

Public Economics and the Household

Patricia Apps and Ray Rees

CAMBRIDGE

CAMBRIDGE

www.cambridge.org/9780521887878

This page intentionally left blank

Public Economics and the Household

Economic models in much of the public economics literature have been slow to reflect the significant changes towards double-income households throughout the developed world. This graduate-level text develops a more sophisticated approach to household economics, one that allows for multiple-income earners and shared decision-making. This approach is used to present a fundamentally new view of consumption. It is then applied to an analysis of tax systems, combining theoretical analysis of optimal taxation and tax reform with a careful empirical study of the characteristics of income tax systems in four different countries: Australia, Germany, the UK and the USA. The book is particularly concerned with analysing, both theoretically and empirically, the impact of taxation on female labour supply, and identifying its effects on work incentives and fairness of income distribution. All this adds up to a fascinating new approach to the economics of households for researchers in both the public and private sectors.

PATRICIA APPS is Professor of Public Economics in the Faculty of Law at the University of Sydney.

RAY REES is Professor of Economics at the University of Munich.

Public Economics and the Household

Patricia Apps

and

Ray Rees



CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore,
São Paulo, Delhi, Dubai, Tokyo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org

Information on this title: www.cambridge.org/9780521887878

© Patricia Apps and Ray Rees 2009

This publication is in copyright. Subject to statutory exception and to the provision of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published in print format 2009

ISBN-13 978-0-511-65033-8 eBook (NetLibrary)

ISBN-13 978-0-521-88787-8 Hardback

ISBN-13 978-0-521-71628-4 Paperback

Cambridge University Press has no responsibility for the persistence or accuracy of urls for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Contents

<i>List of figures</i>	<i>page</i> vii
<i>List of tables</i>	ix
<i>Preface</i>	xi
1 Introduction and overview	1
1.1 Introduction	1
1.2 Labour supply and household production	4
1.3 Household income and welfare rankings	14
1.4 Conclusions	18
2 Time allocation and household production	20
2.1 Introduction	20
2.2 The standard consumer model	20
2.3 The Becker model	23
2.4 The Pollak and Wachter critique	26
2.5 Market substitutes, implicit prices and the value of household production	30
2.6 Conclusions	34
3 Household models: theory	36
3.1 Introduction	36
3.2 Cooperative models	39
3.3 Non-cooperative models	75
3.4 Equilibrium models	80
3.5 The Pareto property and Pareto (in)efficiency	81
4 Empirical household models	88
4.1 Introduction	88
4.2 The household utility function model	88
4.3 Estimation of models on time-use data	94

4.4	Empirical work on the collective model	101
4.5	Conclusions	108
5	Labour supply, consumption and saving over the life cycle	109
5.1	Introduction	109
5.2	Life cycle models	109
5.3	A model of the ‘family life cycle’	129
5.4	Evidence on family life cycle profiles	135
5.5	Conclusions	155
6	Household taxation: introduction	157
6.1	Introduction	157
6.2	Taxation systems	158
6.3	The simple analytics of redistributive tax systems	163
6.4	Income tax systems in practice	181
6.5	Conclusions	194
7	Optimal linear and piecewise linear income taxation	202
7.1	Introduction	202
7.2	Taxation and the within-household income distribution	204
7.3	Optimal linear taxation	208
7.4	Piecewise linear taxation	214
8	Optimal non-linear taxation	221
8.1	Introduction	221
8.2	The two-type case	221
8.3	A continuum of types	234
8.4	Two-person households	246
9	Tax reform	254
9.1	Introduction	254
9.2	Tax reform policies	256
9.3	Tax reform, joint taxation and the tax unit	272
9.4	Conclusions	275
	<i>Bibliography</i>	277
	<i>Index</i>	286

Figures

1.1	All couples: labour supply by gender	<i>page</i>	6
1.2	Couples: pre-children		7
1.3	Couples: children 0–17 years		8
1.4	Couples: post-children		9
1.5	Labour supplies of household types, by primary income		15
1.6	Earnings of household types, by primary income		16
5.1	Life cycle income and consumption paths		113
5.2	Possible equilibria in an imperfect capital market		119
5.3	Income paths under uncertainty		121
5.4	Prudence and the expected marginal utility of income		124
5.5	Equilibrium cash in hand and buffer stock saving		127
5.6	Median net income, consumption and saving, by age		138
5.7	Median net income, consumption and saving, by phase		141
5.8	Male and female hours of market and domestic work		142
5.9	Household and adult full consumption		144
5.10	Median saving by household type		144
5.11	Two-adult total consumption profiles		154
5.12	Perfect capital market model		154
5.13	Imperfect capital market model		155
6.1	Equilibrium without taxation		164
6.2	Equilibrium with linear taxation		165
6.3	Equilibrium with non-linear taxation		167
6.4	Implementing the optimum with taxes		169
6.5	Effects of a convex piecewise linear tax		170
6.6	Determination of the optimal tax threshold		172
6.7	A non-convex tax system		173
6.8	Effects of a non-convex piecewise linear tax		174
6.9	Phasing out a universal benefit		175
6.10	Small differences in type, large differences in income and consumption		180
6.11	Formal income tax system: MTRs by primary income		183
6.12	Formal income tax system: ATRs by primary income		184

6.13	Effective MTRs by primary income	187
6.14	Effective ATRs by primary income	188
6.15	Effective MTRs by household income	191
6.16	Effective ATRs by household income	192
6.17	ATRs on second income	193
6.18	MTRs with social security taxes, by primary income	200
6.19	ATRs with social security taxes, by primary income	200
8.1	Single crossing	223
8.2	Lump sum redistribution	225
8.3	Optimal non-linear taxation	227
8.4	Implementation by taxes	228
8.5	Distortion of the lower type is always optimal	229
8.6	Relative slopes and the degree of distortion	231
8.7	The corner solution	233
8.8	Limits to redistribution	234
8.9	Optimal lump sum redistribution	237
8.10	$\mu(w) < 0, w \in (w_0, w_1)$	240
8.11	No distortion at the top and bottom	241
8.12	Second-order condition binding	243
8.13	Bunching	244
8.14	Bunching at $y = 0$	245
8.15	Directions of possible binding constraints	249
8.16	Illustration of the lemma	250
9.1	A move to Pareto efficiency to reduce social welfare	255

