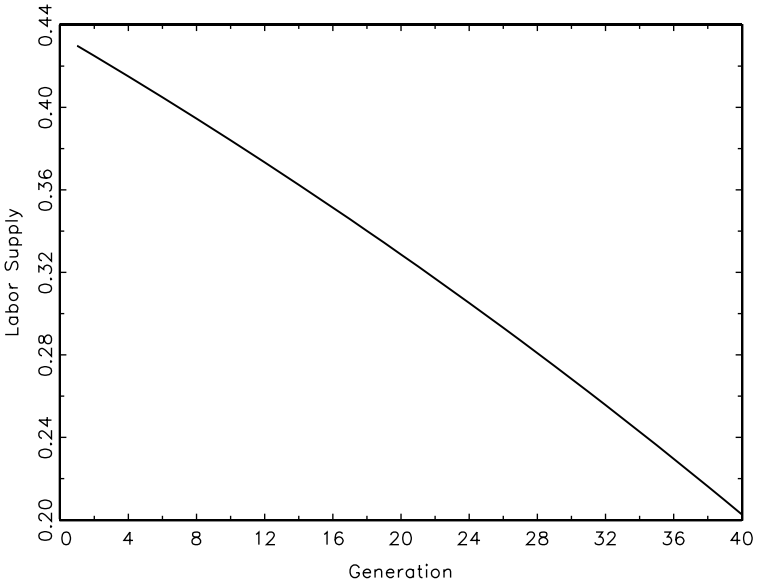

Figure 7.2: Age-labor supply profile

This direct method of computing the policy functions at single values of the state space only is very fast and accurate. However, it has the disadvantage that it is only applicable in economies with perfect foresight. Assume that we introduce an autocorrelated shock to productivity so that next-period income is stochastic and depends on the current period income. For the sake of

the argument, let us assume that we have $n = 5$ different levels of individual income in each period and that the transition between these states is described by a Markov process. In this case, of course, agents build up different levels of wealth during their working life depending on their income history. How many different types would we have to distinguish at age 41? At age 1, all have equal (zero) wealth. At age 2, we have 5 types of agents who differ with regard to their period 1 income and hence with regard to their savings (and wealth in period 2). At age 3, we have $5 \times 5 = 25$ different types of agent. If we continue like this, we have to distinguish 5^{40} different agents at age 41 who all hold different levels of wealth. During retirement, the number of different agents does not increase because all agents receive equal retirement benefits. Therefore, if we compute the optimal savings path recursively with the direct method, we would have to apply Algorithm 7.1.1 $9.09 \cdot 10^{27}$ times![5] In such an economy, of course, the only feasible alternative solution method consists of computing the policy functions over an interval of the state space. This method is described next.

Computation of the policy functions. In order to compute the age-wealth profile of the steady state, $\left\{\bar{k}^s\right\}_{s=1}^{60}$, we may also compute the optimal policy function $k^{s+1}(k^s, \bar{K}, \bar{N})$ for each cohort s over an interval $k^s \in [k^s_{min}, k^s_{max}]$. As we do not know the age-wealth profile in advance, we will start to compute the policy functions for each age over the same interval $[k^s_{min}, k^s_{max}] = [k_{min}, k_{max}]$. In later iterations, we may adapt the state space for each cohort s.[6] Having computed the policy functions, it is easy to find the solution $\left\{\bar{k}^s\right\}_{s=1}^{60}$. We simply start with $k^1 = \bar{k}^1$ and compute $k^2(k^1) = k^2(0) = \bar{k}^2$. Similarly, we compute $\bar{k}^{s+1}$ with the help of $k^{s+1}(\bar{k}^s, \bar{K}, \bar{N})$ for $s = 2, \ldots, T + T^R - 1$.

There are various methods in order to compute the policy functions. We will discuss value function iteration and projection

[5] $5^{40} = 9.0949470 \cdot 10^{27}$.

[6] The adaption of the age-specific asset grid may not be a feasible strategy in the case of heterogeneous agents among the same cohort, a problem that you will encounter in Section 7.2.

methods that you have already encountered in the Chapters 1 and 4 for the solution of the Ramsey problem. We discuss these methods in turn and start with value function iteration in an economy with finite life-time.

Value function iteration. A straightforward method of approximating the value function $V^s(k^s)$, $s = 1, \ldots, T + T^R$, involves tabulating it for a finite number n_k of points on the state space starting in the last period of life, $T + T^R$, and iterating backwards in time to the period $s = 1$. The maximization occurs over the interval $[k_{min}, k_{max}]$, which, in particular, must contain the steady state, a point, which is only determined as the outcome of our computation. In the last period $T + T^R$, the value function $V^{T+T^R}(k^{T+T^R})$ is given by (7.12) with $c^{T+T^R} = (1 + r)k^{T+T^R} + b$. For a given table of values for $V^{s+1}(k^{s+1})$ on the grid $[k_{min}, k_{max}]$, the approximate retired agent's maximum at age s on the right-hand side of (7.11)-(7.12) can be found by choosing the largest value for $V^{s+1}(k^{s+1})$ given k^s, which we store as $k^{s+1}(k^s)$. Together with the two neighboring points on the asset grid, we bracket the maximum and apply a Golden Section Search in order to find the maximum of the Bellman equation (7.11). In order to get values of the value function $V^{s+1}(\cdot)$ off grid points, we interpolate linearly.

At this point, we need to emphasize a crucial difference between finite-horizon and infinite-horizon problems. Differently from value function iteration in infinite-horizon models, we know the value of the agent in the last period of his life, $V^{60} = u(c^{60}, l^{60})$ with $c^{60} = (1 + r)k^{60} + b$ and $l^{60} = 1.0$. As a consequence, we do not have the problem to provide an initial guess for the value function. This feature also holds for the other solution methods of finite-horizon problems, e.g., the projection method presented below. Given the value $V^{60}(k^s, \bar{K}, \bar{N})$ for $k^s \in [k_{min}, k_{max}]$, we can find the value function of the different cohorts, $V^s(\cdot)$, $s = 59, \ldots, 1$ with only one iteration. As a consequence, the computation of the policy functions is much faster in most applications with finite horizons than in infinite-horizon problems. Notice, however, that the need for storage capacity increases as the number of policy functions is multiplied by the number of different age cohorts.

The dynamic programming problem of the working agent (7.11) involves the maximization over an additional control, the labor supply n^s. A standard procedure to solve this kind of problem consists of choosing the largest value over a grid on the labor supply $[n_{min}, n_{max}]$. As a consequence, the optimal next period capital stock together with the optimal labor supply decision is found by iterating over a two-dimensional grid. For reasonable required accuracy, we often find this procedure to already imply prohibitive storage capacity and computing speed in order to be a useful method on personal computers. Instead, we only iterate over a one-dimensional grid of the capital stock and solve the household's Euler equation (7.4) and budget constraint (7.3) for given current and next period capital stock (k^s, k^{s+1}). For our choice of the functional form for utility $u(\cdot)$, we can solve these two equations even directly for c^s and n^s for given k^s and k^{s+1}. Notice that this procedure does not restrict the controls c^s and n^s to lie on any grid.

The solution is computed with the help of value function iteration in the program `ch71v.g`. Concerning our computation details, wealth is bounded below by $k_{min} = 0$, while maximum wealth is set equal to $k_{max} = 12.0$, which is found to never be binding.[7] Furthermore, we choose an equispaced grid over the capital stock $[k_{min}, k_{max}]$ of $n_k = 100$ points. The required storage capacity associated with this algorithm is equal to $(2 * T + T^R) \cdot n_k = 1000$ numbers. The age-wealth profile computed with value function iteration is almost identical to the one displayed in Figure 7.1. The aggregate capital stock and aggregate employment amount to $\bar{K} = 3.181$ and $\bar{N} = 0.2139$, respectively. Notice that due to linear interpolation between grid points, the aggregate capital stock diverges by 0.6% from the aggregate capital stock found with the direct method described above. This difference is a good measure of accuracy as the latter solution can be expected to coincide with the true solution (in the case of direct computation of the steady-state distribution, the accuracy of the non-linear equation

[7] In our model, we abstract from any inequality constraints such as $c \geq 0$ or $k \geq 0$ because these constraints do not bind (except in period 1 with $\bar{k}^1 \equiv 0.0$).

solution and the divergence of $\bar{k}^1$ from zero are both less than 10^{-8}).

Projection methods. Alternatively, we compute the steady state solution $\left\{\bar{k}^s\right\}_{s=1}^{60}$ with the help of projection methods that we introduced in Chapter 4. For this reason, we approximate the consumption function $c^s(k^s, \bar{K}, \bar{N})$, $s = 1, \ldots, 60$, and the labor supply $n^s(k^s, \bar{K}, \bar{N})$, $s = 1, \ldots, 40$, with Chebyshev polynomials of order n_c and n_n over the interval $[k_{min}, k_{max}]$, respectively:

$$\begin{aligned} c^s(k^s, \bar{K}, \bar{N}) &= \frac{1}{2}ac_0^s + \sum_{j=1}^{n_c} ac_j^s T_j(z(k^s)), \\ n^s(k^s, \bar{K}, \bar{N}) &= \frac{1}{2}an_0^s + \sum_{j=1}^{n_n} an_j^s T_j(z(k^s)), \end{aligned}$$

where $z(k^s) = (2k^s - k_{min} - k_{max})/(k_{max} - k_{min})$ is the linear transformation that maps $k^s \in [k_{min}, k_{max}]$ into the interval $[-1, 1]$.[8] We choose orthogonal collocation to compute the coefficients ac_j^s and an_j^s from equations (7.4) and (7.5). In the case of the retired worker, for example, we solve the system of $n_c + 1$ nonlinear equations at the values z that are the $n_c + 1$ (transformed) zeros of the Chebyshev polynomial T_{n_c}. In order to solve the nonlinear-equations problem, we use a quasi-Newton method. The initial guess for the coefficients ac^s and an^s are the coefficients ac^{s+1} and an^{s+1} for $s < T + T^R$. For $s = T + T^R$, we are able to compute the exact values of c^{T+T^R} at $2n_c$ Chebyshev interpolation nodes because we know that the household consumes all his income and his wealth in the last period. Therefore, we can approximate the function in period $t + T^R$ by least squares with the help of Algorithm 8.2.2. For $s = T$, we need to provide an initial guess of the coefficients for labor supply, an^{40}. We use an inelastic labor supply function, $n^{40}(k^s) = 0.3$ in order to initialize the coefficients again making use of Algorithm 8.2.2.

The program `ch71p.g` computes the steady-state solution with the help of projection methods. Concerning our computational details, we chose a degree of approximation n_c equal to 3. By this

[8] See also (8.36) in Chapter 8 where we describe the transformation in more detail.

choice, nc_3 and nc_3/nc_2 are less than 10^{-6} and 10^{-3}, respectively, and we can be confident that our approximation is acceptable. Similarly, we choose $n_n = 3$. We also choose the same interval over the state space, $[k_{min}, k_{max}] = [0, 12]$, as in the case of the value function iteration. The computed Chebyshev coefficients drop off nicely because the decision functions of the household can be described by polynomial functions of small degree quite accurately. Besides, all parameter values are exactly the same as in the case of value function iteration.

The results from the solution of the steady state distribution with the help of the projection method almost coincide with those from the value function iteration. In the former, the aggregate capital stock amounts to $\bar{K} = 3.172$ and employment is equal to $\bar{N} = 0.2148$. So the solution from the projection method is closer to the true value of the aggregate capital but less accurate in the case of employment. The optimal policy functions for consumption differs by less than 0.1% over the range $[k_{min}, k_{max}] = [0, 12]$ between these two methods. Importantly, however, the algorithm based on projection methods is 2,000-3,000 times faster than the one based on value function iteration.

7.1.3 Computation of the Transition Path

In their seminal work, Auerbach and Kotlikoff (1987) have laid the groundwork for the modern analysis of dynamic fiscal policy. Typically, they analyze the question how a particular policy affects the welfare of different generations. For example, how does a change in the pension replacement ratio, i.e. the ratio of pensions to net wage income, affect the life-time utility of present and future cohorts of the population. In their analysis, they assume that the economy is in a steady state in period 0 that, for example, is characterized by a replacement ratio of 30%. At the beginning of period 1, the government announces an unexpected change of pension policy, for example a decrease of the replacement ratio to 20% that becomes effective in period t. Agents have perfect foresight and already adjust their behavior in time period 1 and all subsequent periods. After a certain number of transition periods,

the economy converges to the new steady state. The number of transition periods are taken as approximately 2-3 times the number of generations. Auerbach and Kotlikoff, for example, assume in their 55-overlapping generations model that the economy has reached the new steady state after 150 periods.

In this section, we will compute the transition dynamics associated with a long-run once-and-for-all change of fiscal policy. In particular, we look at an unexpected change of the replacement ratio from 30% to 20% that both is announced in period 1 and also becomes effective in period 1. While Auerbach and Kotlikoff consider a 55-period model in their original work, we will only use 6 generations for illustrative purposes. Of course, the generality of our methods is unaffected by this innocent assumption. The periods in our model can be interpreted as decades. During the first 4 decades, the agents are working, while during the last two decades of their life, they are retired. Besides, the model is exactly the same as the one described in Section 7.1. For your convenience, we have summarized the description of the economy in Example 7.1.1. As we consider decades rather than years, we also need to adjust the calibration of the discount factor β and the depreciation rate δ. The new values are also summarized in Example 7.1.1.

Example 7.1.1

6-period Overlapping Generations Model.

Households live 6 periods. Each generation is of measure 1/6. The first 4 periods, they are working, the last two periods, they are retired and receive pensions. Households maximize life-time utility at age 1 in period t:

$$\sum_{s=1}^{6} \beta^{s-1} u(c^s_{t+s-1}, l^s_{t+s-1}).$$

Instantaneous utility is a function of both consumption and leisure:

$$u(c,l) = \frac{((c+\psi)l^{\gamma})^{1-\eta} - 1}{1-\eta}.$$

The working agent of age s faces the following budget constraint in period t:

$$k_{t+1}^{s+1} = (1 + r_t)k_t^s + (1 - \tau_t)w_t n_t^s - c_t^s, \quad s = 1, \ldots, 4.$$

The budget constraint of the retired worker is given by

$$k_{t+1}^{s+1} = (1 + r_t)k_t^s + b_t - c_t^s, \quad s = 5, 6$$

with $k_t^7 \equiv 0$ and $l_t^5 = l_t^6 = 1$.
Production Y_t is characterized by constant returns to scale and assumed to be Cobb-Douglas:

$$Y_t = F(K_t, N_t) = K_t^\alpha N_t^{1-\alpha}.$$

In a factor market equilibrium, factors are rewarded with their marginal product:

$$\begin{aligned} w_t &= (1 - \alpha)K_t^\alpha N_t^{-\alpha}, \\ r_t &= \alpha K_t^{\alpha-1} N_t^{1-\alpha} - \delta. \end{aligned}$$

Furthermore, the government budget is balanced in every period t:

$$\tau_t w_t N_t = \frac{2}{6} b_t.$$

In equilibrium, individual and aggregate behavior are consistent:

$$\begin{aligned} N_t &= \sum_{s=1}^{4} \frac{n_t^s}{6}, \\ K_t &= \sum_{s=1}^{6} \frac{k_t^s}{6}, \end{aligned}$$

and the goods market clears:

$$K_t^\alpha N_t^{1-\alpha} = \sum_{s=1}^{6} \frac{c_t^s}{6} + K_{t+1} + (1 - \delta)K_t.$$

In period 0, the economy is in the steady state associated with the parameter values $\beta = 0.90$, $\eta = 2.0$, $\gamma = 2.0$, $\alpha = 0.3$, $\delta = 0.40$, and a replacement ratio of pensions relative to net wage earnings equal to $\zeta = \frac{b_t}{(1-\tau)w_t \bar{n}_t} = 30\%$, where $\bar{n}_t$ is the average labor supply in the economy. The small constant ψ is set equal to 0.001. In period $t = 1$, the government announces a change of the replacement ratio to $\zeta = 20\%$, that becomes instantaneously effective in period 1.

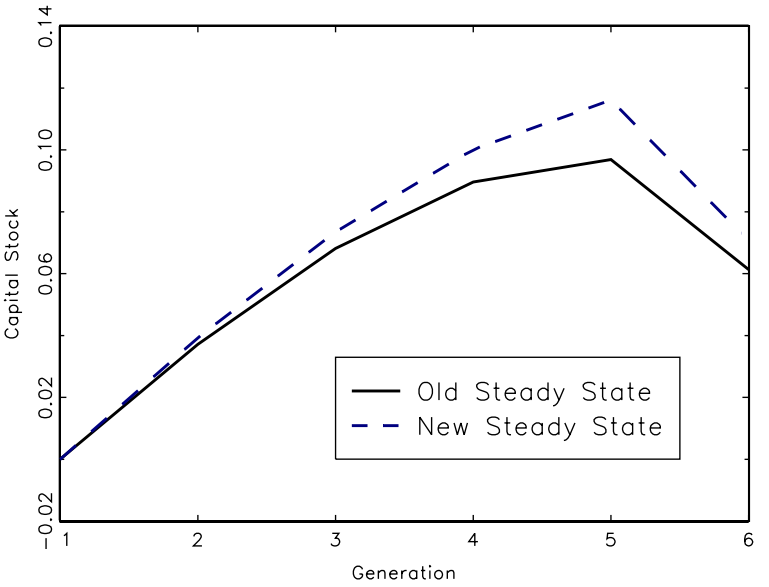

Figure 7.3: Age-capital profile in the new and in the old steady state

This Auerbach-Kotlikoff problem is solved in two steps. First, we compute the old and the new steady state using the methods described above. The age-wealth profile and the age-labor supply profile for the two steady states are displayed in the Figures 7.3 and 7.4, respectively. Notice that savings increase in the new steady state as the government reduces pensions and the agents accumulate private savings for old age. Since the government reduces pensions, it is also able to cut wage taxes in order to keep the government budget balanced. Taxes τ are reduced from 13.04% to 9.09%. Consequently, the labor supply is higher in the new steady state than in the old steady state. The aggregate capital stock and employment amount to 0.0664 (0.0590) and 0.233 (0.227) in the new (old) steady state with replacement ratio of 20% (30%), respectively.

In the third step, we need to compute the transition between the old and the new steady state. For this reason, the two steady states need to be saddle-point stable, an issue we turn to in Section 7.2, and which you are asked to show in the exercises. In order

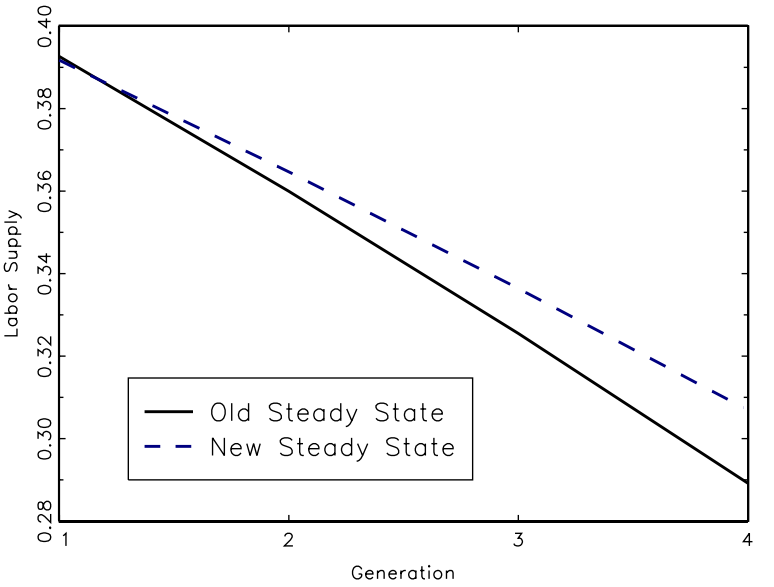

Figure 7.4: Age-labor supply profile in the new and in the old steady state

to compute the transition, we need to assume that the transition only lasts a finite number of periods. Typically, if $T + T^R$ denotes the number of generations, researchers pick a number of transition periods approximately equal to $3 \cdot (T + T^R)$, which is usually found to be sufficient in order to guarantee convergence to the new steady state. We will choose 20 model periods corresponding to 200 years. The computed dynamics of the capital stock are displayed in Figure 7.5. Obviously, the economy has converged from the old to the new steady state aggregate capital stock, from $\bar{K} = 0.0590$ to $\underline{K} = 0.0664$. In period 0, the economy is in the old steady state. All agents have chosen their next-period capital stock k_1^s, $s = 1, \ldots, 6$, assuming that there is no change in the fiscal policy. Consequently, the capital stock of an s-year old generation in period 1, k_1^s, is also equal to the capital stock of the s-year old generation in period 0, k_0^s. Accordingly, the aggregate capital stock is equal in these two periods, $\bar{K} = K_0 = K_1$. Only in period 2 does the capital stock K_t start to change. In period

20, the last period of transition, the capital stock K_{20} is equal to 0.0665 and diverges from the new steady state $\underline{K} = 0.0664$ by less than 0.5%.

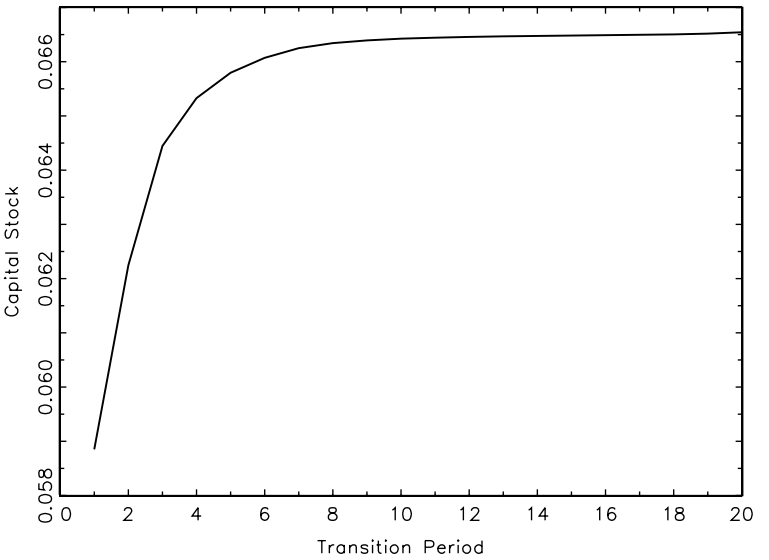

Figure 7.5: Transition from old to new steady state

The transition displayed in Figure 7.5 is computed with the help of the program `ch71AK6.g`. The computation of the dynamics is very similar to the one you got to know in the Chapters 1, 5, and 6. We first provide a guess of the dynamics of the capital stock and employment, $\{K_t, N_t\}_{t=0}^{21}$. We know that $K_0 = 0.0590$ ($N_0 = 0.227$) and $K_{21} = 0.0664$ ($N_{21} = 0.233$). As an initial guess, we simply interpolate linearly between these values. Given the time path of the aggregate state variables, we can compute wages, interest rates, pensions, and the tax rate from the first-order conditions of the firm and the balanced budget constraint. Given the sequence of factor prices, we can compute the capital stock and labor supply of the s-year old household in period t, $s = 1, \ldots, 6$ and $t = 1, \ldots, 20$.

In the program `ch71AK6.g`, we store the individual capital stock in a simple matrix denoted by `kst`:

$$\texttt{kst} = \begin{pmatrix} k_1^6 & k_1^1 & k_1^2 & k_1^3 & k_1^4 & k_1^5 \\ k_2^1 & k_2^2 & k_2^3 & k_2^4 & k_2^5 & k_2^6 \\ & & \cdots & & & \\ k_{17}^4 & k_{17}^5 & k_{17}^6 & k_{17}^1 & k_{17}^2 & k_{17}^3 \\ k_{18}^5 & k_{18}^6 & k_{18}^1 & k_{18}^2 & k_{18}^3 & k_{18}^4 \\ k_{19}^6 & k_{19}^1 & k_{19}^2 & k_{19}^3 & k_{19}^4 & k_{19}^5 \\ k_{20}^1 & k_{20}^2 & k_{20}^3 & k_{20}^4 & k_{20}^5 & k_{20}^6 \end{pmatrix}.$$

This particular form of storage is chosen in order to facilitate the programming. The distribution of capital in period 1 is simply the old steady state distribution, $(k_1^1,\ldots,k_1^6)$=(0, 0.0372, 0.0681, 0.0896, 0.0969, 0.0613), while the distribution in period 21 is given by $(k_{21}^1,\ldots,k_{21}^6)$=(0, 0.0392, 0.0734, 0.1001, 0.1161, 0.0706). Similarly, we store the values of the labor supply of each generation s in each decade t in a matrix `nst` that is arranged exactly the same way as `kst`. As soon as we have computed the matrices of the generational capital stocks and labor supplies, we simply compute the aggregate capital stock and employment as the mean of each row in the matrices `kst` and `nst`, respectively, and update our guess of $\{K_t, N_t\}_{t=0}^{21}$.

In order to compute the matrices `kst` and `nst`, we compute the optimal policy of the generation that is born in the period $t = 20, \ldots, -5$, recursively using the direct method.[9] We start in period $t = 20$. Agents that are born in period $t = 20$ know that they reach the new steady state in the next period, $t + 1 = 21$, so that their capital stock in the next period is given by $k_{21}^2 = 0.0392$. They are also born with no capital, $k_{20}^1 = 0$ and know the wage rate, the interest rate, and the tax rate in this and the next period. From the first-order condition (7.5), it is straightforward to compute the optimal labor supply n_{20}^1.

Next, we consider the generation that is born in period $t = 19$. Again, the generation is born without capital, $k_{19}^1 = 0$, and we

[9] You are asked to compute Example 7.1.1 with the help of value function iteration in the exercises.

know that $k_{21}^3 = 0.0681$. In order to compute k_{20}^2 with the help of (7.4), we need to compute c_{21}^3 with the help of the budget constraint (7.3) and the values $k_{22}^4 = 0.1001$ and $n_{21}^3 = 0.336$. Again, we can compute n_{20}^2 with the help of (7.5). We continue to proceed like this. We compute the capital-age and labor-supply age profile of the generation born in period $t = 18, \ldots, -4$. Notice that for $t \leq 15$, we do not know the capital stock at age $s = 6$ so that, again, we need to use the method described in Section 7.1.2. In particular, we start by providing an educative guess for k_t^6, $t = 15, \ldots, 5$ and iterate backwards in order to find the sequence k_{t-1}^5, k_{t-2}^4, k_{t-3}^3, k_{t-4}^2, k_{t-5}^1, and compare our value for k_{t-5}^1 with the initial capital stock 0. We update k_t^6 until k_{t-5}^1 converges to zero. For $t < 5$, we, again, provide an initial guess for k_t^6 and compute the sequence $k_{t-1}^5, \ldots, k_1^s$ and compare the value k_1^s with the capital stock of the s-year old generation in the old steady state. We update k_t^6 until these two capital stocks are equal. The solution to this problem is the following matrix:

$$\texttt{kst} = \begin{pmatrix} 0.0613 & 0.0000 & 0.0372 & 0.0681 & 0.0896 & 0.0969 \\ 0.0000 & 0.0367 & 0.0705 & 0.0958 & 0.1094 & 0.0613 \\ 0.0378 & 0.0705 & 0.0977 & 0.1133 & 0.0679 & 0.0000 \\ 0.0719 & 0.0979 & 0.1145 & 0.0695 & 0.0000 & 0.0387 \\ 0.0990 & 0.1147 & 0.0700 & 0.0000 & 0.0389 & 0.0728 \\ 0.1154 & 0.0700 & 0.0000 & 0.0390 & 0.0730 & 0.0996 \\ 0.0703 & 0.0000 & 0.0391 & 0.0732 & 0.0998 & 0.1157 \\ & & & \ldots & & \\ 0.1001 & 0.1160 & 0.0705 & 0.0000 & 0.0392 & 0.0734 \\ 0.1160 & 0.0705 & 0.000 & 0.0392 & 0.0734 & 0.1001 \\ 0.0705 & 0.0000 & 0.0392 & 0.0734 & 0.1001 & 0.1160 \\ 0.0000 & 0.0392 & 0.0734 & 0.1001 & 0.1160 & 0.0705 \end{pmatrix}$$

The computation with the direct method is fast and accurate. Notice, however, that this method is only applicable to models with perfect foresight. We will turn to models with uncertainty next. First, we consider a model with individual uncertainty, but aggregate certainty. Thereafter, we solve a model with aggregate uncertainty.

7.2 Life-cycle Economies with Individual or Aggregate Uncertainty

In the following, we will first introduce individual stochastic productivity in the standard OLG model, and, secondly, aggregate stochastic productivity. In the first section, agents have different productivity types. Different from the traditional Auerbach-Kotlikoff models, agents are subject to idiosyncratic shocks and may change their productivity types randomly. As a consequence, the direct computation of policies and transition paths is no longer feasible. As an interesting application, we are trying to explain the empirically observed wealth heterogeneity. In the second section, we introduce aggregate uncertainty and study the business cycle dynamics of the OLG model.

7.2.1 Overlapping Generations Models with Individual Uncertainty

One of the main aims of the heterogeneous-agent literature in the 1990s has been the explanation of the high concentration of wealth. In the US economy, the distribution of wealth is characterized by a Gini coefficient equal to 0.78 according to estimates by Díaz-Giménez et al. (1997). One main explanatory factor, of course, is the unequal distribution of earnings. However, when we added heterogeneous productivity into the Ramsey model in Section 5.3.2, the model failed to replicate the observed wealth concentration. In the present chapter, we add another important determinant of wealth distribution in addition to heterogeneous individual productivity: life-cycle savings. Agents accumulate wealth in order to finance consumption in old age. For this reason, we will consider an overlapping generation model in the following.[10]

Our OLG model for the study of the wealth distribution is characterized by the following features:

[10] As an alternative way to model life-cycle savings, Castañeda et al. (2003) consider the standard Ramsey model with heterogeneous productivity. In addition, they assume that agents retire and die with a certain probability, respectively. In the former case, agents receive pensions which are lower than labor income.

a) life-cycle savings,
b) uncertain lifetime,
c) uncertain earnings,
d) lump-sum pensions,

and follows HUGGETT (1996) closely. Three sectors can be depicted: households, production, and the government. Households maximize discounted life-time utility. They inherit no wealth and leave no bequests. Firms maximize profits. Output is produced with the help of labor and capital. The government provides unfunded public pensions which are financed by a tax on wage income.

Households. Every year, a generation of equal measure is born. The total measure of all generations is normalized to one. As we only study steady-state behavior, we concentrate on the study of the behavior of an individual born at the beginning of period 1. Her first period of life is period 1. A subscript t of a variable denotes the age of the generation, the measure of generation t is denoted by μ_t.

Households live a maximum of $T + T^R$ years. Lifetime is stochastic and agents face a probability s_t of surviving up to age t conditional on surviving up to age $t-1$.[11] During their first T years, agents supply labor $\bar{h}$ inelastically. After T years, retirement is mandatory. Agents maximize life-time utility:

$$E_1 \left[\sum_{t=1}^{T+T^R} \beta^{t-1} \left(\Pi_{j=1}^{t} s_j \right) u(c_t) \right], \tag{7.20}$$

where β and c_t denote the discount factor and consumption at age t, respectively. The instantaneous utility function $u(c)$ is the

[11] The basic reason to introduce stochastic survival is the gain in the model's quality. Empirically, the consumption-age profile is hump-shaped in the US. In an OLG model with stochastic earnings, agents accumulate precautionary savings and, as a consequence, the discount factor β^{-1} is higher than the real interest rate $1+r$. Therefore, if the life-time is certain, consumption increases over lifetime, even into the final years of life. If, however, agents have lower surviving probabilities in old age, consumption is hump-shaped again because future periods of life are discounted at a higher rate.

CRRA (constant relative-risk aversion) function:[12]

$$u(c) = \frac{c^{1-\eta} - 1}{1 - \eta}, \tag{7.21}$$

where η denotes the coefficient of relative risk aversion.

Heterogeneous labor earnings are introduced in a similar way as in Section 5.3. The worker's labor productivity $e(z,t)$ is stochastic and depends on his age t and an idiosyncratic labor productivity shock z. The shock z follows a Markov process and takes only a finite number n_z of possible values in the set $Z = \{z^1 = \underline{z}, \ldots, z^{n_z} = \overline{z}\}$ with $z^i < z^{i+1}$ for $i = 1, \ldots, n_z - 1$. Again, the shocks z are independent across agents so that there is no aggregate uncertainty. The labor productivity process is calibrated in detail below.

Agents are born without wealth, $k_1 = 0$, and do not leave altruistic bequests to their children. All accidental bequests are confiscated by the state and used for public consumption. Agents receive income from capital k_t and labor n_t and face a borrowing limit $k \geq \underline{k}$. The budget constraint of the working agent is given by

$$k_{t+1} = (1 + r)k_t + (1 - \tau)w\bar{h}e(z,t) - c_t, \quad t = 1, \ldots, T, \tag{7.22}$$

where r and w denote the interest rate and the wage rate per efficiency unit of labor, respectively. Wage income is taxed at rate τ.

During retirement, agents receive public pensions b irrespective of their employment history. The budget constraint of the retired agent is given by

$$k_{t+1} = (1 + r)k_t + b - c_t, \quad t = T + 1, \ldots, T + T^R. \tag{7.23}$$

[12] Different from (7.2) , we do not include a small constant ψ in the utility function because we do not consider the case of a zero income.

Production. Firms are of measure one and produce output with effective labor N and capital K. Effective labor N is paid the wage w. Capital K is hired at rate r and depreciates at rate δ. Production Y is characterized by constant returns to scale and assumed to be Cobb-Douglas:

$$Y = K^{\alpha} N^{1-\alpha}. \tag{7.24}$$

In a factor market equilibrium, factors are rewarded with their marginal product:

$$w = (1-\alpha) K^{\alpha} N^{-\alpha}, \tag{7.25}$$

$$r = \alpha K^{\alpha-1} N^{1-\alpha} - \delta. \tag{7.26}$$

Government. The government uses the revenues from taxing labor in order to finance its expenditures on social security:

$$\tau w N = \sum_{t=T+1}^{T+T^R} \mu_t b. \tag{7.27}$$

Following a change in the provision of public pensions b, the labor income tax rate τ adjusts in order to keep the government budget balanced. Government consumption G is exogenous and is financed by accidental bequests.

Stationary Equilibrium. The applied concept of equilibrium uses a recursive representation of the consumer's problem following Stokey and Lucas (1989). Let $V_t(k_t, z_t)$ be the value of the objective function of the t-year old agent with wealth k_t and idiosyncratic productivity level z_t. $V_t(k_t, z_t)$ is defined as the solution to the dynamic program:

$$V_t(k_t, z_t) = \max_{k_{t+1}, c_t} \left\{ u(c_t) + \beta s_{t+1} E\left[V_{t+1}(k_{t+1}, z_{t+1}) | k_t, z_t \right] \right\} \tag{7.28}$$

subject to (7.22) or (7.23) and $k \geq \underline{k}$. Optimal decision rules for consumption $c_t(k, z)$ and next-period capital stock $k_{t+1}(k, z)$ at age t are functions of wealth k and the idiosyncratic productivity shock z.

We further need to describe the distribution of wealth and productivity in our economy. Remember that μ_t is the mass of the t-year old agents and that the total mass of all generations is equal to one. Furthermore, let $F_t(k,z)$ denote the probability distribution of the individual states $(k,z) \in \mathcal{X}$ across age t agents. In our model, the individual capital stock k is also bounded from above as the agents cannot save more than they earn over their finite lifetimes. Let $\bar{k}$ denote the upper bound. Accordingly, the state space $\mathcal{X} = (\underline{k}, \bar{k}) \times Z$ is bounded, too. The t-year old agents with a capital stock and productivity equal or below k and z, respectively, will make up a proportion of $\mu_t F_t(k,z)$ of all agents in the economy. The distribution $F_t(k,z)$ has the property that $F_t(\bar{k},\bar{z}) \equiv 1.0$ for all t. Furthermore, the initial distribution $F_1(k,z)$ is given by the exogenous initial distribution of labor endowment $e(z,1)$, as all agents are born with zero assets. The distribution of individual states across agents is given by the following recursive equation for all $(k_{t+1}, z_{t+1}) \in \mathcal{X}$ and $t = 1, \ldots, T + T^R - 1$:

$$F_{t+1}(k_{t+1}, z_{t+1}) = \sum_{z_t \in Z} \pi\left(z_{t+1}|z_t\right) \cdot F_t\left(\left(k_{t+1}\right)^{-1}\left(k_{t+1}, z_t\right), z_t\right), \tag{7.29}$$

where $\pi(z_{t+1}|z_t)$ denotes the exogenously given transition probability from productivity state z_t to z_{t+1} and $\left(k_{t+1}\right)^{-1}\left(k_{t+1}, z_t\right)$ denotes the inverse of the function for the optimal next-period capital stock $k_{t+1}(k_t, z_t)$ with respect to its first argument k_t.[13] Obviously, our concept of a stationary distribution corresponds closely to the one introduced in Chapter 5 for the infinite-horizon Ramsey model.

We will consider a stationary equilibrium where factor prices and aggregate capital and labor are constant and the distribution of wealth is stationary. Let $(f_1, f_2, \ldots, f_{T+T^R})$ denote the density functions of (k,z) in generation $t = 1, 2, \ldots, T+T^R$. The following properties characterize the stationary equilibrium:

a) Individual and aggregate behavior are consistent:

[13] Please see also footnote 5 in Chapter 5.

$$N = \sum_{t=1}^{T} \sum_{z \in Z} \mu_t \, e(z,t) \, \bar{h} \int_{\underline{k}}^{\bar{k}} f_t(k,z) \, dk,$$

$$K = \sum_{t=1}^{T+T^R} \sum_{z \in Z} \mu_t \int_{\underline{k}}^{\bar{k}} k \, f_t(k,z) \, dk,$$

$$C = \sum_{t=1}^{T+T^R} \mu_t \sum_{z \in Z} \int_{\underline{k}}^{\bar{k}} c_t(k,z) \, f_t(k,z) \, dk.$$

b) Relative prices $\{w, r\}$ solve the firm's optimization problem by satisfying (7.25) and (7.26).
c) Given relative prices $\{w, r\}$ and the government policy b, the individual policy rules $c_t(\cdot)$ and $k_{t+1}(\cdot)$ solve the consumer's dynamic program (7.28).
d) The government budget (7.27) is balanced.
e) Government consumption G equals accidental bequests.
f) The goods market clears:

$$K^{\alpha} N^{1-\alpha} = C + G + \delta K.$$

g) The distributions F_t (and the corresponding density functions f_t), $t = 1, \ldots, T + T^R - 1$, are consistent with individual behavior and follow (7.29).

Calibration. Periods correspond to years. We assume that agents are born at real lifetime age 20 which corresponds to $t = 1$. Agents work $T = 40$ years corresponding to a real lifetime age of 60. They live a maximum life of 60 years ($T^R = 20$) so that agents do not become older than real lifetime age 80.

The sequence of conditional survival probabilities $\{\psi_t\}_{t=1}^{59}$ is set equal to the Social Security Administration's survival probabilities for men aged 20-78 for the year 1994.[14] The survival probabilities decrease with age and are presented in Figure 7.6. ψ_{60} is set equal to zero.

The model parameters are presented in Table 7.1. If not mentioned otherwise, the model parameters are taken from HUGGETT

[14] We thank Mark Huggett and Gustavo Ventura for providing us with the data.