Tables

1.1	All individuals aged 25–59 years: employment rates by gender	<i>page</i> 5
1.2	Couples: employment rates and labour supplies by gender	6
1.3	Market, domestic and child care hours per week	11
1.4	Phase 2: time use with a child aged 0–4 years, hours per week	12
1.5	Phase 2: female time use by employment status	12
1.6	UK and US: household rankings by primary and household income	17
5.1	Median net income, consumption and saving, \$p.a., 1998, by age	138
5.2	Median net income, market consumption and saving, \$p.a., 1998, by phase	140
5.3	Male and female hours of market and domestic work, p.a.	142
5.4	Domestic and full consumption expenditure, \$p.a., 1998	143
5.5	Long-term saving and housing debt	146
5.6	Labour supply, income, saving and taxes, by household type, by age	147
5.7	Reference consumption profiles, by phase	152
5.8	Perfect capital market consumption profiles	153
5.9	Imperfect capital market consumption profiles	153

Preface

The economic model that underlies most of labour economics and public economics, particularly the analysis of taxation, is that of a single consumer/worker dividing their time between market labour supply and leisure. However, many issues in reality are concerned with policies towards families, a majority of which contain two earners. At the same time, quite a large literature has developed over the past three or four decades concerned with developing and testing models of family decision-taking. The purpose of this book therefore is to bring together the economics of multi-person households and those areas of labour supply and public economics for which this generalisation of the standard model seems to us to be most relevant and important. Essentially, it is proposing a new foundation for much of the analysis of labour supply and public economics.

After a general introductory chapter, the first part of the book, chapters 2–5, surveys the main developments in the economics of the household, focusing particularly on the use of the models to analyse labour supply, with emphasis also on models of time allocation, consumption and saving over the life cycle. We present models that can be used in their own work by graduate students and researchers in the areas of the economics of the household, labour supply and public policy.

The second part of the book, chapters 6–9, deals with optimal income taxation and tax reform. We have tried to make the treatment as far as possible self-contained, in that we present graduate-level expositions of the existing models before going on to extend the theory to consider issues such as the taxation of couples, the significance of household production and other determinants of female labour supply, and the implications of income support schemes for the effective marginal tax rates on second earners. Here we present work that has appeared in the recent journal literature as well as the results of new research.

The book therefore sets out to be a combination of a graduate-level textbook in the economics of the household and optimal tax theory, and an account of recent and current research. We hope it will be of interest to teachers and students in graduate courses in economics of the household, labour economics and public economics, as well as to researchers in those areas.

Most of the material has been presented in graduate courses at the universities of Munich and York (Ray Rees), and Essex University, the IZA Summer School in Labour Economics and the Sorbonne (Patricia Apps) over the past five years or so. We are very grateful to participants in these courses for many comments and questions that have stimulated clarification and further development of the ideas we are presenting here.

Patricia Apps

Ray Rees

1 Introduction and overview

1.1 Introduction

The standard ‘workhorse model’ used in economics to study consumption, saving and labour supply decisions, and which also provides the basis for virtually all of public economics, is that of an individual decision-maker, who divides his time between market labour supply and leisure, and allocates the resulting income to consumption goods. There is a vast literature that uses this model to analyse these decisions, both in a static, timeless setting, and within a framework in which consumption and time allocations are chosen over an entire life cycle, with or without uncertainty.

Although this class of models has over the years yielded many valuable insights, household survey data, econometric investigation and theoretical analysis all suggest that it provides an inadequate basis for obtaining a satisfactory understanding of household decisions, and for estimating the behavioural parameters of households formed by two adults, especially if they have children. This therefore limits its usefulness in addressing many of the problems of public economic policy, for which we need both an adequate conceptual framework and robust and reliable estimates of behavioural parameters. In chapters 2, 3 and 4 of this book we expand upon this assertion, which of course may not be readily accepted by at least some of our economist colleagues. In these chapters we first summarise briefly the main results of the model, review the empirical evidence, which generally rejects its implied restrictions on household consumption demand and labour supply functions, and then undertake a comprehensive survey of the alternative models that have been developed over the last three to four decades.

It is fair to say that the theoretical development of these models is well in advance of their empirical application, the main reason for this being data limitations. In some cases, however, there is also a failure to use some very good data that are available, for example, time-use data. Although econometricians are past masters at making bricks without straw, our main contention is that a necessary condition for serious progress in the applications of these models, which are central to the formulation of public economic policy, is greatly improved gathering of data on what actually goes on inside households with respect to consumption, production and time-use decisions.

A first approach to the static analysis of the labour supply decisions of couples, the household utility function model, defines household utility as a function of total consumption and two types of leisure or labour supply. So far this has been the main basis for specifications of male and female labour supply functions that have been estimated for the purpose of tax and welfare reform analysis. We show in chapters 6, 7 and 8, where we apply the model to tax analysis, that it can be useful for theoretical purposes, given that the assumptions allowing its use are clearly understood. However, this model also has severe limitations when it is used for estimating the behavioural responses of two-parent households with dependent children. A key criticism is that, although it can be interpreted as a reduced form of a two-person household with household production, it yields restrictions on consumption demand and labour supply functions which are identical to those of the individual model, and which therefore are similarly rejected by the data. In chapter 3 we develop this point at some length, and in chapter 4 discuss some of the empirical work with this model.

The model has a further major limitation. The data show that in many households, following the arrival of children (though not before), there is a marked division of labour, with the female tending to specialise in home production of goods and services, especially child care. Thus, households are characterised by specialisation and exchange, suggesting that the two-parent household needs to be modelled as a small economy, using the concepts of general equilibrium theory and welfare economics.

The goods and services produced and consumed in the household have close (though in general not perfect) market substitutes, and what the data also show is that there is a marked heterogeneity across households in the extent to which female labour time is divided between producing these goods in the household and working in the market and buying them in. The model we develop should also allow us to explain this heterogeneity, since, as we show in chapters 5 to 9, it plays a crucial role in the analysis of taxation and income redistribution policies. Although it is not hard to develop hypotheses to explain the heterogeneity, going beyond the usual suspects of wage rates and demographic variables (which incidentally explain only a small part of it), much empirical work needs to be done before we really understand what causes it. Again, the problem here seems essentially to be data limitations.

In this introductory chapter we motivate the approach we have taken to modelling the household by presenting data on four countries – Australia, Germany, the UK and the US – that support the three main elements of our modelling approach. These are:

- The redefinition of the categories of time use from the standard two – market work and leisure, defined as time used directly for one's own consumption – to three: market work; leisure; and time spent in household production, i.e. in producing goods and services within the household for consumption by the members of that household. At many points we further refine this to distinguish between time spent on child care and on general household activities that are carried out whether or not children are present in the household.

- The recognition of the role played by wage rates and employment possibilities outside the household in determining the within-household division of labour, rather than restricting attention to individual preferences and the comparative advantage in household production of one partner relative to the other. A wage gap between partners may not only make it rational for the lower-wage partner, typically the female, to specialise in household production, but may also influence the distribution of real income or utility within the household. At the same time, across-household variation of productivities in household production, together with the price of market substitutes, is also given an important role in explaining variation in the utility possibilities of households.
- The redefinition of the life cycle in terms of the phases that the typical household goes through – what we could call the ‘family life cycle’ – rather than in terms of the age of the ‘head of the household’, as in the standard life cycle literature. The importance of this extension is suggested by time-use data that show the dramatic changes that take place, with strongly persistent effects, after the arrival of the first child.

The enormous increase in female labour force participation that took place in the developed economies between the early 1950s and the late 1980s makes the importance of these extensions to the standard models self-evident. Economic and social historians may still be debating the causes of this transformation, but its consequences are clear. The ‘traditional model’ of the household, in which the male head specialises in market work and the female in work within the household, now represents only around a third of families with dependent children in most OECD countries, and fewer in some cases. Moreover, as noted above, households have become highly heterogeneous in respect of the labour supply decision of the female partner. In the UK and the US, for example, roughly 30 per cent of households with dependent children continue to conform to the ‘traditional model’, and in roughly a further 25 per cent in the UK, and 45 per cent in the US, both partners work full-time in the market. The majority of the remaining households have one partner, not always the male, as the primary earner in full-time work, while the second works part-time. As we pointed out earlier, while some of the observed heterogeneity is associated with the age and number of children and, to a lesser extent, with economic variables such as wage rates and non-labour income, much of it remains unexplained after controlling for these variables.

This transformation in work choices has created challenges to the formulation of public policy. Most immediately, it poses the question of how to tax two-earner couples. Different countries have found different solutions, with, for example, the USA and Germany taxing incomes jointly, while many other countries, including the UK, Canada and Australia, tax them separately. The large falls in fertility associated with growth in female labour force participation have been largely responsible for the changes in the age structure of the populations of these countries, and so for the associated problem of funding pay-as-you-go social security and pension systems. In the area of family income support, withdrawal of benefits on the basis of total household income leads to very high marginal tax rates on individual incomes and reduced incentives to work.

Discussion of policy towards child support raises issues of whether this is best done by direct lump-sum payments or by the provision of child care facilities outside the home, and, in the case of the former, whether it matters to which parent the payment is actually made. The basic issues of equity and efficiency that underlie the formulation of these policies become more complex in the context of specialisation in household production and female labour supply heterogeneity. For example, it is no longer self-evident that total household income is an adequate measure of a household's standard of living.

These policy issues, together with discussion of the models we need to analyse them, are the focus of the chapters to follow. We do not address policy solutions to poverty due to long-term unemployment or disability, where specific retraining programmes combined with wage subsidies may offer more effective solutions. In our view, these problems, which the data show affect a relatively small proportion of the population of prime working age in most OECD countries, need to be considered in the context of the economic forces that drive them, and the specific moral hazard problems that may be associated with them. The central question we address in the analysis of policy in this book can be posed as follows: how should the system of income taxation be designed to redistribute income within and across the vast majority of households consisting of couples with children, where at least one partner is fully employed?

In the remainder of this chapter we present detailed empirical evidence on the labour supply, hours of domestic work and child care and the earnings of couples to support our view, first, that the household should be modelled as a small economy, and, second, that we need to define the life cycle in terms of the presence and ages of children, taking account of the demands they create – demands that can be met either by work at home or by the market.

1.2 Labour supply and household production

We draw on data for four comparable OECD countries: Australia, Germany, the United Kingdom (UK) and the United States (US). We use data from the following household and time-use surveys for these countries:

- Australia: The Household, Income and Labour Dynamics in Australia Survey, Wave 5, 2005 (HILDA).¹ Australian Bureau of Statistics (ABS) 2006 Time Use Survey (AU TUS).
- Germany: German Socio-Economic Panel, Wave 22, 2005 (GSOEP).²
- UK: Expenditure and Food Survey, 2005–6, National Statistics (EFS). Time Use Survey 2000, National Statistics, United Kingdom (UK TUS).
- US: Panel Study of Income Dynamics, 2005 Public Release (PSID).³ American Time Use Survey 2005 (ATUS).⁴

¹ See Goode and Watson (2007). ² See DIW Berlin website: www.diw.de/gsoep/.