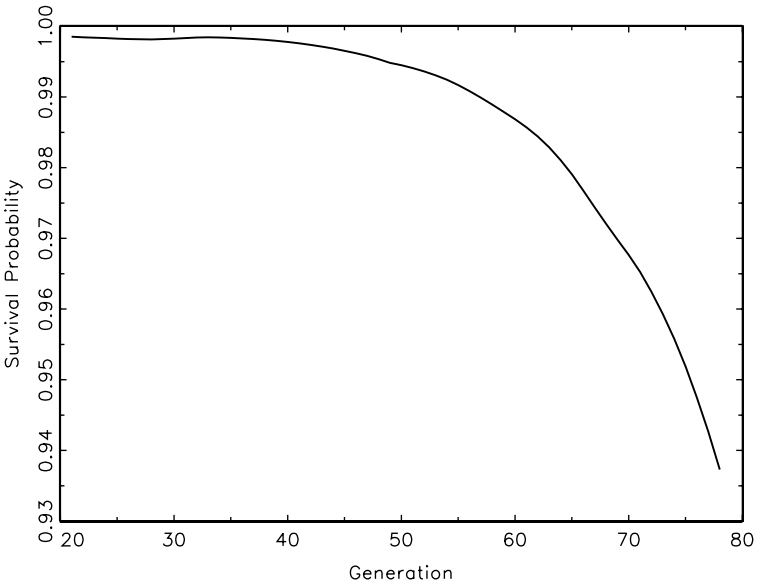

Figure 7.6: Survival probabilities

(1996). The discount rate β is set equal to 1.011. Notice again that, different from infinite-lifetime models like the ones in previous chapters, the discount factor β need not be smaller than one in finite-lifetime models. The credit limit is set at $\underline{k} = 0$. Huggett uses a coefficient of relative risk aversion equal to $\eta = 1.5$. The capital share of output α and the depreciation rate of capital δ are set equal to 0.36 and 0.06, respectively.

Table 7.1

Preferences	Production
$\beta = 1.011$	$\alpha = 0.36$, $\delta = 0.06$
$\eta = 1.5$	$\bar{h} = 0.30$

The labor endowment process is given by $e(z,t) = e^{z_t + \bar{y}_t}$, where $\bar{y}_t$ is the mean log-normal income of the t-year old. The mean efficiency index $\bar{y}_t$ of the t-year-old worker is taken from HANSEN

(1993), and interpolated to in-between years. As a consequence, the model is able to replicate the cross-section age distribution of earnings of the US economy. Following İMROHOROĞLU ET AL. (1995), we normalize the average efficiency index to one. The mean efficiency of the t-year old agents $e^{\bar{y}_t}$ is displayed in Figure 7.7. Notice that the age-productivity profile is hump-shaped and earnings peak at age 50.

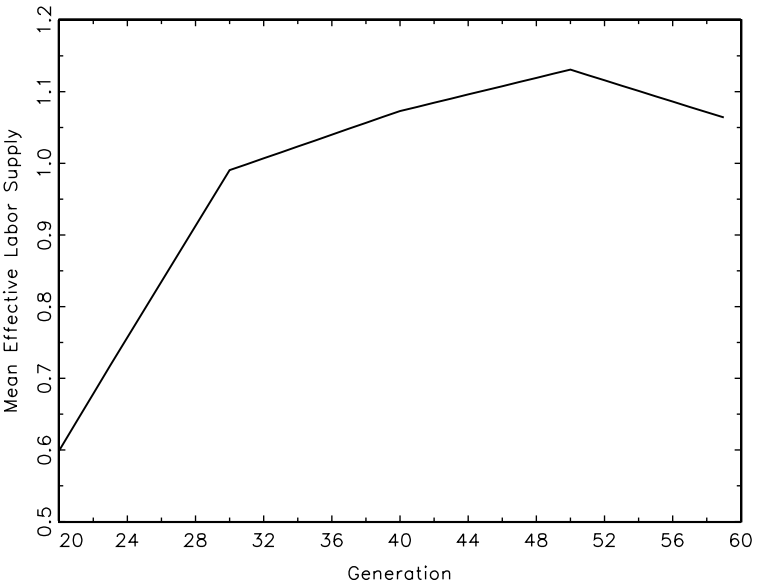

Figure 7.7: Age-productivity profile

The idiosyncratic productivity shock z_t follows a Markov process given by:

$$z_t = \rho z_{t-1} + \epsilon_t, \tag{7.31}$$

where $\epsilon_t \sim N(0, \sigma_\epsilon)$. Huggett uses $\rho = 0.96$ and $\sigma_\epsilon = 0.045$. Furthermore, we follow HUGGETT (1996) and choose a log-normal distribution of earnings for the 20-year old with $\sigma_{y_1} = 0.38$ and mean $\overline{y_1}$. As the log endowment of the initial generation of agents is normally distributed, the log efficiency of subsequent agents will continue to be normally distributed. This is a useful property of

the earnings process, which has often been described as log-normal in the literature.

We discretize the state space Z using $n_z = 9$ values. The states z are equally spaced and range from $-2\sigma_{y_1}$ to $2\sigma_{y_1}$. The probability of having productivity shock z_1 in the first period of life is computed by integrating the area under the normal distribution. The transition probabilities are computed using the Algorithm 9.2.1. As a consequence, the efficiency index $e(z,t)$ follows a finite Markov chain. Furthermore, we set the shift length equal to $\bar{h} = 0.3$.

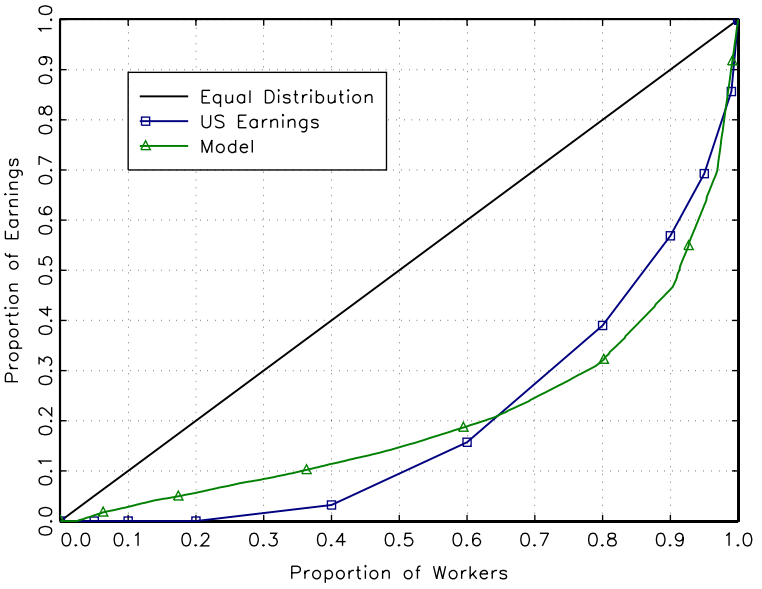

Figure 7.8: Lorenz curve of US and model earnings

The earnings process is exogenous in our model. The Lorenz curve for the earnings in our model and for the US are displayed in Figure 7.8. The inequality of earnings for the model economy and for the US are similar. In our model, the lowest quintile of earners has a higher labor income share than observed empirically, while the top quintile earns a higher income share than those in the

US.[15] In our model, the Gini coefficient of labor income is equal to 0.413 and matches empirical values as reported in Chapter 5 quite closely.

Finally, the government provides pensions b. The replacement ratio of pensions relative to average net wages is set equal to 30%.

Computation. The solution algorithm follows Algorithm 7.1.1 closely and consists of the following steps:

Step 1: Parameterize the model and compute aggregate employment N.
Step 2: Make initial guesses of the steady state values of the aggregate capital stock K and the social security contribution rate τ.
Step 3: Compute the values w, r, and b, which solve the firm's Euler equation and the government budget.
Step 4: Compute the household's decision functions by backward induction.
Step 5: Compute the optimal path for consumption and savings for the new-born generation by forward induction given the initial capital stock $k_1 = 0$.
Step 6: Compute the aggregate capital stock K.
Step 7: Update K and N and return to step 3 until convergence.

The algorithm is implemented in the program `ch72.g`. In step 4, a finite-time dynamic programming problem is to be solved. Again, we solve this problem with value function iteration with linear interpolation between grid points. We choose an equispaced grid with $\underline{k} = 0$, $\bar{k} = 40$, and $n_k = 601$. Associated with every optimal next period capital stock $k_{t+1}(k_t, z_t)$ and $k_{t+1}(k_t)$ is an optimal consumption policy $c_t(k_t, z_t) = (1+r)k_t + b - k_{t+1}(k_t, z_t)$ for the retired agent and $c_t(k_t, z_t) = (1+r)k_t + (1-\tau)w\bar{h}e(z,t) - k_{t+1}(k_t, z_t)$ for the working agent, respectively.

In step 5, we compute the endogenous wealth distribution in every generation over an equispaced grid of the asset space $[\underline{k}, \bar{k}] \times Z$ with $2n_k n_z$ points. We start with the newborn generation at $t = 1$ with zero wealth. Furthermore, we know the distribution

[15] The source of the US data is described in the previous chapter.

of the idiosyncratic productivity at age 1. Given the distribution of the capital stock and the productivity at age t, we can compute the distribution at age $t+1$ using the optimal decision functions of the agents, $k_{t+1}(k_t, z_t)$, and the transition probabilities for the idiosyncratic productivities. We continue to compute the distribution for $t = 2, \ldots, 60$. Notice that, different from the Ramsey model, we can compute the stationary distribution in the OLG model with one iteration only.

Finally, in step 7, we use extrapolation to stabilize the sequence, i.e. let K^i and K^* denote the starting value in the i-th iteration and the computed endogenous value of the capital stock, respectively, then $K^{i+1} = \phi K^i + (1-\phi)K^*$. We set ϕ equal to 0.8.

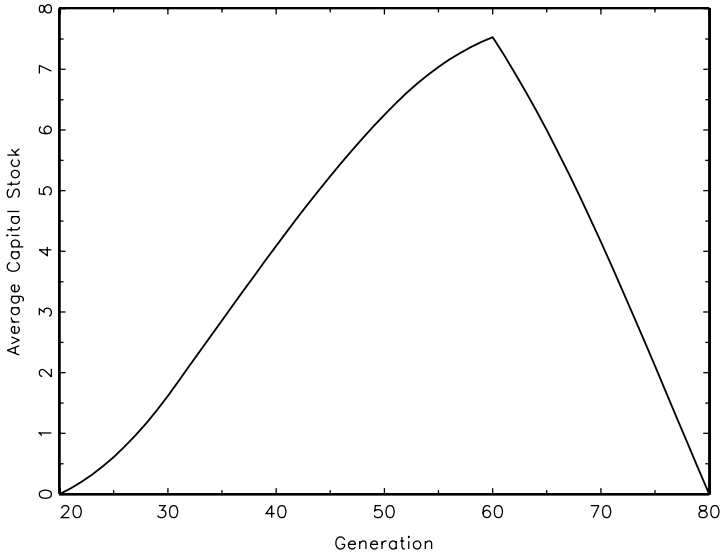

Figure 7.9: Age-wealth profile

Results. Figure 7.9 displays the average wealth over the lifetime. As you know by now, the hump-shape of the profile is typical for the life-cycle model. Agents build up savings during working life, and assets start to fall after retirement. Therefore, wealth heterogeneity is higher in the OLG model than in the standard Ramsey model as agents have different savings at different ages.

Furthermore, we have many agents who are liquidity constraint and only have zero wealth, especially the young agents with low productivity. In the Ramsey model of the previous chapter, all agents hold strictly positive wealth and wealth heterogeneity is also much lower for this reason.

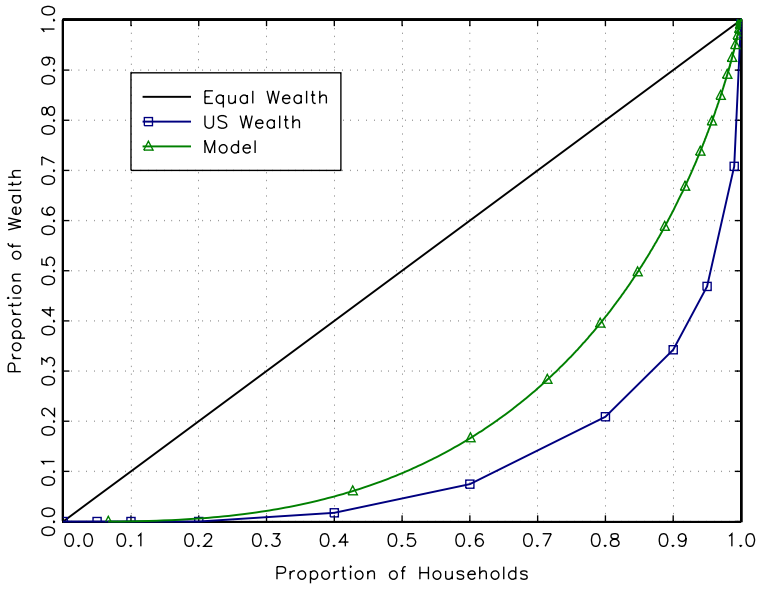

Figure 7.10: Lorenz curve of US and model wealth

Average wealth in the economy amounts to $K = 3.94$. For our choice of the earnings process, aggregate effective employment is equal to $N = 0.386$ so that the interest rate equals $r = 2.15\%$. The equilibrium social security contribution rate amounts to $\tau = 9.75\%$. The wealth distribution of the model economy and the US economy are displayed in Figure 7.10. The model economy is characterized by a much more equal wealth distribution than the US economy. The Gini coefficient of the wealth distribution in our model is equal to 0.587 and is below the one for the US economy (that is approximately equal to 0.78). However, wealth is much more concentrated than earnings on the one hand, and, on the other hand, the model generates more wealth heterogeneity than

the Ramsey model with heterogeneous productivity presented in Section 5.3.

There are numerous reasons why the endogenous wealth heterogeneity of our model is smaller than observed empirically:

a) Pensions are not related to the earnings history of the recipient. If the earnings-rich agents get higher pensions, one might suppose that wealth heterogeneity would also be higher. However, as earnings-poor agents also know they only receive little pensions, they will save more for precautionary reasons.[16]
b) We neglect any asset-based means tests of social security. Hubbard et al. (1995) show that, in the presence of social insurance programs with means tests, low-income households are likely to hold virtually no wealth across lifetime. Unemployment and asset-based social insurance would imply a much higher proportion of agents with zero or near-zero wealth.
c) Furthermore, agents are not unemployed. Heer (2003) studies a life-cycle economy with endogenous search unemployment. Working agents may loose their job at an exogenous rate; higher search effort increases the job finding probability, but searching for a job also causes a disutility for the agent. Heer shows that the replacement rate of unemployment insurance has only a very small effect on wealth heterogeneity. Even though income is distributed from the income-rich agents to the income-poor workers with the help of unemployment insurance, higher unemployment insurance also increases endogenous unemployment so that the number of unemployment recipients increases. As a consequence, the wealth Gini coefficient changes by less than one percentage point if the replacement ratio of unemployment insurance increases from 0% to 50% or even a 100%; for a replacement ratio exceeding 70%, wealth heterogeneity even starts to increase again.
d) We neglect bequests. For example, Kessler and Masson (1989), considering France, find that only 36% of the house-

[16] In our own research, we only encountered applications where the introduction of earnings-related benefits decreased wealth heterogeneity as measured by the Gini coefficient.

holds receive any inheritances and those who do are about 2.4 times richer than the representative household. Heer (2001b) considers an OLG model where parents leave altruistic and accidental bequests to their children. He, however, finds, that bequests are able to explain only a small fraction of observed wealth heterogeneity. The main reasons are that i) poor agents may also receive bequests and ii) agents who expect a high inheritance in the future also spend more on consumption. Importantly, however, Heer (2001b) only considers intergenerational transfers of physical wealth, but not transfers of human wealth. Rich parents may have rich kids because they may invest in their college education, for example. Loury (1981) analyzes parental human capital investment in their offspring. The allocation of training and hence the earnings of the children depend on the distribution of earnings among the parents. Becker and Tomes (1979) present a model framework comprising both human and non-human capital transfers from parents to children. The introduction of human capital transfers in an OLG model in order to explain the observed wealth heterogeneity is a promising question for future research.

e) In our model, agents are not allowed to borrow against anticipated bequests implying a credit limit $k \geq 0$. For lower binding constraints, $k < 0$, wealth heterogeneity increases as demonstrated by Huggett (1996). In particular, the proportion of agents holding zero and negative assets increases.

Accounting for these features in our model is likely to result in an increase of wealth inequality for agents characterized by low to high asset holdings. However, we are sceptical as to whether it proves successful in reproducing the observed wealth concentration among the very rich. As one of the very few exceptions to these modelling choices (known to us),[17] Quadrini (2000) presents a promising approach in order to explain the high con-

[17] Quadrini and Ríos-Rull (1997) present a review of recent studies of wealth heterogeneity in computable general equilibrium models with uninsurable idiosyncratic exogenous shocks to earnings, including business ownership, higher rates of return on high asset levels, and changes in health and marital status, among others.

centration of wealth among the very rich agents. He introduces entrepreneurship into a dynamic general equilibrium model.

7.2.2 Aggregate Uncertainty

In this section, we introduce aggregate uncertainty in the standard OLG model. We will illustrate the numerical and analytical methods with the help of a 60-period OLG model that is described in Example 7.2.1. The model is a simple extension of Example 7.1.1. In particular, we rather consider 60 than 6 different generations and we add a technology shock ϵ_t to production. The (logarithmic) aggregate technology level follows the AR(1) process:

$$\ln Z_t = \rho \ln Z_{t-1} + \epsilon_t, \tag{7.32}$$

where ϵ_t is i.i.d., $\epsilon_t \sim N(0, \sigma^2)$. Production, therefore, is given by

$$Y_t = Z_t K_t^{\alpha} N_t^{1-\alpha}.$$

In addition, the household forms rational expectations about future income and future prices and maximizes expected life-time utility. Besides, the model is identical to the one described in Example 7.1.1.

Example 7.2.1

60-period Overlapping Generations Model with Aggregate Uncertainty.

Households live 60 periods. Each generation is of measure 1/60. The first 40 periods, they work, the last 20 periods, they are retired and receive pensions. Households maximize expected life-time utility at age 1 in period t:

$$E_t \sum_{s=1}^{60} \beta^{s-1} u(c^s_{t+s-1}, l^s_{t+s-1}).$$

Instantaneous utility is a function of both consumption and leisure:

$$u(c, l) = \frac{(cl^{\gamma})^{1-\eta}}{1-\eta}.$$

The working agent of age s faces the following budget constraint in period t:

$$k_{t+1}^{s+1} = (1 + r_t)k_t^s + (1 - \tau_t)w_t n_t^s - c_t^s, \quad s = 1, \ldots, 40.$$

The budget constraint of the retired worker is given by

$$k_{t+1}^{s+1} = (1 + r_t)k_t^s + b - c_t^s, \quad s = 41, \ldots, 60,$$

with $k_t^{61} \equiv 0$ and $l_t^{51} = l_t^{52} = \ldots = l_t^{60} \equiv 1.0$. Pensions b are constant.

Production Y_t is characterized by constant returns to scale and assumed to be Cobb-Douglas:

$$Y_t = Z_t F(K_t, N_t) = Z_t K_t^{\alpha} N_t^{1-\alpha},$$

where $\ln Z_t$ follows an AR(1) process:

$$\ln Z_t = \rho \ln Z_{t-1} + \epsilon_t,$$

and ϵ_t is i.i.d., $\epsilon_t \sim N(0, \sigma^2)$.

In a factor market equilibrium, factors are rewarded with their marginal product:

$$\begin{aligned} w_t &= (1 - \alpha) Z_t K_t^{\alpha} N_t^{-\alpha}, \\ r_t &= \alpha Z_t K_t^{\alpha - 1} N_t^{1-\alpha} - \delta. \end{aligned}$$

Furthermore, the government budget is balanced in every period t:

$$\tau_t w_t N_t = \frac{20}{60} b.$$

In equilibrium, individual and aggregate behavior are consistent:

$$\begin{aligned} N_t &= \sum_{s=1}^{40} \frac{n_t^s}{60}, \\ K_t &= \sum_{s=1}^{60} \frac{k_t^s}{60}, \end{aligned}$$

and the goods market clears:

$$Z_t K_t^{\alpha} N_t^{1-\alpha} = \sum_{s=1}^{60} \frac{c_t^s}{60} + K_{t+1} - (1 - \delta) K_t.$$

The parameter values are chosen as follows: $\beta = 0.99$, $\eta = 2.0$, $\gamma = 2.0$, $\alpha = 0.3$, $\delta = 0.04$ and a non-stochastic replacement ratio of pensions relative to net wage earnings equal to $\zeta = \frac{b}{(1-\tau)w\bar{n}} = 30\%$, where $\bar{n}$ is the average labor supply in the non-stochastic steady state of the economy. The parameters of the AR(1) process for the technology are set equal to $\rho = 0.814$ and $\sigma = 0.024$. These parameters correspond to annual frequencies by a quarterly AR(1) process for the Solow residual with parameter 0.95 and 0.00763, which are the parameters in Prescott (1986). ———

For the economy described in Example 7.2.1, we can compute the non-stochastic steady state with the help of the methods described in Section 7.1. The non-stochastic steady state is characterized by a constant technology level, $Z_t = Z = 1$. Furthermore, all individual and aggregate variables are constant, too, and are denoted by letters without time indices. For example, k^s and K denote the non-stochastic steady-state capital stock of the individual at age s and the non-stochastic steady-state aggregate capital stock, respectively. For our calibration, we compute the following economy-wide values: $K = 1.856$, $N = 0.2293$, $b = 0.1175$, $\tau = 13.04$, $w = 1.311$, $r = 2.938\%$.

Log-linearization. In Section 7.1, we analyzed the transition dynamics of a 6-period OLG model. We refrained from showing that the economy displays saddle-point stability, even though our analysis requires determinacy and stability. In the following, we will show that, indeed, the model in Example 7.2.1 is stable. Therefore, we first need to log-linearize the equations characterizing the economy around the non-stochastic steady state applying the methods of Chapter 2. These equations, in particular, consist of the first-order equations of the households and the firm, the budget constraint of the households, and the government budget constraint.

The first-order conditions of the households for $s =, 1 \ldots, 60$ in period t are analogous to the equations (7.4) and (7.5) for labor supply and next-period capital stock, respectively:

$$u_l(c_t^s, l_t^s) = \gamma \left(c_t^s\right)^{1-\eta} \left(1 - n_t^s\right)^{\gamma(1-\eta)-1} = \lambda_t^s (1 - \tau_t) w_t,$$

$$\lambda_t^s = (c_t^s)^{-\eta} (l_t^s)^{\gamma(1-\eta)},$$
$$\frac{1}{\beta} = E_t \left\{ \frac{\lambda_{t+1}}{\lambda_t} [1 + r_{t+1}] \right\}.$$

Log-linearization of these equations around the non-stochastic steady state results in:

$$(1-\eta)\widehat{c_t^s} + (1-\gamma(1-\eta)) \frac{n^s}{1-n^s} \widehat{n_t^s} = \widehat{\lambda_t^s} - \frac{\tau}{1-\tau} \widehat{\tau}_t + \widehat{w}_t, \quad s = 1, \ldots, 40, \quad (7.34)$$

$$\widehat{\lambda_t^s} = -\eta \widehat{c_t^s} - \gamma(1-\eta) \frac{n^s}{1-n^s} \widehat{n_t^s}, \quad s = 1, \ldots, 60, \quad (7.35)$$

$$\widehat{\lambda_t^s} = E_t \widehat{\lambda_{t+1}^{s+1}} + \frac{r}{1+r} E_t \widehat{r_{t+1}}, \quad s = 1, \ldots, 59. \quad (7.36)$$

Furthermore, we need to log-linearize the working household's budget constraint (7.3) around the steady state for the one-year old with $k^1 \equiv 0$:

$$k^2 \widehat{k_{t+1}^2} = -\tau w n^1 \widehat{\tau}_t + (1-\tau) w n^1 \widehat{w}_t + (1-\tau) w n^1 \widehat{n_t^1} - c^1 \widehat{c_t^1}, \quad (7.37)$$

and for $s = 2, \ldots, 40$:

$$k^{s+1} \widehat{k_{t+1}^{s+1}} = (1+r) k^s \widehat{k_t^s} + r k^s \widehat{r}_t - \tau w n^s \widehat{\tau}_t + (1-\tau) w n^s \widehat{w}_t + (1-\tau) w n^s \widehat{n_t^s} - c^s \widehat{c_t^s}. \quad (7.38)$$

Log-linearization of the retired agent's budget constraint (7.6) around the non-stochastic steady state results in:

$$k^{s+1} \widehat{k_{t+1}^{s+1}} = (1+r) k^s \widehat{k_t^s} + r k^s \widehat{r}_t - c^s \widehat{c_t^s}, \quad s = 41, \ldots, 59. \quad (7.39)$$

Finally, consumption at age $s = 60$ is given by:

$$c^{60} \widehat{c_t^{60}} = (1+r) k^{60} \widehat{k_t^{60}} + r k^{60} \widehat{r}_t. \quad (7.40)$$

Summarizing, we have 60 controls c_t^s, $s = 1, \ldots, 60$, 40 controls n_t^s, $s = 1, \ldots, 40$, 60 costates λ_t^s, $s = 1, \ldots, 60$, and 59 predetermined variables k_t^s, $s = 2, \ldots, 60$. We also have $60+40+60+59 = 219$ equations. We have three further endogenous variables w_t, r_t, and τ_t.

The wage rate is given by the marginal product of labor:

$$w_t = (1-\alpha)Z_t K_t^{\alpha} N_t^{-\alpha} = (1-\alpha)Z_t \left(\sum_{s=2}^{60} \frac{k_t^s}{60}\right)^{\alpha} \left(\sum_{s=1}^{40} \frac{n_t^s}{60}\right)^{-\alpha}.$$

Log-linearization results in:

$$\widehat{w}_t = \hat{Z}_t + \alpha \sum_{s=2}^{60} \frac{k^s}{K}\frac{1}{60}\widehat{k_t^s} - \alpha \sum_{s=1}^{40} \frac{n^s}{N}\frac{1}{60}\widehat{n_t^s}. \tag{7.41}$$

Similarly, we derive the percentage deviation of the interest rate, $\widehat{r_t}$, from its non-stochastic steady state $r = \alpha K^{\alpha-1}N^{1-\alpha}$:

$$\widehat{r_t} = \hat{Z}_t - (1-\alpha) \sum_{s=2}^{60} \frac{k^s}{K}\frac{1}{60}\widehat{k_t^s} + (1-\alpha) \sum_{s=1}^{40} \frac{n^s}{N}\frac{1}{60}\widehat{n_t^s}. \tag{7.42}$$

The government budget $\tau wN = (1/3)b$ is the remaining equation that we need to approximate locally around the non-stochastic steady state:

$$\widehat{\tau_t} = -\widehat{w_t} - \sum_{s=1}^{40} \frac{n^s}{N}\frac{1}{60}\widehat{n_t^s}. \tag{7.43}$$

(7.41)-(7.43) constitute three further equations in the three endogenous variables w_t, r_t, and τ_t. Finally, we have the law of motion for the exogenous state variable Z_t:

$$\hat{Z}_{t-1} = \rho \hat{Z}_{t+1} + \epsilon_t \tag{7.44}$$

Local Stability of the Non-stochastic Steady State. Our log-linearization of the 60-period Auerbach-Kotlikoff model in Example 7.2.1 is described by equations (7.34)-(7.44). In order to conduct a local stability analysis, it is convenient to express our system of stochastic difference equations in the form (2.36) of Chapter 2:

$$C_u \mathbf{u}_t = C_{x\lambda} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} + C_z \mathbf{z}_t, \tag{7.45a}$$

$$D_{x\lambda} E_t \begin{bmatrix} \mathbf{x}_{t+1}, \\ \boldsymbol{\lambda}_{t+1} \end{bmatrix} + F_{x\lambda} \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} = D_u E_t \mathbf{u}_{t+1} + F_u \mathbf{u}_t + D_z E_t \mathbf{z}_{t+1} + F_z \mathbf{z}_t. \tag{7.45b}$$

Therefore, we define:

$$\mathbf{u}_t = \begin{bmatrix} \widehat{c_t^1} \\ \widehat{c_t^2} \\ \vdots \\ \widehat{c_t^{60}} \\ \widehat{n_t^1} \\ \widehat{n_t^2} \\ \vdots \\ \widehat{n_t^{40}} \\ \widehat{r_t} \\ \widehat{w_t} \\ \widehat{\tau_t} \end{bmatrix}, \quad \mathbf{x}_t = \begin{bmatrix} \widehat{k_t^2} \\ \widehat{k_t^3} \\ \vdots \\ \widehat{k_t^{60}} \end{bmatrix}, \quad \boldsymbol{\lambda}_t = \begin{bmatrix} \lambda_t^1 \\ \lambda_t^2 \\ \vdots \\ \lambda_t^{59} \end{bmatrix}, \quad \mathbf{z}_t = \hat{Z}_t.$$

Notice that λ_t^{60} is not a costate variable because it is implied by (7.40) and (7.33), and, for this reason, rather constitutes a control variable than a costate. If we included λ^{60} in $\mathbf{u}_t$, the matrix $D_{x\lambda} - D_u C_u^{-1} C_{x\lambda}$ would not be invertible!

In the contemporary equation (7.45a), the first 40 equations represent the first-order condition of the household with respect to labor, (7.34), the next 59 equations represent the equality of the costate variable λ_t^s, $s = 1, \ldots, 59$, and the marginal utility of consumption at age s, (7.35), the next equation is given by the consumption in period 60, (7.40), and the last three equations are the equations (7.41), (7.42) and (7.43) for the wage rate, the interest rate, and the government budget, respectively. For the matrix C_u, we get the following convenient partition:

$$C_u = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{21} & A_{22} & A_{23} \end{bmatrix},$$

with the 40×60 submatrix A_{11} and the 40×40 submatrix A_{12}:

$$A_{11} = \begin{bmatrix} (1-\eta) & 0 & \dots & 0 \\ 0 & (1-\eta) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (1-\eta) \end{bmatrix},$$

$$A_{12} = \begin{bmatrix} (1-\gamma(1-\eta))\frac{n^s}{1-n^s} & 0 & \dots \\ 0 & (1-\gamma(1-\eta))\frac{n^s}{1-n^s} & \dots \\ \vdots & \vdots & \ddots \end{bmatrix},$$

and the 40×3 submatrix A_{13}:

$$A_{13} = \begin{bmatrix} 0 & -1 & \frac{\tau}{1-\tau} \\ 0 & -1 & \frac{\tau}{1-\tau} \\ & \vdots & \\ 0 & -1 & \frac{\tau}{1-\tau} \end{bmatrix}.$$

The 60×60 submatix A_{21} and the 60×40 submatrix A_{22} together with the 60×3 submatrix A_{23} represent (7.35) and (7.40):

$$A_{21} = \begin{bmatrix} -\eta & 0 & \dots & 0 \\ 0 & -\eta & \dots & 0 \\ & \vdots & & \\ 0 & 0 & \dots & c^{60} \end{bmatrix},$$

$$A_{22} = \begin{bmatrix} -\gamma(1-\eta)\frac{n^s}{1-n^s} & 0 & \dots & 0 \\ 0 & -\gamma(1-\eta)\frac{n^s}{1-n^s} & \dots & 0 \\ & \vdots & & \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_{23} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ & \vdots & \\ -rk^{60} & 0 & 0 \end{bmatrix}.$$

The last three rows are given by the 3×60 submatrix $A_{31} = \mathbf{0}$, the 3×40 submatrix A_{32} and the 3×3 submatrix A_{33}:

$$A_{32} = \begin{bmatrix} \alpha\frac{n^1}{N}\frac{1}{60} & \alpha\frac{n^2}{N}\frac{1}{60} & \dots & \alpha\frac{n^{40}}{N}\frac{1}{60} \\ -(1-\alpha)\frac{n^1}{N}\frac{1}{60} & -(1-\alpha)\frac{n^2}{N}\frac{1}{60} & \dots & -(1-\alpha)\frac{n^{40}}{N}\frac{1}{60} \\ \frac{n^1}{N}\frac{1}{60} & \frac{n^2}{N}\frac{1}{60} & \dots & \frac{n^{40}}{N}\frac{1}{60} \end{bmatrix},$$

$$A_{33} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

The matrix $C_{x\lambda}$ is presented by:

$$\begin{bmatrix} 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ & & \vdots & & & & & \\ 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ & & \vdots & & & & & \\ 0 & 0 & \dots & (1+r)k^{60} & 0 & 0 & \dots & 0 \\ \alpha\frac{k^2}{K}\frac{1}{60} & \alpha\frac{k^3}{K}\frac{1}{60} & \dots & \alpha\frac{k^{60}}{K}\frac{1}{60} & 0 & 0 & \dots & 0 \\ -(1-\alpha)\frac{k^2}{K}\frac{1}{60} & -(1-\alpha)\frac{k^3}{K}\frac{1}{60} & \dots & -(1-\alpha)\frac{k^{60}}{K}\frac{1}{60} & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \end{bmatrix},$$

and, finally, the remaining matrix C_z from (7.45a) is equal to:

$$C_z = (0, 0, \dots, 0, 0, \dots, 0, 1, 1, 0)^T.$$

The dynamic equation system (7.45b) is characterized by the matrices $D_{x\lambda}$, $D_{x\lambda}$, D_u, F_u, D_z, and F_z. The first 59 rows of these matrices represent the first-order conditions of the household with respect to k_{t+1}^{s+1}, $s = 1, \dots, 59$, as described by (7.36). In the 59th equation, we have replaced $\widehat{\lambda_{t+1}^{60}}$ by the percentage deviation of the marginal utility of consumption of the 60-year old, $-\eta\widehat{c_{t+1}^{60}}$. Again, this is necessary as λ_t^{60} is a control rather than a costate variable in the finite-lifetime model. The remaining 59 equations are the budget constraints of the household at age $s = 1, \dots, 59$:

$$D_{x\lambda} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & -1 & \dots & 0 \\ & & & & & \vdots & & & & & \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ k^2 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & k^3 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ & & & & & \vdots & & & & & \\ 0 & 0 & \dots & 0 & k^{42} & \dots & 0 & 0 & 0 & \dots & 0 \\ & & & & & \vdots & & & & & \end{bmatrix},$$

$$F_{x\lambda} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ & & & & \vdots & & & & & & \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ -(1+r)k^2 & 0 & \dots & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ & & & & \vdots & & & & & & \\ 0 & 0 & \dots & -(1+r)k^{41} & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ & & & & \vdots & & & & & & \end{bmatrix},$$

$$D_u = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & \frac{r}{1+r} & 0 & 0 \\ & & & & & \vdots & & & & & \\ 0 & 0 & \dots & -\eta & 0 & 0 & \dots & 0 & \frac{r}{1+r} & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ & & & & & \vdots & & & & & \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & 0 \\ & & & & & \vdots & & & & & \end{bmatrix},$$

and $D_z = F_z = \mathbf{0}$. For expositional reasons, we decompose the matrix F_u into convenient submatrices:

$$F_u = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{11} & B_{12} & B_{13} \end{bmatrix},$$

with the 59×60 submatrix $B_{11} = \mathbf{0}$, 59×40 submatrix $B_{12} = \mathbf{0}$, and 59×3 submatrix $B_{13} = \mathbf{0}$. The submatrix B_{21} of dimension 59×60 is given by:

$$B_{21} = \begin{bmatrix} -c^1 & 0 & \dots & 0 & 0 \\ 0 & -c^2 & \dots & 0 & 0 \\ & & \vdots & & \\ 0 & 0 & \dots & -c^{59} & 0 \end{bmatrix}.$$

The submatrix B_{21} of dimension 59×40 describes the coefficients of the terms $\hat{n}_t$:

$$B_{22} = \begin{bmatrix} (1-\tau)wn^1 & 0 & \dots & 0 \\ 0 & (1-\tau)wn^2 & \dots & 0 \\ & & \vdots & \\ 0 & 0 & \dots & (1-\tau)wn^{40} \\ 0 & 0 & \dots & 0 \\ & & \vdots & \\ 0 & 0 & \dots & 0 \end{bmatrix}.$$

Finally, the 59×3 submatrix B_{23} is as follows:

$$B_{23} = \begin{bmatrix} 0 & (1-\tau)wn^1 & -\tau wn^1 \\ rk^2 & (1-\tau)wn^2 & -\tau wn^2 \\ rk^3 & (1-\tau)wn^3 & -\tau wn^3 \\ & \vdots & \\ rk^{40} & (1-\tau)wn^{40} & -\tau wn^{40} \\ rk^{41} & 0 & 0 \\ & \vdots & \\ rk^{59} & 0 & 0 \end{bmatrix}.$$

This system of equations can be solved with the methods introduced in Chapter 2. In particular, we can reduce the system to the equation systems (2.40):

$$\begin{aligned} E_t \begin{bmatrix} \mathbf{x}_{t+1} \\ \boldsymbol{\lambda}_{t+1} \end{bmatrix} &= W \begin{bmatrix} \mathbf{x}_t \\ \boldsymbol{\lambda}_t \end{bmatrix} + R\mathbf{z}_t, \\ W &= -\left(D_{x\lambda} - D_u C_u^{-1} C_{x\lambda}\right)^{-1} \left(F_{x\lambda} - F_u C_u^{-1} C_{x\lambda}\right), \\ R &= \left(D_{x\lambda} - D_u C_u^{-1} C_{x\lambda}\right)^{-1} \\ &\quad \times \left[\left(D_z + D_u C_u^{-1} C_z\right) \Pi + \left(F_z + F_u C_u^{-1} C_z\right)\right]. \end{aligned}$$

In our problem, the matrix $D_{x\lambda} - D_u C_u^{-1} C_{x\lambda}$ is invertible and we can compute the matrix W. The economy displays local saddle-point stability around the non-stochastic steady state if and only if the number of eigenvalues of W inside the unit circle is equal to the number of predetermined variables. As it turns out, the matrix W of the dynamic system (7.45a)-(7.45b) has 59 eigenvalues with absolute value less than one which is exactly the number of predetermined variables k_t^s, $s = 2, \ldots, 59$.[18] Therefore, our economy is locally stable.[19]

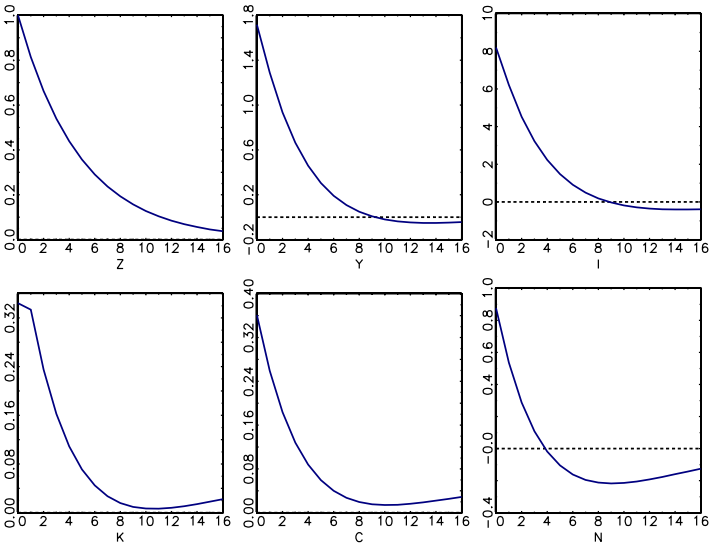

Figure 7.11: Impulse responses in the OLG model

Business Cycle Dynamics. How do the business cycle dynamics of the OLG model compare with those of the standard Ramsey model?[20] For this reason, we look at the impulse response functions of our OLG model (see Figure 7.11) and also compute arti-

[18] In the exercises, you are asked to show that Example 7.1.1 also displays local stability.

[19] See Laitner (1990) for a detailed analysis of local stability and determinancy in Auerbach-Kotlikoff models.

[20] As the only other work on this subject known to us, Ríos-Rull (1996) also considers the dynamics in a stochastic life-cycle model. Different from

ficial time series for output Y_t, investment $I_t = K_{t+1} - (1-\delta)K_t$, consumption C_t, working hours N_t, and the real wage w_t. The impulse responses and statistics (see Table 7.2) are computed with the help of program `ch722.g`. The impulse responses of the technology level Z and the endogenous variables Y, I, K, C, and N are presented in Figure 7.11. Notice that we used years rather than quarters as time periods which is usually not the case in business cycle studies that are based on the stochastic Ramsey model. The impulse responses are similar to those in the benchmark model. Interestingly, however, employment even undershoots its long-run steady state value along the adjustment path after 4 years.

Table 7.2

Variable	s_x	r_{xy}	r_x
Output	0.70	1.00	0.74
	(2.23)	(1.00)	
Investment	2.65	0.56	0.30
	(8.66)	(0.82)	
Consumption	0.70	0.77	0.74
	(1.69)	(0.86)	
Hours	0.70	0.41	0.75
	(1.88)	(0.94)	
Real Wage	1.69	-0.12	0.34

Notes: s_x:=standard deviation of HP-filtered simulated series of variable x, r_{xy}:=cross correlation of variable x with output, r_x:=first order autocorrelation of variable x.

Analogously to our procedure in Section 2.4, we use the linear law of motion for $\mathbf{x}_{t+1}$, $\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t + (0, \ldots, 0, \epsilon_t)^T$, and the GAUSS random number generator `rndn` in order to simulate time

us, however, he concentrates on the analysis of a pareto-optimal economy and studies the problem of a central planner. In particular, he uses the Hansen-Prescott Algorithm presented in Section 2.1 in order to compute the solution of this model. Our approach also allows for the computation of the dynamics in a stochastic decentralized economy.

series data. The logs of the results are passed to the HP-Filter (with $\lambda = 100$ for annual data) in order to get percentage deviations from the stationary solution. Table 7.2 presents the second moments from the filtered series for the OLG model. Be careful not to compare the results obtained from the annual periods of the OLG model with those presented in Table 2.1 from quarterly periods for the stochastic Ramsey model.

The cyclical behavior of the US economy during 1956-87 is presented in parentheses in Table 7.2.[21] Table 7.2 shows that all variables in our model are not volatile enough. Furthermore, consumption in our model varies as much as output, while this is obviously not the case for the US economy. The correlations of investment, hours, and consumption with output in our model, however, are in good accordance with the US data.

[21] The data is taken from Table 4 in Ríos-Rull (1996).

Problems

7.1 **Earnings-related pensions**
Consider the steady state of the model in Section 7.1 with a constant wage and interest rate. Pension payments b are assumed to be lump-sum irrespective of the individual's earnings history and contributions to social security. As a more realistic description of pension payments in modern industrial countries, let pension payments depend on average life-time earnings. In addition, the government provides minimum social security b_{min} in old age. More formally, for an individual with earnings history $\{n^s w\}_{s=1}^T$, annual pension payments are calculated by the formula

$$b = \epsilon \sum_{s=1}^{T} \frac{n^s w}{T} + b_{min}, \quad 0 \leq \epsilon < 1.$$

As a consequence, the individual state variables of the value function of the retired agent are given by individual wealth k and annual pension payments b, while the individual state variables of the young agent are given by his wealth k and his accumulated earnings.
Furthermore, the working agent maximizes his labor supply taking the intertemporal effect on his entitlement to pension payments into account. Accordingly, his first-order condition with respect to labor supply in the steady state is given by

$$\frac{u_l(c^s, l^s)}{u_c(c^s, l^s)} = (1-\tau)w + \beta^{T+1-s} \frac{\partial V^{T+1}(k^T, b, K, N)}{\partial b} \epsilon \frac{w}{T},$$

where the second additive term on the right-hand side of the equation reflects the increase of old-age utility from an increase in labor supply through its effect on pensions.
Compute the stationary equilibrium and show that an initial increase of ϵ from 0 to 0.1 increases employment and the welfare as measured by the value of the newborn generation $V^1(0, 0, K, N)$.

7.2 Recompute Example 7.1.1. Use value function iteration with linear interpolation in order to compute `kst`.

7.3 Show that the economy described in Example 7.1.1 is saddlepoint stable.

7.4 **Concentration of wealth**
Consider the model described in Section 7.2.1.

a) Recompute the model for a less stricter borrowing constraint where the agent can borrow up to the average wage in the economy, $(1-\tau)w\bar{h}N/\sum_{t=1}^{T}\mu_t$. How does this affect the Gini coefficient of wealth?

b) Compute the effect of higher public pensions on the wealth heterogeneity. For this reason, increase the replacement ratio to 50%.
c) Compute the model assuming that all accidental bequests are transferred lump-sum to the households in equal amounts. How does this affect the concentration of wealth as measured by the Gini coefficient?
d) Compute the model assuming that labor supply is endogenous. Use the utility function and calibration presented in Example 7.2.1.

Part III

Tools

Chapter 8

Numerical Methods

8.1 A Quick Refresher in Linear Algebra

In this section we provide some elementary and some more advanced, but very useful concepts and techniques from linear algebra. Most of the elementary material gathered here is found in any undergraduate textbook on linear algebra as, e.g., Lang (1987). For the more advanced subjects Bronson (1989) as well as Golub and van Loan (1996) are good references. In addition, many texts on econometrics review matrix algebra, as, e.g., Greene (2003), Appendix A, or Judge et al. (1982), Appendix A.

8.1.1 Complex Numbers

Remember that a complex number c is an object of the form $c = \alpha + i\beta$. The i stands for the imaginary unit, whose square is defined to equal minus unity,i.e., $i^2 = -1$. Complex numbers are points in the Gaussian plane (see Figure 9.1 on page 490). The complex conjugate of c, is denoted by $\bar{c}$ and given by $\bar{c} = \alpha - i\beta$. Addition and multiplication of complex numbers $c_1 = \alpha_1 + i\beta_1$ and $c_2 = \alpha_2 + i\beta_2$ are defined by the following formulas:

$$\begin{aligned} c_1 + c_2 &= \alpha_1 + \alpha_2 + i(\beta_1 + \beta_2), \\ c_1 c_2 &= (\alpha_1 + i\beta_1)(\alpha_2 + i\beta_2) \\ &= (\alpha_1\alpha_2) - (\beta_1\beta_2) + i(\alpha_1\beta_2 + \alpha_2\beta_1). \end{aligned}$$

The symbol $\mathbb{C}$ denotes the set of all complex numbers.

8.1.2 Vectors

A real (complex) vector of dimension n is a n-tuple of numbers $x_i \in \mathbb{R}$ ($x_i \in \mathbb{C}$) $i = 1, 2, \ldots n$, denoted by

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

The space of all n-tuples is $\mathbb{R}^n$ ($\mathbb{C}^n$). Vector addition and scalar multiplication are defined by

$$\mathbf{y} = a + b\mathbf{x} = \begin{bmatrix} a + bx_1 \\ a + bx_2 \\ \vdots \\ a + bx_n \end{bmatrix}.$$

8.1.3 Norms

Norms are measures of distance. Since the distance of $\mathbf{x}$ from the zero vector is the length of $\mathbf{x}$, norms are also measures of vector length. More formally, a norm on $\mathbb{R}^n$ (and similarly on $\mathbb{C}^n$) is a real valued function $\|\mathbf{x}\|$ that obeys:

$$\begin{aligned} &\|\mathbf{x}\| \geq 0 \text{ for all } \mathbf{x} \in \mathbb{R}^n, \text{ and } \|\mathbf{x}\| = 0 \text{ if and only if } \mathbf{x} = \mathbf{0} \in \mathbb{R}^n, \\ &\|a\mathbf{x}\| = |a| \cdot \|\mathbf{x}\| \text{ for all } \mathbf{x} \in \mathbb{R}^n \text{ and } a \in \mathbb{R}, \\ &\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \text{ for all } \mathbf{x}, \mathbf{y} \in \mathbb{R}^n. \end{aligned} \tag{8.1}$$

The most common examples of norms on $\mathbb{R}^n$ are

- the ℓ_∞ or sup norm: $\|\mathbf{x}\|_\infty := \max_{1 \leq i \leq n} |x_i|$, where $|x_i|$ denotes the absolute value of x_i.
- the ℓ_2 or Euclidean norm: $\|\mathbf{x}\|_2 := \left(\sum_{i=1}^n x_i^2\right)^{1/2}$.

A set of n vectors $\mathbf{x}_i$, $i = 1, 2, \ldots n$ is linearly independent if and only if the solution to

$$\mathbf{0} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_n\mathbf{x}_n$$

is $a_1 = a_2 = \cdots = a_n = 0$. A set $\mathcal{B} := \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ of n linearly independent vectors is a basis $\mathcal{B}$ for $\mathbb{R}^n$, since any element $\mathbf{x} \in \mathbb{R}^n$ can be represented by a linear combination of the elements of $\mathcal{B}$, i.e.,

$$\mathbf{x} = \sum_{i=1}^{n} a_i\mathbf{v}_i \quad \forall \mathbf{x} \in \mathbb{R}^n.$$

8.1.4 Matrices

A real (complex) matrix A with typical element $a_{ij} \in \mathbb{R}$ ($a_{ij} \in \mathbb{C}$) is the following n-by-m array of real numbers:

$$A = (a_{ij}) := \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{bmatrix}.$$

If $n = m$, A is called a square matrix. Other special matrices encountered in the main text are:

$$\underbrace{\begin{bmatrix} a_{11} & 0 & \ldots & 0 \\ 0 & a_{22} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{bmatrix}}_{\text{diagonal matrix}}, \quad \underbrace{\begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ 0 & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & a_{nn} \end{bmatrix}}_{\text{upper triangular matrix}}, \quad \underbrace{\begin{bmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{bmatrix}}_{\text{identity matrix}}.$$

If we consider the matrix $A = (a_{ij})$ as the row vector

$$[\underbrace{a_{11}, a_{21}, \ldots a_{n1}}_{\text{column 1}}, \underbrace{a_{12}, \ldots, a_{n2}}_{\text{column 2}}, \ldots, \underbrace{a_{1n} \ldots, a_{nn}}_{\text{column } n}],$$

we may apply the definition of any vector norm to this "long" vector to find the corresponding matrix norm. For instance, the ℓ_2 norm of A is

$$\|A\| = \left(\sum_{j=1}^{n} \sum_{i=1}^{n} a_{ij}^2 \right)^{1/2}.$$

Matrix addition and scalar multiplication are defined componentwise:

$$C = A + dB = \begin{bmatrix} a_{11} + db_{11} & a_{12} + db_{12} & \dots & a_{1m} + db_{1m} \\ a_{21} + db_{21} & a_{22} + db_{22} & \dots & a_{2m} + db_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} + db_{n1} & a_{n2} + db_{n2} & \dots & a_{nm} + db_{nm} \end{bmatrix} \tag{8.2}$$

for $A, B, C \in \mathbb{R}^{n \times m}$ and $d \in \mathbb{R}$. Thus, matrix addition obeys the following rules:

$$A + B = B + A, \tag{8.3a}$$

$$A + (B + C) = (A + B) + C. \tag{8.3b}$$

The product of two matrices $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^{m \times n}$, is the $n \times n$ matrix $C = (c_{ij})$, defined by

$$c_{ij} = \sum_{k=1}^{m} a_{ik} b_{kj}. \tag{8.4}$$

The Kronecker product $\otimes$ of two matrixes A and B is the following expression:

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \dots & a_{1m}B \\ a_{21}B & a_{22}B & \dots & a_{2m}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1}B & s_{n2}B & \dots & a_{nm}B \end{bmatrix}. \tag{8.5}$$

For suitable matrices A, B, C, and D matrix multiplication satisfies the following rules

$$AB \neq BA, \tag{8.6a}$$

$$A(B + C) = AB + AC, \tag{8.6b}$$

$$A(BC) = (AB)C, \tag{8.6c}$$

$$A(B + C)D = ABD + ACD. \tag{8.6d}$$

The trace of a square matrix A is the sum of the elements of its main diagonal, i.e.,

$$\operatorname{tr} A = \sum_{i=1}^{n} a_{ii}. \tag{8.7}$$

The determinant of a 2×2 matrix A, denoted by either $|A|$ or $\det(A)$, is defined by

$$|A| = a_{11}a_{22} - a_{12}a_{21}. \tag{8.8}$$

There is a recursive formula to compute the determinant of an arbitrary square matrix of dimension n. Using an arbitrary row (say i) or column (say j), the formula is:

$$|A| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|A_{ij}| = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}|A_{ij}|, \tag{8.9}$$

where A_{ij} is the matrix obtained from A by deleting the i-th row and j-th column. This expansion gives the determinant of A in terms of a sum of determinants of $n-1$ matrices. These can be reduced further to determinants of $n-2$ matrices and so forth until the summands are 2×2 matrices, computed from equation (8.8).

The rank of an arbitrary $n \times m$ matrix A is the maximal number of linearly independent rows of A. This also equals the maximal number of linearly independent columns of A.

The transpose of A, denoted by A' or A^T, is the $m \times n$ matrix obtained by interchanging the rows and columns of A:

$$A' = (a'_{ij}) = (a_{ji}) = \begin{bmatrix} a_{11} & a_{21} & \dots & a_{n1} \\ a_{12} & a_{22} & \dots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1m} & a_{2m} & \dots & a_{nm} \end{bmatrix}.$$

In the case of a complex matrix, we use the prime $'$ or the superscript T to denote conjugate complex transposition. Thus, A'

is the matrix whose element in the ij-th position is the complex conjugate of a_{ji}.

The inverse of a square matrix A, denoted $A^{-1} = (a^{ij})$ (note that we use superscripts to indicate the typical element of an inverse matrix) solves the problem $AA^{-1} = I$. If it exists, the inverse is unique and given by

$$a^{ij} = \frac{a_{ij}(-1)^{i+j}|A_{ji}|}{|A|}. \tag{8.10}$$

If $|A| = 0$, the inverse does not exist. It is an implication of the expansion formula (8.9) that matrices with a row (or column) of zeros or with linearly dependent rows (or columns) have no inverse. In general, an invertible (non-invertible) matrix is named non-singular (singular).

A square matrix A is symmetric, if it equals its transpose: $A = A'$. The transpose operator obeys the following rules:

$$(A')' = A, \tag{8.11a}$$
$$(A + B)' = A' + B', \tag{8.11b}$$
$$(AB)' = B'A', \tag{8.11c}$$
$$(A^{-1})' = (A')^{-1}. \tag{8.11d}$$

8.1.5 Linear and Quadratic Forms

Let $\mathbf{a} = (a_1, a_2, \ldots, a_n)'$ and $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$ denote two n-dimensional column vectors. The dot product

$$z = \mathbf{a}'\mathbf{x} = \sum_{i=1}^{n} a_i x_i \tag{8.12}$$

with given $\mathbf{a}$ is called a linear form. The column vector of partial derivatives of z with respect to x_i, $i = 1, 2, \ldots, n$, denoted by ∇z, is obviously given by:

$$\nabla z := \frac{\partial \mathbf{a}'\mathbf{x}}{\partial \mathbf{x}} = \mathbf{a} = (\mathbf{a}')'. \tag{8.13}$$

Since $z = z' = \mathbf{x}'\mathbf{a}$ we also have

$$\frac{\partial \mathbf{x}'\mathbf{a}}{\partial \mathbf{x}} = \mathbf{a}. \tag{8.14}$$

A direct application of these findings are the following two rules:

$$\frac{\partial \mathbf{u}'B\mathbf{x}}{\partial \mathbf{x}} = (\mathbf{u}'B)' = B'\mathbf{u}, \tag{8.15a}$$

$$\frac{\partial \mathbf{u}'B\mathbf{x}}{\partial \mathbf{u}} = B\mathbf{x}, \tag{8.15b}$$

where $\mathbf{u} \in \mathbb{R}^m$, $B \in \mathbb{R}^{m\times n}$, and $\mathbf{x} \in \mathbb{R}^n$.

Let $A = (a_{ij})$ denote a $n \times n$ square matrix and $\mathbf{x} = (x_1, x_2, \ldots, x_n)'$ a n-dimensional column vector. The expression

$$q = \mathbf{x}'A\mathbf{x}, \quad q \in \mathbb{R}, \tag{8.16}$$

is a quadratic form. If $q \geq 0$ ($q \leq 0$) for each non-zero vector $\mathbf{x}$, the matrix A is said to be positive (negativ) semi-definite. If $q > 0$ ($q < 0$), A is positive (negative) definit. Let $B \in \mathbb{R}^{n\times m}$, $\mathbf{x} \in \mathbb{R}^n$, and $\mathbf{v} = B\mathbf{x}$. Since

$$\mathbf{v}'\mathbf{v} = \sum_{i=1}^{m} v_i^2 = \mathbf{x}'B'B\mathbf{x} \geq 0 \; \forall \mathbf{x},$$

the square matrix $A := B'B$ is obviously positive definite. Using the rule for matrix multiplication given in (8.4), equation (8.16) can be written in several ways:

$$\begin{aligned} q &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j, \\ &= \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij}x_i x_j, \\ &= \sum_{i=1}^{n} a_{ii}x_i^2 + \sum_{i=1}^{n}\sum_{j=1+i}^{n} (a_{ij} + a_{ji})x_i x_j. \end{aligned}$$

Setting $\tilde{a}_{ij} = \tilde{a}_{ji} \equiv (a_{ij} + a_{ji})/2$, it is obvious that we can assume without loss of generality that the matrix A is symmetric. Using

this assumption, it is easy to show that the column vector of first partial derivatives of q with respect to x_i, $i = 1, 2, \ldots, n$, is given by

$$\nabla q := \frac{\partial \mathbf{x}' A \mathbf{x}}{\partial \mathbf{x}} = (A + A')\mathbf{x} = 2A\mathbf{x}. \tag{8.17}$$

8.1.6 Eigenvalues and Eigenvectors

Let $A \in \mathbb{R}^{n \times n}$. A right eigenvector of A is a vector $\mathbf{v}$ that solves

$$A\mathbf{v} = \lambda \mathbf{v} \quad \Leftrightarrow \quad (A - \lambda I)\mathbf{v} = \mathbf{0}. \tag{8.18}$$

Similarly, the solution of $\mathbf{v}'A = \lambda \mathbf{v}'$ is named a left eigenvector of A. The system of n linear equations (8.18) has non-trivial solutions $\mathbf{v} \neq \mathbf{0}$, if the determinant $|A - \lambda I|$ vanishes. The condition $|A - \lambda I| = 0$ results in a polynomial of degree n in λ. It is well known from the Fundamental Theorem of Algebra (see, e.g., Hirsch and Smale (1974), pp. 328ff.) that this polynomial has n roots, which may be real, complex, or multiples of each other. These roots are the eigenvalues of the matrix A. Solving equation (8.18) for a given λ_i gives the associated eigenvector $\mathbf{v}_i$. Thus, eigenvectors are vectors that either stretch or shrink when multiplied by A. If $\mathbf{v}_i$ solves (8.18) and c is an arbitrary scalar, then $c\mathbf{v}_i$ also solves (8.18). Therefore, eigenvectors are unique up to a scalar multiple and, thus, may be normalized to have unit length.

There are two important relations between the elements of A and its eigenvalues:

$$\sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} a_{ii}, \tag{8.19a}$$

$$\prod_{i=1}^{n} \lambda_i = |A|. \tag{8.19b}$$

In words: the sum of the eigenvalues of A equals the trace of A, and the determinant of A equals the product of the n eigenvalues.

Note that equation (8.18) is a special case of

$$(A - \lambda I)^m \mathbf{v}_m = \mathbf{0}$$

for $m = 1$. If there are non-trivial solutions $\mathbf{v}_m$ for $m \geq 2$ but not for $m - 1$, the vector $\mathbf{v}_m$ is called a generalized right eigenvector of rank m for the square matrix A. The space spanned by the (generalized) eigenvectors of A is called the eigenspace of A. The eigenspace can be partitioned in three subspaces formed by generalized eigenvectors that belong to the eigenvalues with

- modulus less than one (stable eigenspace, E^s),
- modulus equal to one (center eigenspace, E^c),
- modulus greater than one (unstable eigenspace, E^u).

8.1.7 Matrix Factorization

Matrix factorizations play an important role in the solution of systems of linear difference equations. The are also used to solve systems of linear equations. Here, we touch on the Jordan, the Schur, the LU, and the Cholesky factorization.

Jordan Factorization. Consider the case of n distinct real eigenvalues and associated eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$ of a square matrix A. The matrix

$$P = [\mathbf{v}_1, \mathbf{v_2}, \ldots, \mathbf{v}_n]$$

transforms A into a diagonal matrix Λ with the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ on its main diagonal:

$$\Lambda = P^{-1} A P.$$

In the general case of real and complex eigenvalues, possibly with multiplicity $m > 1$, it may not be possible to diagonalize A. Yet there exists a modal matrix M of a set of linearly independent generalized eigenvectors (which is not unique) that puts A in Jordan canonical form:

$$A = MJM^{-1}, \quad J = \begin{bmatrix} D & 0 & \ldots & 0 \\ 0 & J_1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & J_K \end{bmatrix}. \tag{8.20}$$

Here, $D \in \mathbb{C}^{n_1}$ is diagonal matrix of those n_1 eigenvalues that appear once, and the Jordan blocks $J_k \in \mathbb{C}^{m\times m}$ given by

$$J_k = \begin{bmatrix} \lambda & 1 & 0 & 0 & \dots & 0 \\ 0 & \lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & \lambda \end{bmatrix},$$

refer to eigenvalues λ with multiplicity $m > 1$. There is also a real Jordan factorization of A, where each complex root $\lambda_j = \alpha_j + i\beta_j$ either in D or J_k is represented by a matrix

$$\begin{bmatrix} \alpha_j & -\beta_j \\ \beta_j & \alpha_j \end{bmatrix}.$$

Schur Factorization. The Schur factorization of a square matrix A is given by

$$A = TST^{-1}. \tag{8.21}$$

The complex matrix S is upper triangular with the eigenvalues of A on the main diagonal. It is possible to choose T such that the eigenvalues appear in any desired order along the diagonal of S. The transformation matrix T has the following properties:

- its complex conjugate transpose T' equals the inverse of T,
- therefore: $TT' = TT^{-1} = I$, i.e., T is an unitary matrix,
- all eigenvalues of T have absolute value equal to 1.

LU and Cholesky Factorization. Consider a system of linear equations

$$\left.\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ \vdots &= \vdots, \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= b_n \end{aligned}\right\} \Leftrightarrow\ A\mathbf{x} = \mathbf{b}. \tag{8.22}$$

We assume that the square matrix A has full rank, i.e., there are no linearly dependent rows or columns in A. In this case it is possible to factorize A as follows

$$LU = A,$$
$$L = \begin{bmatrix} 1 & & 0 & 0 & \dots & 0 \\ l_{21} & 1 & 0 & 0 & \dots & 0 \\ l_{31} & l_{32} & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & l_{n4} & \dots & 1 \end{bmatrix}, \quad U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \dots & u_{1n} \\ 0 & u_{22} & u_{23} & \dots & u_{2n} \\ 0 & 0 & u_{33} & \dots & u_{3n} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & u_{nn} \end{bmatrix}. \tag{8.23}$$

If A is symmetric and positive definite, its Cholesky factor is the lower triangular matrix L that solves

$$LL' = A. \tag{8.24}$$

Both the LU and the Cholesky factorization can be used to solve the linear system (8.22). Let $\tilde{\mathbf{x}} := U\mathbf{x}$. Then it is easy to solve the system

$$L\tilde{\mathbf{x}} = \mathbf{x}$$

by forward substitution:

$$\begin{aligned} \tilde{x}_1 &= b_1, \\ \tilde{x}_2 &= b_2 - l_{21}\tilde{x}_1, \\ \tilde{x}_3 &= b_3 - l_{31}\tilde{x}_1 - l_{32}\tilde{x}_2, \\ \vdots &= \vdots \end{aligned}$$

Given the solution for $\tilde{\mathbf{x}}$, one gets the desired solution for $\mathbf{x}$ via backward substitution from $U\mathbf{x} = \tilde{\mathbf{x}}$:

$$\begin{aligned} x_n &= \frac{\tilde{x}_n}{u_{nn}}, \\ x_{n-1} &= \frac{1}{u_{n-1\,n-1}} \left(\tilde{x}_{n-1} - u_{n-1\,n} x_n\right), \\ x_{n-2} &= \frac{1}{u_{n-2\,n-2}} \left(\tilde{x}_{n-2} - u_{n-2\,n-1} x_{n-1} - u_{n-2\,n} x_n\right), \\ \vdots &= \vdots \end{aligned}$$

The solution of a system of linear equations via its LU or Cholesky factorization is the strategy that underlies linear equations solvers. For instance, the LAPACK routine `dgesv.for` uses this procedure as well as the IMSL subprogram `DLSARG`. In GAUSS you must first get the LU factorization from the command `{L,U}=LU(A)` and then use the command `x=LuSol(b,L,U)`.

It is beyond the scope of this text to deal with algorithms that compute any of the above mentioned factorizations.[1] In GAUSS the command `{S,T}=Schtoc(Schur(A))` can be used to get the matrices S and T from the matrix A. In Fortran the subroutine `ZGEES` from LAPACK can be used for the same purpose.[2] Whereas `ZGEES` has an option to order the eigenvalues on the main diagonal of S, the `Schtoc` command returns S with unordered eigenvalues. However, Givens rotations may be used to order the eigenvalues in S.

8.1.8 Givens Rotation

Consider the symmetric matrix $G \in \mathbb{C}^n$ defined by

$$G := \begin{bmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \\ 0 & 0 & \dots & b & c & \dots & 0 \\ 0 & 0 & \dots & -\bar{c} & \bar{b} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 \end{bmatrix} \begin{matrix} \\ \\ \\ \leftarrow \text{row } i \\ \leftarrow \text{row } i{+}1 \\ \\ \\ \end{matrix}$$

$$\begin{matrix} & & & \uparrow & \uparrow & & \\ & & & \text{column } i,i{+}1 & & & \end{matrix}$$

and the upper triangular matrix $S \in \mathbb{C}^n$. Choose the complex numbers b and c as follows:

$$b = \frac{s_{ii+1}}{r},\ c = \frac{s_{i+1i+1} - s_{ii}}{r},\ r := \sqrt{s_{ii+1}^2 + (s_{i+1i+1} - s_{ii})^2}.$$

[1] A good reference on this subject is Golub and van Loan (1996).

[2] The LAPACK Fortran 77 routines can be downloaded from `www.netlib.com/lapack` for free. They are also included in the CXML library being shipped with Compaq's Digital Fortran compiler.

In this case G is also an unitary matrix, $GG' = I$, so that

$$A = T(GG')S(GG')T = (TG)(G'SG)(G'T').$$

As an exercise you may want to verify that pre-multiplying S with G' and post-multiplying S with G interchanges s_{ii} and s_{i+1i+1}. The new transformation matrix putting A into the newly ordered matrix $G'SG$ is given by TG. Via a number of such pairwise Givens rotations the eigenvalues on the main diagonal of S can be brought into any desired order.

8.2 Function Approximation

There are numerous instances where we need to approximate functions of one or several variables. In some cases we need a local approximation around a given point x_0. For instance, in Chapter 2, the linear-quadratic approximation method requires a quadratic approximation of the return function at the stationary equilibrium. In other cases we must approximate functions over a given interval. Think of the value function in Chapter 1, or the policy function in Chapter 4.

Usually, local approximations rest on Taylor's theorem, which we review in Section 8.2.1. The simplest case of function approximation over a given interval is linear interpolation, which we discuss in Section 8.2.2. Linear interpolation is a special case of polynomial approximation dealt with in Section 8.2.3. Among the various families of polynomials, orthogonal polynomials have very desirable properties. From this class, we consider Chebyshev polynomials in Section 8.2.4. Finally, we briefly touch on neural networks.

8.2.1 Taylor's Theorem

Consider a function f of a single variable $x \in U$, where U is an open subset of $\mathbb{R}$. Taylor's theorem states the following:[3]

[3] Statements of this theorem appear in any calculus textbook and in most mathematics for economists texts. Judd (1998), p. 23, states the theorem for the singe variable case; p. 239 of the same book presents the formula for the n-variable case.

Theorem 8.2.1 *Let $f : [a, b] \to \mathbb{R}$ be a $n+1$ times continuously differentiable function on (a, b), let $\bar{x}$ be a point in (a, b). Then*

$$f(\bar{x}+h) = f(\bar{x}) + f^{(1)}(\bar{x})h + f^{(2)}(\bar{x})\frac{h^2}{2} + \cdots + f^{(n)}(\bar{x})\frac{h^n}{n!} + f^{(n+1)}(\xi)\frac{h^{n+1}}{(n+1)!}, \quad \xi \in (\bar{x}, \bar{x}+h).$$

In this statement $f^{(i)}$ is the i–th derivative of f evaluated at the point $\bar{x}$. The derivative that appears in the rightmost term of this formula is evaluated at some unknown point between $\bar{x}$ and $\bar{x}+h$. When we neglect this term, the formula approximates the function f at $\bar{x}$ and the approximation error is of order $n+1$. By this we mean that the error is proportional to h^{n+1} where the constant of proportionality is given by $C = f^{n+1}(\xi)/(n+1)!$.

There is also a version of this theorem for the n-variable case. To present it, we need a fair amount of additional notation. Let $\mathbf{a} := [a_1, a_2, \ldots, a_n]$, define

$$\begin{aligned}
|\mathbf{a}| &:= \sum_{i=1}^{n} a_i, \ \forall a_i = 0, 1, \ldots, \\
\mathbf{a}! &:= a_1! a_2! \ldots a_n!, \\
\mathbf{x}^{\mathbf{a}} &:= x_1^{a_1} x_2^{a_2} \ldots x_n^{a_n}, \\
D_i f(\mathbf{x}) &:= \frac{\partial f(\mathbf{x})}{\partial x_i}, \\
D_i^{a_i} &:= \underbrace{D_i D_i \ldots D_i}_{a_i \text{ times}}, \\
D^{\mathbf{a}} f(\mathbf{x}) &:= D_1^{a_1} D_2^{a_2} \ldots D_n^{a_n}.
\end{aligned}$$

Then, the following holds:

Theorem 8.2.2 *Let $U \subset \mathbb{R}^n$ be an open subset, $\mathbf{x} \in U$, $\mathbf{h} \in \mathbb{R}^n$ so that $\mathbf{x}+t\mathbf{h} \in U$ for all $t \in [0.1]$. Assume that $f : U \to \mathbb{R}$ is $(k+1)$-times continuously differentiable. Then, there is a $\lambda \in [0, 1]$, so that*

$$f(\mathbf{x}+\mathbf{h}) = \sum_{|\mathbf{a}| \le k} \frac{D^{\mathbf{a}} f(\mathbf{x})}{\mathbf{a}!} \mathbf{h}^{\mathbf{a}} + \sum_{|\mathbf{a}| = k+1} \frac{D^{\mathbf{a}} f(\mathbf{x}+\lambda\mathbf{h})}{\mathbf{a}!} \mathbf{h}^{\mathbf{a}}.$$

Note that $D_i^0 \equiv 1$ and that summation is over all n-tuples $[a_1, a_2, \ldots, a_n]$, which sum to $0, 1, 2, \ldots, k$ (or $k+1$).

An immediate corollary of this theorem is the quadratic approximation of f at $\mathbf{x}$: assume that $k = 2$. Let

$$\nabla f(\mathbf{x}) := [D_1 f(\mathbf{x}), D_2 f(\mathbf{x}), \ldots, D_n f(\mathbf{x})]'$$

denote the column vector of first partial derivatives and let

$$H(\mathbf{x}) := \begin{bmatrix} D_1 D_1 f(\mathbf{x}) & D_1 D_2 f(\mathbf{x}) & \ldots & D_1 D_n f(\mathbf{x}) \\ D_2 D_1 f(\mathbf{x}) & D_2 D_2 f(\mathbf{x}) & \ldots & D_2 D_n f(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ D_n D_1 f(\mathbf{x}) & D_n D_2 f(\mathbf{x}) & \ldots & D_n D_n f(\mathbf{x}) \end{bmatrix}$$

be the Hesse matrix of second partial derivatives (which is symmetric if f is two times continuously differentiable). Then

$$f(\mathbf{x}+\mathbf{h}) \approx f(\mathbf{x}) + [\nabla f(\mathbf{x})]'\mathbf{h} + \frac{1}{2}\mathbf{h}' H(\mathbf{x})\mathbf{h}, \tag{8.25}$$

where the approximation error $\phi(\mathbf{h})$, with $\phi(\mathbf{0}) = 0$, has the property

$$\lim_{\substack{\mathbf{h}\to\mathbf{0} \\ \mathbf{h}\neq 0}} \frac{\phi(\mathbf{h})}{\|\mathbf{h}\|^2} = 0.$$

Similarly, the linear approximation is given by:

$$f(\mathbf{x}+\mathbf{h}) \approx f(\mathbf{x}) + [\nabla f(\mathbf{x})]'\mathbf{h}, \tag{8.26}$$

where the error $\phi(\mathbf{h})$ now obeys

$$\lim_{\substack{\mathbf{h}\to\mathbf{0} \\ \mathbf{h}\neq 0}} \frac{\phi(\mathbf{h})}{\|\mathbf{h}\|} = 0.$$

8.2.2 Linear Interpolation

Linear interpolation is simple and shape preserving. This property is important, if we use interpolation to approximate the value function, which is known to be concave and increasing.[4]

Consider Figure 8.1 that depicts the graph of a given function f. Suppose we want to approximate $f(x)$ at a given point x with the property $x_1 < x < x_2$. Linear interpolation uses the point

$$\hat{f}(x) := f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1). \tag{8.27}$$

Thus, f is approximated by the line through $(x_1, f(x_1))$ and $(x_2, f(x_2))$.

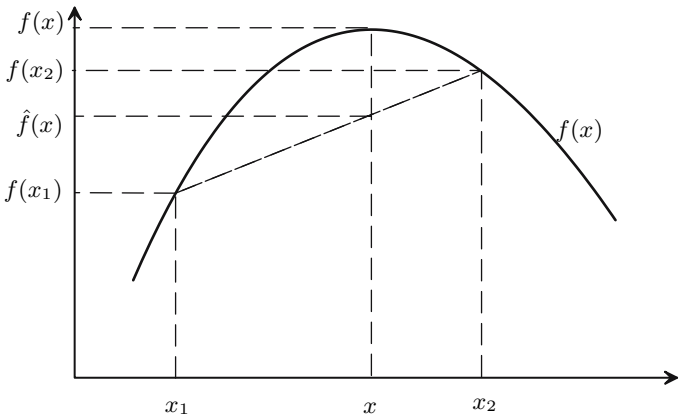

Figure 8.1: Linear Interpolation

In many applications (such as value function iteration with interpolation between grid points) the function f is known only at a given number of points $x_1, x_2, \ldots$. In this case it is helpful to have a procedure that finds the neighboring points $x_i < x < x_{i+1}$ and returns $\hat{f}(x)$. Our GAUSS procedure `LinInter` does this. It takes the vector of grid points $\mathbf{x} = [x_1, x_2, \ldots, x_n]$ and the

[4] There are so-called *shape-preserving* methods which, however, have been produced for mainly one-dimensional problems. Costantini and Fontanella (1990) consider shape-preserving bivariate interpolation.

corresponding vector $\mathbf{y} = [f(x_1), f(x_2), \ldots, f(x_n)]$ as well as the point x as input, locates x in the grid, and returns $\hat{f}(x)$ given by (8.27).

8.2.3 Families of Polynomials

Bases for Function Spaces. The formula given in (8.27) is a special case of the more general formula

$$f(x) = \sum_{i=0}^{n} \alpha_i \varphi_i(x),$$

with $n = 1$, $\varphi_i(x) := x^i$, $\alpha_0 = [x_2 f(x_1) - x_1 f(x_2)]/(x_2 - x_1)$, and $\alpha_1 = [f(x_2) - f(x_1)]/(x_2 - x_1)$. To understand the idea behind this formula, remember that a vector $\mathbf{x} \in \mathbb{R}^n$ can be represented as a linear combination of n linearly independent vectors $\mathcal{B} := \{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n\}$ of $\mathbb{R}^n$. The collection of vectors $\mathcal{B}$ is said to build a base of the vector space $\mathbb{R}^n$. If the members of $\mathcal{B}$ are mutually orthogonal (i.e., $\mathbf{v}_i' \mathbf{v}_j = 0$ for $i \neq j$) and normal (i.e., $\mathbf{v}_i' \mathbf{v}_i = 1$) the base is called an orthonormal base.

Now, consider the set of all continuous functions that map the interval $[a, b]$ to the real line. Like $\mathbb{R}^n$, this set, denoted by $C[a, b]$, is a vector space. The monomials $x^i, i = 0, 1, \ldots$ build a base $\mathcal{B}$, for this space, i.e., every member of $C[a, b]$ can be represented by

$$\sum_{i=0}^{\infty} \alpha_i x^i.$$

For this reason it is common to use a linear combination of the first p members of this base to approximate a continuous function $f(x) \in C[a, b]$:

$$f(x) \simeq \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_p x^p.$$

Yet, this may not always be a good choice. For instance, if we use a regression of $y = f(x)$ on $(1, x, x^2, ..., x^p)$ to determine $\boldsymbol{\alpha} := (\alpha_0, \alpha_1, \ldots, \alpha_p)'$, as we actually do in the parameterized expectations approach considered in Chapter 3, we may face the

problem of multicollinearity (i.e., nearly linear dependence among the x^i), since for large i, x^i and x^{i+1} may be difficult to distinguish. Bases that consists of polynomials that are – in an appropriate sense – orthogonal circumvent this problem.

Orthogonal Polynomials. To motivate the notion of orthogonality in a function space, consider the following problem. Assume we want to approximate $f(x) \in C[a, b]$ by

$$\hat{f}(x) := \sum_{i=1}^{n} \alpha_i \varphi_i(x),$$

where $\mathcal{P} := (\varphi_0(x), \varphi_1(x), \dots)$ is some family of polynomials. Suppose further that there is a weight function $w(x)$ on $[a, b]$. A weight function is a function that has a finite integral

$$\int_a^b w(x)dx < \infty$$

and that is positive almost everywhere on $[a, b]$.[5] Our goal is to choose the parameters $\alpha_i, i = 0, 1, \dots n$ such that the weighted sum of squared errors $R(\boldsymbol{\alpha}, x) := f(x) - \hat{f}(\boldsymbol{\alpha}, x)$ over all $x \in [a, b]$ attains a minimum:[6]

$$\min_{\boldsymbol{\alpha}} \int_a^b w(x) \left[f(x) - \sum_{i=0}^{n} \alpha_i \varphi_i(x) \right]^2 dx. \tag{8.28}$$

The first order conditions for this problem are

$$0 = 2 \int_a^b w(x) \left[f(x) - \sum_{j=0}^{n} \alpha_j \varphi_j(x) \right] \varphi_i(x)dx, \ i = 0, 1, \dots, n,$$

which may be rewritten as

[5] Intuitively, the qualifier 'almost everywhere' allows $w(x)$ to be non-positive on a very small set of points. This set must be so small that its size – technically, its measure – equals zero.