³ See PSID website: <http://psidonline.isr.umich.edu/>.

⁴ See US Bureau of Labor Statistics and US Census Bureau website www.bls.gov/tus/.

Table 1.1 *All individuals aged 25–59 years: employment rates by gender*

	AUSTRALIA		GERMANY		UK		US	
	M	F	M	F	M	F	M	F
% Employed	88.1	72.1	84.2	67.5	84.3	70.7	88.4	72.8
% Full-time	80.2	38.4	81.1	44.4	77.8	37.9	84.0	56.0

When male and female employment rates for countries such as Australia, Germany, the UK and the US are compared over time, we observe a large measure of convergence since the 1950s, especially in the latter two countries, due not only to growing female employment, but also to declining male employment. Table 1.1 shows for each of the countries the percentage of males and females of prime working age who are employed, based on data for all individuals aged 25 to 59 in HILDA, GSOEP, EFS and the ATUS.⁵ In each country the gap between male and female employment rates is less than 17 percentage points. The male rate ranges from 84.2 per cent in Germany to 88.4 per cent in the US, and the female rate from 67.5 per cent in Germany to 72.8 per cent in the US.

Broad comparisons of employment rates of this kind are sometimes assumed to show that the labour supplies of males and female are converging in the same way as employment rates. This is a mistake. Employment and participation rates can give a misleading picture of the true relation between the characteristics of labour supply of men and women, and this is in fact the case for these four countries. This is because there is a large gap between rates of full-time employment of men and women, as shown by the figures in the second row of table 1.1.⁶

When we select the data for couples, we find an even larger gap between the male and female full-time rate. This is because most singles work full-time, irrespective of gender, and therefore excluding them has this effect.

In the remainder of this chapter we focus on couples.⁷ Table 1.2 presents couples' employment rates and average hours of work by gender.⁸ In the US, Germany and Australia over 85 per cent of married men are employed full-time, and in the UK 81.7 per cent. Apart from the US, only 34–36 per cent of married women are employed full-time. As a result, there is a large gender difference in average hours of market work, even in the US where the female full-time rate is 50.4 per cent.

⁵ The numbers of male and female records in this age category drawn from each of the data files are as follows: HILDA: 3,652 males and 4,064 females; GSOEP: 4,478 males and 4,928 females; EFS: 3,532 males and 3,942 females; and ATUS: 3,743 males and 4,841 females.

⁶ The full-time employment rate is computed as the percentage of records showing 35 hours of work or more per week.

⁷ The samples include couples in single-family households with no others except children less than 15, dependent students or non-dependent children older than 15.

⁸ The numbers of male and female records drawn from the couples samples are: HILDA, 2,084 male and 2,140 female; GSOEP, 3,544 male and 3,910 female; EFS, 2,675 male and 2,808 female; and ATUS, 2,604 male and 3,023 female.

Table 1.2 *Couples: employment rates and labour supplies by gender*

	AUSTRALIA		GERMANY		UK		US	
	M	F	M	F	M	F	M	F
% Employed	91.8	72.8	87.5	64.7	87.9	72.8	91.5	70.0
% Full-time	85.3	34.2	85.3	36.0	81.7	36.1	87.9	50.4
Hours per week	42.2	22.4	46.6	23.7	35.4	21.0	43.2	25.6

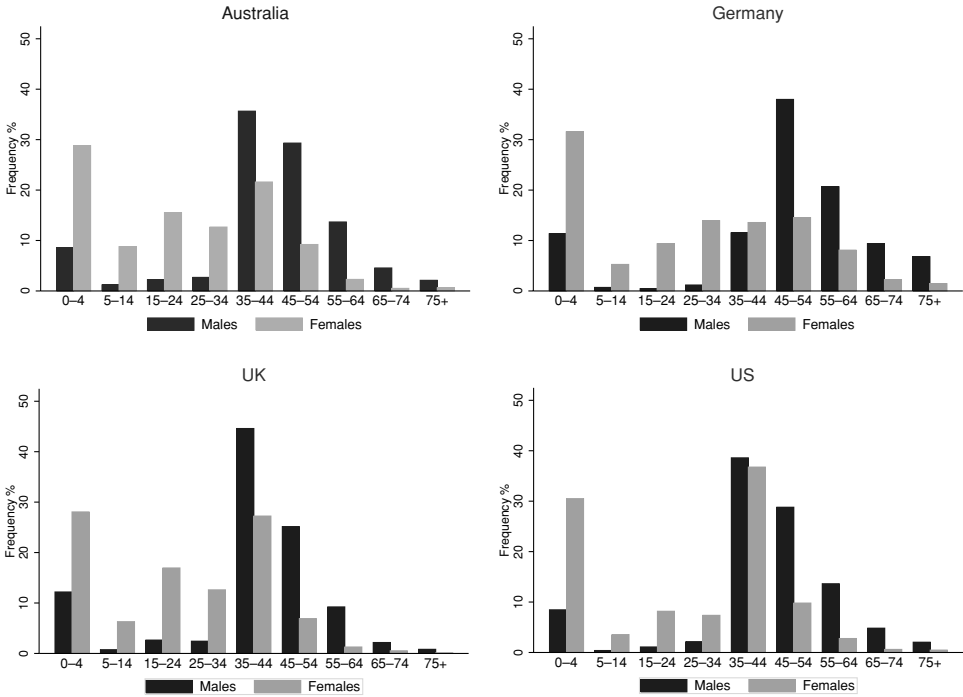


Figure 1.1 All couples: labour supply by gender

The employment rates in table 1.2 also show the very high degree of heterogeneity in female labour supply. While male labour supply shows very little variation, with almost all men working full-time, females are distributed more evenly between zero hours and full-time work.