[6] This is called a continuous least squares approximation of $f(x)$.

$$\int_a^b w(x)f(x)\varphi_i(x)dx = \sum_{j=0}^{n} \alpha_j \int_a^b w(x)\varphi_j(x)\varphi_i(x)dx, \quad i = 0, 1, \ldots, n. \tag{8.29}$$

If the integral on the rhs of (8.29) vanishes for $i \neq j$ and equals a constant ζ_j for $i = j$, it will be easy to compute α_i from

$$\alpha_i = \frac{1}{\zeta_i} \int_a^b w(x)f(x)\varphi_i(x)dx.$$

This motivates the following definition of orthogonal polynomials: A set of functions $\mathcal{P}$ is called orthogonal with respect to the weight function $w(x)$ if and only if:

$$\int_a^b w(x)\varphi_i(x)\varphi_j(x)dx = \begin{cases} 0 & \text{if } i \neq j, \\ \zeta_i & \text{if } i = j. \end{cases} \tag{8.30}$$

If in addition $\zeta_i = 1 \, \forall i$, the set of functions is said to be orthonormal. Among the families of orthonormal polynomials are the Chebyshev polynomials used extensively in Chapter 4.

8.2.4 Chebyshev Polynomials

Definition. The domain of Chebyshev polynomials is the interval $[-1, 1]$, and the i-th member of this family is defined by

$$T_i(x) = \cos(i \arccos x). \tag{8.31}$$

The weight function for which $\int_{-1}^{1} w(x)T_i(x)T_j(x)dx = 0$ for $i \neq j$ is given by

$$w(x) := \frac{1}{\sqrt{1 - x^2}}. \tag{8.32}$$

In particular, the following holds:

$$\int_{-1}^{1} \frac{T_i(x)T_j(x)}{\sqrt{1 - x^2}} \, dx = \begin{cases} 0 & \text{if } i \neq j, \\ \frac{\pi}{2} & \text{if } i = j \geq 1, \\ \pi & \text{if } i = j = 0. \end{cases} \tag{8.33}$$

Thus, if we use

$$\hat{g}(\boldsymbol{\alpha}, x) := \frac{1}{2}\alpha_0 + \sum_{i=1}^{n} \alpha_i T_i(x) \tag{8.34}$$

to approximate $g(x) \in C[-1, 1]$ the coefficients of the continuous least squares approximation are given by

$$\alpha_i = \frac{2}{\pi} \int_{-1}^{1} \frac{g(x)T_i(x)}{\sqrt{1-x^2}} dx. \tag{8.35}$$

Notice that in (8.34), α_0 is multiplied by the factor $1/2$ so that (8.35) also holds for $i = 0$ where the integral (8.33) for $i = j = 0$ is equal to $\zeta_0 = \pi$.

Most often, of course, we are interested in approximating a function f on the interval $[a, b]$ where a and b do not necessarily coincide with the values -1 and 1, respectively. Suppose that we have a function $f(z)$, $f : [a, b] \to R$ and want to compute a polynomial approximation over $[a, b]$ that corresponds to the Chebyshev approximation over $x \in [-1, 1]$ with weighting function (8.32). This can simply be done by defining the transformation

$$X(z) = \frac{2z}{b-a} - \frac{a+b}{b-a}, \quad z \in [a, b] \tag{8.36}$$

and the reverse transformation

$$Z(x) = \frac{(x+1)(b-a)}{2} + a, \quad x \in [-1, 1]. \tag{8.37}$$

With these transformations, we can define the function $g(x) = f(Z(x))$ on the interval $[-1, 1]$ with approximation

$$\hat{f}(z; \alpha) = \sum_{i=0}^{n} \alpha_i T_i\left(X(z)\right). \tag{8.38}$$

The coefficients of the continuous least squares approximation are then given by

$$\alpha_j = \frac{2}{\pi} \int_{a}^{b} \frac{f(z)T_j(X(z))}{\sqrt{1-(X(z))^2}}\, dz. \tag{8.39}$$

Properties. Chebyshev polynomials (8.31) have many other properties which make them a prime candidate in numerical applications. Importantly, we can compute a Chebyshev polynomial of order $i+1$, if we know the values of the Chebyshev polynomials of order i and $i-1$.[7] This property of the Chebyshev polynomial family helps economizing on computational time. In particular, the Chebyshev polynomials satisfy the following recursive scheme:

$$T_{i+1}(x) = 2xT_i(x) - T_{i-1}(x). \tag{8.40}$$

The recurrence relation (8.40) can be shown by introducing the substitution $\theta = \arccos x$. With this substitution, (8.31) can be rewritten as

$$T_i(\theta(x)) = T_i(\theta) = \cos(i\theta). \tag{8.41}$$

Furthermore, applying the trigonometric identities

$$\begin{aligned} T_{i+1}(\theta) &= \cos((i+1)\theta) = \cos(i\theta)\cos\theta - \sin(i\theta)\sin\theta, \\ T_{i-1}(\theta) &= \cos((i-1)\theta) = \cos(i\theta)\cos\theta + \sin(i\theta)\sin\theta, \end{aligned}$$

we get

$$T_{i+1}(\theta) = 2\cos(i\theta)\cos\theta - T_{i-1}(\theta),$$

or

$$T_{i+1} = 2xT_i(x) - T_{i-1}(x). \tag{8.42}$$

The first four Chebyshev polynomials are

$$\begin{aligned} T_0(x) &= \cos(0 \cdot \arccos x) = 1, \\ T_1(x) &= \cos(1 \cdot \arccos x) = x, \\ T_2(x) &= 2xT_1(x) - T_0(x) = 2x^2 - 1, \\ T_3(x) &= 2xT_2(x) - T_1(x) = 4x^3 - 3x. \end{aligned}$$

Notice that the Chebyshev polynomial $T_i(x)$ is a polynomial of degree i with leading coefficient 2^{i-1}. The Chebyshev polynomials T_1, T_2, and T_3 are displayed in Figure 8.2.

[7] There exist different families which satisfy such a recursive scheme, e.g. the Legendre or Hermite polynomials.

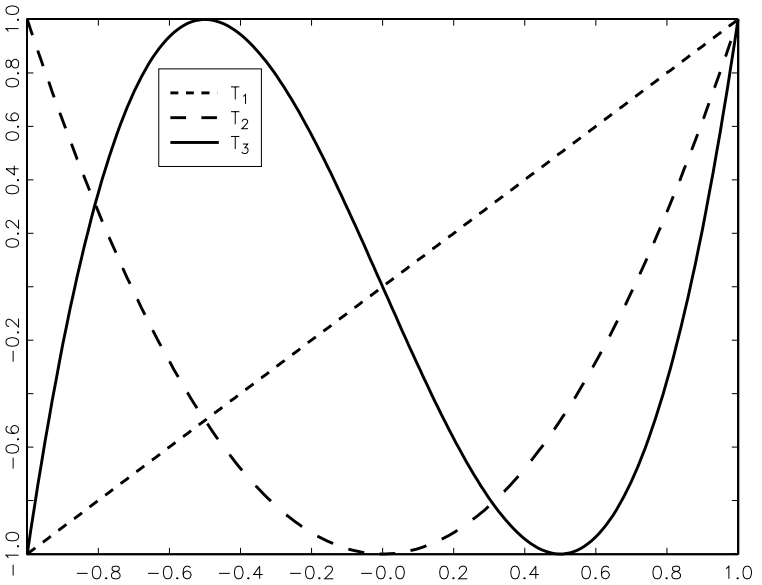

Figure 8.2: Chebyshev polynomials T_1, T_2 and T_3

The recursive formula (8.40) yields an efficient way to evaluate the polynomial at given point x, which we present in the following algorithm:

Algorithm 8.2.1 (Chebyshev Evaluation)

Purpose: *Evaluate a n-th degree Chebyshev polynomial at x*

Steps:

Step 1: Initialize: use the $n+1$*-vector* $\mathbf{y}$ *to store the values of* $T_i(x)$ *for* $i = 0, 1, \ldots, n$*. Put* $y[1] = 1$ *and* $y[2] = x$*.*

Step 2: For $i = 2, 3, \ldots, n-1$ *compute:*

$$y_{i+1} = 2xy_i - y_{i-1}.$$

Step 3: Return $\sum_{i=0}^{n} a_i y_i$*.*

In our GAUSS toolbox file you will find the procedure `ChebEval1` that implements this algorithm.

Zeros. Consider the sequence of points

$$\bar{x}_k := \cos\left(\frac{2k-1}{2n}\pi\right), \quad k = 1, 2, \ldots, n. \tag{8.43}$$

Using the substitution $\theta_k = \arccos(x_k)$ we get $n\theta_k = \pi(2k - 1)/2$. Since the cosine function cuts the abscissa at $\pi/2$, $(3/2)\pi$, $(5/2)\pi, \ldots$, the points $\bar{x}_k$ are the zeros of the n-th degree Chebyshev polynomial. Now, for $i, j < n$, Chebyshev polynomials satisfy a discrete version of the orthogonality relation:

$$\sum_{k=1}^{n} T_i(\bar{x}_k)T_j(\bar{x}_k) = \begin{cases} 0 & i \neq j, \\ m/2 & i = j \neq 0, \\ m & i = j = 0. \end{cases} \tag{8.44}$$

Computation of the Chebyshev coefficients. There are three possible ways. First, we may compute the integral on the rhs of equation (8.39) using the techniques presented in 8.3.2. Second, we may choose the n-dimensional vector $\boldsymbol{\alpha}$ such that f and $\hat{f}$ coincide at n points. Third, we can determine the coefficients from a regression of $f(x_i)$ on x_i, $i = 1, 2, \ldots, m$, using $m > n$ points. Since the second approach is a special case of the third for $m = n$, we consider the latter.

It can be shown that the maximal interpolation error attains a minimum, if the n interpolation nodes coincide with the n zeros of the n-th degree Chebyshev polynomial $T_n(x)$:[8] Using the $m > n$ zeros of $T_m(x)$ in a regression produces an even smoother approximation. Assume we want to approximate $f(x) \in C[a, b]$ by a n-th degree Chebyshev polynomial. Let $\bar{x} = [\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_m]$ denote the m zeros of $T_m(x)$. The corresponding points in the interval $[a, b]$ are $\bar{z}_k := Z(\bar{x}_k)$ (see (8.37)), so that $\bar{y}_k = f(\bar{z}_k)$. We choose $\boldsymbol{\alpha} = [\alpha_0, \alpha_1, \ldots, \alpha_{n-1}]$ to solve

$$\min_{\boldsymbol{\alpha}} \quad \sum_{k=1}^{m} \left[\bar{y}_k - \sum_{j=0}^{n-1} \alpha_j T_j(X(\bar{z}_k))\right]^2 .$$

[8] For a formal proof of this minimax property of Chebyshev zeros see any introductory textbook on numerical analysis such as Burden/Faires (1993).

The respective first order conditions are:

$$\begin{aligned}
\sum_{k=1}^{m} T_0 \bar{y}_k &= \sum_{j=0}^{n-1} \alpha_j \sum_{k=1}^{m} T_0 T_j(\bar{x}_k) = m\alpha_0, \\
\sum_{k=1}^{m} \bar{y}_k T_1(\bar{x}_k) &= \sum_{j=0}^{n-1} \alpha_j \sum_{k=1}^{m} T_1(\bar{x}_k) T_j(\bar{x}_k) = (m/2)\alpha_1, \\
&\vdots = \vdots, \\
\sum_{k=1}^{m} \bar{y}_k T_{n-1}(\bar{x}_k) &= \sum_{j=0}^{n-1} \alpha_j \sum_{k=1}^{m} T_{n-1}(\bar{x}_k) T_j(\bar{x}_k) = (m/2)\alpha_{n-1},
\end{aligned} \tag{8.45}$$

where the respective rhs follow from (8.44). This provides the following algorithm:

Algorithm 8.2.2 (Chebyshev Regression)

Purpose: *Approximate* $f(z) \in C[a,b]$ *with* $\sum_{j=0}^{n-1} \alpha_j T_j(x)$.

Steps:

Step 1: Choose the degree $n-1$ *of the approximating Chebyshev polynomial. Compute* $m \geq n$ *Chebyshev interpolation nodes* $\bar{x}_k$ *from* (8.43) *and adjust the nodes to the interval* $[a,b]$ *using* (8.37).

Step 2: For $k = 1, 2, \ldots, m$, *compute* $\bar{y}_k = f(\bar{z}_k)$.

Step 3: Compute the Chebyshev coefficients: $\alpha_0 = (1/m) \sum_{k=1}^{m} \bar{y}_k$. *For* $i = 1, 2, \ldots, n-1$ *the coefficients are given by:*

$$\alpha_i = \frac{2}{m} \sum_{k=1}^{m} \bar{y}_k T_i(\bar{x}_k).$$

The GAUSS program `ChebCoef` implements this approach.

Examples. In order to demonstrate the performance of the computation of Chebyshev coefficients we compute the Chebyshev approximation of $f(x) = e^x$ and of

$$g(x) = \begin{cases} 0 & \text{if } x < 1, \\ (x-1) & \text{if } x \geq 1. \end{cases}$$

The latter type of function might often be encountered in economic problems with constraints. In Chapter 3, gross investment is assumed to be nonnegative. As another example, assume that agents supply labor elastically, instead. If the wage is below unemployment compensation, they do not work and labor supply is equal to zero. For a wage exceeding unemployment compensation, they supply labor and, if the income effect is less than the substitution effect, labor supply increases with the wage rate. The optimal labor supply may look similar to the function $g(x)$.

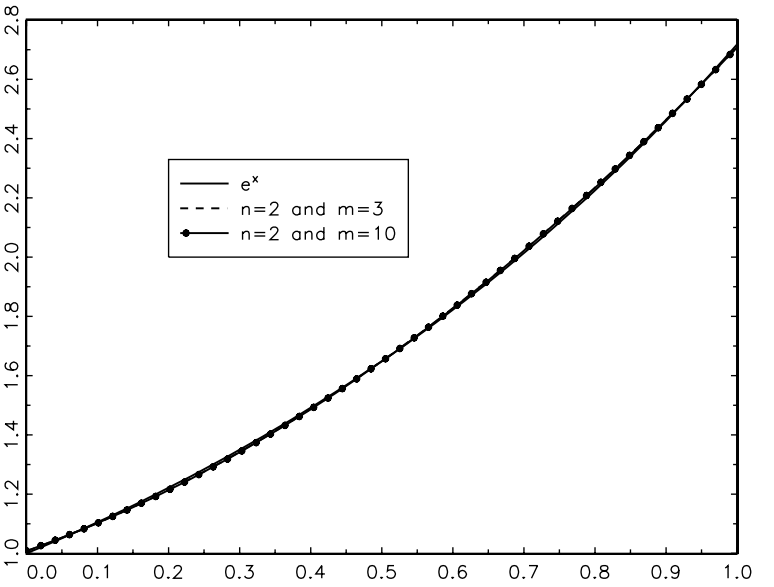

Figure 8.3: Chebyshev Approximation of e^x

Figure 8.3 depicts e^x and its second degree Chebyshev polynomial approximation. If we use the 3 zeros of T_3 as interpolation nodes, the maximum absolute error between e^x and its approximation in the interval $[0, 1]$ is 0.0099. With 10 data points it drops to 0.0093. However, if we use the fifth degree approximation with 10 data points the error drops below $(1.17) \times 10^{-6}$. In this case, the

graph of e^x and the graph of its approximation virtually coincide in Figure 8.3.

Table 8.1

Coefficient	$n = 2$, $m = 10$	$n = 5$, $m = 10$
α_0	1.753	1.753
α_1	0.850	0.850
α_2	0.105	0.105
α_3		0.0087
α_4		0.0005
α_5		0.000027

Table 8.1 shows that the value of the Chebyshev coefficients drop off rapidly. This result is not surprising. In fact, one can show the following theorem:[9]

Theorem 8.2.3 *If $f \in C^k[-1, 1]$ has a Chebyshev representation $f(x) = \sum_{i=1}^{\infty} \alpha_i T_i(x)$, then there is a constant c such that*

$$|\alpha_i| \leq \frac{c}{i^k}, \quad i \geq 1. \tag{8.46}$$

This theorem also gives a hint for the choice of the degree n of the approximating polynomial. If the α_i are falling rapidly and if α_n is small, then we can be more confident to ignore higher-order polynomial terms. Notice further that the values of the first 4 coefficients do not change if we increase the degree of the Chebyshev polynomial from $n = 2$ to $n = 5$. Of course, this is obvious from (8.45) if we keep m constant.

As we have just learned, smooth functions can be approximated quite accurately by Chebyshev polynomials. However, Chebyshev approximation is less apt for the approximation of functions displaying a kink, like the function $g(x) = \max(0, x - 1)$, or step functions. The function $g(x)$ is not differentiable at $x = 1$. The approximation by Chebyshev interpolation (i.e., $m = n + 1$) and

[9] See JUDD (1998), Theorem 6.4.2.

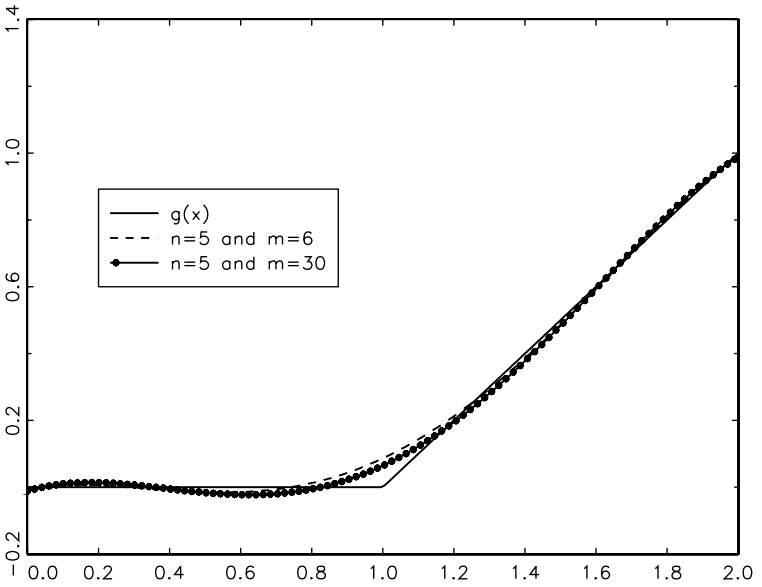

Figure 8.4: Chebyshev Approximation of $\max\{0, x-1\}$

Chebyshev regression are displayed in Figure 8.4. Notice that with regression, we are better able to approximate the kink at $x = 1$ than with interpolation. As the degree of the Chebyshev polynomial increases, the approximation is getting closer.

8.2.5 Multidimensional Approximation

Choice of Bases. Even in the simple stochastic growth model the domain of the policy function is already a subset of $\mathbb{R}^2$. So how can we generalize the polynomial approximation to the n-dimensional case? One approach is to use the n-fold tensor product base.Let x_i denote the i-th element of $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and use $\varphi_k(x_i)$ for the k-th member of a family of polynomials. For instance $\varphi_k(x_i) \equiv x_i^k$. The set

$$\mathscr{T} := \left\{ \prod_{i=1}^{n} \varphi_{k_i}(x_i) | k_i = 0, 1, \ldots p \right\}$$

is the n-fold tensor product base. The linear combination of the $(1+p)^n$ elements of $\mathscr{T}$ can be used to approximate $f(\mathbf{x})$. For

instance, in the stochastic growth model $\mathbf{x} = (K, Z)$, and for $p = 2$ and $\varphi_k(x_i) = x_i^k$ the set $\mathscr{T}$ is given by

$$\{K^0Z^0, K^0Z^1, K^0Z^2, K^1Z^0, K^1Z^1, K^1Z^2, K^2Z^0, K^2Z^1, K^2Z^2\}$$
$$\equiv \{1, Z, Z^2, K, KZ, KZ^2, K^2, K^2Z, K^2Z^2\}.$$

This set grows exponentially with the dimension n of $\mathbf{x}$. A smaller set that delivers as good an approximation (in terms of asymptotic convergence) as the tensor product base is the complete set of polynomials of degree p in n variables, denoted $\mathscr{P}_p^n$. This set is derived by considering the p-th order approximation of $f(\mathbf{x})$. As we know from Theorem 8.2.2, this approximation involves the following products:

$$\mathscr{P}_p^n = \left\{ (x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}) \middle| \sum_{i=1}^{n} k_i = j, k_i \geq 0, j = 0, 1, 2, \ldots, p \right\}$$

Thus, for $\mathbf{x} = (K, Z)$ and $p = 2$, the set is

$$\mathscr{P}_2^2 = \{\underbrace{K^0Z^0}_{j=0}, \underbrace{K^1Z^0, K^0Z^1}_{j=1}, \underbrace{K^2Z^0, K^1Z^1, K^0Z^2}_{j=2}\},$$
$$= \{1, K, Z, K^2, KZ, Z^2\}.$$

More generally, we can build a complete set of polynomials of total degree p in n variables from any family of polynomials by replacing $x_i^{k_i}$ with $\varphi_{k_i}(x_i)$ in the definition of $\mathscr{P}_p^n$.

Chebyshev Approximation in Two Dimensions. The algorithm to compute Chebyshev approximation in multidimensional space is derived in a straightforward manner with the help of least squares and we will illustrate it for the two-dimensional case only. Let $f(z_1, z_2)$ be a function on $[a, b] \times [c, d]$ that we would like to approximate by a two-dimensional Chebyshev polynomial[10]

$$\sum_{j_1=0}^{n1} \sum_{j_2=0}^{n2} \alpha_{j_1 j_2} T_{j_1}(X(z_1)) T_{j_2}(X(z_2)). \tag{8.47}$$

[10] Remember, $X(z)$ is the linear transformation of points in $[a, b]$ or $[c, d]$ to $[-1, 1]$ defined in (8.36).

We need $m_1 \geq n_1 + 1$ and $m_2 \geq n_2 + 1$ points and choose them as the zeros of the m_1-dimensional and m_2-dimensional Chebyshev polynomial, adjusted to the interval $[a, b]$ and $[c, d]$, respectively. Let

$$\bar{y}_{k_1 k_2} := f(Z(\bar{x}_{1k_1}), Z(\bar{x}_{2k_2})),$$

and consider the least squares criterion

$$\min \quad \sum_{k_1=1}^{m_1} \sum_{k_2=1}^{m_2} \left[\bar{y}_{k_1 k_2} - \sum_{j_1=0}^{n_1} \sum_{j_2=0}^{n_2} \alpha_{j_1 j_2} T_{j_1}(\bar{x}_{1k_1}) T_{j_2}(\bar{x}_{2k_2}) \right]^2 .$$

The first order condition with respect to $\alpha_{j_1 j_2}$ yields:

$$\begin{aligned}
&\sum_{k_1=1}^{m_1} \sum_{k_2=1}^{m_2} \bar{y}_{k_1 k_2} T_{j_1}(\bar{x}_{1k_1}) T_{j_2}(\bar{x}_{2k_2}) \\
&= \sum_{k_1=1}^{m_1} \sum_{k_2=1}^{m_2} \sum_{i_1=0}^{n_1} \sum_{i_2=0}^{n_2} \alpha_{i_1 i_2} T_{i_1}(\bar{x}_{1k_1}) T_{i_2}(\bar{x}_{2k_2}) T_{j_1}(\bar{x}_{1k_1}) T_{j_2}(\bar{x}_{2k_2}) \\
&= \sum_{i_1=0}^{n_1} \sum_{i_2=0}^{n_2} \alpha_{i_1 i_2} \underbrace{\sum_{k_1=1}^{m_1} T_{i_1}(\bar{x}_{1k_1}) T_{j_1}(\bar{x}_{1k_1})}_{\substack{= 0 \quad \text{if} \quad j_1 \neq i_1 \\ = m_1/2 \ \text{if} \ j_1 = i_1 \neq 0 \\ = m_1 \quad \text{if} \ j_1 = i_1 = 0}} \underbrace{\sum_{k_2=1}^{m_2} T_{i_2}(\bar{x}_{2k_2}) T_{j_2}(\bar{x}_{2k_2})}_{\substack{= 0 \quad \text{if} \quad j_2 \neq i_2 \\ = m_2/2 \ \text{if} \ j_2 = i_2 \neq 0 \\ = m_2 \quad \text{if} \ j_2 = i_2 = 0}} .
\end{aligned}$$

Therefore, we get the following estimator:

$$\begin{aligned}
\alpha_{00} &= \frac{1}{m_1 m_2} \sum_{k_1=1}^{m_1} \sum_{k_2=1}^{m_2} \bar{y}_{k_1 k_2}, \\
\alpha_{0 j_2} &= \frac{1}{m_1} \frac{2}{m_2} \sum_{k_1=1}^{m_1} \sum_{k_2=1}^{m_2} \bar{y}_{k_1 k_2} T_{j_2}(\bar{x}_{2k_2}), \\
\alpha_{j_1 0} &= \frac{2}{m_1} \frac{1}{m_2} \sum_{k_1=1}^{m_1} \sum_{k_2=1}^{m_2} \bar{y}_{k_1 k_2} T_{j_1}(\bar{x}_{1k_1}), \\
\alpha_{j_1 j_2} &= \frac{2}{m_1} \frac{2}{m_2} \sum_{k_1=1}^{m_1} \sum_{k_2=1}^{m_2} \bar{y}_{k_1 k_2} T_{j_1}(\bar{x}_{1k_1}) T_{j_2}(\bar{x}_{2k_2}).
\end{aligned} \tag{8.48}$$

The extension to more than two-dimensional state spaces should be straightforward.

8.2.6 Neural Networks

Instead of linear combinations of polynomials neural networks use non-linear approximation schemes. A single-layer neural network is a function of the form

$$\Phi(\mathbf{a}, \mathbf{x}) := h\left(\sum_{i=1}^{n} a_i g(x_i)\right),$$

where h and g are scalar functions. In the left panel of Figure 8.5 the first row of nodes represents the function g processing the inputs x_i. The result is aggregated via summation, as indicated by the arrows to the single node which represents the function h that delivers the final output y. In the single hidden-layer feedforward network displayed in the right panel of Figure 8.5 the function g delivers its output to a second row of nodes. There, this input is processed by another function G, say, before it is aggregated and passed on to the function h.

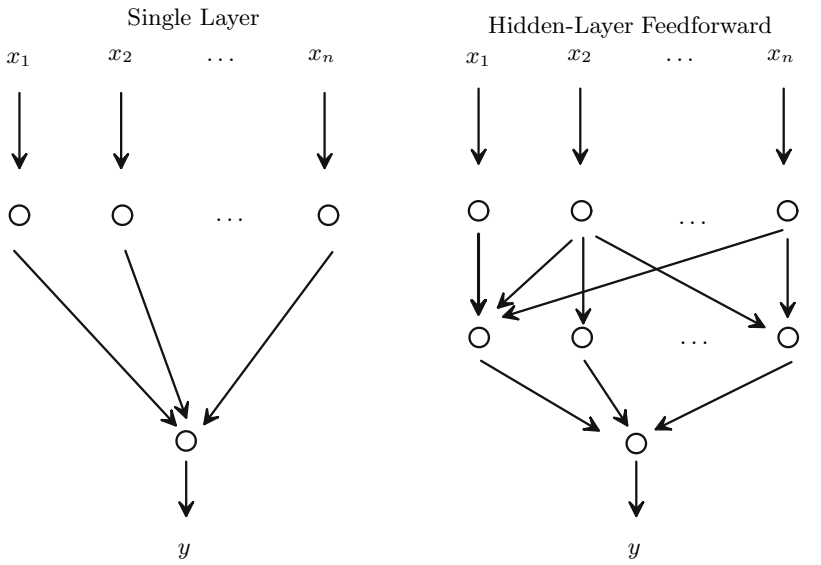

Figure 8.5: Neural Networks

Formally, the single hidden-layer feedforward network is given by:

$$\Phi(\mathbf{a}, \mathbf{b}, \mathbf{x}) := h\left(\sum_{j=1}^{m} b_j G\left[\sum_{i=1}^{n} a_{ij} g(x_i)\right]\right).$$

The function G is called the hidden-layer activation function. A common choice for G is the sigmoid function

$$G(x) = \frac{1}{1 + e^{-x}}.$$

Neural networks are efficient functional forms for approximating multidimensional functions. Often they require less parameters for a given accuracy than polynomial approximations.[11]

In the case of the stochastic growth model, which we considered in Section 3.3.1, Duffy and McNelis (2001) approximate the conditional expectation that appears in the Euler equation (3.13a) by the following single hidden-layer feedforward neural network with six parameters:

$$\Phi(\boldsymbol{\gamma}, K, Z) = \gamma_1 + \frac{\gamma_2}{1 + e^{-\gamma_3 K - \gamma_4 Z}} + \frac{1}{1 + e^{-\gamma_5 K - \gamma_6 Z}}.$$

8.3 Numerical Differentiation and Integration

8.3.1 Differentiation

Many of our algorithms require derivatives. Think of the Newton-Raphson algorithm or the linear quadratic approximation method. In the simplest case of a real valued function of one variable, $y = f(x)$, the obvious choice is the analytical derivative $f'(x)$. Yet, if $f'(x)$ is given by a complicated formula, mistakes easily sneak into the computer code. It may even be impossible to derive an explicit expression for the derivative at all. Think of the sum of squares in the case of the parameterized expectations approach. If the Jacobian of a vector valued function or the Hesse

[11] See Sargent (1993), p. 58f and the literature cited there.

matrix of a function of many independent variables are required, analytical derivatives – if available at all – require many lines of computer code; something which is failure-prone. Therefore, we use numerical derivatives in almost all of our computer programs.

This section provides some background on numerical differentiation and presents two algorithms that approximate the Jacobian matrix of a vector valued function and the Hesse matrix of a real valued function, respectively. The related program code can be used in place of built-in routines, as, e.g., `Gradp` and `Hessp` in GAUSS or `DFDJAC` and `DFDHES` from the IMSL library of Fortan subroutines.

First Difference Formulas. The basis of numerical derivative formulas is Taylor's Theorem. Consider Theorem 8.2.1 for the case $n = 1$. We get

$$f(\bar{x} + h) = f(\bar{x}) + f^{(1)}(\bar{x})h + f^{(2)}(\xi)\frac{h^2}{2}. \tag{8.49}$$

Thus, we may approximate the first derivative by the formula[12]

$$D_{FD}f(\bar{x}, h) := \frac{f(\bar{x} + h) - f(\bar{x})}{h}. \tag{8.50}$$

This is known as the forward difference formula. The approximation error is proportional to h, since from (8.49):

$$\left|D_{FD}f(\bar{x}, h) - f^{(1)}(\bar{x})\right| = \left|f^{(2)}(\xi)/2\right| h.$$

Thus, the error is of first order. The backward difference formula derives from Taylor's theorem for $-h$ in place of h. Its error is also of first order. Now consider Taylor's theorem for $n = 2$, h, and $-h$:

$$f(\bar{x} + h) = f(\bar{x}) + f^{(1)}(\bar{x})h + f^{(2)}(\bar{x})\frac{h^2}{2} + f^{(3)}(\xi_1)\frac{h^3}{6}, \tag{8.51a}$$

$$f(\bar{x} - h) = f(\bar{x}) - f^{(1)}(\bar{x})h + f^{(2)}(\bar{x})\frac{h^2}{2} - f^{(3)}(\xi_2)\frac{h^3}{6}, \tag{8.51b}$$

[12] As in Section 8.2.1, we use the symbol D^i to denote the i-th derivative of f. Subscripts denote the kind of approximation.

and subtract the second line from the first. The quadratic term disappears and from

$$f(\bar{x}+h) - f(\bar{x}-h) = 2f^{(1)}(\bar{x}) + \left(f^{(3)}(\xi_1) + f^{(3)}(\xi_2)\right)\frac{h^3}{6}$$

we find the approximation

$$D_{CD}f(\bar{x},h) := \frac{f(\bar{x}+h) - f(\bar{x}-h)}{2h} \tag{8.52}$$

known as central difference formula. Letting C denote the maximum of $(f^{(3)}(\xi_1)+f^{(3)}(\xi_2))/6$ in $[\bar{x},\bar{x}+h]$, we see that the approximation error is proportional to Ch^2 and, thus, of second order. When we add equation (8.51a) to equation (8.51b) the first derivative terms cancel and we get the central difference formula for the second derivative:

$$D^2_{CD}f(\bar{x},h) := \frac{f(\bar{x}+h) + f(\bar{x}-h) - 2f(\bar{x})}{h^2}, \tag{8.53}$$

whose approximation error is bound by Ch and, thus, of first order.

Choice of h. From the previous discussion it might seem to be a good idea to choose h as small as possible. But remember the finite precision of computer arithmetic. Suppose your PC is able to represent, say, the first 10 digits to the right of the decimal point of any floating point number correctly. If h is too small, the first and second term in the numerator of equation (8.50) may differ only in the eleventh digit and the computed derivative is highly unreliable.

Suppose the error in computing $f(\bar{x})$ and $f(\bar{x}+h)$ is $\bar{e}$ and e_h, respectively. At least, $\bar{e}$ and e_h equal the machine epsilon ϵ, i.e., the smallest positive number for which the statement $1+\epsilon > 1$ is true on your machine. However, if $f(x)$ is the result of complicated and involved computations, the actual error may be much larger. We want to find an upper bound on the total error $E(\delta,h)$ that results when we use $\tilde{f}(\bar{x}) := f(\bar{x})+\bar{e}$ and $\tilde{f}(\bar{x},h) := f(\bar{x}+h)+e_h$ to compute $D_{FD}f(\bar{x},h)$, where $\bar{e}, e_h \leq \delta$ for some $\delta \geq \epsilon$.

$$
\begin{aligned}
E(\delta, h) &:= \left| f'(\bar{x}) - \frac{\tilde{f}(\bar{x}) - \tilde{f}(\bar{x}+h)}{h} \right| \\
&\leq \underbrace{|f'(\bar{x}) - D_{FD} f(\bar{x}, h)|}_{\leq Ch} + \underbrace{\frac{|\bar{e} - e_h|}{h}}_{\leq 2\delta/h}, \\
&\leq Ch + \frac{2\delta}{h}, \quad C := \max_{\xi \in [\bar{x}, \bar{x}+h]} \frac{f^2(\xi)}{2}.
\end{aligned}
$$

Setting the derivative of this upper bound with respect to h to zero and solving for h gives the step size that provides the smallest upper bound:

$$
h^* = \sqrt{\frac{2\delta}{C}}. \tag{8.54}
$$

If we perform the same exercise with respect to the central difference formulas (8.52) and (8.53) we find that the optimal choice of h is

$$
h^{**} = \sqrt[3]{\frac{2\delta}{C}}, \quad C := \max_{\xi_1, \xi_2 \in [\bar{x}, \bar{x}+h]} \frac{f^{(3)}(\xi_1) + f^{(3)}(\xi_2)}{6}. \tag{8.55}
$$

Computation of the Jacobian. It is easy to apply the above results to a vector valued function $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$. Let $f^i(\mathbf{x}), \mathbf{x} = [x_1, x_2, \ldots, x_n]$ denote the i-th component function of $\mathbf{f}$. Using the central difference formula (8.52) we may approximate the element f^i_j of the Jacobian matrix at the point $\bar{\mathbf{x}}$ by

$$
f^i_j := \frac{\partial f^i(\bar{\mathbf{x}})}{\partial x_j} \simeq \frac{f(\bar{x} + \mathbf{e}_j h) - f(\bar{x} - \mathbf{e}_j h)}{2h}, \tag{8.56}
$$

where $\mathbf{e}_j$ is the unit (row) vector with one in the j–th position and zeros elsewhere.

If the x_i differ considerably in size, we set h proportional to x_j using

$$
h_j = h^{**} \max\{|x_j|, 1\}. \tag{8.57}
$$

Our program `CDJac` uses equation (8.56) together with this choice of h_j (and $h^{**} = \sqrt[3]{\epsilon}$ as default) to compute the Jacobian of a user supplied routine that evaluates $f^i(\bar{x}), i = 1, 2, \ldots, m$ at $\bar{x}$.

Computation of the Hesse Matrix. Suppose we want to compute the elements of the Hesse matrix of $f : \mathbb{R}^n \to \mathbb{R}$ given by

$$H(\bar{x}) := \begin{bmatrix} h_{11} & h_{12} & \dots & h_{1n} \\ h_{21} & h_{22} & \dots & h_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ h_{n1} & h_{n2} & \dots & h_{nn} \end{bmatrix}, \quad h_{ij} := \frac{\partial^2 f(\bar{x})}{\partial x_i \partial x_j}.$$

There are two possibilities. Note, that the Hesse matrix equals the Jacobian of the gradient of f. Thus, if an analytic expression for the gradient of f is easy to program, one can use this as an input to a procedure that approximates the Jacobian. This gives a better approximation than the use of difference formulas for second partial derivatives.[13] If this is not an option, one can apply the central difference formula for the second derivative of a function in one variable to compute the diagonal elements of H.[14] This gives:

$$h_{ii} \simeq \frac{f(\bar{x} + \mathbf{e}_i h_i) + f(\bar{x} - \mathbf{e}_i h_i) - 2f(\bar{x})}{h_i^2}. \tag{8.58}$$

There are several choices for the off-diagonal elements of H. From a third order expansion of f at $\bar{x}$, we can get the following equations:[15]

$$\begin{aligned} f(\bar{x} + h_i + h_j) &= f(\bar{x}) + f_i(\bar{x})h_i + f_j(\bar{x})h_j + (1/2)h_{ii}h_i^2 \\ &\quad + (1/2)h_{jj}h_j^2 + h_{ij}h_ih_j + C_1, \\ f(\bar{x} + h_i - h_j) &= f(\bar{x}) + f_i(\bar{x})h_i - f_j(\bar{x})h_j + (1/2)h_{ii}h_i^2 \\ &\quad + (1/2)h_{jj}h_j^2 - h_{ij}h_ih_j + C_2, \\ f(\bar{x} - h_i + h_j) &= f(\bar{x}) - f_i(\bar{x})h_i + f_j(\bar{x})h_j + (1/2)h_{ii}h_i^2 \\ &\quad + (1/2)h_{jj}h_j^2 - h_{ij}h_ih_j + C_3, \\ f(\bar{x} - h_i - h_j) &= f(\bar{x}) - f_i(\bar{x})h_i - f_j(\bar{x})h_j + (1/2)h_{ii}h_i^2 \\ &\quad + (1/2)h_{jj}h_j^2 + h_{ij}h_ih_j + C_4, \end{aligned}$$

[13] The error of the central difference formula for the first derivative is of second order, whereas the error from the central difference formula for the second derivative is of first order.

[14] In the following we use h_i proportional to $\max\{|x_i|, 1\}$ as in (8.57).

[15] See, e.g., JUDD (1998), p. 239, for Taylor's formula in the case of many independent variables.

where $C_k, k = 1, 2, 3, 4$ are sums of the mixed third partial derivatives of f and of third order.[16] If we add the first and the last equation and subtract the second and third equation from this sum, we find the following four-point formula:

$$h_{ij} \simeq \frac{1}{4h_ih_j}\Big[f(\bar{x} + \mathbf{e}_ih_i + \mathbf{e}_jh_j) - f(\bar{x} - \mathbf{e}_ih_i + \mathbf{e}_jh_j) \\ - f(\bar{x} + \mathbf{e}_ih_i - \mathbf{e}_jh_j) + f(\bar{x} - \mathbf{e}_ih_i - \mathbf{e}_jh_j)\Big], \tag{8.59}$$

whose approximation error is bound by $Ch := h\sum_k |C_k|$. Our GAUSS procedure `CDHesse` uses (8.58) and (8.59). This agrees with the suggestion of Hansen and Prescott (1995) for the quadratic approximation of the current period return function.

8.3.2 Numerical Integration

Newton-Cotes Formulas. Basically, there are two different approaches to compute an integral $\int_a^b f(x)dx$ numerically.[17] The first idea is to approximate the function $f(x)$ by piecewise polynomials and integrate the polynomials over subdomains of $[a, b]$. For example, the Trapezoid rule evaluates the function $f(x)$ at the end points $x = a$ and $x = b$ and uses the linear Lagrange polynomial

$$P_1(x) = \frac{x-b}{a-b}f(a) + \frac{x-a}{b-a}f(b) \tag{8.60}$$

to approximate $f(x)$. Integration of P_1 over $[a, b]$ results in the formula

$$\int_a^b f(x)dx \approx \frac{b-a}{2}\left[f(a) + f(b)\right]. \tag{8.61}$$

[16] For instance,

$$C_1 := (1/6)(f_{iii}h_i^3 + f_{jjj}h_j^3 + f_{iij}h_i^2h_j + f_{jji}h_ih_j^2),$$

where the third partial derivatives are evaluated at some point between $\bar{x}$ and $\bar{x} + \mathbf{e}_ih_i + \mathbf{e}_jh_j$ (see Theorem 8.2.2). Since one can always choose $h_i = ah, h_j = bh$ for some h, all terms on the rhs of the previous equation are of third order.

[17] In fact, there is a third approach that we do not pursue here. It considers the related problem to solve an ordinary differential equation.

If we use higher-order polynomials or a higher number of subdomains, more generally, we derive a Newton-Cotes formula for the approximation of the integral which evaluates the integral at a number of points:

$$\int_a^b f(x)dx \approx \sum_{i=1}^{n} a_i f(x_i). \tag{8.62}$$

Gaussian Formulas. In Newton-Cotes formulas, the coefficients a_i are chosen so that the local approximation is correct and the nodes x_i are chosen arbitrarily; usually, the x_i are equidistant. The second approach, which we will pursue in all quadrature applications of this book, is to choose both the weights a_i and the nodes x_i optimally in order to provide a good approximation of $\int_a^b f(x)dx$. It is obvious that we increase the degrees of freedom at our disposal if we choose both the nodes x_i and the weights a_i simultaneously rather than just the weights a_i in order to get a good approximation. Essentially, the resulting Gaussian quadrature formulas have twice the order than the Newton-Cotes formulas for the same number of function evaluations.[18]

The following theorem highlights the importance of orthogonal families of polynomials for numerical integration:[19]

Theorem 8.3.1 *Suppose that* $\{\varphi_i(x)\}_{i=0}^{\infty}$ *is an orthonormal family of polynomials with respect to the weight function* $w(x)$ *on* $[a, b]$ *with* $\varphi_i(x) = q_i x^i + q_{i-1}x^{i-1} + \ldots + q_0$. *Let* $\bar{x}_i$, $i = 1, \ldots, n$, *be the* n *zeros of* φ_n. *Then* $a < \bar{x}_1 < \cdots < \bar{x}_n < b$, *and if* $f \in C^{(2n)}[a, b]$, *then*

$$\int_a^b w(x)f(x)dx = \sum_{i=1}^{n} \omega_i f(\overline{x}_i) + \frac{f^{2n}(\zeta)}{q_n^2(2n)!}, \tag{8.63}$$

for some $\zeta \in [a, b]$, *where*

$$\omega_i = -\frac{q_{n+1}/q_n}{\varphi_n'(\bar{x}_i)\varphi_{n+1}(\bar{x}_i)} > 0.$$

[18] Notice, however, that higher order does not always translate into higher accuracy.

[19] See also Judd (1998), Theorem 7.2.1.

Accordingly, we can evaluate the integral of a polynomial of degree $2n-1$ exactly by applying formula (8.63). Usually, one does not have to compute the nodes and weights, since they are kept in tables. It is a nice property of the Chebyshev polynomials that their weights ω_i are constant. The Gauss-Chebyshev quadrature formula for a function $f(x)$ on the interval $[-1, 1]$ is defined by:

$$\int_{-1}^{1} \frac{f(x)}{\sqrt{1-x^2}} dx = \frac{\pi}{n} \sum_{i=1}^{n} f(\overline{x}_i) + \frac{\pi}{2^{2n-1}} \frac{f^{(2n)}(\zeta)}{(2n)!}$$

for some $\zeta \in [-1, 1]$, where the quadrature nodes $\overline{x}_i$ are the zeros of the Chebyshev polynomial $T_n(x)$ as presented in (8.43).

For integrals of the form $\int_a^b f(z)dz$, we use (8.37) to adjust $[-1, 1]$ to $[a, b]$. Since the linear transformation (8.37) implies

$$dz = \frac{b-a}{2} dx,$$

we can derive the following approximation:

$$\begin{aligned} \int_a^b f(z)dz &= \int_{-1}^{1} f(Z(x)) \frac{b-a}{2} \frac{\sqrt{1-x^2}}{\sqrt{1-x^2}} dx \\ &\approx \frac{\pi(b-a)}{2n} \sum_{i=1}^{n} f\left(\frac{(\overline{x}_i + 1)(b-a)}{2} + a \right) \sqrt{1-\overline{x}_i^2}, \end{aligned} \tag{8.64}$$

where the $\overline{x}_i$, again, are the Chebyshev zeros from (8.43).

Very often we have to compute (conditional) expectations. In these instances it is natural to refer to the Hermite polynomials, since their weight function is given by $w(x) := e^{-x^2}$. Suppose z is distributed normally with mean μ and variance σ^2. Then,

$$E(f(z)) := (2\pi\sigma^2)^{-1/2} \int_{-\infty}^{\infty} f(z) e^{[(z-\mu)^2/2\sigma^2]} dz.$$

Since

$$x = \frac{z-\mu}{\sqrt{2}\sigma}$$

has a standard normal distribution (i.e., $E(x) = 0$ and $\text{var}\,(x) = 1$), we get

$$E(f(z)) = \pi^{-1/2} \int_{-\infty}^{\infty} f\left(\sqrt{2}\sigma x + \mu\right) e^{-x^2} dx,$$

where we used

$$dz = \sqrt{2}\sigma dx.$$

This integral can be approximated by the Gauss-Hermite quadrature formula

$$E(f(z)) \simeq \pi^{-1/2} \sum_{i=1}^{n} \omega_i f\left(\sqrt{2}\sigma x + \mu\right). \tag{8.65}$$

For $n = 2, \ldots, 5$ the integration nodes and weights are given in Table 8.2.

Table 8.2

n	x_i	ω_i
2	−0.7071067811	0.8862269254
	0.7071067811	0.8862269254
3	−0.01224744871	0.2954089751
	0.0000000000	0.0118163590
	0.01224744871	0.2954089751
4	−0.01650680123	0.08131283544
	−0.5246476232	0.8049140900
	0.01650680123	0.08131283544
	0.5246476232	0.8049140900
5	−0.0202018287	0.01995324205
	−0.9585724646	0.3936193231
	0.0000000000	0.9453087204
	0.0202018287	0.01995324205
	0.9585724646	0.3936193231

Source: JUDD (1998), Table 7.4

Multidimensional Integration. Even the stochastic growth model has a state space of dimension two. When we use least squares or Galerkin projection to solve this model or models whose state space has an even higher dimension, we must compute multiple integrals. The most natural way to do this, is to use product rules. For instance, if we want to compute the integral of $f(z_1, \ldots, z_n)$ over the n-dimensional rectangle $[a_1, b_1] \times [a_2, b_2] \times \ldots [a_n, b_n]$ the Gauss-Chebyshev quadrature (8.64) implies

$$\begin{aligned} &\int_{a_1}^{b_1} \ldots \int_{a_n}^{b_n} f(z_1, \ldots, z_n) dz_1 \ldots dz_n \\ \simeq &\frac{\pi^n (b_1 - a_1) \ldots (b_n - a_n)}{(2m)^n} \sum_{i_1=1}^{m} \cdots \sum_{i_n=1}^{m} f(Z(\bar{x}_1), \ldots, Z(\bar{x}_n)) \\ &\times \sqrt{1 - \bar{x}_{i_1}^2} \ldots \sqrt{1 - \bar{x}_{i_n}^2}. \end{aligned} \tag{8.66}$$

In this formula $Z(x_i)$ denotes the linear transformation given in equation (8.37), and $\bar{x}_i$ is the i-th the zero of the Chebyshev polynomial of degree m. The problem with product rules is the curse of dimension. It requires m^n function evaluations to compute the approximate integral. If f itself is time consuming to evaluate, computational time becomes a binding constraint, even for n as small as 4. Monomial formulas are derived from the problem to exactly integrate a weighted product of monomials over a subset D of $\mathbb{R}^n$. A large number of specific formulas derive from this approach. A good source for those formulas is Strout (1971). A formula that is particularly helpful to compute the expectation $E(f(\mathbf{x}))$ if $\mathbf{x} \in \mathbb{R}^n$ has a multivariate standard normal distribution, is the following:

$$(2\pi)^{-n/2} \int_{\mathbb{R}^n} f(\mathbf{x}) e^{-\sum_{i=1}^n x_i^2} d\mathbf{x} \simeq \frac{1}{2n} \sum_{i=1}^{n} f(\pm\sqrt{n/2}\mathbf{e}_i), \tag{8.67}$$

where $\mathbf{e}_i = (0, \ldots, 0, 1, 0, \ldots 0)$ denotes the i-th unit vector. To apply this rule to the general case of a random normal vector

$\mathbf{z}$ with mean $\boldsymbol{\mu}$ and covariance matrix Σ, we use the change of variable technique.[20] The linear transformation

$$\mathbf{x} = \frac{\Sigma^{-1/2}}{\sqrt{2}}(\mathbf{z} - \boldsymbol{\mu})$$

implies that $E(f(\mathbf{z}))$ can be expressed as an integral function of $\mathbf{x}$:

$$\begin{aligned} E(f(\mathbf{z})) &= (2\pi)^{-n/2}|\Sigma|^{-1/2} \int_{\mathbb{R}^n} f(\mathbf{z}) e^{\frac{-1}{2}(\mathbf{z}-\boldsymbol{\mu})'\Sigma^{-1}(\mathbf{z}-\boldsymbol{\mu})} d\mathbf{z} \\ &= \pi^{-n/2} \int_{\mathbb{R}^n} f\left(\sqrt{2}\Sigma^{1/2}\mathbf{x} + \boldsymbol{\mu}\right) e^{-\sum_{i=1}^n x_i^2} d\mathbf{x}. \end{aligned}$$

This integral can be approximated by formula (8.67).

8.4 Stopping Criteria for Iterative Algorithms

Most, if not all of our algorithms are iterative. Think of the value function iterations presented in Chapter 1 or the fixed point iterations used in Chapter 3. For those algorithms we must know when to stop them.

Stopping criteria can be based on two questions:[21]

1. "Have we solved the problem?" and
2. "Have we ground to a halt?"

Consider the problem to find the root of the system of non-linear equations $\mathbf{f}(\mathbf{x}) = \mathbf{0}$. To answer the first question we must decide when $\mathbf{f}(\mathbf{x})$ is close to zero. To answer the second question we must decide when two successive points $\mathbf{x}^{s+1}$ and $\mathbf{x}^s$ are close together so that we can reasonably assume the sequence is near its limit point.

To tackle both problems we need measures of distance, or, more generally, vector norms as defined in equation (8.1). Given a vector norm, we will be speaking of a vector sequence $(\mathbf{x}^s)_{s=1}^{\infty}$ converging

[20] See, e.g., Theorem 7.5.3 in Judd (1998).
[21] See Dennis and Schnabel (1983), p. 159.

to a point $\mathbf{x}^*$, if $\lim_{s\to\infty} \|\mathbf{x}^s - \mathbf{x}^*\| = 0$. A key property of a convergent series is the rate at which it converges to its limit. We say that $\mathbf{x}^s$ converges at rate q to $\mathbf{x}^*$, if there exists a constant $c \in [0, 1)$ and an integer $\bar{s}$ such that

$$\|\mathbf{x}^{s+1} - \mathbf{x}^*\| \leq c\|\mathbf{x}^s - \mathbf{x}^*\|^q \text{ for all } s \geq \bar{s}. \tag{8.68}$$

If q in (8.68) equals 1 (2), we say that the vector series converges linearly (quadratically). If there is a sequence $(c_s)_{s=1}^{\infty}$ that converges to zero,

$$\|\mathbf{x}^{s+1} - \mathbf{x}^*\| \leq c_s\|\mathbf{x}^s - \mathbf{x}^*\|,$$

then we say the sequence $(\mathbf{x}_s)_{s=1}^{\infty}$ converges superlinearly.

With these definitions at hand we may accept $\mathbf{x}^c$ as a solution of $\mathbf{f}(\mathbf{x}) = 0$ if

$$\|\mathbf{f}(\mathbf{x}^c)\|_\infty < \epsilon,\ \epsilon \in \mathbb{R}. \tag{8.69}$$

Care must be taken with respect to the scaling of $\mathbf{f}$. For example, if $f^j \in [10^{-5}, 10^{-4}]\ \forall j$ and $\epsilon = 10^{-3}$, any $\mathbf{x}$ will cause us to stop. If the f^j differ greatly in magnitude, applying (8.69) may be overly restrictive. Therefore, before applying (8.69) the x_i should be scaled so that the f^j have about the same magnitude at points not near the root.

An answer to the second question can be based on the rule

$$\frac{|x_i^s - x_i^{s+1}|}{1 + |x_i^s|} \leq \epsilon\ \forall\ i = 1, 2, \ldots, n,\ \epsilon \in \mathbb{R}_{++}. \tag{8.70}$$

It tests whether the change in the i–th coordinate of $\mathbf{x}$ is small relative to the magnitude of x_i^s. To circumvent the possibility of $x_i \simeq 0$, $1 + |x_i^s|$ instead of $|x_i^s|$ is used in the denominator. However, if $\forall i\ x_i$ is much smaller than unity this criterium indicates convergence too early. Therefore, if the typical value of x_i, *typ* x_i, say, is known, Dennis and Schnabel (1983), p. 160 recommend

$$\frac{|x_i^s - x_i^{s+1}|}{\max\{|x_i^s|, |typ\ x_i|\}} \leq \epsilon\ \forall i = 1, 2, \ldots, n, \epsilon \in \mathbb{R}_{++}. \tag{8.71}$$

In some cases, like, e.g., in iterations over the value function, it is known that

$$\|\mathbf{x}^{s+1} - \mathbf{x}^*\| \leq c\|\mathbf{x}^s - \mathbf{x}^*\|, 0 \leq c < 1 \text{ for all } s \geq 1.$$

Thus the properties of norms given in (8.1) imply

$$\|\mathbf{x}^s - \mathbf{x}^*\| \leq \frac{\|\mathbf{x}^s - \mathbf{x}^{s+1}\|}{1-c}.$$

Using

$$\|\mathbf{x}^s - \mathbf{x}^{s+1}\| \leq \epsilon(1-c),\ \epsilon \in \mathbb{R}_{++} \tag{8.72}$$

as stopping rule secures that the error $\|\mathbf{x}^s - \mathbf{x}^*\|$ in accepting $\mathbf{x}^{s+1}$ as solution is always bounded from above by ϵ.

In Section 8.5.2 we present a globally convergent extension of the Newton-Raphson algorithm 8.5.1. It is based on finding the minimizer $\mathbf{x}^* \in \mathbb{R}^n$ of a real valued function $f(\mathbf{x})$. Therefore, in addition to the stopping criteria discussed so far, we need criteria that tell us, when we are close to the minimum of $f(\mathbf{x})$. A necessary condition for any minimum is

$$f_i(\mathbf{x}^*) = 0, \quad i = 1, 2, \ldots, n,$$

where $f_i(\mathbf{x}^*)$ denotes the partial derivative of f with respect to its i-th argument evaluated at $\mathbf{x}^*$. Let

$$\nabla f := [f_1, f_2, \ldots, f_n]'$$

denote the column vector of partial derivatives (the gradient of f). Then, a natural choice seems to stop, if at the k-th step

$$\|\nabla f(\mathbf{x}^k)\| \leq \epsilon$$

for a small positive number ϵ. However, this criterium is sensitive with respect to the scale of f. To see this, suppose $f(\mathbf{x})$ are the costs of producing the quantities x_i, $i = 1, 2, \ldots, n$ in US \$. Now, if instead we measure costs in thousands of dollars, so that $\tilde{f} = S_f f$ with $S_f = 10^{-3}$ the algorithm would already stop if $S_f\|\nabla f(\mathbf{x}^s)\| \leq \epsilon$. To circumvent this problem, we could use

$$\frac{\|\nabla f(\mathbf{x}^k)\|}{\max\{|typ\ f|, |f|\}} \leq \epsilon, \tag{8.73}$$

which is independent of the scale of f. Here, again, we use $\max\{|typ\ f|, |f|\}$ instead of $1+|f|$ in the denominator to allow for the typical value of f, $typ\ f$, to be much smaller than 1. However, (8.73) is not independent of the scale of x_i, $i = 1, 2, \ldots, n$. For instance, let $f(x) := (S_x x)^2$, where S_x is a scaling factor for x and assume $|typf| \equiv 1 < (S_x x^k)^2$. In this case the lhs of (8.73) yields

$$\frac{2}{|S_x x^k|}.$$

Here, again, the algorithm stops the sooner the larger is the scale – tons instead of kilos – for x. An obvious measure that is free of both the scale of f and of x_i is the partial elasticity of f with respect to x_i:

$$\frac{f_i(\mathbf{x}^k)x_i}{f(\mathbf{x}^k)}.$$

To account for either $x_i^k \simeq 0$ or $f(\mathbf{x}^k) \simeq 0$, Dennis and Schnabel (1983), p. 160 recommend the following stopping criterium:

$$\left| \frac{f_i(\mathbf{x}^k) \max\{|x_i|, typ\ x_i\}}{\max\{|f(\mathbf{x}^k)|, typ\ f\}} \right| < \epsilon, \quad \forall i = 1, 2, \ldots, n. \tag{8.74}$$

8.5 Non Linear Equations

There are many algorithms that require us to solve non-linear equations. For instance, in the stochastic growth model with endogenous labor supply considered in 1.4.4 we must solve the first order condition (1.51) for working hours N. Or, think of the deterministic extended path algorithm presented in Chapter 1, that boils down to solving a large system of non-linear equations. In this section we first describe one well-known method that works well locally. Thereafter we explain a line search method that is able to cope with problems, where the initial guess is far from the true solution.

8.5.1 The Newton-Raphson Method

The Zero of a Real Valued Function. Suppose we want to locate the zero of a real valued function $f : [a, b] \to \mathbb{R}$ in the interval $[a, b]$. In Figure 8.6 the point x^* solves the problem $f(x) = 0$. There, the domain of the function $f(x)$ is the set of non-negative real numbers $\mathbb{R}_+$. Consider the point x_0. We approximate f linearly around x_0. This gives $g^0(x) := f(x_0) + f'(x_0)(x - x_0)$, where $f'(x_0)$ is the slope of f at x_0. The root of $g^0(x)$ is given by

$$0 = f(x_0) + f'(x_0)(x_1' - x_0) \quad \Rightarrow \quad x_1' = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Yet, $x_1' < 0$ where $f(x)$ is not defined. Hence, we choose a point x_1 between x_1' and x_0. Approximating f at x_1 and solving for the root of $g^1(x) = f(x_1) + f'(x_1)(x - x_1)$ takes us close to x^*.

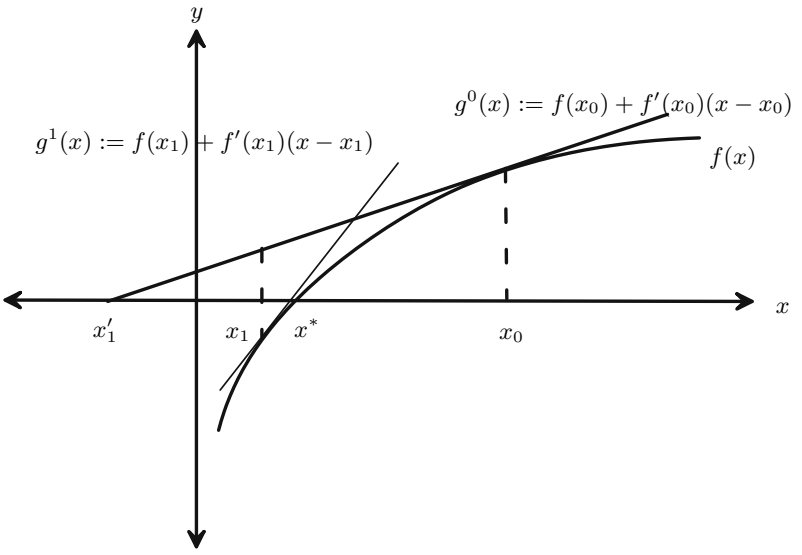

Figure 8.6: Modified Newton-Raphson Method to Locate the Zero of a Single Valued Function

The method where one iterates over

$$x_{s+1} = x_s - \frac{f(x_s)}{f'(x_s)} \tag{8.75}$$

until $f(x_{s+1}) \approx 0$ is called the Newton-Raphson method. The modified Newton-Raphson method takes care of regions where f is not defined and backtracks from x_{s+1} along the direction $f'(x_s)$ to a point x'_{s+1} at which f can be evaluated.

There are a problems where it is impossible or very costly (in terms of computational time) to compute the derivative of f. For instance, in Section 5.3.1 we compute the stationary distribution of an exchange economy with credit constraints. In this problem, there is no analytical expression for the function that relates the economy's interest rate to average asset holdings. In these situations we use the slope of the secant that connects two points $(x_s, f(x_s))$ and $(x_{s+1}, f(x_{s+1}))$ in place of $f'(x_s)$ in (8.75) (see Figure 8.7). This provides the secant method:

$$x_{s+2} = x_{s+1} - \frac{x_{s+1} - x_s}{f(x_{s+1} - f(x_s)}. \tag{8.76}$$

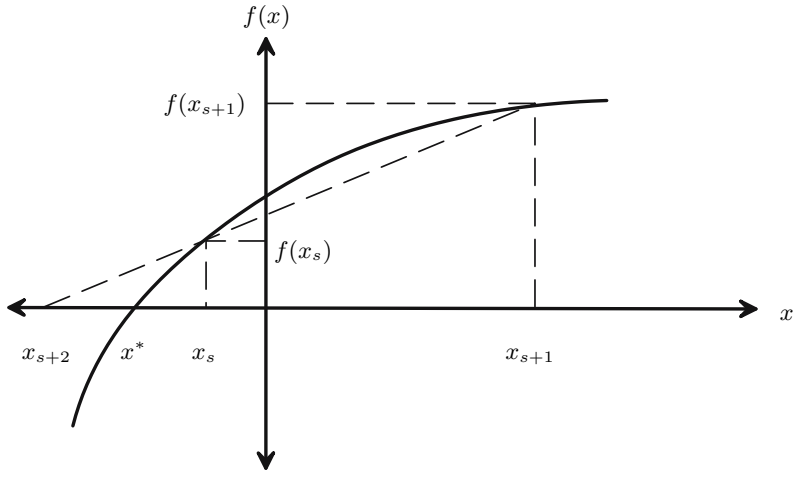

Figure 8.7: Secant Method

It can be shown that both the iterative scheme (8.75) and the (8.76) converge to the solution x^* under suitable conditions.[22] Fur-

[22] See, e.g., Dennis and Schnabel (1983), Theorem 2.4.3 and Theorem 2.6.3

thermore, they are easily generalized to the multi-variable framework.

The Zero of a System of Real Valued Functions. Assume we want to solve the system of n equations

$$\left.\begin{array}{l} 0{=}f^1(x_1, x_2, \ldots, x_n), \\ 0{=}f^2(x_1, x_2, \ldots, x_n), \\ \vdots{=} \quad \vdots, \\ 0{=}f^n(x_1, x_2, \ldots, x_n), \end{array}\right\} \Longleftrightarrow \mathbf{0} = \mathbf{f}(\mathbf{x}) \tag{8.77}$$

in the unknowns $\mathbf{x} = [x_1, x_2, \ldots, x_n]$. The equivalent to $f'(x)$ in the multi-variable case is the Jacobian matrix $J(\mathbf{x})$ of partial derivatives of $\mathbf{f} = [f^1, f^2, \ldots, f^n]'$ with respect to $x_i, i = 1, 2, \ldots n$. We use the notation

$$f^i_j := \frac{\partial f^i(\mathbf{x})}{\partial x_j}$$

to denote the partial derivative of f^i with respect to its j–th argument evaluated at the point $\mathbf{x}$. Thus, the Jacobian is given by

$$J(\mathbf{x}) := \begin{bmatrix} f^1_1 & f^1_2 & \ldots & f^1_n \\ f^2_1 & f^2_2 & \ldots & f^2_n \\ \vdots & \vdots & \ddots & \vdots \\ f^n_1 & f^n_2 & \ldots & f^n_n \end{bmatrix}. \tag{8.78}$$

The Jacobian is itself a function J that maps points from $\mathbb{R}^n$ to points in $\mathbb{R}^{n\times n}$. Let $\|\cdot\|$ denote a given vector or matrix norm, depending on the respective context.[23] We define an open ball with center $\mathbf{x}_0$ and radius r, $\mathcal{N}(\mathbf{x}_0, r)$, as the collection of all $\mathbf{x} \in \mathbb{R}^n$ whose distance to $\mathbf{x}_0$ is less than r:

$$\mathcal{N}(\mathbf{x}_0, r) := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{x}_0\| < r\}.$$

J is said to be Lipschitz on $\mathcal{N}(\mathbf{x}_0)$ of $\mathbf{x}_0$ with constant γ if for $\mathbf{x}_1, \mathbf{x}_2 \in \mathcal{N}(\mathbf{x}_0)$ the following condition holds

[23] See Section 8.1 on the definition of vector and matrix norms.

$$\|J(\mathbf{x}_1) - J(\mathbf{x}_2)\| \leq \gamma \|\mathbf{x}_1 - \mathbf{x}_2\|.$$

This is a stronger condition than the continuity of J. It can be shown that a sufficient condition for J to be Lipschitz is that J is continuously differentiable, i.e., that it belongs to the class of C^1 functions.[24] In the one-dimensional case $f(x) = 0$ this requires the function f to be twice continuously differentiable.

The linear approximation of $\mathbf{f}$ at $\mathbf{x}_0$ is

$$\mathbf{g}(\mathbf{x}) := \mathbf{f}(\mathbf{x}_0) + J(\mathbf{x}_0)\mathbf{dx}, \ \mathbf{dx} := (\mathbf{x} - \mathbf{x}_0),$$

with the Jacobian J defined in (8.78). Solving for $\mathbf{g}(\mathbf{x}_1) = \mathbf{0}$ gives

$$\mathbf{x}_1 = \mathbf{x}_0 - J(\mathbf{x}_0)^{-1}\mathbf{f}(\mathbf{x}_0). \tag{8.79}$$

The following theorem, taken from Dennis and Schnabel (1983), p. 90, shows that the sequence of points $\mathbf{x}_0, \mathbf{x}_1, \ldots$ converges quadratically to $\mathbf{x}^*$, if $\mathbf{x}_0$ is sufficiently close to $\mathbf{x}^*$.

Theorem 8.5.1 *Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable in an open convex set $\mathcal{D} \subset \mathbb{R}^n$. Assume that there exists $\mathbf{x}^* \in \mathbb{R}^n$ and $r, \beta > 0$, such that $\mathcal{N}(\mathbf{x}^*, r) \subset \mathcal{D}$, $\mathbf{f}(\mathbf{x}^*) = \mathbf{0}$, $J(\mathbf{x}^*)^{-1}$ exists with $\|J(\mathbf{x}^*)^{-1}\| \leq \beta$, and J Lipschitz with constant γ on $\mathcal{N}(\mathbf{x}^*, r)$. Then there exists $\epsilon > 0$ such that for all $\mathbf{x}_0 \in \mathcal{N}(\mathbf{x}^*, \epsilon)$ the sequence $\mathbf{x}_1, \mathbf{x}_2, \ldots$ generated by*

$$\mathbf{x}_{k+1} = \mathbf{x}_k - J(\mathbf{x}_k)^{-1} f(\mathbf{x}_k), \quad k = 0, 1, \ldots,$$

is well defined, converges to $\mathbf{x}^*$, and obeys

$$\|\mathbf{x}_{k+1} - \mathbf{x}^*\| \leq \beta\gamma \|\mathbf{x}_k - \mathbf{x}^*\|^2, \quad k = 0, 1, \ldots.$$

If the initial guess $\mathbf{x}_0$ is not as close to the final solution as required by this theorem, the algorithm may hit points for which $\mathbf{f}$ is not defined (as the point x_1' in Figure 8.6). To circumvent this case, we specify upper and lower bounds $[\underline{\mathbf{x}}, \bar{\mathbf{x}}]$ such that $\mathbf{f}$ is well defined for all $\mathbf{x} \in [\underline{\mathbf{x}}, \bar{\mathbf{x}}]$.

[24] A prove of this statement can be found in Hirsch and Smale (1974), p. 163f.

Putting all pieces together, provides the following algorithm:

Algorithm 8.5.1 (Modified Newton-Raphson)

Purpose: *Solve* $\mathbf{0} = \mathbf{f}(\mathbf{x})$, *where* $\mathbf{f} = [f^1, f^2, .., f^n]'$ *and* $\mathbf{x} \in \mathbb{R}^n$.

Steps:

Step 1: Initialize: choose $\mathbf{x}_0 \in [\underline{\mathbf{x}}, \bar{\mathbf{x}}]$.

Step 2: Compute $J(\mathbf{x}_0)$ *the Jacobian of* f *at* $\mathbf{x}_0$ *and solve* $J(\mathbf{x}_0)\mathbf{dx} = -\mathbf{f}(\mathbf{x}_0)$. *If* $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{dx} \notin [\underline{\mathbf{x}}, \bar{\mathbf{x}}]$ *choose* $\lambda \in (0, 1)$ *such that* $\mathbf{x}_2 = \mathbf{x}_0 + \lambda\mathbf{dx} \in [\underline{\mathbf{x}}, \bar{\mathbf{x}}]$ *and set* $\mathbf{x}_1 = \mathbf{x}_2$.

Step 3: Check for convergence: if $\|\mathbf{f}(\mathbf{x}_1)\|_\infty < \epsilon$ *and/or* $|x_i^1 - x_i^0|/(1 + |x_i^0|) \leq \epsilon \forall i$ *for a given tolerance* $\epsilon \in \mathbb{R}_{++}$ *stop, else set* $\mathbf{x}_0 = \mathbf{x}_1$ *and return to step 2.*

We implemented the single equation version of this algorithm in the GAUSS procedure `FixpMN` in the file `ToolBox.src`. For the multi-equation version we provide two implementations. The procedure `FixvMN1` can be used, if you have no a priori knowledge about the domain of $\mathbf{f}$ (as is the case, for instance, in the parameterized expectations approach presented in Chapter 3). In this case, the procedure returning $\mathbf{f}(\mathbf{x})$ must return a GAUSS missing value code if it is not possible to evaluate $\mathbf{f}$ at $\mathbf{x} + d\mathbf{x}$. The procedure than backtracks from $\mathbf{x} + d\mathbf{x}$ towards $\mathbf{x}$. If you know the boundaries (as in the application of the deterministic extended path approach of Chapter 1) you can use `FixvMN2`. Both implementations use `CDJac` (see Section 8.3.1) to compute the Jacobian. You may choose between two methods to solve for $d\mathbf{x}$. First, you may use the inverse of $J(\mathbf{x}_0)$, and use (8.79). Since equation (8.79) may also be written as a system of linear equations

$$(\mathbf{x}_1 - \mathbf{x}_0)J(\mathbf{x}_0) \equiv d\mathbf{x}J(\mathbf{x}_0) = f(\mathbf{x}_0)$$

you may want to apply a linear equations solver. As mentioned in Section 8.1.7 GAUSS contains procedures that solve linear systems via the LU-factorization.

If the initial guess is bad, it may happen that the algorithm diverges. The line search method presented in the next subsection can cope with this case.

8.5.2 A Globally Convergent Newton-Raphson Method

Here we describe a strategy that forces the algorithm to converge to the solution from any starting point in the domain of the vector valued function $\mathbf{f}$. It is based on two observations: firstly, the solution to $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ is also a minimizer of

$$g(\mathbf{x}) := (1/2)\mathbf{f}(\mathbf{x})'\mathbf{f}(\mathbf{x}) = (1/2)\sum_{i=1}^{n}(f^i(x_1, \ldots, x_n))^2.$$

Secondly, the Newton-Raphson step at $\mathbf{x}_0$,

$$d\mathbf{x} = -J(\mathbf{x_0})^{-1}\mathbf{f}(\mathbf{x}_0),$$

is a descent direction for g. To see this, note that the linear approximation (see equation (8.26)) of g at $\mathbf{x}_0$ is given by

$$g(\mathbf{x}_1) \simeq g(\mathbf{x}_0) + [\nabla g(\mathbf{x}_0)]' \underbrace{(\mathbf{x}_1 - \mathbf{x}_0)}_{d\mathbf{x}},$$

where ∇g denotes the gradient of g, i.e., the column vector of first partial derivatives of g:

$$\nabla g(\mathbf{x}_0) = J(\mathbf{x}_0)'\mathbf{f}(\mathbf{x}_0).$$

Therefore

$$\begin{aligned} g(\mathbf{x}_1) - g(\mathbf{x}_0) \simeq [\nabla g(\mathbf{x}_0)]'d\mathbf{x} &= \mathbf{f}(\mathbf{x_0})'J(\mathbf{x}_0)(-J(\mathbf{x}_0)^{-1}\mathbf{f}(\mathbf{x}_0)) \\ &= -\mathbf{f}(\mathbf{x_0})'\mathbf{f}(\mathbf{x}_0) < 0. \end{aligned}$$

The idea is, thus, to move in the Newton-Raphson direction, and check, whether going all the way actually reduces g. If not, we move back towards $\mathbf{x}_0$ until we get a sufficient reduction in g. The details of this procedure are from Dennis and Schnabel (1983), who show that this algorithm converges to a minimum of g except in rare cases (see their Theorem 6.3.3 on p. 121).

Let

$$h(s) := g(\mathbf{x}_0 + sd\mathbf{x})$$

denote the restriction of g to the line through $\mathbf{x}_0$ in the direction $d\mathbf{x}$. We look for a step of size $s \in (0, 1]$ that reduces $g(\mathbf{x}_0)$ at least by $s\alpha\nabla g(\mathbf{x}_0)d\mathbf{x}$ for a small $\alpha \in (0, 1/2)$, i.e.,

$$g(\mathbf{x}_0 + sd\mathbf{x}) \leq g(\mathbf{x}_0) + s\alpha\nabla g(\mathbf{x}_0)d\mathbf{x}. \tag{8.80}$$

Dennis and Schnabel (1983) recommend $\alpha = 10^{-4}$.

At first we try the full Newton-Raphson step, and hence, put $s_1 = 1$. If s_1 fails to satisfy (8.80), we approximate h by a parabola,

$$y := as^2 + bs + c$$

and choose s_2 as the minimizer of this function:

$$s_2 = -\frac{b}{2a}.$$

We get a and b from:

$$\left.\begin{aligned} h(0) &= g(\mathbf{x}_0) =: c, \\ h(1) &= g(\mathbf{x}_0 + d\mathbf{x}) =: a + b + c, \\ h'(0) &= \nabla g(\mathbf{x}_0)d\mathbf{x} =: b. \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} a &= h(1) - h(0) - h'(0), \\ b &= h'(0), \\ c &= h(0). \end{aligned}\right. \tag{8.81}$$

Therefore:

$$s_2 = -\frac{b}{2a} = \frac{-h'(0)}{2(h(1) - h(0) - h'(0)}. \tag{8.82}$$

Note that $s_2 < (1/2)$ if $g(\mathbf{x}_0 + d\mathbf{x}) > g(\mathbf{x}_0)$ and $s_2 = 1/[2(1-\alpha)]$. Since too small or too large steps can prevent the algorithm from converging to the minimum of g, we require $s_2 \in [0.1, 0.5]$.[25]

If the quadratic approximation was not good, s_2 may still violate (8.80). In this case we approximate h by a cubic function:

$$y := as^3 + bs^2 + cs + d.$$

The parameters of this approximation must solve the following system of equations

[25] See Dennis and Schnabel (1983) for examples.

$$\begin{aligned} as_1^3 + bs_1^2 + cs_1 + d &= h(s_1) = g(\mathbf{x}_0 + s_1 d\mathbf{x}), \\ as_2^3 + bs_2^2 + cs_2 + d &= h(s_2) = g(\mathbf{x}_0 + s_2 d\mathbf{x}), \\ c &= h'(0) = \nabla g(\mathbf{x}_0) d\mathbf{x}, \\ d &= h(0) = g(\mathbf{x}_0), \end{aligned} \tag{8.83}$$

and the minimizer of y is the solution to

$$s_3 = \frac{-b + \sqrt{b^2 - 3ac}}{3a}. \tag{8.84}$$

If $\alpha < (1/4)$ this solution is always real.[26] Here, again we avoid too large or too small steps by restricting s_3 to

$$s_3 \in [0.1 s_2, 0.5 s_2].$$

If s_3 still violates (8.80) we approximate h at the points $\mathbf{x}_0$, $\mathbf{x}_0 + s_2 d\mathbf{x}$, and $\mathbf{x}_0 + s_3 d\mathbf{x}$, solve (8.83) and (8.84) for s_4 and continue this procedure until s_k satisfies (8.80). To prevent the line search to get trapped in an endless loop, we check at each step whether s_k is larger than some minimal value s_{min}. We choose s_{min} so that $s < s_{min}$ implies convergence according to the parameter convergence criterium chosen. For example, in the case of (8.71), define

$$\Delta_i := \frac{|x_i^s - x_i^{s+1}|}{\max\{|x_i^s|, |typ\ x_i|\}}$$

and $\Delta = \arg\max\{\Delta_1, \Delta_2, \ldots, \Delta_n\}$. Then $s_{min} = \epsilon/\Delta$. If the line search is used in a pure minimization routine, where (8.74) is used to stop the algorithm, $s < s_{min}$ should never occur. If it nevertheless does, this usually indicates that the ϵ used in (8.71) is to large relative to the ϵ used in (8.74). If $s < s_{min}$ occurs in a non-linear equation solver, the calling program should verify whether the minimum of g as defined above, is also a zero of f.

We summarize this algorithm in the following statement:

Algorithm 8.5.2 (Line Search)

Purpose: *Find a step size that achieves a sufficient decrease in the value of a function to be minimized.*

[26] See Dennis and Schnabel (1983), p. 129.