We illustrate this heterogeneity graphically in figure 1.1 for all couples. The figure presents histograms of ‘usual weekly hours of work’ for partners aged 25 to 59 years, with the first band representing 0 to 4 hours, and subsequent bands increments of 10 hours. The vast majority of men in all four countries work between 35 and 54 hours a week, with a second much smaller modal frequency at 0–4 hours. The distribution for women is tri-modal in Australia and the UK, with roughly equal frequencies at 0–4 and 35–44 hours per week, and a small peak at 15–24 hours per week. In the US, female

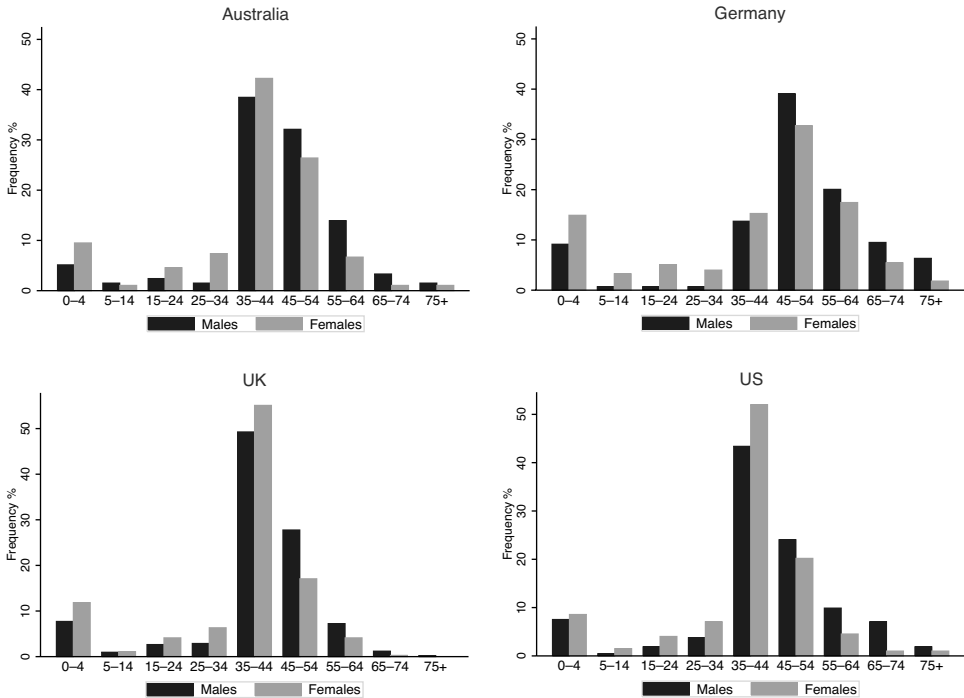


Figure 1.2 Couples: pre-children

labour supply is more nearly bimodal, although there is a significant proportion of married women working part-time. Germany has the highest percentage of prime-aged married women in the 0–4 hours band, and those who are employed tend to be spread evenly across a very wide band of part-time and full-time hours.

Much of the observed heterogeneity in female labour supply is strongly associated with children, as we would expect. However, exactly how and to what extent children play a role, and precisely what that role is, is not well explained by the existing empirical literature. After controlling for demographics, as well as for wage rates and non-labour incomes, female labour supply heterogeneity remains high, suggesting that there is a great deal of room for additional explanatory variables.

We find it useful to explore these points further by taking a very broad-brush life cycle approach. We classify households into three phases:⁹ a pre-children phase; a phase in which children aged under 18 years are present in the household; and a post-children phase, in which there are no longer children under 18 years present. The usual weekly hours of work of partners in these phases are depicted graphically across the four countries in figures 1.2 to 1.4.¹⁰

⁹ In chapter 5 we refine this classification considerably. We also deal there with two obvious questions about this approach, the first stemming from the endogeneity of the decision to have children, the second from the possibility of cohort effects (since these data are taken from cross-sections).

¹⁰ The numbers of male and female records in phases 1 to 3 are as follows:

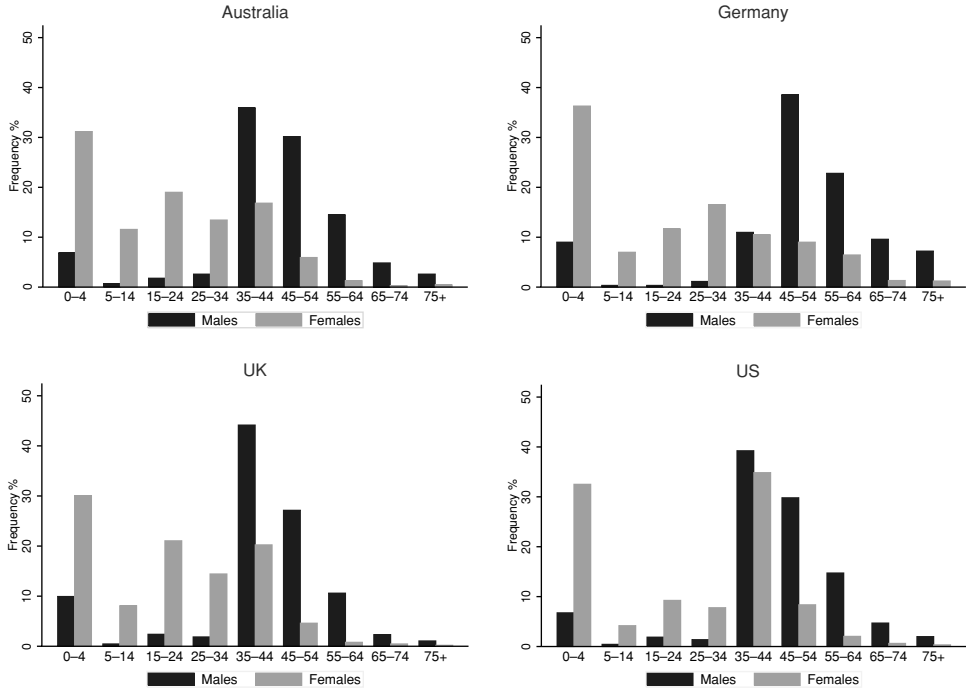


Figure 1.3 Couples: children 0–17 years

In the first phase, shown in figure 1.2, the profiles closely match – partners of prime working age tend to work full-time, and for the same hours, in all four countries.¹¹ In the with-children phase, the proportion of men working full-time remains about the same, while that of women falls dramatically. At the same time, the heterogeneity in female labour supply emerges. Around a third of women still work full-time, and over a third work part-time. The association between children and female labour supply is also highly persistent.¹² In the post-children phase female labour supply remains well below its pre-children level. In fact, there is little change in most of the countries, and much of what there is would seem to involve a shift from part-time to full-time work.