Steps:

Step 1. *Initialize: Choose* $\alpha = 10^{-4}$*, compute* s_{min}*, put* $s_k = 1$*, and* $k = 1$*.*

Step 2. *If* s_k *satisfies* (8.80) *stop and return* s_k*, else increase* k *by 1 and proceed to the next step.*

Step 3. *If* $k = 2$ *solve* (8.82) *for* s_2*, yet restrict the solution to the interval* $s_2 \in [0.1, 0.5]$*.*
If $k > 2$ *solve* (8.83) *and* (8.84) *using the two most recent values of* s*, say* s_{k-1} *and* s_{k-2}*, and restrict the solution to the interval* $s_k \in [0.1s_{k-1}, 0.5s_{k-1}]$*.*
In any case put $s = s_k$*. If* $s > s_{min}$ *return to step 2, else stop and let the calling program know that no further decrease of* g *can be achieved within the given parameter tolerance* ϵ*.*

8.6 Numerical Optimization

There are some algorithms where we must find the minimizer of a given function.[27] Think of the non-linear least squares problem encountered in the parameterized expectations approach of Chapter 4 or think of the maximization step as part of Algorithm 1.2.3. In other algorithms we are free to choose whether to solve the system of first order conditions that characterizes the optimal solution or to employ numerical optimization tools. Sometimes one line of attack may work while the other performs poorly. Here we describe three well known tools from numerical optimization. The golden section search is a simple means of locating the maximizer of a single valued function in a given interval $[a, b]$. The Gauss-Newton approach is tailored to non-linear least squares problems, while the BFGS quasi-Newton method is suitable to a wide class of unconstrained maximization problems. Finally, we consider stochastic algorithms.

[27] Since the maximizer of $-f(\mathbf{x})$ is identical to the minimizer of $f(\mathbf{x})$, we can restrict ourselves to minimization problems.

8.6.1 Golden Section Search

This method locates the maximum of a single peaked function $f(x)$ in the interval $I = [A, D]$. The idea is to shrink the interval around the true maximizer x^* in successive steps until the midpoint of the remaining interval is a good approximation to x^* (see Figure 8.8).

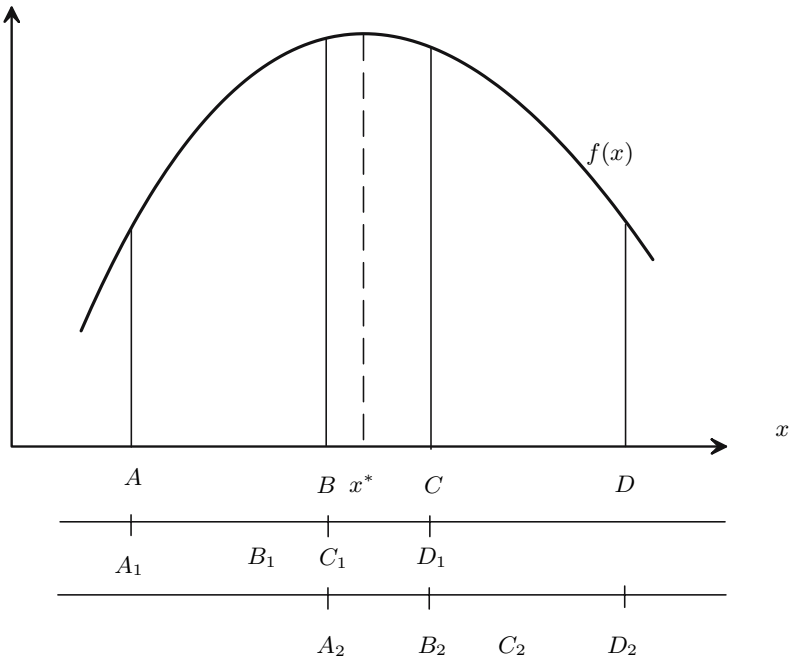

Figure 8.8: Golden Section Search

Assume we have two more function evaluations at points B and C, respectively. It is obvious from Figure 8.8 that for $f(B) > f(C)$ the maximum lies in the shorter interval $[A, C]$. In the opposite case $f(B) < f(C)$ the maximizer is located in $[B, D]$. The question is, how should we choose B and C?

There are two reasonable principles that will guide our choice. First, note that we do not know in advance whether we end up with $[A, C]$ or $[B, D]$. Our aim is to reduce the interval as much as possible. The unfavorable case is to end up with the larger of

the two intervals. We exclude this possibility by choosing B and C so that both intervals are of the same size:

$$\overline{AC} = \overline{BD} \;\Rightarrow\; \overline{AB} = \overline{CD}. \tag{8.85}$$

Consider what happens, if $[A_1, D_1] = [A, C]$ is the new interval. Since we know $f(B)$, we need only one more function evaluation to find the next smaller interval. The second principle that we employ is that $[A_1, D_1]$ is a scaled down replication of $[A, D]$, i.e., the points B_1 and C_1 divide $[A_1, D_1]$ the same way as did B and C with $[A, D]$:

$$p := \frac{\overline{AC}}{\overline{AD}} = \frac{\overline{A_1C_1}}{\overline{A_1D_1}} \quad \Rightarrow \frac{\overline{AC}}{\overline{AD}} = \frac{\overline{AB}}{\overline{AC}}, \tag{8.86a}$$

$$\frac{\overline{AB}}{\overline{AC}} = \frac{\overline{A_1B_1}}{\overline{A_1C_1}}. \tag{8.86b}$$

Symmetrically, if it turns out that the new interval is $[A_2, D_2] = [B, D]$, we demand:

$$p = \frac{\overline{AB}}{\overline{AC}} = \frac{\overline{B_2D_2}}{\overline{A_2D_2}}, \tag{8.87a}$$

$$\frac{\overline{CD}}{\overline{BD}} = \frac{\overline{C_2D_2}}{\overline{B_2D_2}}. \tag{8.87b}$$

Equation (8.85) and (8.86a) (as well as (8.85) and (8.87a)) imply the condition

$$\frac{1-p}{p} = p.$$

Solving this quadratic equation in p delivers:

$$p = \frac{\sqrt{5}-1}{2} \approx 0.618. \tag{8.88}$$

This is the fraction by which we are able to shrink the interval in successive iterations. It divides each interval into the so called

golden sections. Thus, in the first step we choose points B and C according to

$$B = A + (1-p)\overline{AD},$$
$$C = A + p\overline{AD}.$$

In the next step we choose $[A_1, D_1] = [A, C]$ if $f(B) > f(C)$ and $[A_2, D_2] = [B, D]$ otherwise. In the first case, we put $A_1 = A$, $C_1 = B$, and $D_1 = C$. Condition (8.86b) gives

$$B_1 = pC_1 + (1-p)A_1.$$

In the second case we put $A_2 = B$, $B_2 = C$, and $D_2 = D$. The new point C_2 is given by (8.87b):

$$C_2 = pB_2 + (1-p)D_2.$$

Summarizing, we can construct the following iterative scheme to bracket x^*:

Algorithm 8.6.1 (Golden Section Search)

Purpose: *Find the maximizer of a single peaked function $f(x)$ in the interval $[\underline{x}, \overline{x}]$.*

Steps:

Step 1. Initialize: Set $A = \underline{x}$, $D = \overline{x}$ and compute

$$B = pA + (1-p)D,$$
$$C = (1-p)A + pD,$$
$$p = +(\sqrt{5}-1)/2,$$

and store $f(B)$ in fB and $f(C)$ in fC.

Step 2. If $fB > fC$ replace D by C, C by B, and fC by fB. Find the new B from $B = pC + (1-p)A$ and store $f(B)$ in fB.
Otherwise: replace A by B, B by C, and fB by fC. Find the new C from $C = pB + (1-p)D$ and store $f(C)$ in fC.

Step 3. Check for convergence: if $|D - A| < \epsilon \max\{1, |B| + |C|\}$ *stop and return* B*, else repeat the previous step.*

Our GAUSS procedure `GSS` implements this algorithm. Its inputs are the pointer to the procedure that returns $f(x)$ and the boundaries of the interval in which the maximum lies.

8.6.2 Gauss-Newton Method

Algorithms that solve non-linear least squares problems are adapted from procedures that solve the more general problem of finding the minimizer of a real valued function. The solution that we propose is known as the damped Gauss-Newton method.[28] To introduce this algorithm we return to the more common notion of seeking to minimize

$$S(\boldsymbol{\gamma}) := \frac{1}{T}\sum_{i=1}^{T}(y_i - f(\boldsymbol{\gamma}, \mathbf{x}_i))^2, \ \mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{in}). \tag{8.89}$$

with respect to the parameter vector $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, \ldots, \gamma_p)'$.

The minimizer $\boldsymbol{\gamma}^*$ must solve the set of first order conditions

$$\begin{aligned} &\frac{\partial S}{\partial \gamma_j} = \frac{-2}{T}\sum_{i=1}^{T}(y_i - f(\boldsymbol{\gamma}^*, \mathbf{x}_i))\frac{\partial f}{\partial \gamma_j}(\boldsymbol{\gamma}^*, \mathbf{x}_i) = 0, \\ &\quad j = 1, 2, \ldots, p. \end{aligned} \tag{8.90}$$

Instead of solving this system of p non-linear equations in $\boldsymbol{\gamma}$ the simple Gauss-Newton method operates on a linearized minimization problem. Suppose we have an initial guess $\boldsymbol{\gamma}_s$ and consider the linear approximation of f at this vector:[29]

$$f(\boldsymbol{\gamma}, \mathbf{x}_i) \simeq f(\boldsymbol{\gamma}_s, \mathbf{x}_i) + [\nabla f(\boldsymbol{\gamma}_s, \mathbf{x}_i)]'(\boldsymbol{\gamma} - \boldsymbol{\gamma}_s),$$

where $\nabla f(\cdot)$ is the column vector of the first partial derivatives of f with respect to $\gamma_j, j = 1, 2, \ldots, p$ evaluated at the given $\boldsymbol{\gamma_s}$. Put

[28] See Dennis and Schnabel (1983), Chapter 10.
[29] See equation (8.26).

$$\bar{y}_i := y_i - f(\boldsymbol{\gamma}_s, x_i),$$
$$\bar{\mathbf{x}}_i := \nabla f(\boldsymbol{\gamma}_s, \mathbf{x}_i),$$
$$\bar{\boldsymbol{\gamma}} = \boldsymbol{\gamma} - \boldsymbol{\gamma}_s.$$

The solution to the linear least squares problem

$$\min_{\bar{\boldsymbol{\gamma}}} \sum_{i=1}^{T} [\bar{y}_i - \bar{\mathbf{x}}_i \bar{\boldsymbol{\gamma}}]^2$$

is provided by the well known formula

$$\bar{\boldsymbol{\gamma}} = (\bar{X}'\bar{X})^{-1}\bar{X}'\bar{\mathbf{y}},$$
$$\bar{X} := [\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \ldots, \bar{\mathbf{x}}_T], \bar{\mathbf{y}} = [\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_T]'.$$

The simple Gauss-Newton method chooses

$$\boldsymbol{\gamma}_{s+1} = \boldsymbol{\gamma}_s + \underbrace{(\bar{X}'\bar{X})^{-1}\bar{X}'\bar{\mathbf{y}}}_{=:d\boldsymbol{\gamma}}$$

as the next value of $\boldsymbol{\gamma}$. Along $d\boldsymbol{\gamma}$ the sum of squares S is decreasing. To see this, note that

$$\nabla S(\boldsymbol{\gamma}) := \frac{-2}{T}\bar{X}'\bar{\mathbf{y}}$$

is the (column) vector of partial derivatives of S evaluated at $\boldsymbol{\gamma}$. Therefore,

$$[\nabla S(\boldsymbol{\gamma}_s)]' d\boldsymbol{\gamma} = \frac{-2}{T} \underbrace{\bar{\mathbf{y}}'\bar{X}}_{=:\mathbf{z}'} \underbrace{(\bar{X}'\bar{X})^{-1}}_{=:A} \underbrace{\bar{X}'\bar{\mathbf{y}}}_{=:\mathbf{z}} < 0.$$

This follows from the fact that the matrix $\bar{X}'\bar{X}$ and thus its inverse A is positive definite.[30]

If $\boldsymbol{\gamma}_{s+1}$ is not the minimizer of S, f is linearized at the new value of $\boldsymbol{\gamma}_{s+1}$ and the related linear least squares problem is solved again to deliver $\boldsymbol{\gamma}_{s+2}$. These steps are repeated until convergence.

If the initial value of $\boldsymbol{\gamma}$ is not near the (local) minimizer, this method may fail to converge, much like the Newton-Raphson

[30] See Section 8.1.5 on definite quadratic forms.

method considered in Algorithm 8.5.1. The damped Gauss-Newton method uses the line search from Algorithm 8.5.2 to force the iterations downhill towards a local minimum. Indeed, since the sum of squares (8.89) is bounded from below and since the gradient of a polynomial is continuously differentiable and thus Lipschitz, these iterations satisfy the conditions of Theorem 6.3.3 from DENNIS and SCHNABEL (1983). As a consequence, using the damped Gauss-Newton method will take us to a local minimum of $S(\boldsymbol{\gamma})$. We use the stopping rule (8.74) (see 8.4) to terminate the algorithm.[31]

Taking all pieces together, the damped Gauss-Newton algorithm proceeds as follows:

Algorithm 8.6.2 (Damped Gauss-Newton)

Purpose: *Find the mimizier of the non-linear least squares problem* (8.89)

Steps:

Step 1: Initialize: Choose a vector $\boldsymbol{\gamma}_0$ *and stopping criteria* $\epsilon_1 \in \mathbb{R}_{++}$ *and* $\epsilon_2 \in \mathbb{R}_{++}$, $\epsilon_1 >> \epsilon_2$. *Put* $s = 0$.

Step 2: Linearize $f(\boldsymbol{\gamma}, \mathbf{x}_i)$ *at* $\boldsymbol{\gamma}_s$ *and put*

$$\bar{y}_i = y_i - f(\boldsymbol{\gamma}_s, \mathbf{x}_i), \quad \bar{\mathbf{y}} = [\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_T]',$$
$$\bar{\mathbf{x}}_i = [\nabla f(\boldsymbol{\gamma}_s, \mathbf{x}_i)]', \quad \bar{X} = [\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \ldots, \bar{\mathbf{x}}_T].$$

Step 3: Compute $\boldsymbol{\gamma}_{s+1}$*: Solve the linear system*

$$\bar{X}'\bar{X}\bar{\boldsymbol{\gamma}} = \bar{X}'\bar{y}$$

for $\bar{\boldsymbol{\gamma}}$. *Use Algorithm 8.5.2 with* ϵ_2 *to find the step length* d *and put*

[31] Indeed, since Theorem 6.3.3. from DENNIS and SCHNABEL (1983) establishes convergence of

$$\frac{\nabla S(\boldsymbol{\gamma}_s)'(\boldsymbol{\gamma}_{s+1} - \boldsymbol{\gamma}_s)}{\|\boldsymbol{\gamma}_{s+1} - \boldsymbol{\gamma}_s\|_2}$$

but not of $\boldsymbol{\gamma}_s$, it makes sense, to try criterion (8.74) at first. Our line search procedure will warn us, if it is not possible to decrease S further, even if (8.74) is not met.

$$\boldsymbol{\gamma}_{s+1} = \boldsymbol{\gamma}_s + d\bar{\boldsymbol{\gamma}}.$$

Step 4: Check for convergence: Use criterion (8.74) *with* ϵ_1 *to see whether the algorithm is close to the minimizer. If so, stop. If not, and if the line search was successful, increase* s *by one and return to Step 2. Otherwise stop and report convergence to a nonoptimal point.*

We provide implementations of this algorithm in GAUSS as well as in Fortran. Look for the subroutine `GaussNewton` in the file `ToolBox.src` and the file `GaussNewton.for`, respectively. Both procedures allow the user to either supply his own routine for the computation of the gradient of f or to use built-in forward difference methods (or our routines described in Section 8.3.1) to approximate the gradient. Note that the matrix $\bar{X}$ is the Jacobian of the vector valued function

$$\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_T \end{bmatrix} = \begin{bmatrix} f(\boldsymbol{\gamma}, \mathbf{x}_1) \\ f(\boldsymbol{\gamma}, \mathbf{x}_2) \\ \vdots \\ f(\boldsymbol{\gamma}, \mathbf{x}_T) \end{bmatrix}.$$

Thus, if you write a subroutine that returns the vector $\mathbf{z} = [z_1, z_2, \ldots, z_T]'$ and pass this routine to another routine that approximates the Jacobian of a vector valued function, as, e.g., the `gradp` routine in GAUSS, the output of this routine is $\bar{X}$.

8.6.3 Quasi-Newton

In this section we introduce the so called BFGS method to locate the minimizer of a function of several variables. This method derives from Newton's method, which we describe next.

Newton's Method. Suppose you want to minimize $y = f(\mathbf{x})$ on an open subset U of $\mathbb{R}^n$. Newton's method solves this problem by considering the quadratic approximation (see equation (8.25))

$$\hat{f}(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + [\nabla f(\mathbf{x}_0)]'\mathbf{h} + \frac{1}{2}\mathbf{h}' H(\mathbf{x}_0)\mathbf{h}.$$

In this formula $\nabla f(\mathbf{x}_0)$ is the column vector of first partial derivatives of f with respect to x_i, $i = 1, 2, \ldots, n$, and H is the Hesse matrix of second partial derivatives. Minimizing $\hat{f}$ with respect to the vector $\mathbf{h}$ requires the following first order conditions to hold:[32]

$$\nabla f(\mathbf{x}_0) + H(\mathbf{x}_0)\mathbf{h} = 0.$$

Solving for $\mathbf{x}_1 = \mathbf{x}_0 + \mathbf{h}$ provides the following iterative formula:

$$\mathbf{x}_1 = \mathbf{x}_0 - H(\mathbf{x}_0)^{-1}\nabla f(\mathbf{x}_0). \tag{8.91}$$

It is well known that iterations based on this formula converge quadratically to the minimizer $\mathbf{x}^*$ of $f(\mathbf{x})$, if the initial point $\mathbf{x}_0$ is sufficiently close to the solution $\mathbf{x}^*$.[33] Note, that the second order conditions for a local minimum require the Hesse matrix to be positive semidefinite in a neighborhood of $\mathbf{x}^*$.[34] Furthermore, using $\nabla f(\mathbf{x}_0) = -H(\mathbf{x}_0)\mathbf{h}$ in the quadratic approximation formula gives

$$\hat{f}(\mathbf{x}_1) - f(\mathbf{x}_0) = -(1/2)\mathbf{h}'H(\mathbf{x}_0)\mathbf{h}.$$

Thus, if the Hessian is positive definite (see Section 8.1 on definite matrices), the Newton direction is always a decent direction.

The computation of the Hesse matrix is time consuming. Furthermore, there is nothing that ensures this matrix to be positive definite far away from the solution. So called quasi-Newton methods tackle these problems by providing secant approximations to the Hesse matrix. In addition, they implement line search methods that direct the algorithm downhill and, thus, help to ensure almost global convergence. The secant method that has proven to be most successful was discovered independently by Broyden, Fletcher, Goldfarb, and Shanno in 1970. It is known as the BFGS update formula.

[32] See Section 8.1.5 on the differentiation of linear and quadratic forms.

[33] This follows from Theorem 8.5.1, since the iterative scheme (8.91) derives from the Newton-Raphson method applied to the system of first order conditions $\nabla f(\mathbf{x}) = \mathbf{0}$.

[34] See, e.g., SUNDARAM (1996), Theorem 4.3.

BFGS Secant Update. The BFGS quasi-Newton method replaces the Hessian in (8.91) by a positive definite matrix H_k that is updated at each iteration step k. The identity matrix I_n can be used to initialize the sequence of matrices. Consider the following definitions:

$$\mathbf{x}_{k+1} - \mathbf{x}_k = -H_k^{-1} \nabla f(\mathbf{x}_k), \tag{8.92a}$$

$$\mathbf{w}_k := \mathbf{x}_{k+1} - \mathbf{x}_k, \tag{8.92b}$$

$$\mathbf{z}_k := \nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k), \tag{8.92c}$$

$$H_{k+1} := H_k + \frac{\mathbf{z}_k \mathbf{z}_k'}{\mathbf{z}_k' \mathbf{w}_k} - \frac{H_k \mathbf{w}_k \mathbf{w}_k' H_k'}{\mathbf{w}_k' H_k \mathbf{w}_k}, \tag{8.92d}$$

where the last line defines the BFGS update formula for the secant approximation of the Hesse matrix.

The following theorem provides the foundation of the BFGS method:[35]

Theorem 8.6.1 *Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable in an open convex set $D \subset \mathbb{R}^n$, and let $H(\mathbf{x})$ be Lipschitz. Assume there exists $\mathbf{x}^* \in D$ such that $\nabla f(\mathbf{x}^*) = \mathbf{0}$ and $H(\mathbf{x}^*)$ is nonsingular and positive definite. Then there exist positive constants ϵ, δ such that if $\|\mathbf{x}_0 - \mathbf{x}^*\|_2 \leq \epsilon$ and $\|H_0 - H(\mathbf{x}^*)\| \leq \delta$, then the positive definite secant update* (8.92) *is well defined, $\{\mathbf{x}_k\}_{k=1}^{\infty}$ remains in D and converges superlinearly to $\mathbf{x}^*$.*

Instead of updating the approximate Hessian H_k one can also start with a positive definite approximation of the inverse of H_k, say $A_k := H_k^{-1}$. The next iterate of $\mathbf{x}_k$ is then given by

$$\mathbf{x}_{k+1} = \mathbf{x}_k - A_k \nabla f(\mathbf{x}_k).$$

This involves only vector addition and matrix multiplication, whereas (8.91) requires the solution of a system of linear equations. The BFGS update formula for A_k is given by (see Press et Al. (2001), p. 420):

[35] See Theorems 9.1.2 and 9.3.1 of Dennis and Schnabel (1983).

$$A_{k+1} = A_k + \frac{\mathbf{w}_k\mathbf{w}_k'}{\mathbf{w}_k'\mathbf{z}_k} - \frac{(A_k\mathbf{z}_k)(A\mathbf{z}_k)'}{\mathbf{z}_k'A_k\mathbf{z}_k} + (\mathbf{z}_kA_k\mathbf{z}_k)\mathbf{u}_k\mathbf{u}_k', \qquad \mathbf{u}_k := \frac{\mathbf{w}_k}{\mathbf{w}_k'\mathbf{z}_k} - \frac{A_k\mathbf{z}_k}{\mathbf{z}_k'A_k\mathbf{z}_k}. \tag{8.93}$$

Yet another approach is to use the fact that a positive definite matrix H_k has a Cholesky factorization $L_kL_k' = H_k$, where L_k is a lower triangular matrix. Using this factorization, it is easy to solve the linear system $(L_kL_k')(\mathbf{x}_1 - \mathbf{x}_0) = -\nabla f(\mathbf{x}_0)$ by forward and backward substitution (see Section 8.1.7). Thus, instead of updating H_k, one may want to update L_k. Goldfarb (1976) provides the details of this approach, which underlies the GAUSS routine `QNewton`.

The BFGS iterations may be combined with the line search algorithm 8.5.2 to enhance global convergence. Indeed, if the Hesse matrix (not its approximation!) of $f(\mathbf{x})$ is positive definite for all $\mathbf{x} \in \mathbb{R}^n$ a Theorem due to Powell (see Dennis and Schnabel (1983), Theorem 9.51 on p. 211) establishes global convergence.[36]

Taking all pieces together, the minimization algorithm consists of these steps:

Algorithm 8.6.3 (BFGS Quasi-Newton)

Purpose: *Minimize $f(\mathbf{x})$ in $U \subset \mathbb{R}^n$.*

Steps:

Step 1: Initialize: Choose $\mathbf{x}_0$, stopping criteria $\epsilon_1 \in \mathbb{R}_{++}$ and $\epsilon_2 \in \mathbb{R}_{++}$, $\epsilon_1 >> \epsilon_2$, and either $A_0 = I_n$ or $H_0 = I_n$. Put $k = 0$.

Step 2: Compute the gradient $\nabla f(\mathbf{x}_k)$ and solve for $\mathbf{w}_k$ either from

$$H_s\mathbf{w}_k = -\nabla f(\mathbf{x}_s)$$

or from

[36] Note, this does not imply that a computer coded algorithm does indeed converge. Finite precision arithmetic accounts for differences between the theoretical gradient, the theoretical value of f and the approximate Hessian.

$$\mathbf{w}_k = -A_k \nabla f(\mathbf{x}_k).$$

Step 3: Use Algorithm 8.5.2 with ϵ_2 to find the step length s, and put

$$\mathbf{x}_{k+1} = \mathbf{x}_k + s\mathbf{w}_k.$$

Step 4: Check for convergence: Use criterion (8.74) *with ϵ_1 to see whether the algorithm is close to the minimizer. If so, stop. If not, and if the line search was successful, proceed to Step 5. Otherwise stop and report convergence to a nonoptimal point.*

Step 5: Use either (8.92d) *or* (8.93) *to get A_{k+1} or H_{k+1}, respectively. Increase k by one and return to Step 2.*

The BFGS quasi-Newton algorithm is available in the GAUSS command `QNewton` and the ISML subroutine `DUMINF`. Our versions of this algorithm are the Fortran subroutines `QuasiNewton1` (secant update of H_k) and `QuasiNewton2` (secant update of A_k) (both in the file `QN.FOR`) and the GAUSS program `QuasiNewton` in the file `Toolbox.src`.

8.6.4 Genetic Algorithms

The Gauss-Newton as well as the BFGS quasi-Newton method start from a given initial guess and move uphill along the surface of the objective function until they approach a maximizer. Thus, they may not be able to find the global maximizer. Genetic algorithms, instead, search the set of possible solutions globally.

Terminology. Genetic algorithms (GAs) use operators inspired by natural genetic variation and natural selection to evolve a set of candidate solutions to a given problem. The terminology used to describe GAs is from biology. The set of candidate solutions is called a population, its members are referred to as chromosomes, and each iteration step results in a new generation of candidate solutions. In binary-coded GAs chromosomes are represented by bit strings of a given length l. Each bit is either on (1) or off (0).

In real-coded GAs a chromosome is a point in an m-dimensional subspace of $\mathbb{R}^m$. A chromosomes fitness is its ability to solve the problem at hand. In most problems the fitness is determined by a real valued objective function that assigns higher numbers to better solutions.

Basic Structure. The evolution of a population of chromosomes consists of four stages:

1) selection of parents,
2) creation of offspring (crossover),
3) mutation of offspring,
4) and the final selection of those members of the family that survive to the next generation.

The encoding of the problem (binary or floating point) and the operators used to perform selection, crossover, and mutation constitute a specific GA. The many different choices that one can make along these dimensions give raise to a variety of specific algorithms that are simple to describe and program. Yet, at the same time, this variety is a major obstacle to any general theory that is able to explain why and how these algorithms work.[37] Intuitively, and very generally, one can think of GAs as contractive mappings operating on metric spaces whose elements are populations.[38] A mapping f is contractive if the distance between $f(x)$ and $f(y)$ is less than the distance between x and y. Under a contractive mapping an arbitrary initial population will converge to a population where each chromosome achieves the same (maximum) fitness value that is the global solution to the problem at hand. The problem with this statement is that it gives no hint as to how fast this convergence will take place, and whether specific operators accelerate or slow down convergence. Therefore, many insights in the usefulness of specific GAs come from simulation studies.

In the following we restrict ourselves to real-coded GAs. This is motivated by the kind of problems to which we apply GAs. The

[37] Mitchell (1996) as well as Michalewicz (1999) review the theoretical foundations of genetic algorithms.

[38] See, Michalewicz (1999), p.68ff.

methods presented in Chapter 3 and Chapter 4 rest on the approximation of unknown functions by linear combinations of members of a family of polynomials. The problem is to find the parameters γ_i that constitute this approximation. Usually, we have no idea about the domain of γ_i. Therefore, it is difficult to decide about the length l of the binary strings, which determines the precision of the solution. Furthermore, using floating point numbers avoids the time consuming translation to and from the binary alphabet. Yet another advantage of real-coded GAs is their capacity for the local fine tuning of the solutions.[39]

Choice of Initial Population. The initial population of a real-coded GA is chosen at random. If there are no a priori restrictions on the candidate solutions one can use a random number generator to perform this task. In our applications we use draws from the standard normal. When we pass a randomly chosen chromosome to the routine that evaluates the candidate's fitness it may happen that the chromosome violates the model's restrictions. For instance, in the parameterized expectations approach, a time path may become infeasible. In this case, the program returns a negative number and our initialization routine discards the respective chromosome. Alternatively, one may want to assign a very small fitness number to those chromosomes. After all, bad genes can mutate or generate reasonable good solutions in the crossover process.

Selection of Parents. There are many different ways to choose parents from the old generation to produce offspring for the new generation. The most obvious and simplest approach is sampling with replacement, where two integers from the set $1, 2, \ldots, n$ that index the n chromosomes of the population are drawn at random. More in the spirit of natural selection, where fitter individuals usually have a better chance to reproduce, is the concept of fitness-proportionate selection. Here, each chromosome $i = 1, 2, \ldots, n$

[39] Advantages and disadvantages of real-coded GAs vis-á-vis binary-coded GAs are discussed by Herrera et al. (1998). In their experiments most real-coded GAs are better than binary-coded GAs in minimizing a given function.

has a chance to reproduce according to its relative fitness $p(i) = f(i)/\sum_i f(i)$, where $f(i)$ denotes the fitness of chromosome i. The following code implements this selection principle:

Algorithm 8.6.4 (Fitness-proportionate Selection)

Purpose: *Choose a chromosome from the old generation for reproduction*

Steps:

Step 1: *For* $i = 1, 2, \ldots, n$ *compute* $p(i) = f(i)/\sum_i f(i)$.
Step 2: *Use a random number generator that delivers random numbers uniformly distributed in* $[0, 1]$ *and draw* $y \in [0, 1]$.
Step 3: *For* $i = 1, 2, \ldots, n$ *compute* $q(i) = \sum_{j=1}^{i} p(j)$. *If* $q(i) \geq y$ *select* i *and stop.*

In small populations the actual number of times an individual is selected as parent can be far from its expected value $p(i)$. The concept of stochastic universal sampling avoids this possibility and gives each chromosome a chance to be selected as parent that is between the floor and the ceiling of $p(i)n$.[40] Rather than choosing one parent after the other stochastic universal sampling selects n parents at a time. Each member of the old generation is assigned a slice on a roulette wheel, the size of the slice being proportionate to the chromosomes fitness $f(i)$. There are n equally spaced pointers and the wheel is spun ones. For instance, in Figure 8.9 the chromosome $i = 1$ with relative fitness $p(1)$ is not selected, whereas chromosome 4 is selected twice. Stochastic universal sampling can be implemented as follows[41]

Algorithm 8.6.5 (Stochastic Universal Sampling)

Purpose: *Choose* n *parents from the old generation for reproduction.*

[40] The floor of x is the largest integer i_1 with the property $i_1 \leq x$ and the ceiling is the smallest integer i_2 with the property $x \leq i_2$.
[41] See Mitchell (1996), p. 167.

Steps:

Step 1: For $i = 1, 2, \ldots, n$ *compute the relative fitness* $r(i) = f(i)/(\sum_i f(i)/n)$ *so that* $\sum_i r(i) = n$.
Step 2: Use a random number generator that delivers random numbers uniformly distributed in $[0, 1]$ *and draw* $y \in [0, 1]$.
Step 3: Put $i = 1$.
Step 4: Compute $q(i) = \sum_{j=1}^{i} r(j)$.
Step 5: If $q(i) > y$ *select* i *and increase* y *by 1.*
Step 6: Repeat Step 5 until $q(i) \leq y$.
Step 7: Terminate if $i = n$, *otherwise increase* i *by 1 and return to Step 4.*

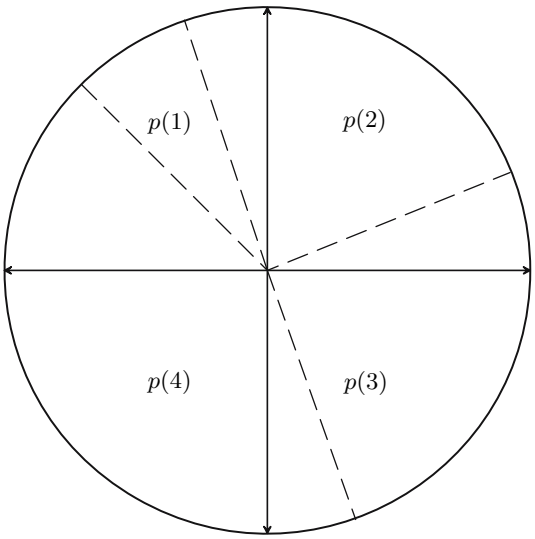

Figure 8.9: Stochastic Universal Sampling

The major problem with both fitness proportionate and stochastic universal sampling is "premature convergence". Early in the search process the fitness variance in the population is high, and under both selection schemes the small number of very fit chromosomes reproduces quickly. After a few generations they and their descendants build a fairly homogenous population that

limits further exploration of the search space. A selection scheme that deals with this problem is sigma scaling. Let σ denote the standard deviation of fitness and $\bar{f} = \sum_{i=1}^{n} f(i)/n$ the average fitness. Under sigma scaling chromosome i is assigned a probability of reproduction according to

$$p(i) := \begin{cases} 1 + \frac{f(i)-\bar{f}}{2\sigma} & \text{if } \sigma(t) \neq 0, \\ 1 & \text{if } \sigma(t) = 0. \end{cases}$$

An addition to many selection methods is "elitism": the best chromosome in the old generation replaces the worst chromosome in the new generation irrespective of whether it was selected for reproduction.[42]

Crossover. In nature the chromosomes of most species are arrayed in pairs. During sexual reproduction these pairs split and the child's chromosomes are the combination of the chromosomes of its two parents. Crossover operators mimic this process. To describe some of these operator for real-codes GAs we use the following notation: $P_1 = (p_1^1, p_1^2, \ldots, p_1^m)$ and $P_2 = (p_2^1, p_2^2, \ldots, p_2^m)$ denote the chromosomes of two parents, and $C_1 = (c_1^1, c_1^2, \ldots, c_1^m)$ and $C_2 = (c_2^1, c_2^2, \ldots, c_2^m)$ are their children.[43]

- **Simple crossover:** A position $i = 1, 2, \ldots, n-1$ is randomly chosen. The two children are:

  $$C_1 = (p_1^1, p_1^2, \ldots, p_1^i, p_2^{i+1}, \ldots, p_2^m),$$
  $$C_2 = (p_2^1, p_2^2, \ldots, p_2^i, p_1^{i+1}, \ldots, p_1^m).$$

- **Shuffle crossover:** For each position $i = 1, 2, \ldots, n$ draw a random number $\lambda \in [0, 1]$. If $\lambda < 0.5$ put

  $$c_1^i = p_1^i,$$
  $$c_2^i = p_2^i,$$

  else put

  $$c_1^i = p_2^i,$$
  $$c_2^i = p_1^i.$$

[42] See Mitchell (1996) for further selection schemes.
[43] See Herrera et al. (1998), p. 288ff.

- **Linear crossover:** Three offspring are built according to

$$c_1^i = \frac{1}{2}p_1^i + \frac{1}{2}p_2^i,$$
$$c_2^i = \frac{2}{3}p_1^i - \frac{1}{2}p_2^i,$$
$$c_3^i = -\frac{1}{2}p_1^i + \frac{3}{2}p_2^i.$$

  The two most promising offspring of the three are retained for the next generation.
- **Arithmetical crossover:** A scalar $\lambda \in [0, 1]$ is randomly chosen (or given as constant) and the chromosomes of child 1 and child 2 are build from

$$c_1^i = \lambda p_1^i + (1 - \lambda)p_2^i,$$
$$c_2^i = (1 - \lambda)p_1^i + \lambda p_2^i.$$

- **BLX-α crossover:** One child is generated, where for each $i = 1, 2, \ldots, m$ the number c^i is randomly (uniformly) chosen from the interval

$$[p_{min} - \alpha\Delta, p_{max} + \alpha\Delta],$$

  $p_{max} := \max\{p_1^i, p_2^i\}$, $p_{min} := \min\{p_1^i, p_2^i\}$, $\Delta := p_{max} - p_{min}$. Herrera et al. (1998) report good results for $\alpha = 0.5$.

Mutation. In nature, mutations, i.e., sudden changes of the genetic code, result either from copying mistakes during sexual reproduction or are triggered in the living organism by external forces as, e.g., by radiation. In binary-coded GAs the mutation operator randomly selects a position in a bit string and changes the respective bit from 0 to 1 or vice versa. Mutation operators designed for real-codes GAs also randomly select an element of a chromosome and either add or subtract another randomly selected number. Non-uniform operators decrease this number from generation to generation towards zero and, thus, allow for the local fine-tuning of the candidate solutions. The experiments of Herrera et al. (1998) show that non-uniform mutation is very

appropriate for real-coded GAs. In our algorithm we use the following operator suggested by Michalewicz (1999), p. 128. Let c_i denote the i-th element in a child chromosome selected for mutation and c_i' the mutated element. The operator selects

$$c_i' = \begin{cases} c_i + \Delta(t) & \text{if a random binary digit is 0} \\ c_i - \Delta(t) & \text{if a random binary digit is 1} \end{cases} \tag{8.94}$$
$$\Delta(t) := y(1 - r^{(1-(t/T)^b)}),$$

where y is the range of c_i and $r \in [0, 1]$ is a random number. t is the current generation and T the maximal number of iterations. Since we do not know the range of the parameters of the expectations function in advance, we draw y from a standard normal distribution. The parameter b defines the degree of non-uniformity. Michalewicz (1999) suggests $b = 2$ and Herrera et al. (1998) use $b = 5$.

Final Selection. Whether a final selection among children and parents is undertaken depends upon the choice of selection method of parents. If parents are chosen at random with replacement from generation $\mathcal{P}(t)$ one needs a final fitness tournament between parents and children to exert selection pressure. In this case the initial heterogeneity in the population decreases quickly and reasonable good solutions emerge within a few generations. However, this tight selection pressure may hinder the algorithm to sample the solution space more broadly so that only local optima are found. Therefore, there is a trade-off between tight selection and short run-time on the one hand and more precise solutions and a longer run-time on the other hand.

Implementation. This sketch of the building blocks of GAs, which is by no means exhaustive, demonstrates that the researcher has many degrees of freedom in developing his own implementation. Therefore, it is a good idea to build on GAs that have performed good in previous work.

Duffy and McNelis (2002) used a genetic algorithm to find the parameters of the approximate expectations function.[44] They

[44] In the notation used in Section 3.1.2, their solution is the minimizer of

choose four parents at random (with replacement) from the old generation. With a probability of 0.95 the best two of the four will have two children. With equal probability of selection crossover is either arithmetical, single point, or shuffle. The probability of mutations in generation T, $\pi(t)$, is given by

$$\pi(t) = \mu_1 + \mu_2/t, \tag{8.95}$$

where $\mu_1 = 0.15$ and $\mu_2 = 0.33$. Mutations are non-uniform as given in (8.94) with $b = 2$, and there is a final fitness tournament between parents and children. The two members of the family with the best fitness pass to the new generation. In addition, the best member of the old generation replaces the worst member of the new generation (elitism).

The Fortran 95 subroutine `Search1.for` implements this GA. The user can supply the following parameters in the file `GSP1.txt`:

- `npop`: the size of the population,
- `ngen`: the number of iterations (generations),
- `probc`: the probability of crossover,
- `mu1`: the first parameter in (8.95),
- `mu2`: the second parameter in (8.95),
- `mu3`: the parameter b in (8.94).