- HILDA: phase 1: 330 male and 284 female; phase 2: 1,339 male and 1,331 female; phase 3: 415 male and 525 female.
- GSOEP: phase 1: 266 male and 274 female; phase 2: 2,021 male and 2,023 female; phase 3: 1,257 male and 1,613 female.
- EFS: phase 1: 414 male and 363 female; phase 2: 1,499 male and 1,496 female; phase 3: 762 male and 949 female.
- ATUS: phase 1: 201 male and 202 female; phase 2: 1,840 male and 2,099 female; phase 3: 563 male and 722 female.

¹¹ There are relatively few records in this phase because there are relatively few young married couples without children. However, when we include singles who have not yet had children, and who are therefore essentially in the same life cycle phase, we obtain similar results from much larger samples for all four countries. Almost all men and all women not in higher education work full-time prior to having children.

¹² This is consistent with the results of panel data studies for the US. See, for example, Shaw (1989, 1994).

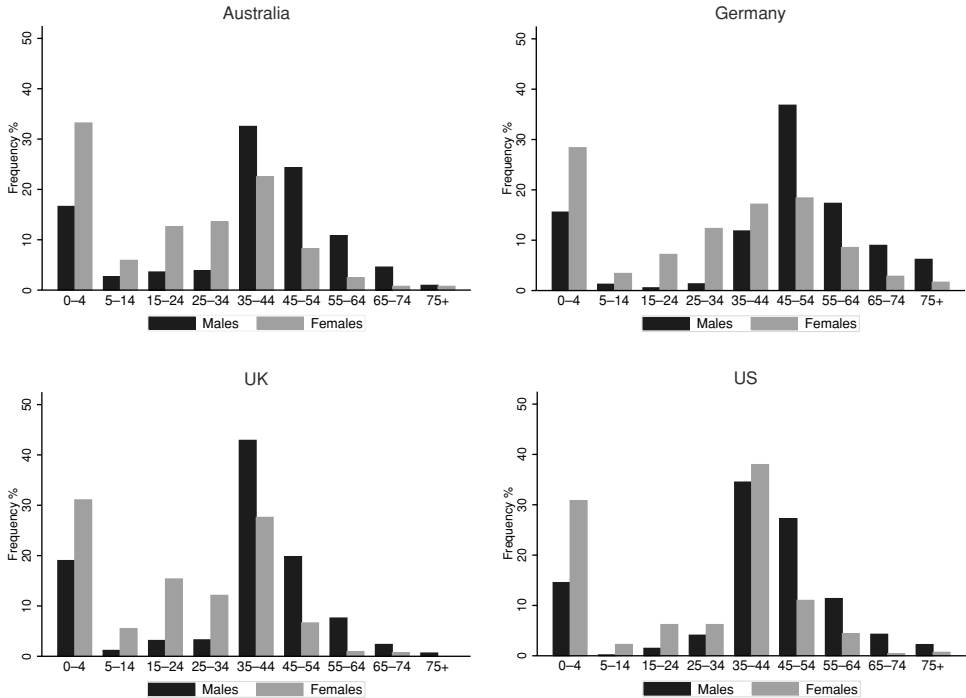


Figure 1.4 Couples: post-children

While the dramatic change in the profile of female hours from phase 1 to phase 2 indicates the strong association between the decision to have children and the labour supply decisions of married women, we would argue that the relationship is far more subtle and complex than the simple view of unidirectional causation that says: children ('demographics') cause the reduction in female labour supply. Given the decision to have children, the observed changes in female labour supply are driven by the economics of investment in the care and education of children in phase 2, much of which is directly influenced by government policy. The effects of the change in female labour supply on current and future income and employment possibilities will, however, then feed back on the household's decision on the timing and number of births. Thus there is a relationship of simultaneity between the fertility and female labour supply decisions, while both are jointly conditioned by underlying factors such as the cost, quality and availability of market child care, on the one hand, and the opportunity costs of parental time, on the other. These latter depend on: net of tax wage rates; productivity in household production; the flexibility of working arrangements in relation to the time demands of child care; the rate of depreciation of work-related human capital when out of the labour force;¹³ and the probability of future re-employment, all of which are or can be strongly influenced by public policy.

¹³ An extensive literature on work-related human capital accumulation includes the contributions of Eckstein and Wolpin (1989), Altug and Miller (1998) and, more recently, Imai and Keane (2004) and Olivetti (2006), among others.

The argument is straightforward. In phase 1 of the life cycle there is a low demand for home-produced goods and services because there are few of the kinds of goods and services couples in this phase consume for which there are not good, affordable market substitutes, and so there is a low demand for domestic labour in this phase. Put simply, there is nothing much to do in the home, and so it would make no sense for either partner to specialise in household production, or for singles who have not yet had children to do so. This explains why almost all males and females who have not yet had children, whether single or married, work full-time and have close to the same average weekly hours.

The arrival of children creates a very large demand for their care and for investment in their education. Governments of the four countries we are highlighting have taken over much of the role of investing in the education of children once they reach school age. However, all four countries tend to have child care and education sectors for those under school age that are expensive, poorly developed, and frequently difficult to access.