In the Fortran 95 subroutine `Search2.for` we provide a more flexible implementation of a GA to solve for the parameters of the expectations function. As in `Search1.for` the fitness criterium is the smallest absolute value of (3.9). The user can choose between two selection methods: stochastic universal sampling and the method used in `Search1.for`. We do not provide an option

$$\frac{1}{T}\sum_{t=1}^{T}[\phi(\boldsymbol{\gamma}, \mathbf{u}_{t+1}) - \psi(\boldsymbol{\gamma}, \mathbf{x}_t)]^2,$$

which does not correspond to the definition of the PEA solution found in the theoretical work of Marcet and Marshall (1992), (1994), which we use in Section 3.1.2.

Another application of a GA to the stochastic growth model is the paper of Gomme (1997), who solves for the policy function. His procedure replaces the worst half of solutions with the best half, plus some noise.

for sigma scaling, since, from our experiments, we learned that a sufficiently high probability of mutation prohibits population heterogeneity from shrinking too fast. In addition to arithmetical, single point, and shuffle crossover we allow for BLS-α and linear crossover. This is motivated by the good results obtained for these two operators in the experiments of HERRERA et al. (1998). The user can decide either to use a single operator throughout or to apply all of the operators with equal chance of selection. The program uses the same mutation operator as `Search1.for`. If stochastic universal sampling is used, there is no final fitness tournament. The two children always survive, except they provide invalid solutions (i.e., if it is not possible to compute the sequence $\{u_{t+1}(\boldsymbol{\gamma})\}_{t=0}^{T}$. If this happens, they are replaced by their parents.

The basic structure of both implementation is summarized in the following algorithm.

Algorithm 8.6.6 (Genetic Algorithm)

Purpose: *Find the minimum of a user defined objective function.*

Steps:

Step 1: Initialize: Set $t = 1$. *Choose at random an initial population of candidate solutions* $\mathcal{P}(t)$ *of size* n.

Step 2: Find a new set of solutions $\mathcal{P}(t+1)$*: Until the size of* $\mathcal{P}(t+1)$ *is* n*, repeat these steps:*

Step 2.1: Select two parents from the old population $\mathcal{P}(t-1)$.

Step 2.2: Produce two offspring (crossover).

Step 2.3: Perform random mutation of the new offspring.

Step 2.4: Depending upon the selection method in Step 2.1, either evaluate the fitness of parents and offspring and retain the two fittest or pass the two children to the next generation.

Step 3: If t=`ngen` *terminate, otherwise return to Step 2.*

Chapter 9

Various Other Tools

9.1 Difference Equations

Dynamic models are either formulated in terms of difference or differential equations. Here we review a few basic definitions and facts about difference equations.

9.1.1 Linear Difference Equations

Consider a function x that maps $t \in \mathbb{R}$ into $x(t) \in \mathbb{R}$. In practice, we do not observe economic variables x at every instant of time t. Most economic data are compiled at a yearly, quarterly, or monthly frequency. To account for that fact, we consider the function x only at equally spaced points in time: $x(t), x(t+h), x(t+2h), \ldots$, and usually we normalize $h \equiv 1$. It is then common to write x_t instead of $x(t)$.

The first difference of x_t, Δx_t, is defined as

$$\Delta x_t := x_t - x_{t-1},$$

and further differences are computed according to

$$\begin{aligned}
\Delta^2 x_t &:= \Delta x_t - \Delta x_{t-1} = x_t - 2x_{t-1} + x_{t-2},\\
\Delta^3 x_t &:= \Delta^2 x_t - \Delta^2 x_{t-1} = x_t - 3x_{t-1} + 3x_{t-2} - x_{t-3},\\
&\vdots,\\
\Delta^n x_t &:= \Delta^{n-1} x_t - \Delta^{n-1} x_{t-1}.
\end{aligned}$$

A difference equation of order n relates the function x to its n differences. The simplest of these equations is

$$\Delta x_t = x_t - x_{t-1} = a x_{t-1}, \quad a \in \mathbb{R}. \tag{9.1}$$

In this equation x_{t-1} and its first difference Δx_t are linearly related (only addition and scalar multiplication are involved). Furthermore, the coefficient at x_{t-1} does not depend on t. Therefore, equation (9.1) is a called first order linear difference equation with constant coefficient. The unknown in this equation is the function x. For this reason equations like (9.1) are known as functional equations.

Assume we know the time $t = 0$ value x_0. We can then determine all future (or past) values of x_t by iterating forwards (or backwards) on (9.1):

$$\begin{aligned} x_1 &= \lambda x_0, \quad \lambda := 1 + a, \\ x_2 &= \lambda x_1 = \lambda^2 x_0, \\ x_3 &= \lambda x_2 = \lambda^3 x_0, \\ &\vdots, \\ x_t &= \lambda x_{t-1} = \lambda^t x_0. \end{aligned}$$

In most applications we are interested in the limiting behavior of x as $t \to \infty$. The previous derivations show that x approaches zero for every initial $x_0 \in \mathbb{R}$ if and only if $|\lambda| < 1$. This behavior is called asymptotic stability.

Now, consider the generalization of (9.1) to n variables $\mathbf{x} := [x_1, x_2, \ldots, x_n]' \in \mathbb{R}^n$:

$$\mathbf{x}_t = A\mathbf{x}_{t-1}. \tag{9.2}$$

To give you an idea of how we can proceed, suppose the matrix possesses n distinct real eigenvalues and associated eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$. The matrix

$$P = [\mathbf{v}_1, \mathbf{v_2}, \ldots, \mathbf{v}_n]$$

transforms A into a diagonal matrix Λ with the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ on its main diagonal (see 8.1):

$$\Lambda = P^{-1}AP.$$

Thus, we can transform the original system (9.2) to new coordinates $\mathbf{y} = P^{-1}\mathbf{x}$:

$$\mathbf{y}_t = P^{-1}\mathbf{x}_t = P^{-1}A\mathbf{x}_{t-1} = P^{-1}AP\mathbf{y}_{t-1} = \Lambda\mathbf{y}_{t-1},$$

with solution

$$\mathbf{y}_t = \Lambda^t\mathbf{y}_0.$$

Since

$$\Lambda^t = \begin{bmatrix} \lambda_1^t & 0 & \dots & 0 \\ 0 & \lambda_2^t & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n^t \end{bmatrix},$$

a necessary and sufficient condition for asymptotic stability of $\mathbf{y}_t$ and $\mathbf{x}_t = P\mathbf{y}_t$ is that all eigenvectors λ_i have absolute value of less than unity.

Now assume (without loss of generality) that the first n_1 eigenvalues are less than one in absolute value while the remaining $n_2 = n - n_1$ eigenvalues exceed one in absolute value. If we are free to choose the initial point $\mathbf{y}_0$, we can pick

$$\bar{\mathbf{y}}_0 := [y_1, y_2, \dots, y_{n_1}, \underbrace{0, 0, \dots, 0}_{n_2 \text{ elements}}]',$$

for arbitrary $y_i \in \mathbb{R}, i = 1, 2, \dots, n_1$. By this choice we are able to ensure that $\bar{\mathbf{y}}_0$ converges to zero. This is also true for all $\bar{\mathbf{x}}_0 = P\bar{\mathbf{y}}_0$ of the original system. The set of all such starting points constitutes the stable eigenspace of the difference equation (9.2). In the two-dimensional case, we refer to this space sometimes as the saddle path.

Now suppose that one of the eigenvalues is a complex number $\lambda := \alpha + \beta i$, (see Section 8.1.1). In this case one of the columns of P is a complex vector and the respective new coordinate, say y, is a complex number, too. For a given initial y_0 the solution is nevertheless given by

$$(y_{1t} + iy_{2t}) = (\alpha + i\beta)^t(y_{10} + iy_{20}). \tag{9.3}$$

How shall we interpret this? Remember that any complex number $y = y_1 + y_2 i$ represents a vector in the Gaussian plane (see Figure 9.1). Consider the matrix

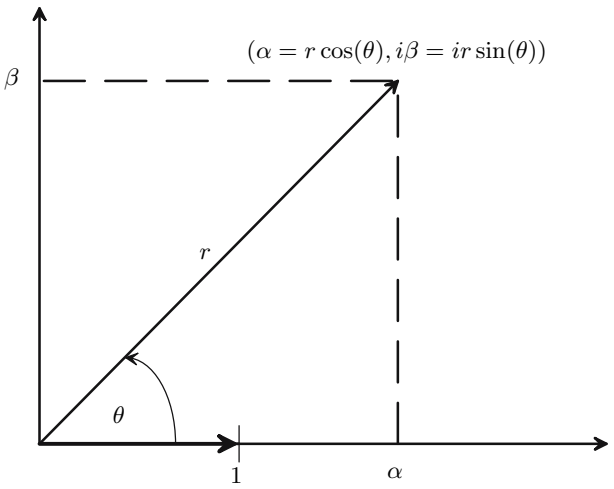

Figure 9.1: Complex Number in the Gaussian Plane

$$A := \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix}$$

and the vector

$$\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = A \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

It is the result of two simultaneous operations: the unit vector $\mathbf{e}_1 = (1, 0)'$ is rotated counter clock wise by θ degrees and stretched to a length of $r = \sqrt{\alpha^2 + \beta^2}$. Therefore, we may associate the product of two complex numbers

$$\lambda y = (\alpha + i\beta)(y_1 + iy_2) = (\alpha y_1 - \beta y_2) + i(\alpha y_2 + \beta y_1)$$

with the matrix-vector product

$$\begin{bmatrix} \alpha y_1 - \beta y_2 \\ \alpha y_2 + \beta y_1 \end{bmatrix} = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} r\cos(\theta) & -r\sin(\theta) \\ r\sin(\theta) & r\cos(\theta) \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}. \tag{9.4}$$

Thus, (9.3) may equivalently be stated as:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = A^t \begin{bmatrix} y_{10} \\ y_{20} \end{bmatrix} = r^t \begin{bmatrix} \cos(t\theta) & -\sin(t\theta) \\ \sin(t\theta) & \cos(t\theta) \end{bmatrix} \begin{bmatrix} y_{10} \\ y_{20} \end{bmatrix}. \tag{9.5}$$

This formula may be verified by induction. It holds obviously for $t = 1$ (see equation (9.4)). Suppose it is true for $t > 1$. Then

$$\begin{aligned} A^t A &= r^t \begin{bmatrix} \cos(t\theta) & -\sin(t\theta) \\ \sin(t\theta) & \cos(t\theta) \end{bmatrix} r \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \\ &= r^{t+1} \begin{bmatrix} \underbrace{\cos(t\theta)\cos(\theta) - \sin(t\theta)\sin(\theta)}_{\cos((t+1)\theta)} & \underbrace{-\cos(t\theta)\sin(\theta) - \sin(t\theta)\cos(\theta)}_{-\sin((t+1)\theta)} \\ \underbrace{\sin(t\theta)\cos(\theta) + \sin(\theta)\cos(t\theta)}_{\sin((t+1)\theta)} & \underbrace{\cos(t\theta)\cos(\theta) - \sin(t\theta)\sin(\theta)}_{\cos((t+1)\theta)} \end{bmatrix} \\ &= A^{t+1} \end{aligned}$$

by two (probably not so well known) formulas from trigonometry.[1] Thus, if the absolute value r of the complex eigenvalue λ is smaller than one, the complex number y_t approaches zero if $t \to \infty$. Geometrically, the initial vector y_0 approaches the origin of the Gaussian plane along a spirally shaped path.

These insights extend to the case of mixed real and complex eigenvalues of multiplicity $k \geq 2$, as the following theorem shows:[2]

Theorem 9.1.1 *Let A be a real square matrix. Then*

$$\lim_{t \to \infty} A^t = 0_{n \times n}$$

if and only if every eigenvalue of A is less than unity in modulus.

9.1.2 Non–Linear Difference Equations

Let $f^i : \mathbb{R}^n \to \mathbb{R}, i = 1, 2, \ldots, n$ denote arbitrary differentiable functions. A system of n non–linear first–order difference equations is defined by the following set of equations:

$$\begin{bmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{nt} \end{bmatrix} = \begin{bmatrix} f^1(x_{1\,t-1}, x_{2\,t-1}, \ldots, x_{n\,t-1}) \\ f^2(x_{1\,t-1}, x_{2\,t-1}, \ldots, x_{n\,t-1}) \\ \vdots \\ f^n(x_{1\,t-1}, x_{2\,t-1}, \ldots, x_{n\,t-1}) \end{bmatrix}. \tag{9.6}$$

[1] See, e.g., Sydsæter, Strøm, and Berck (1999), p. 15, formulas 2.56 and 2.57, respectively.

[2] See, e.g., Murata (1977), p. 85.

In more compact notation this can be written as

$$\mathbf{x}_t = \mathbf{f}(\mathbf{x}_{t-1}).$$

Assume there is a point $\bar{\mathbf{x}}$ that satisfies

$$\bar{\mathbf{x}} = \mathbf{f}(\bar{\mathbf{x}}).$$

Such a point is called a rest point or, in more economic terms, a stationary equilibrium. A powerful theorem says that the behavior of the dynamical system (9.6) near its rest point can be inferred from the properties of the linear system

$$\mathbf{x}_t = J(\bar{\mathbf{x}})\mathbf{x}_{t-1}.$$

Here, J denotes the Jacobian matrix of $\mathbf{f}$ at $\bar{\mathbf{x}}$, i.e., the matrix of partial derivatives

$$J(\bar{\mathbf{x}}) := \begin{pmatrix} f_1^1(\bar{\mathbf{x}}) & f_2^1(\bar{\mathbf{x}}) & \dots & f_n^1(\bar{\mathbf{x}}) \\ f_1^2(\bar{\mathbf{x}}) & f_2^2(\bar{\mathbf{x}}) & \dots & f_n^2(\bar{\mathbf{x}}) \\ \vdots & \vdots & \ddots & \vdots \\ f_1^n(\bar{\mathbf{x}}) & f_2^n(\bar{\mathbf{x}}) & \dots & f_n^n(\bar{\mathbf{x}}) \end{pmatrix}, \tag{9.7}$$

where

$$f_j^i(\bar{\mathbf{x}}) := \frac{\partial f^i(\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)}{\partial x_j}.$$

You will find a precise statement of this theorem for instance in Grandmont (1988), Theorems B.5.1 and B.5.2. We do not reproduce it here, since it requires additional definitions (as, for instance that of topologically conjugate maps) that are beyond the scope of this short overview. Yet, it means that if $J(\bar{\mathbf{x}})$ has n_1 eigenvalues of modulus less than one and $n_2 = n - n_1$ eigenvalues of modulus greater than one and if $J(\bar{\mathbf{x}})$ is invertible, (i.e., there is no eigenvalue equal to zero) then there is a local subspace W^s and a local subspace W^u of $\mathbb{R}^n$ that are tangent to the respective eigenspaces E^s and E^u, (see Section 8.1 on the notion of the eigenspace) and that each $\mathbf{x}_0$ in W^s near $\bar{\mathbf{x}}$ stays in W^s and converges linearly to $\bar{\mathbf{x}}$.

9.2 Markov Processes

Markov processes are an indispensable ingredient of stochastic DGE models. They preserve the recursive structure that these models inherit from their deterministic relatives. In this section we review a few results about these processes that we have used repeatedly in the development of solution methods and in applications.

Stochastic Processes. A stochastic process is a time sequence of random variables $\{Z_t\}_{t=0}^{\infty}$. If the members Z_t of this sequence have a countable number of realizations $Z_t \in \{z_1, z_2, \ldots, z_n\}$ the process is discrete valued as opposed to a continuous valued process whose realizations are taken from an interval of the real line $Z_t \in [a, b] \subseteq \mathbb{R}$. This interval is known as the support of the process. We say that the elements of a stochastic process are identically and independently distributed (iid for short), if the probability distribution is the same for each member of the process Z_t and independent of the realizations of other members of the process $Z_{t+s}, s \neq 0$. In this case the probability of the event $[Z_1 = z_1, Z_2 = z_2, \ldots Z_T = z_T]$ is given by

$$\begin{aligned} &Prob[Z_1 = z_1, Z_2 = z_2, \ldots Z_T = z_T] \\ &= Prob(Z_1 = z_1) \times Prob(Z_2 = z_2) \times \cdots \times Prob(Z_T = z_T). \end{aligned}$$

A stochastic process has the Markov property, if the probability distribution of Z_{t+1} only depends upon the realization of Z_t.

The AR(1) Process. An example of a Markov process is the first order autoregressive process (AR(1) for short)

$$Z_t = (1 - \varrho)\bar{Z} + \varrho Z_{t-1} + \epsilon_t, \quad \varrho \in [0, 1), \epsilon_t \sim N(0, \sigma^2). \tag{9.8}$$

The random variable ϵ_t, the so called innovations of the AR(1) process, are iid draws from a normal distribution with mean 0 and variance σ^2. Given Z_t, next period's shock Z_{t+1} is normally distributed with mean $E(Z_{t+1}|Z_t) = (1 - \varrho)\bar{Z} + \varrho Z_t$ and variance $\text{var}\,(Z_{t+1}|Z_t) = \sigma^2$. Since higher order autoregressive processes can be reduced to first order vector autoregressive processes the

first order process plays a prominent role in the development of stochastic Ramsey models. As an example, consider the second order autoregressive process

$$Z_t = \varrho_1 Z_{t-1} + \varrho_2 Z_{t-2} + \epsilon_t. \tag{9.9}$$

Defining $X_t = Z_{t-1}$ equation (9.9) can be written as

$$\begin{aligned} Z_{t+1} &= \varrho_1 Z_t + \varrho_2 X_t + \epsilon_t, \\ X_{t+1} &= Z_t, \end{aligned}$$

which is a first order vector autoregressive process in $(Z_t, X_t)'$ with innovations $(\epsilon_t, 0)'$.

Markov Chains. Markov chains are discrete valued Markov processes. They are characterized by three objects:

(1) The column vector $\mathbf{z} = [z_1, z_2, \ldots, z_n]'$ summarizes the n different realizations of Z_t.
(2) The probability distribution of the initial date $t = 0$ is represented by the vector $\boldsymbol{\pi}_0 = [\pi_{01}, \pi_{02}, \ldots, \pi_{0n}]'$, where π_{0i} denotes the probability of the event $Z_0 = z_i$.
(3) The dynamics of the process is represented by a transition matrix $P = (p_{ij})$, where p_{ij} denotes the probability of the event $Z_{t+1} = z_j | Z_t = z_i$, i.e., the probability that next period's state is z_j given that this period's state is z_i. Therefore, $p_{ij} \geq 0$ and $\sum_{j=1}^{m} p_{ij} = 1$.

Thus, given $Z_t = z_i$ the conditional expectation of Z_{t+1} is $E(Z_{t+1}|Z_t = z_i) = P_i\mathbf{z}$, where P_i denotes the i–th row of P and the conditional variance is $\text{var}\,(Z_{t+1}|Z_t = z_i) = \sum_j P_{ij}(z_j - P_i\mathbf{z})^2$. The probability distribution of Z_t evolves according to

$$\boldsymbol{\pi}'_{t+1} = \boldsymbol{\pi}'_t P. \tag{9.10}$$

Computation of the Ergodic Distribution. The limit of (9.10) for $t \to \infty$ is the time invariant, stationary, or ergodic distribution of the Markov chain $(\mathbf{z}, P, \boldsymbol{\pi}_0)$. It is defined by

$$\boldsymbol{\pi}' = \boldsymbol{\pi}' P \quad \Leftrightarrow (I - P')\boldsymbol{\pi} = \mathbf{0}. \tag{9.11}$$

Does this limit exist? And if it exists, is it independent of the initial distribution $\boldsymbol{\pi}_0$? The answer to both questions is yes, if either all $p_{ij} > 0$ or, if for some integer $k \geq 1$ all elements of the matrix

$$P^k := \underbrace{P \times P \cdots \times P}_{k-\text{elements}}$$

are positive, i.e., $p_{ij}^k > 0$ for all (i, j). This latter condition states that it is possible to reach each state j in at least k steps from state i.[3] Obviously, this is a weaker condition than $p_{ij} > 0$ for all (i, j). As an example, consider the transition matrix

$$P = \begin{pmatrix} 0.0 & 1.0 \\ 0.9 & 0.1 \end{pmatrix},$$

for which

$$P^2 = P \times P = \begin{pmatrix} 0.9 & 0.1 \\ 0.81 & 0.91 \end{pmatrix}.$$

We need to compute the invariant distribution in many applications. For instance, in Section 5.2 we must solve for the stationary distribution of employment in oder to find the stationary distribution of assets. The states of the respective Markov chain are $z_1 = e$ and $z_2 = u$, where e (u) denotes (un)employment, and π_{01} ($\pi_{02} = 1 - \pi_{01}$) is the probability that a randomly selected agent from the unit interval is employed in period $t = 0$. The transition matrix P is given by

$$P = \begin{pmatrix} p_{uu} & p_{ue} \\ p_{eu} & p_{ee} \end{pmatrix} = \begin{pmatrix} 0.5000 & 0.5000 \\ 0.0435 & 0.9565 \end{pmatrix},$$

where p_{uu} (p_{ue}) denotes the probability that an unemployed agent stays unemployed (becomes employed).

One obvious way to find the stationary distribution is to iterate over equation (9.10) until convergence. When we start with an arbitrary fraction of unemployed and employed agents of $(0.5, 0.5)$, say, and iterate over (9.10) we get the sequence in Table 9.1, which

Table 9.1

Iteration No.	π_u	π_e
0	0.500000	0.500000
1	0.271750	0.728250
2	0.167554	0.832446
3	0.119988	0.880012
4	0.098275	0.901725
5	0.088362	0.911638
10	0.080202	0.919798
20	0.080037	0.919963

converges quickly to $(0.08, 0.92)$, the stationary probabilities of being (un)employed.

Another procedure to compute the stationary distribution of a Markov chain is by means of Monte Carlo simulations. For the two-state chain of the previous example this is easily done: assume an initial state of employment z_{0i}, for example $z_{02} = e$. Use a uniform random number generator with the support $[0, 1]$. If the random number is less than 0.9565, $z_{12} = e$, otherwise the agent is unemployed in period 1, $z_{11} = u$. In the next period, the agent is either employed or unemployed. If employed, the agent remains employed if the random number of this period is less than 0.9565 and becomes unemployed otherwise. If unemployed, the agent remains unemployed if the random number of this period is less than 0.5 and becomes employed otherwise. Continue this process for T periods and count the number of times the agent is either employed or unemployed. The relative frequencies will converge slowly to the ergodic distribution according to the Law of Large Numbers. In our computation, we get the simulation results displayed in Table 9.2. Notice that this procedure converges very slowly. Furthermore, if the Markov chain has more than $n = 2$ states this becomes a very cumbersome procedure. For this reason, we will usually employ a third, more direct way to compute the ergodic distribution. Observe that the definition of the in-

[3] See, e.g., Ljungqvist and Sargent (2000), Theorem 1 and Theorem 2.

Table 9.2

Iteration No.	π_u	π_e
10	0.10	0.90
100	0.12	0.88
1000	0.063	0.937
10000	0.0815	0.9185
100000	0.0809	0.9191
500000	0.0799	0.9201

variant distribution (9.11) implies that $\boldsymbol{\pi}$ is an eigenvector to the eigenvalue of one of the matrix $-P'$, where $\boldsymbol{\pi}$ has been normalized so that $\sum_{i=1}^{n} \pi_i = 1$. Solving the eigenvalue problem for the matrix given above gives $\pi_1 = 0.0800$ and $\pi_2 = 0.920$. An equivalent procedure uses the fact that the matrix $I - P'$ has rank $n-1$ (given that P' has rank n) and that the π_i must sum to one. Therefore, the vector $\boldsymbol{\pi}$ must solve the following system of linear equations:

$$\boldsymbol{\pi}' \begin{bmatrix} p_{11}-1 & p_{12} & \dots & p_{1,n-1} & 1 \\ p_{21} & p_{22}-1 & \dots & p_{2,n-1} & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{n-1,1} & p_{n-1,2} & \dots & p_{n-1,n-1}-1 & 1 \\ p_{n1} & p_{n2} & \dots & p_{n,n-1} & 1 \end{bmatrix} = (0, \dots, 0, 1).$$

Markov Chain Approximations of AR(1) Processes. In Section 1.3.3 we extend the value function iteration method from Section 1.2.3 to solve the stochastic Ramsey model when the productivity shock is a finite state Markov chain. Empirically, however, the shift parameter of the production function resembles an AR(1) process. Fortunately, TAUCHEN (1986) develops a method for choosing values for the realizations and the transition matrix so that the resulting Markov chain closely mimics the underlying continuous valued autoregressive process.

Consider the process

$$Z_{t+1} = \varrho Z_t + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2).$$

The unconditional mean and variance of this process are 0 and $\sigma_Z^2 = \sigma_\epsilon^2/(1-\varrho^2)$.[4] TAUCHEN (1986) proposes to choose a grid $\mathscr{Z} = [z_1, z_2, \ldots, z_m]$ of equidistant points $z_1 < z_2, \ldots, < z_m$, whose upper end point is a multiple, say λ, of the unconditional standard deviation of the autoregressive process, $z_m = \lambda\sigma_Z$ and whose lower end point is $z_1 = -z_m$. For a given realization $z_i \in \mathscr{Z}$ the variable $z := \varrho z_i + \epsilon$ is normally distributed with mean ϱz_i and variance σ_ϵ^2. Let dz denote half of the distance between two consecutive grid points. The probability that z is in the interval $[z_j - dz, z_j + dz]$ is given by

$$prob(z_j - dz \leq z \leq z_j + dz) = \pi(z_j + dz) - \pi(z_j - dz)$$

where $\pi(\cdot)$ denotes the cumulative distribution function of the normal distribution with mean ϱz_i and variance σ^2. Equivalently, the variable $v := (z - \varrho z_i)/\sigma_\epsilon$ has a standard normal distribution. Thus, the probability to switch from state z_i to state z_j for $j = 2, 3, ..., m-1$, say p_{ij}, is given by the area under the probability density function of the standard normal distribution in the interval

$$\left[\frac{z_j - \varrho z_i - dz}{\sigma_\epsilon}, \frac{z_j - \varrho z_i + dz}{\sigma_\epsilon}\right].$$

The probability to arrive at state z_1 is the area under the probability density in the interval $[-\infty, z_1 + dz]$. Since $\sum_j p_{ij} = 1$, the probability to go from any state i to the upper bound z_m is simply $p_{im} = 1 - \sum_{j=1}^{m-1} p_{ij}$.

We summarize this method in the following steps:

Algorithm 9.2.1 (Markov Chain Approximation)

Purpose: *Finite state Markov chain approximation of first order autoregressive process*

Steps:

Step 1: Compute the discrete approximation of the realizations: Let ϱ and σ_ϵ denote the autoregressive parameter and

[4] See, e.g., HAMILTON (1994), p. 53.

the standard deviation of innovations, respectively. Select the size of the grid by choosing $\lambda \in \mathbb{R}_{++}$ *so that* $z_1 = -\lambda\sigma_\epsilon/\sqrt{1-\varrho^2}$. *Choose the number of grid points* m. *Put* $step = -2z_1/(m-1)$ *and for* $i = 1, 2, ..., m$ *compute* $z_i = z_1 + (i-1)step$.

Step 2: Compute the transition matrix $P = (p_{ij})$*: Let* $\pi(\cdot)$ *denote the cumulative distribution function of the standard normal distribution. For* $i = 1, 2, \ldots, m$ *put*

$$\begin{aligned} p_{i1} &= \pi\left(\tfrac{z_1-\varrho z_i}{\sigma_\epsilon} + \tfrac{step}{2\sigma_\epsilon}\right), \\ p_{ij} &= \pi\left(\tfrac{z_j-\varrho z_i}{\sigma_\epsilon} + \tfrac{step}{2\sigma_\epsilon}\right) - \pi\left(\tfrac{z_j-\varrho z_i}{\sigma_\epsilon} - \tfrac{step}{2\sigma_\epsilon}\right), \\ j &= 2, 3, \ldots, m-1, \\ p_{im} &= 1 - \textstyle\sum_{j=1}^{m-1} p_{ij}. \end{aligned}$$

Tauchen (1986) reports the results of Monte Carlo experiments that show that choosing $m = 9$ and $\lambda = 3$ gives an adequate representation of the underlying AR(1) process. Our GAUSS procedure `MarkovAR` in the file `ToolBox.src` implements the above algorithm. It takes ϱ, σ_ϵ, λ, and m as input and returns the vector $\mathbf{z} = [z_1, z_2, ..., z_m]'$ and the transition matrix P.

9.3 DM-Statistic

In this section we consider the DM-statistic proposed by Den Haan and Marcet (1994). It is measure of the accuracy of an approximate solution of a stochastic DGE model, based on the residuals of the model's Euler equations.

Single Equation. For the sake of concreteness let us return to the stochastic growth model in Example 1.3.2. The residual of the Euler equation (3.1b),

$$C_t^{-\eta} = \beta E_t\left[C_{t+1}^{-\eta}(1-\delta+\alpha Z_{t+1}K_{t+1}^{\alpha-1})\right],$$

is defined by

$$y_t := \beta\left[C_{t+1}^{-\eta}(1-\delta+\alpha Z_{t+1}K_{t+1}^{\alpha-1})\right] - C_t^{-\eta}.$$

Any deviation of y_t from zero is due to forecast errors. Since the household's expectations are rational, the forecast errors are uncorrelated with any variable whose realization is known to the household prior to period $t+1$. Put differently, it should not be possible to predict y_t from past information on, say, consumption C and the productivity shock Z.

Let us state this proposition in more formal terms. Consider the linear regression model

$$y_t = \sum_{i=1}^{n} a_i x_{ti} + \epsilon_t, \quad t = 1, 2, \ldots, T. \tag{9.12}$$

The x_{it} are the n variables that we use to test our proposition. For instance, this list may include consumption and the productivity shock at various lags. The error term ϵ_t captures all deviations of y_t from zero that the household cannot predict from the information conveyed by the row vector $\mathbf{x}_t := [x_{t1}, x_{t2}, \ldots, x_{tn}]'$. When we say y_t is unpredictable, we posit $\mathbf{a} = [a_1, a_2, \ldots, a_n]' = \mathbf{0}$. A bad solution, however, should violate this condition. Using the usual econometric notation, $\mathbf{y} := [y_1, y_2, \ldots, y_T]'$ and $X = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T]$, the normal equations of the least squares estimator of $\mathbf{a}$, denoted by $\hat{\mathbf{a}}$, may be written as

$$X'X\hat{\mathbf{a}} = X'\mathbf{y}.$$

Thus, $E(\hat{\mathbf{a}}) = \mathbf{a} = \mathbf{0}$, is equivalent to

$$E(X'\mathbf{y}) = \mathbf{0}.$$

The sample analog of $E(X'\mathbf{y})$ is

$$\mathbf{q} := \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} y_t x_{t1} \\ \frac{1}{T}\sum_{t=1}^{T} y_t x_{t2} \\ \vdots \\ \frac{1}{T}\sum_{t=1}^{T} y_t x_{tn} \end{bmatrix}. \tag{9.13}$$

Of course, given any time series of y_t and $\mathbf{x}_t$ computed from a solution of the model, the vector $\mathbf{q}$ is never precisely equal to the

zero vector, as is the estimate $\hat{\mathbf{a}}$. DEN HAAN and MARCET (1994) propose a Wald-type statistic to test, whether any deviations of (9.13) from zero are only due to sampling variability. Specifically, they propose the statistic

$$\mathrm{DM}(n) := T\mathbf{q}' \left[\hat{\mathrm{var}}\,(\mathbf{q})\right]^{-1} \mathbf{q},$$

where $\hat{\mathrm{var}}\,(\mathbf{q})$ is a consistent estimate of the variance of $\mathbf{q}$. This variance is given by:

$$\begin{aligned}\mathrm{var}\,(\mathbf{q}) &:= E[(\mathbf{q} - E(\mathbf{q}))(\mathbf{q} - E(\mathbf{q}))'] = E[\mathbf{q}\mathbf{q}'], \\ &= E[X'\mathbf{y}\mathbf{y}'X] = E[X'(X\mathbf{a} + \boldsymbol{\epsilon})(\boldsymbol{\epsilon}' + \mathbf{a}'X')X], \\ &= E[(X'X\mathbf{a} + X'\boldsymbol{\epsilon})(\boldsymbol{\epsilon}'X + \mathbf{a}'X'X)] = E[X'\boldsymbol{\epsilon}\boldsymbol{\epsilon}'X].\end{aligned}$$

It is well known from WHITE (1980) that a consistent estimator of $E(X'\boldsymbol{\epsilon}\boldsymbol{\epsilon}'X)$ is given by

$$\hat{\mathrm{var}}\,(\mathbf{q}) = \frac{1}{T}X'\hat{\Sigma}X, \quad \hat{\Sigma} = \begin{bmatrix} \hat{\epsilon}_1^2 & 0 & 0 & \dots & 0 \\ 0 & \hat{\epsilon}_2^2 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \hat{\epsilon}_T^2 \end{bmatrix}, \tag{9.14}$$

where $\hat{\epsilon}_t = y_t - X\hat{\mathbf{a}}$ is the estimated error for observation t. This estimate is based on the assumption that the errors in the regression (9.12) are not autocorrelated. This assumption can be violated, if the conditional expectation on the rhs of the Euler equation includes variables dated $t + 2$ and later. Consistent covariance estimators for this case can be found in the literature on the generalized method of moments estimator.[5] Note, however, that non of our applications belongs to this class of problems. The asymptotic distribution of the DM statistic is χ^2 with n degrees of freedom.[6] Replacing $\hat{\mathrm{var}}\,(\mathbf{q})$ by the estimator (9.14) the DM-statistic can be rewritten as

$$\mathrm{DM}(n) = \mathbf{y}'X \left[\sum_t \mathbf{x}_t\mathbf{x}_t'\hat{\epsilon}_t^2\right]^{-1} X'\mathbf{y}. \tag{9.15}$$

[5] See, e.g., NEWEY and WEST (1987). .
[6] See DEN HAAN and MARCET (1994).

Any approximate solution never exactly satisfies the condition $E(X'\mathbf{y}) = \mathbf{0}$. Hence, if the researcher uses a very large sample size T the statistic will discover this and reject the null. Therefore, the DM-test is the more stringent, the larger T is.

To reduce the type I error (rejection of the null when it is true) Den Haan and Marcet (1994) propose the following procedure: For a given sequence of shocks compute the approximate solution for a large T; use this solution, draw a new sequence of shocks for a sample size T_2 much smaller than T, compute the respective time path of the model's variables and calculate the DM-statistic for these observations. Repeat this very often[7] and compute the percentage of the DM-statistic that is below the lower or above the upper 2.5 percent critical values of the $\chi^2(m)$ distribution, respectively. If these fractions differ markedly from the theoretical 5 percent, this indicates an inaccurate solution.

Multiple Equations. The DM-statistic is also applicable to models with more than one Euler equation. The simplest thing to do, of course, is to compute this statistic for every single equation. However, this neglects the fact that the respective equations are interrelated and it does not provide an answer to the question whether the approximate solution in general is sufficiently accurate. Fortunately it is not difficult to generalize the presentation from the previous paragraph to the case of m Euler equations.

Let $\mathbf{y}_j := [y_{j1}, y_{j2}, \ldots, y_{jT}]$ denote the residual computed from the model's j-th Euler equation, $j = 1, 2, \ldots, m$. Then we wish to test, whether the nm column vector

$$\mathbf{q} := \begin{bmatrix} X'\mathbf{y}_1/T \\ X'\mathbf{y}_2/T \\ \vdots \\ X'\mathbf{y}_m/T \end{bmatrix} \tag{9.16}$$

is close to the zero vector. Note that we use the same set of explanatory variables X in each of the m regressions. The variance of $\mathbf{q}$ is given by

[7] In their examples, Den Haan and Marcet (1994) compute 500 realizations of DM with a sample size of T_2=3,000 and $T = 29,000$.

$$\begin{aligned}\operatorname{var}(\mathbf{q}) &= E\begin{bmatrix} X'\mathbf{y}_1\mathbf{y}_1'X & \dots & X'\mathbf{y}_1\mathbf{y}_m'X \\ \vdots & \ddots & \vdots \\ X'\mathbf{y}_m\mathbf{y}_1'X & \dots & X'\mathbf{y}_m\mathbf{y}_m'X \end{bmatrix}, \\ &= E\begin{bmatrix} X'\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_1'X & \dots & X'\boldsymbol{\epsilon}_1\boldsymbol{\epsilon}_m'X \\ \vdots & \ddots & \vdots \\ X'\boldsymbol{\epsilon}_m\boldsymbol{\epsilon}_1'X & \dots & X'\boldsymbol{\epsilon}_m\boldsymbol{\epsilon}_m'X \end{bmatrix},\end{aligned}$$

where $\boldsymbol{\epsilon}_j = [\epsilon_{1j}, \epsilon_{2j}, \dots, \epsilon_{Tj}]'$ is the vector of errors in the j-th regression of $\mathbf{y}_j$ on the vector $\mathbf{a}_j$ and X. A consistent estimate of this matrix in the case of heteroscedastic but serially uncorrelated errors is given by

$$\hat{\operatorname{var}}(\mathbf{q}) := \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} \hat\epsilon_{1t}^2 & \hat\epsilon_{1t}\hat\epsilon_{2t} & \dots & \hat\epsilon_{1t}\hat\epsilon_{mt} \\ \hat\epsilon_{2t}\hat\epsilon_{1t} & \hat\epsilon_{2t}^2 & \dots & \hat\epsilon_{2t}\hat\epsilon_{mt} \\ \vdots & \vdots & \ddots & \vdots \\ \hat\epsilon_{mt}\hat\epsilon_{1t} & \hat\epsilon_{mt}\hat\epsilon_{2t} & \dots & \hat\epsilon_{mt}^2 \end{bmatrix} \otimes \mathbf{x}_t'\mathbf{x}_t, \qquad (9.18)$$

where $\otimes$ denotes the Kronecker product. The statistic

$$\mathrm{DM}(nm) = T\mathbf{q}'\left[\hat{\operatorname{var}}(\mathbf{q})\right]^{-1}\mathbf{q}$$

with $\mathbf{q}$ and $\hat{\operatorname{var}}(\mathbf{q})$ as defined in (9.16) and (9.18), respectively, is asymptotically distributed as a $\chi^2(nm)$ random variable.

9.4 The HP-Filter

In this section we consider the Hodrick-Prescott or for short the HP-filter that has been used in numerous studies to derive the cyclical component of a time series. This filter is proposed in a discussion paper by Robert Hodrick and Edward Prescott that circulated in the nineteen eighties and which was recently published.[8]

Let $(y_t)_{t=1}^T$ denote the log of a time series that may be considered as realization of a non-stationary stochastic process. The

[8] See Hodrick and Prescott (1980) and (1997), respectively.

growth component $(g_t)_{t=1}^T$ of this series as defined by the HP-Filter is the solution to the following minimization problem:

$$\min_{(g_t)_{t=1}^T} \sum_{t=1}^{T}(y_t - g_t) + \lambda \sum_{t=2}^{T-1}[(g_{t+1} - g_t) - (g_t - g_{t-1})]^2. \tag{9.19}$$

The parameter λ must be chosen by the researcher. Its role can be easily seen by considering the two terms to the right of the minimization operator. If λ were equal to zero, the obvious solution to (9.19) is $y_t = g_t$, i.e., the growth component were set equal to the original series. As λ gets large and larger it becomes important to keep the second term as small as possible. Since this term equals the growth rate of the original series between two successive periods, the ultimate solution for $\lim_\lambda \to \infty$ is a constant growth rate g. Thus, by choosing the size of the weight λ the filter returns anything between the original time series and a linear time trend.

The first order conditions of the minimization problem imply the following system of linear equations:

$$A\mathbf{g} = \mathbf{y}, \tag{9.20}$$

where $\mathbf{g} = [g_1, g_2, \ldots, g_T]'$, $\mathbf{y} = [y_1, y_2, \ldots, y_T]'$, and A is the tridiagonal matrix

$$\begin{bmatrix}
1+\lambda & -2\lambda & \lambda & 0 & 0 & \ldots & 0 & 0 & 0 \\
-2\lambda & 1+5\lambda & -4\lambda & \lambda & 0 & \ldots & 0 & 0 & 0 \\
\lambda & -4\lambda & 1+6\lambda & -4\lambda & \lambda & \ldots & 0 & 0 & 0 \\
0 & \lambda & -4\lambda & 1+6\lambda & -4\lambda & \ldots & 0 & 0 & 0 \\
0 & 0 & \lambda & -4\lambda & 1+6\lambda & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & \ldots & 1+6\lambda & -4\lambda & \lambda \\
0 & 0 & 0 & 0 & 0 & \ldots & -4\lambda & 1+5\lambda & -2\lambda \\
0 & 0 & 0 & 0 & 0 & \ldots & \lambda & -2\lambda & 1+\lambda
\end{bmatrix}.$$

Note, that A can be factored in[9]

[9] See Brandner and Neusser (1990), p. 5.

$$A = I + \lambda K'K,$$

$$K = \begin{bmatrix} 1 & -2 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & -2 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & 1 & -2 & 1 \end{bmatrix},$$

which shows that A is positive definite.[10] Linear algebra routines that use spares matrix methods[11] can be used to solve the system (9.20). These methods require considerably less workspace than methods that operate on the matrix A. For instance, the Fortran subroutine `DLSLQS` available in the `IMSL` library requires only the main and the two upper codiagonals of A, i.e., a $3 \times T$-matrix, whereas general linear system solvers require the full $T \times T$ matrix A. Our implementation of the HP-Filter in the GAUSS procedure `HPFilter` in the file `Toolbox.src` uses the command `bandsolpd` to solve (9.20).

The cyclical component of $\mathbf{y}$,

$$\mathbf{c} = \mathbf{y} - \mathbf{g} = [I - A^{-1}]\mathbf{y},$$

remains unchanged if a linear time trend

$$\mathbf{a} := \begin{bmatrix} a_1 + a_2 \\ a_1 + 2a_2 \\ \vdots \\ a_1 + Ta_2 \end{bmatrix}$$

[10] A matrix A is called positive definite, if for each vector $\mathbf{x} \neq \mathbf{0}$

$$\mathbf{x}'A\mathbf{x} > 0.$$

The matrix $I + \lambda K'K$ clearly satisfies this requirement, since

$$\mathbf{x}'[I + \lambda K'K]\mathbf{x} = \sum_{i=1}^{T} x_i^2 + \lambda \sum_{i=1}^{T} z_i^2, \ \mathbf{z} := K\mathbf{x}.$$

[11] A sparse matrix is a matrix that has most of its entries set to zero. This can be used to reduce the size of the memory in which the matrix is stored on a computer and to develop fast algorithms that operate with these matrices.

is added to the time series $(y_t)_{t=1}^T$. To see this, note that[12]

$$\mathbf{c} = [I - A^{-1}][\mathbf{y} + \mathbf{a}] = [I - A^{-1}]\mathbf{y} + \underbrace{[I - A^{-1}]\mathbf{a}}_{=\mathbf{0}}.$$

The usual choice of the filter weight is $\lambda = 1600$ for quarterly data. It rests on the observation that with this choice the filter "removes all cycles longer than 32 quarters leaving shorter cycles unchanged" (Brandner and Neusser (1990), p. 7). For yearly data Ravn and Uhlig (2001) propose $\lambda = 6.5$ whereas Baxter and King (1999) advocate for $\lambda = 10$.

[12] This statement can be proven by noting that

$$[I - A^{-1}]\mathbf{a} = \mathbf{0} \Leftrightarrow A^{-1}[A - I]\mathbf{a} = \mathbf{0} \Leftrightarrow [A - I]\mathbf{a} = \mathbf{0},$$

and considering the product on the rightmost side of this statement.

Epilogue

In this book, we present a number of algorithms to compute dynamic general equilibrium models. In Chapters 1 to 4 we introduce you to six different methods to solve representative agent models and in Chapters 5 to 7 we consider methods and applications from the active field of research on the income and wealth distribution. So at the end of this book you may probably ask two questions: "Is there a favorite method to solve representative agent models?" and "What are future areas of research?" We consider each in turn.

Probably the most easy to use and widely applicable method is log-linearization. It is not restricted to models whose competitive equilibrium is Pareto optimal and hence solves a social planning problem (as the linear-quadratic method) or that have only a small number of states (as Galerkin, least squares projection methods and value function iteration). Its possible shortcomings are that it ignores the model's non-linearities and that the policy functions do not incorporate the information from higher moments of the stochastic processes that drive the model. Our results for the benchmark model show that these deficiencies do not surface in the second moments that the researcher usually is most interested in. Our results in Tables 1.2, 2.1, and 3.2 show no noteworthy difference across the various methods, and this is confirmed by a more elaborate study in Heer and Maussner (2004a). Of course, this does not prove that log-linearization provides the correct answers for all models one might imagine. Therefore, a reasonable strategy might be to start with the log-linear solution and to check its adequacy by considering any of the other methods that suits your needs. In terms of ease of implementation we advocate for the parameterized expectations approach.

The dynamics of heterogeneous-agent economies is an area of active and ongoing research. With the advances in the computer technology, numerical methods can be applied to an increasingly number of dimensions of the state space. While this book went to press, KEN JUDD was organizing a conference at Stanford University where the leading researchers in the field of computational economics came together in order to discuss advances in the computation of heterogeneous-agent economies. A follow-up conference volume is planned to be published in the *Journal of Economic Dynamics and Control* in 2005 or later.[13] Two problems are considered: First, the solution of a model with a discrete number of heterogeneous agent, and second, a model with a continuum of heterogeneous agents. We would like to conclude this chapter by commenting on the first problem set, which might naturally arise in the study of OLG models. In this chapter, we only studied aggregate uncertainty in the log-linearized version of the OLG model. If we were trying to approximate the non-linear dynamics of a 60 period OLG model with aggregate uncertainty, we would face severe problems as the state space is equal to 60 variables (the capital stock of the 59 generations[14] and the aggregate technology). As you learned, methods that rely upon the discretization of the state space run into problems if the model has more than two or three individual state variables. The computational time becomes exorbitant with current computer technology. In Chapters 3 and 4, you were introduced to the Parameterized Expectations and Projection methods, which were able to handle problems with a higher number of state variables. At this moment, we do not know of any method that is able to compute the non-linear dynamics with 60 state variables, with the possible exception of parameterized expectations, even though we are unaware of any application yet. Our own research, for example, indicates that projection methods can reasonably be applied to the computation of models with a number of approximately 15-20 of individual states (see HEER and MAUSSNER (2004b)). Similar

[13] The reader is encouraged to look out for this volume of the journal.
[14] The capital stock of the first generation is equal to zero.

results are obtained by KRUEGER and KUEBLER (2004) who use the so-called *Smolyak's Algorithm* in their collocation projection in order to compute the nonlinear dynamics in an overlapping generations model with stochastic aggregate production. At this moment, however, we conjecture that it will not take long until we will be able to solve such problems.

References

Aiyagari, Rao S. 1994. Uninsured Idiosyncratic Risk and Aggregate Saving. *Quarterly Journal of Economics*. Vol. 109. pp. 659-84.

Aiyagari, Rao S. 1995. Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints and Constant Discounting. *Journal of Political Economy*. Vol. 103. pp. 1158-75.

Ambler, Steve and Alain Paquet. 1996. Fiscal spending shocks, endogenous government spending, and real business cycles. *Journal of Economic Dynamics and Control*. Vol. 20. pp. 237-56.

Atkinson, Anthony B., Francois Bourguignon, and Christian Morrisson. 1992. Empirical Studies of Earnings Mobility, Chur: Harwood

Auerbach, Alan J. and Laurence J. Kotlikoff. 1987. Dynamic Fical Policy, Cambridge: Cambridge University Press.

Auerbach, Alan J., Laurence J. Kotlikoff, and J. Skinner. 1983. The Efficiency Gains from Dynamic Tax Reform. *International Economic Review*. Vol. 24. pp. 81-100.

Baxter, Marianne and Robert G. King. 1993. Fiscal policy in general equilibrium. *American Economic Review*. Vol. 83. pp. 315-334.

Baxter, Marianne and Robert G. King. 1999. Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series. *Review of Economics and Statistics*. pp. 575-593.

Becker, Gary S. and Nigel Tomes. 1979. An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy*. Vol. 87. pp. 1153-89.

Benhabib, Jess, Richard Roberson, and Randall Wright. 1991. Homework in Macroeconomics: Household Production and

Aggregate Fluctuations, *Journal of Political Economy*, Vol. 99, pp. 1166-87

Bergin, Paul R. and Robert C. Feenstra. 2000. Staggered Price Setting, Translog Preferences, and Endogenous Persistence. *Journal of Monetary Economics*. Vol. 45. pp. 657-680.

Blanchard, Oliver J. and Charles M. Kahn. 1980. The Solution of Linear Difference Models Under Rational Expectations. *Econometrica*. Vol. 48. pp. 1305-1311.

Boldrin, Michele, Lawrence C. Christiano and Jonas D.M. Fisher. 2001. Habit Persistence, Asset Returns and the Business Cycle. *American Economic Review*. Vol. 91. pp. 149-166.

Bomsdorf, Eckart. 1989. Die personelle Vermögensverteilung in der Bundesrepublik Deutschland 1973, 1978 und 1983. *Vierteljahreshefte zur Wirtschaftsforschung*. Heft 4.

Brandner, Peter and Klaus Neusser. 1990. Business Cycles in Open Economies: Stylized Facts for Austria and Germany. Austrian Institute for Economic Research. Working Papers No. 40.

Brock, William A. and Leonard Mirman. 1972. Optimal Economic Growth and Uncertainty: The Discounted Case. *Journal of Economic Theory*. Vol. 4. pp. 479-513.

Bronson, Richard. 1989. Theory and Problems of Matrix Operations. New York: McGraw-Hill.

Brooks, Robin. 2002. The equity premium and the baby boom, Research Department. IMF. unpublished manuscript.

Browning, Martin, Lars Peter Hansen and James J. Heckman. 1999. Micro Data and General Equilibrium Models. In John B. Taylor and Michael Woodfors (Eds.). Handbook of Macroeconomics. Vol. 1A. Amsterdam: North-Holland. pp. 543-633.

Budría, Santiago, Javier Díaz-Giménez, Vincenzo Quadrini, and José-Victor Ríos-Rull. 2001. New Facts on the distributions of Earnings, Income and Wealth. Mimeo.

Burkhauser, Richard V., Douglas Holtz-Eakin and Stephen E. Rhody. 1997. Labor Earnings Mobility and Inequality in the United States and Germany during the Growth Years of the 1980s. *International Economic Review*. Vol. 38. pp. 775-94.

Burnside, Craig. 1999. Real Business Cycle Models: Linear Approximation and GMM Estimation. Mimeo. Revision 7, May 1.

Calvo, Guillermo A. 1983. Staggered Prices in a Utility-Maximizing Framework. *Journal of Monetary Economics*. Vol. 12. pp. 383-398.

Cass, David. 1965. Optimum Growth in an Aggregative Model of Capital Accumulation. *Review of Economic Studies*. Vol. 32. pp. 233-240.

Castañeda, Ana, Javier Díaz-Giménez and José-Victor Ríos-Rull. 1998a. Earnings and Wealth Inequality and Income Taxation: Quantifying the Trade-Offs of Switching to a Proportional Tax in the US. Mimeo.

Castañeda, Ana, Javier Díaz-Giménez and José-Victor Ríos-Rull. 1998b. Exploring the income distribution business cycle dynamics. *Journal of Monetary Economics*. Vol. 42. pp. 93-130.

Castañeda, Ana, Javier Díaz-Giménez and José-Victor Ríos-Rull. 2004. Earnings and Wealth Inequality and Income Taxation: Quantifying the Trade-Offs of Switching to a Proportional Tax in the US. *Journal of Political Economy*. Forthcoming.

Caucutt, Elisabeth M., Ayse İmrohoroğlu and Krishna B. Kumar. 2003. Growth and welfare analysis of tax progressivity in a heterogeneous-agent model. *Review of Economic Dynamics*. Vol. 6. pp. 546-77.

Chari, Varadaraian V., Patrick J. Kehoe and Ellen R. McGrattan. 2000. Sticky Price Models of the Business Cycle: Can the Contract Multiplier Solve the Persistence Problem? *Econometrica*. Vol. 68. pp. 1151-1179.

Chiu , W.H. and E. Karni, 1998, Endogenous Adverse Selection and Unemployment Insurance, *Journal of Political Economy*, vol. 106, 806-27.

Cho, Jang-Ok and Thomas F. Cooley. 1995. The Business Cycle with Nominal Contracts. *Economic Theory*. Vol. 6. pp. 13-33.

Chow, Gregory C. 1997. Dynamic Economics. Optimization by the Lagrange Method. New York, Oxford: Oxford University Press.

Costain, James S. 1997. Unemployment Insurance with Endogenous Search Intensity and Precautionary Saving. Universitat Pompeu Fabra Economics Working Paper No. 243.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 1997. Sticky Price and Limited Participation Models of Money: A Comparison. *European Economic Review*. Vol. 41. pp. 1201-1249.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans. 1999. Monetary Policy Shocks: What Have We Learned and to What End? In John B. Taylor and Michael Woodford (Eds.). Handbook of Macroeconomics. Vol. 1A. Amsterdam: Elsevier. pp. 65-148.

Christiano, Lawrence J. and Jonas D.M. Fisher. 2000. Algorithms for Solving Dynamic Models with Occasionally Binding Constraints. *Journal of Economic Dynamics and Control*. Vol. 24. pp. 1179-1232.

Christiano, Lawrence J. and Richard M. Todd. 1996. Time to Plan and Aggregate Fluctuations. *Federal Reserve Bank of Minneapolis Quarterly Review*, Vol. 20, 1996, pp. 14-27

Cochrane, John H. 1998. What do the VARs Mean? Measuring the Output Effects of Monetary Policy. *Journal of Monetary Economics*. Vol. 41. pp. 277-300.

Constantini, Paolo and Fabio Fontanella. 1990. Shape-preserving Bivariate Interpolation. *SIAM Journal of Numerical Analysis*. Vol. 27. pp. 488-506.

Cooley, Thomas F., Gary D. Hansen. 1989. The Inflation Tax in a Real Business Cycle Model. *American Economic Review*. 79. pp. 733-748.

Cooley, Thomas F. and Gary D. Hansen. 1995. Money and the Business Cycle. In Thomas F. Cooley (Ed.). Frontiers of Business Cycle Research. Princeton, NJ: Princeton University Press. pp. 175-216.

Cooley, Thomas F. and Gary D. Hansen. 1998. The Role of Monetary Shocks in Equilibrium Business Cycle Theory: Three Examples. *European Economic Review*. Vol. 42. pp. 605-617.

Cooley, Thomas F. and Edward C. Prescott. 1995. Economic Growth and Business Cycles. In Thomas F. Cooley (Ed.).

Frontiers of Business Cycle Research. Princeton, NJ: Princeton University Press. pp. 1-38.

Deaton, Angus. 1991. Saving and liquidity constraints. *Econometrica*. Vol. 59. pp. 1221-48.

Debreu, Gerard. 1954. Valuation Equilibrium and Pareto Optimum. *Proceedings of the National Academy of Science*. Vol. 40. pp. 588-592.

De Nardi, Mariachristina, Selahattin İmrohoroğlu and Thomas J. Sargent. 1999. Projected US Demographics and Social Security. *Review of Economic Dynamics*. Vol. 2. pp. 575-615.

Den Haan, Wouter J. and Albert Marcet. 1990. Solving the Stochastic Growth Model by Parameterizing Expectations. *Journal of Business & Economic Statistics*. Vol. 8. pp. 31-34.

Den Haan, Wouter J. and Albert Marcet. 1994. Accuracy in Simulations. *Review of Economic Studies*. Vol. 61. pp. 3-17.

Den Haan, Wouter J. 1997. Solving Dynamic Models with Aggregate Shocks and Heterogenous Agents. *Macroeconomic Dynamics*. Vol. 1. pp. 335-386.

Den Haan, Wouter J., Garey Ramey and Joel C. Watson. 2000. Job Destruction and the Propagation of Shocks. *American Economic Review*. Vol. 90. pp. 482-98.

Dennis, John E. and Robert B. Schnabel. 1983. Numerical Methods for Unconstrained Optimization and Nonlinear Equations: Lehrbuch. Englewood Cliffs, NJ: Prentice-Hall.

Diamond, Peter. 1965. National Debt in a Neoclassical Growth Model. *American Economic Review*. Vol. 55. pp. 1126-50.

Díaz-Giménez, Javier, Vincenzo Quadrini and José-Victor Ríos-Rull. 1997. Dimensions of Inequality: Facts on the US Distributions of Earnings, Income, and Wealth. *Federal Reserve Bank of Minneapolis Quarterly Review*. Vol. 21. pp. 3-21.

Duffy, John and Paul D. McNelis. 2001. Approximating and simulating the stochastic growth model: Parameterized expectations, neural networks, and the genetic algorithm. *Journal of Economic Dynamics and Control*. Vol. 25. pp. 1273-1303.

Eichenbaum, Martin and Jonas D. M. Fisher. 2004. Evaluating the Calvo Model of Sticky Prices. NBER Working Paper No. W10617.

Evans, George W. and Seppo Honkapohja. 2001. Learning and Expectations in Macroeconomics. Princeton, Oxford: Pinceton University Press.

Evans, Owen J. 1983. Tax Policy, the Interest Elasticity of Saving, and Capital Accumulation: Numerical Analysis of Theoretical Models. *American Economic Review*. Vol. 83. pp. 398-410.

Fair, Ray C. and John B. Taylor. 1983. Solution and Maximum Likelihood Estimation of Dynamic Nonlinear Rational Expectations Models. *Econometrica*. Vol. 51. pp. 1169-1185.

Farmer, Roger E. A. 1993. The Macroeconomics of Self-Fulfilling Prophecies. Cambridge, MA, London: MIT Press.

Finn, Mary G. 1995. Variance Properties of Solow's Productivity Residual and their Cyclical Implications. *Journal of Economic Dynamics and Control*. Vol. 19. pp. 1249-1281.

Gagnon, Joseph E. 1990. Solving the Stochastic Growth Model by Deterministic Extended Path. *Journal of Business & Economic Statistics*. Vol. 8. pp. 35-36.

Goldfarb, Donald. 1976. Factorized Variable Metric Methods for Unconstrained Optimization *Mathematics of Computation*. Vol. 30. pp. 796-811.

Golub, Gene H. and Charles F. Van Loan. 1996. Matrix Computations. 3rd. Ed. Baltimore and London: The Johns Hopkins University Press.

Gomme, Paul. 1997. Evolutionary Programming as a Solution Technique for the Bellman Equation. Federal Reserve Bank of Cleveland Working Paper No. 9816.

Grandmont, Jean-Michel. 1988. Nonlinear Difference Equations, Bifurcations and Chaos: An Introduction, Centre D'Etudes Prospectives D'Economie Mathematique Appliquees A La Planification (CEPREMAP) #8811.

Greene, William H. 2003. Econometric Analysis. 5th Ed. Upper Saddle River, NJ: Prentice Hall.

Greenwood, Jeremy, Zvi Hercowitz and Gregory W. Huffman. 1988. Investment, Capacity Utilization, and the Real Business Cycle. *American Economic Review*. Vol. 78. pp. 402-417.

Grüner, Hans Peter and Burkhard Heer. 2000. Optimal Flat-rate Taxes on Capital, A reexamination of Lucas' Supply Side Model', *Oxford Economic Papers*, Vol. 52, pp. 289-305

Hairault, Jean-Oliver and Franck Portier. 1993. Money, New-Keynesian Macroeconomics and the Business Cycle. *European Economic Review*. Vol. 37. pp. 1533-1568.

Hamilton, James D. 1994. Time Series Analysis. Princeton, NJ: Princeton University Press.

Hansen, Gary D. 1985. Indivisible Labor and the Business Cycle *Journal of Monetary Economics*. Vol. 16. pp. 309-327.

Hansen, Gary D. 1993. The cyclical and secular behavior of the labor input: comparing efficiency units and hours worked. *Journal of Applied Econometrics*. Vol. 8. pp. 71-80.

Hansen, Gary D. and Edward C. Prescott. 1995. Recursive Methods for Computing Equilibria of Business Cycle Models. In Thomas F. Cooley (Eds.). Frontiers of Business Cycle Research. Princeton, NJ: Princeton University Press. pp. 39-64.

Hansen, Gary D. and Ayse İmrohoroğlu. 1992. The role of Unemployment Insurance in an Economy with Liquidity Constraints and Moral Hazard. *Journal of Political Economy*. Vol. 100. no. 1. pp. 118-42.

Harris, Milton. 1987. Dynamic Economic Analysis. Oxford u.a.: Oxford University Press.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998. Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents. *Review of Economic Dynamics*. Vol. 1. pp. 1-58.

Heer, Burkhard. 2001a. On the welfare gain from stabilizing cyclical fluctuations. *Applied Economic Letters*. Vol. 8. pp. 331-34.

Heer, Burkhard. 2001b. Wealth Distribution and Optimal Inheritance Taxation in Life-Cycle Economies. *Scandinavian Journal of Economics*. Vol. 103. pp. 445-65.

Heer, Burkhard. 2003. Employment and Welfare Effects of a Two-Tier Unemployment Compensation System. *International Tax and Public Finance*. Vol. 10. pp. 147-68.

Heer, Burkhard. 2004. Nonsuperneutrality of money in the Sidrauski model with heterogenous agents. *Economics Bulletin*. Vol. 5. pp. 1-6.

Heer, Burkhard and Mark Trede. 2003. Efficiency and Distribution Effects of a Revenue-neutral Income Tax Reform. *Journal of Macroeconomics*. Vol. 25. pp. 87-107.

Heer, Burkhard and Alfred Maußner. 2004a. Computation of Business Cycle Models: A Comparision of Numerical Methods. CESifo Working Paper No. 1207.

Heer, Burkhard, and Alfred Maußner. 2004b. Projection methods and the curse of dimensionality, mimeo

Herrera, Fransisco, Manuel Lozano and José L. Verdegay. 1998. Tackling Real-Coded Genetic Algorithms: Operators and Tools for Behavioural Analysis. *Artificial Intelligence Review*. Vol. 12. pp. 265-319.

Hirsch, Morris W. and Stephen Smale. 1974. Differntial Equations, Dynamical Systems, and Linear Algebra. New York: Academic Press.

Hodrick, Robert J. and Edward C. Prescott. 1980. Post-War U.S. Business Cycles: An Empirical Investigation. University of Warwick, Discussion Paper No. 451.

Hodrick, Robert J. and Edward C. Prescott. 1997. Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit, and Banking*. Vol. 29. pp. 1-16.

Hoffman, Dennis L., Robert H. Rasche and Margie A. Tieslau. 1995. The Stability of Long-Run Money Demand in Five Industrial Countries. *Journal of Monetary Economics*. Vol. 35. pp. 317-339.

Holland, Allison, and Andrew Scott. 1998. The Determinants of UK Business Cycles, *Economic Journal*, Vol. 108, pp. 1067-1092

Hubbard, Robert G. and Kenneth L. Judd. 1987. Social security and individual welfare: precautionary saving, borrowing constraints, and the payroll tax. *American Economic Review*. Vol. 77. pp. 630-46.

Hubbard, Robert G., Jonathan Skinner, and Stephen P. Zeldes. 1995. Precautionary Saving and Social Insurance. *Journal of Political Economy*. Vol. 103. pp. 360-99.

Huggett, Mark. 1993. The risk-free rate in heterogenous-agent incomplete-insurance economies. *Journal of Economic Dynamics and Control*. Vol. 17. pp. 953-69.

Huggett, Mark. 1996. Wealth Distribution in Life-Cycle Economies. *Journal of Monetary Economics*. Vol. 38. pp. 469-94.

Huggett, Mark and Gustavo Ventura. 2000. Understanding why high income households save more than low income households. *Journal of Monetary Economics*. Vol. 45. pp. 361-397.

Huggett, M. and Sandra Ospina. 2001. On aggregate precautionary savings: When is the third derivative irrelevant? *Journal of Monetary Economics*. Vol. 48. pp. 373-96.

İmrohoroğlu, Ayse. 1989. Cost of Business Cycle with Indivisibilities and Liquidity Constraints. *Journal of Political Economy*. Vol. 97. pp. 1364-83.

İmrohoroğlu, Ayse. 1992. Welfare Costs of Inflation under Imperfect Insurance. *Journal of Economic Dynamics and Control*. Vol. 16. pp. 79-91.

İmrohoroğlu, Ayse, Selahattin İmrohoroğlu, and Douglas H. Joines. 1995. A life cycle analysis of social security. *Economic Theory*. Vol. 6. pp. 83-114.

İmrohoroğlu, Ayse, Selahattin İmrohoroğlu, and Douglas .H. Joines. 1998. The Effect of Tax-Favored Retirement Acccounts on Capital Accumulation. *American Economic Review*. Vol. 88. pp. 749-68.

İmrohoroğlu, Selahattin. 1998. A Quantitative Analysis of Capital Income Taxation. *International Economic Review*. Vol. 39. pp. 307-28.

Jagannathan, Ravi, Ellen R. McGrattan and Anna Scherbina. 2001. The Declining U.S. Equity Premium. NBER Working Paper No. 8172.

Jermann, Urban J. 1998. Asset Pricing in Production Economies. *Journal of Monetary Economics*. Vol. 41. pp. 257-275.

Judd , Kenneth L., 1992, Projection methods for aggregate growth models, *Journal of Economic Theory*, vol. 58, 410-52.

Judd, Kenneth L. 1998. Numerical Methods in Economics. Cambridge, MA, London: MIT Press.

Judge, George G., R. Carter Hill, William E. Griffiths and Helmut Lütkepohl. 1988. Introduction to the Theory and Practice of Econometrics. 2nd. Ed. New York: John Wiley & Sons.

Kamihigashi, Takashi. 2002. A Simple Proof of the Necessity of the Transversality Condition. *Journal of Economic Theory*. forthcoming.

Kessler, Denis and André. Masson. 1989. Bequest and Wealth Accumulation: Are Some Pieces of the Puzzle Missing? *Journal of Economic Perspectives*. Vol. 3. pp. 141-52.

King, Robert G., Charles I. Plosser and Sergio Rebelo. 1988. Production, Growth and Business Cycles I, The Basic Neoclassical Model. *Journal of Monetary Economics*. Vol. 21. pp. 195-232.

King, Robert G., Charles I. Plosser and Sergio Rebelo. 1988. Production, Growth and Business Cycles II, New Directions. *Journal of Monetary Economics*. Vol. 21. pp. 309-341.

King, Robert G. and Sergio T. Rebelo. 1993. Low Frequency Filtering and Real Business Cycles. *Journal of Economic Dynamics and Control*. Vol. 17. pp. 207-231.

King, Robert G. and Sergio T. Rebelo. 1993. Low Frequency Filtering and Real Business Cycles. *Journal of Economic Dynamics and Control*. Vol. 17. pp. 207-231.

King, Robert G. and Mark W. Watson. 1998. The Solution of Singular Linear Difference Systems Under Rational Expectations. *International Economic Review*. Vol. 39. pp. 1015-1026.

King, Robert G. and Mark W. Watson. 2002. System Reduction and Solution Algorithms for Singular Linear Difference Systems under Rational Expectations. *Computational Economics*. Vol. 20. pp. 57-86.

Kocherlakota, Narayana R. 1996. The Equity Premium: It's Still a Puzzle. *Journal of Economic Literature*. Vol. 34. pp. 42-71.

Koopmans, Tjalling C. 1965. On the Concept of Optimal Economic Growth. In The Econometric Approach to Development Planning. Amsterdam: North-Holland. pp. 225-287.

Kremer, Jana. 2001. Arbeitslosigkeit, Lohndifferenzierung und wirtschaftliche Entwicklung. Lohmar, Köln: Josef Eul Verlag.

Krueger, D. and F. Kubler, 2004, Computing Equilibrium in OLG Models with Stochastic Production, *Journal of Economic Dynamics and Control*, forthcoming.

Krusell, Per and Anthony A. Smith. 1998. Income and Wealth Heterogeneity in the Macroeconomy. *Journal of Political Economy*. Vol. 106. pp. 867-96.

Kydland, Finn E. and Edward C. Prescott. 1982. Time to Built and Aggregate Fluctuations. *Econometrica*. Vol. 50. pp. 1345-1370.

Laitner, John. 1990. Tax changes and phase diagrams for an Overlapping Generations model. *Journal of Political Economy*. Vol. 98. pp. 193-220.

Lang, Serge. 1986. Introduction to Linear Algebra. New York: Springer Verlag.

Leeper, Eric M. and Christopher A. Sims. 1994. Toward a Modern Macroeconomic Model Usable for Policy Analysis. NBER, Macroeconomics Annual.

Lillard, Lee A. and Robert J. Willis. 1978. Dynamic aspects of earnings mobility. *Econometrica*. Vol. 46. pp. 985-1012.

Linnemann, Ludger. 1999. Sectoral and Aggregate Estimates of the Cyclical Behavior of Markups: Evidence from Germany. *Weltwirtschaftliches Archiv*. Vol. 135. pp. 480-500.

Linnemann, Ludger and Andreas Schabert. 2003. Fiscal policy in the New Neoclassical Synthesis. *Journal of Money, Credit, and Banking*. Vol. 35. pp. 911-29.

Ljungqvist, Lars and Thomas J. Sargent. 2000. Recursive Macroeconomic Theory. Cambridge, MA, London: MIT Press.

Long, John B. and Charles I. Plosser. 1983. Real Business Cycles. *Journal of Political Economy*. Vol. 91. pp. 39-69.

Loury, Glenn C. 1981. Intergenerational transfers and the distribution of earnings. *Econometrica*. Vol. 49. pp. 843-67.

Lucas, Robert E. 1987. *Models of Business Cycles*. New York: Blackwell.

Lucas, Robert E. 1988. On the Mechanics of Economic Development. *Journal of Monetary Economics*. Vol. 22. pp. 3-42.

Lucas, Rober. E. 1990. Supply-Side Economics: An Analytical Review, *Oxford Economic Papers*, Vol. 42, pp. 3-42

Lucke, Bernd. 1998. Theorie und Empirie realer Konjunkturzyklen. Heidelberg: Physica-Verlag.

MaCurdy, Thomas E. 1982. The use of time series processes to model the error structure of earnings in longitudinal data analysis. *Journal of Econometrics*. Vol. 18. pp. 83-114.

Mankiw, Gregory N. 2000. Macroeconomics. 4th Ed. New York: Worth Publishers.

Marcet, Albert and David A. Marshall. 1992. Convergence of Approximate Model Solutions to Rational Expectations Equilibria Using the Method of Parameterized Expectations. Kellogg Graduate School of Management. Working Paper No. 73.

Marcet, Albert and David A. Marshall. 1994. Solving Nonlinear Rational Expectations Models by Parameterized Expectations: Convergence to Stationary Solutions. Manuscript.

Marcet, Albert and Guido Lorenzoni. 1999. The Parameterized Expectations Approach: Some Practical Issues. In Ramon Marimon and Andrew Scott (Eds.). Computational Methods for the Study of Dynamic Economies. Oxford: Oxford University Press. pp. 143-171.

Mas-Colell, Andreu, Michael D. Whinston and Jerry R. Green. 1995. Microeconomic Theory. New York, Oxford: Oxford University Press.

Maußner, Alfred. 1999. Unvollkommene Gütermärkte, Konjunktur und Wachstum. Wolfgang Franz et al. (Eds.). Trend und Zyklus. Tübingen: Mohr Siebeck. pp. 121-151.

McCallum, Bennett T. 1989. Real Business Cycle Models. In Robert J. Barro (Ed.). Modern Business Cycle Theory. Oxford: Basil Blackwell. pp. 16-50.

McGrattan , Ellen R., 1996, Solving the stochastic growth model with a finite element method, *Journal of Economic Dynamics and Control*, vol. 20, 19-42.

McGrattan, Ellen R. 1999. Application of Weighted Residual Methods to Dynamic Economic Models. In Ramon Marimon and Andrew Scott. Computational Methods for the Study of Dynamic Economies. Oxford and New York: Oxford University Press. pp. 114-142.

Mehra, Rajnish and Edward C. Prescott. 1985. The Equity Premium: A Puzzle. *Journal of Monetary Economics*. Vol.15(2). pp. 145-62.

Michalewicz, Zbigniew 1996. Genetic Algorithms+Data Structures=Evolution Programs. 3rd. Ed. Berlin, Heidelberg und New York: Springer.

Mitchell, Melanie. 1996. An Introduction to Genetic Algorithms. Cambridge, MA, London: MIT Press.

Moré, Jorge J. and Stephen J. Wright. 1993. Optimization Software Guide. Philadelphia: Siam.

Murata, Yasuo. 1977. Mathematics for Stability and Optimization of Economic Systems. New York: Academic Press.

Nelson, Charles R. and Charles I. Plosser. 1982. Trends and Random Walks in Macroeconomic Time Series, Some Evidence and Implications. *Journal of Monetary Economics*. Vol. 10. pp. 139-162.

Newey, Whitney K. and Kenneth D. West. 1987. A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*. Vol. 55. pp. 703-708.

Plosser, Charles I. 1989. Understanding Real Business Cycles. *Journal of Economic Perspectives*. Vol. 3. pp. 51-77.

Prescott, Edward C. 1986. Theory ahead of business cycle measurement. *Carnegie Rochester Series on Public Policy*. Vol. 25. pp. 11-44.

Quadrini, Vincenzo. 2000. Entrepreneurship, Saving and Social Mobility. *Review of Economic Dynamics*. Vol. 3. pp. 1-40.

Quadrini, Vincenzo and José-Victor Ríos-Rull. 1997. Understanding the U.S. Distribution of Wealth. *Federal Reserve Bank of Minneapolis Quarterly Review*. Vol. 21. pp. 22-36.

Ramsey, Frank. 1928. A Mathematical Theory of Saving. *Economic Journal*. Vol. 38. pp. 543-559.

Ravn, Morten and Harald Uhlig. 2001. On Adjusting the HP-Filter for the Frequency of Observations. CEPR Discussion Paper No. 2858.

Reddy , J.N., 1993, An Introduction to the Finite Element Method, McGraw-Hill: New York.

Ríos-Rull, José-Victor. 1996. Life-Cycle Economies and Aggregate Fluctuations. *Review of Economic Studies*. Vol. 63. pp. 465-89.

Ríos-Rull, José-Victor. 1999. Computation of Equilibria in Heterogenous-Agent Models. In Marimon, R., and A. Scott (eds.). *Computational Methods for the Study of Dynamic Economies*. Oxford: Oxford University Press.

Romer, Paul. 1991. Increasing Returns and New Developments in the Theory of Growth. In Wiliam A. Barnett et al. (Eds.). Equilibrium Theory and Applications. Cambridge, MA.: Cambridge University Press. pp. 83-110.

Rotemberg, Julio J. 1987. The New Keynesian Microfoundations. NBER, Macroeconomics Annual. Vol. 2. pp. 69-104.

Rotemberg, Julio J. and Michael Woodford. 1995. Dynamic General Equilibrium Models with Imperfectly Competitive Product Markets. In Thomas F. Cooley (Ed.). Frontiers of Business Cycle Research. Princeton, NJ: Princeton University Press. pp. 243-293.

Samuelson, Paul A. 1958. An exact consumption-loan model of interest, with or without the social contrivance of money. *Journal of Political Economy*. Vol. 66. pp. 467-82.

Sargent, Thomas J. 1987. Macroeconomic Theory, 2. Aufl. San Diego: Academic Press.

Sargent, Thomas J. 1993. Bounded Rationality in Macroeconomics. Oxford: Clarendon Press.

Shorrocks, Anthony F. 1976. Income Mobility and the Markov Assumption. *Economic Journal*. Vol. 86. pp. 566-578.

Sidrauski, Miguel. 1967. Rational Choice and Patterns of Growth in a Monetary Economy. *American Economic Review*. Vol. 57. pp. 534-44.

Sieg, Holger. 2000. Estimating a dynamic model of household choices in the presence of income taxation. *International Economic Review*. Vol. 43. no. 3. pp. 637-68.

Simon, Herbert A. 1957. Models of Man: Social and Rational; Mathematical Essays on Rational Human Behavior in Society Setting. New York: John Wiley.

Solow, Robert M. 1988. Growth Theory, An Exposition. New York, Oxford: Oxford University Press.

Starr, Ross M. 1997. General Equilibrium Theory, An Introduction. Cambridge: Cambridge University Press.

Stokey, Nancy, Robert E. Lucas and Edward C. Prescott. 1989. *Recursive methods in economic dynamics*. Cambridge, MA.: Harvard University Press.

Storesletten, Kjetil, Chris Telmer and Amir Yaron. 2001. Asset pricing with idiosyncratic risk and overlapping generations. Stockholm University, Institute for International Economic Studies. Seminar Paper No. 703.

Strout, A.H. 1971. Approximate Calculation of Multiple Integrals. Englewood Cliffs, NJ: Prentice-Hall

Summers, Lawrence H. 1981. Capital Taxation and Accumulation in a Life Cycle Growth Model. *American Economic Review*. Vol. 71. pp. 533-44.

Sundaram, Rangarajan K. 1996. A First Course in Optimization Theory. Cambridge: Cambridge University Press.

Sydsæter, Knut, Arne Strøm and Peter Berck. 1999. Economists' Mathematical Manual. Third Edition. Berlin: Springer.

Takayama, Akira. 1985. Mathematical Economics. 2nd Ed. Cambridge: Cambridge University Press.

Tauchen, George. 1986. Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions. *Economics Letters*. Vol. 20. pp. 177-181.

Taylor, John B. and Harald Uhlig. 1990. Solving Nonlinear Stochastic Growth Models: A Comparision of Alternative Solution Methods. *Journal of Business & Economic Statistics*. Vol. 8. pp. 1-17.

Turnovsky, Stephen J. 2000.Methods of Macroeconomic Dynamics, Second Edition. Cambridge, MA., London: MIT Press.

Uhlig, Harald. 1999. A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily. In Ramon Marimon and Andrew Scott (Eds.). Computational Methods for the Study of Dynamic Economies. Oxford: Oxford University Press.

Ventura, Gustavo. 1999. Flat-Rate Tax Reform: A Quantitative Exploration. *Journal of Economic Dynamics and Control*. Vol. 23. pp. 1425-58.

Walsh, Carl E. 2003. Monetary Theory and Policy. 2nd Ed. Cambridge, MA., London: MIT Press.

White, Halbert. 1980. A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, *Econometrica*, Vol. 48, pp. 817-838

Author Index

Subject Index