Child care can be provided by some combination of parental time and services bought in from the market. The opportunity cost of parental child care is determined by the present value of the current and future market income forgone, which depends on the factors listed above, and may well differ between the parents because of differences in wage rates and career types. The more costly and difficult it is to access market child care, the more of it will be provided at home, other things being equal. The demand for child care then implies a large demand for household production and introduces a fundamental change in the work choices of couples, which will reflect the relative costs of each partner's time.

This can of course reinforce labour market discrimination, since a lower wage for women in the expectation that they will leave the labour force to look after the children is self-fulfilling. The loss of human capital and career possibilities resulting from leaving full-time work offers an explanation for the strong persistence of female labour supply decisions made in phase 2 into the post-children phase.¹⁴

Time-use data provide evidence of close substitution between market labour supply and household production, consisting mostly of child care. Table 1.3 reports data means for average weekly hours allocated to these activities by couples in each of the phases. The time input to household production in phases 1 and 3 is computed as time spent on domestic work (washing, cooking, cleaning, etc.) and shopping. In phase 2, the time input to household production is split into domestic work and child care.¹⁵

In the pre-children phase, both partners typically work full-time in the market, and close to the same average hours. Even in Germany, which has the largest gender gap

¹⁴ See also Attanasio *et al.* (2003).

¹⁵ Market hours are computed from the data on 'usual weekly hours of work' in HILDA, the GSOEP, EFS and PSID. Domestic and child care hours for Germany are based on the time-use data available in Wave 5 of the GSOEP. For Australia, the UK and US we merge time-use data on domestic and child care hours in the AU TUS, UK TUS and ATUS with the HILDA, EFS and PSID samples, respectively, using regression models with the dependent variable of the domestic hours equation specified as the ratio of domestic hours to leisure hours, and of child care hours as the ratio of child care hours to domestic hours. For further detail, see chapter 5.

Table 1.3 *Market, domestic and child care hours per week*

	AUSTRALIA		GERMANY		UK		US	
	M	F	M	F	M	F	M	F
Phase 1								
Market hours	43.1	37.7	47.8	40.2	36.2	31.9	42.4	37.3
Domestic hours	11.2	14.2	8.4	18.3	14.2	18.9	13.5	18.2
Phase 2								
Market hours	43.2	19.2	48.4	18.9	37.0	19.0	44.1	24.0
Domestic hours	12.8	25.5	6.7	26.9	15.1	19.1	12.9	24.5
Child care hours	11.7	27.0	14.1	42.1	9.0	26.5	17.0	28.9
Phase 3								
Market hours	36.7	22.1	43.4	26.0	32.5	21.2	40.1	27.0
Domestic hours	15.5	28.5	7.2	23.0	16.6	28.1	15.2	24.0

in hours in this phase, females work around 85 per cent of the hours of males. The allocation of time to household production involves only domestic work since there are no children present, and is relatively low. Both partners tend to work the same total hours, with the small difference in their market hours being offset by an opposite difference in their domestic hours.

The time allocation decisions of partners in phase 2 are very different. Female market hours drop to around 50 to 60 per cent of their phase 1 level, and there is a very large increase in time spent on household production, more than half of which is child care.¹⁶ Males also spend more time on household production, with the increase due almost entirely to child care. Unlike married females, who substitute away from both market work and leisure, married males tend to substitute away from leisure only. In the post-child phase, the increase in female market hours tends to be marginal in all four countries, and hours of domestic work remain significantly higher than in phase 1. A possible explanation for the latter is that it is a result of the effect of a loss of market human capital in reducing the female market wage. As the wage of the partner specialising in household production falls over time, the opportunity cost or implicit prices of domestic goods and services fall relative to those of their market substitutes.

If we select households in which the youngest child is under 5,¹⁷ the fall in female labour supply and rise in child care hours are both much greater, as shown in table 1.4. The largest fall is in Germany, where female hours are only 9.7 per week. Interestingly,

¹⁶ Child care hours for Germany need to be interpreted cautiously. Wave 5 of the GSOEP does not provide information on the primary v. secondary status of an activity, and so it has not been possible to weight the allocation of time to home child care appropriately. According to most surveys that do collect these data, child care is usually a secondary activity.

¹⁷ The samples of couples with dependent children aged under 5 years contain around 35 to 40 per cent of the records in phase 2 across the four countries.

Table 1.4 *Phase 2: time use with a child aged 0–4 years, hours per week*

	AUSTRALIA		GERMANY		UK		US	
	M	F	M	F	M	F	M	F
Market hours	43.2	16.9	49.5	9.7	36.0	15.4	44.2	20.9
Domestic hours	14.0	26.2	7.5	27.0	13.8	24.8	12.6	23.8
Child care hours	13.9	35.7	21.6	71.9	14.8	32.6	21.8	38.2

Table 1.5 *Phase 2: female time use by employment status*

	Domestic hours p.w.	Child care hours p.w.	Market hours p.w.	Percentage child 0–4	No. of children	Percentage of records
Australia:						
Full-time	22.3	16.1	42.7	29.4	1.88	34.3
Part-time	27.9	23.0	18.9	43.3	2.06	38.6
Not employed	36.5	32.0	0.0	56.5	2.31	27.1
Germany:						
Full-time	22.1	25.7	48.6	11.7	1.5	23.3
Part-time	26.1	40.8	20.5	23.1	1.7	37.0
Not employed	30.4	52.9	0.0	41.7	1.8	39.7
UK:						
Full-time	21.7	12.3	38.9	28.1	1.62	24.1
Part-time	26.8	18.9	20.7	32.7	1.76	46.3
Not employed	31.4	26.7	0.0	49.5	2.18	29.6
US:						
Full-time	18.4	22.8	42.3	36.3	1.8	46.3
Part-time	25.3	29.7	20.0	41.0	2.1	21.9
Not employed	32.6	37.1	0.0	50.0	2.2	31.9

female hours of domestic work change very little. Child care hours rise by a third or more above the average of phase 2.¹⁸ Male market hours are much the same as the overall average for phase 2, in all four countries.

The data means for time allocations reported in tables 1.3 and 1.4 give averages across all families with children and therefore conceal the heterogeneity indicated by the average weekly hours profiles in figure 1.2 for phase 2. To bring out this heterogeneity, table 1.5 reports data means for female hours of domestic work, home child care and market work in phase 2, disaggregated by full-time and part-time employment status, to give an indication of the variation in time use by married mothers in all four countries.

¹⁸ In contrast to the results for Germany, child care hours for the UK may be biased downwards because the UK TUS does not include travel with a young child as a secondary child care activity, which is likely to be significant for an employed mother of a pre-school child.

In each case, hours of market work are inversely related to hours of household production, especially child care. For example, in the UK, mothers who are employed full-time work an average of 38.9 hours per week in the market, and they spend 21.7 hours on domestic work and 12.3 hours on child care. For those who work entirely at home, time spent on domestic work rises to 31.4 hours, and on child care to 26.7 hours. The figures are broadly similar for the US and Australia, apart from longer hours on child care than in the UK.¹⁹

Table 1.5 also reports the percentage of families with a child aged 0–4 years, and the average number of children under 18 years in each employment status category. The final column gives the percentage of records in each category. The results clearly show that the age and number of children present in the household have a significant impact. Married mothers who work full-time tend to have older and fewer children. This is especially evident in Germany, which has the lowest percentage of mothers in full-time work, and of those, the lowest percentage with a child under 5.

Observations of this kind suggest that the parameter estimates of a labour supply model specifying the age and number of children as exogenous variables will indicate they are strongly significant. Nevertheless, there will still be considerable variation in female labour supply that remains unexplained, which, in such models, is usually attributed to unobserved fixed effects, costs of working or preference heterogeneity. In the US, for example, 46.3 per cent of married mothers with a dependent child work full-time, and of these 36.3 per cent have a child under 5. Note that there is very little variation in family size by employment status. Clearly, in the US, as in the other countries, there are many families with identical demographic characteristics making very different time-use choices.²⁰

These data have implications for the modeling approach needed for estimation of the parameters of behavioural responses to policy changes, as we shall discuss at some length in chapter 4. These parameters are of course necessary for evaluation of the efficiency effects of policy changes. The data also suggest a different view of the issue of evaluating the effects on the across-household welfare distribution. Neither an income tax nor a consumption tax can be applied to a tax base that includes household production. As a result, households with the same wage rates and demographic characteristics will pay different amounts of tax, depending on the second earner's choices between market and domestic work. Those who substitute household production for market work avoid paying tax on the implicit income derived from domestic labour, and they avoid consumption taxes on the output. Under a progressive individual income tax, a two-earner household can pay twice as much tax as the single-earner household working only half the market hours, and with the second partner working full-time at home. Under a joint income tax, as in Germany and the US, the two-earner household pays

¹⁹ As already noted, the UK figures for child care hours may be too low because they omit travel time with a child.

²⁰ As we will see in the section to follow, gross hourly earnings would also appear to explain relatively little of the observed heterogeneity in female labour supply.

more than twice as much tax. This may not only make no sense in terms of efficiency if the second earner has a high wage elasticity because there are close home substitutes for her market work, but it can also make no sense in terms of equity, unless we make highly specialised assumptions about her home productivity or the price of bought-in child care. The problem is that household income is no longer a reliable measure of household welfare in the presence of domestic and market labour supply heterogeneity. We now investigate this issue in more detail.

1.3 Household income and welfare rankings

If families with the same wage rates and demographic characteristics were observed to make the same time allocation decisions, then, all else being equal, we could reasonably expect to find a strong correlation between observed market incomes and household welfare. However, with heterogeneity in the labour supply of one parent, this is no longer the case. Furthermore, the potential for errors in welfare rankings defined on market income can be especially serious when, as the analysis to follow will show, the profile of wage earnings for full-time work by the partner with the higher income, typically the male, is relatively flat across the middle of the distribution and then rises sharply towards the top.

In this section we illustrate the sensitivity of a family's position in a ranking by household income to the second earner's market v. domestic work choice, by comparing rankings of single and two-earner households based first on the income of the 'primary earner', defined as the partner with the higher private income, and then on household income. For the purpose of the analysis we use household survey data for two-parent families selected on the criteria that the male partner is aged from 25 to 59 years, at least one child under 18 is present, and at least one parent is employed and earning a positive income.²¹ The samples for Australia, Germany, the UK and the US are drawn from HILDA, GSOEP, EFS and PSID, respectively.²²

The analysis assumes a model of the household in which the labour supply of the primary earner is fixed. Only the labour supply of the partner with the lower income, the 'second earner', is variable.²³ In a model of this kind, tax systems based on household income do not discriminate formally on the basis of gender, but according to the within-household primary v. second income status. We therefore compare rankings of single and two-earner household 'types' defined in terms of the employment status of the second earner, rather than that of the female partner.²⁴

²¹ Records in the top 1 per cent of income and those with negative incomes are also excluded.

²² The HILDA sample contains 1,223 records, the GSOEP 2,129 records, the EFS 1,141 records, and the PSID 1,928 records.

²³ The model is set out formally in chapter 4.

²⁴ Households in which the primary earner is the female represent 14.3 per cent of the HILDA sample, 15.8 per cent of the GSOEP sample, 17.1 per cent of the EFS sample and 18.2 per cent of the PSID sample. In each country, households with a female primary earner tend to be concentrated in the lower quintiles. Note that, if there is imperfect 'assortative matching', a decline in the gender wage gap can be expected to lead to an increase in the proportion of households in which the female partner has the higher income.

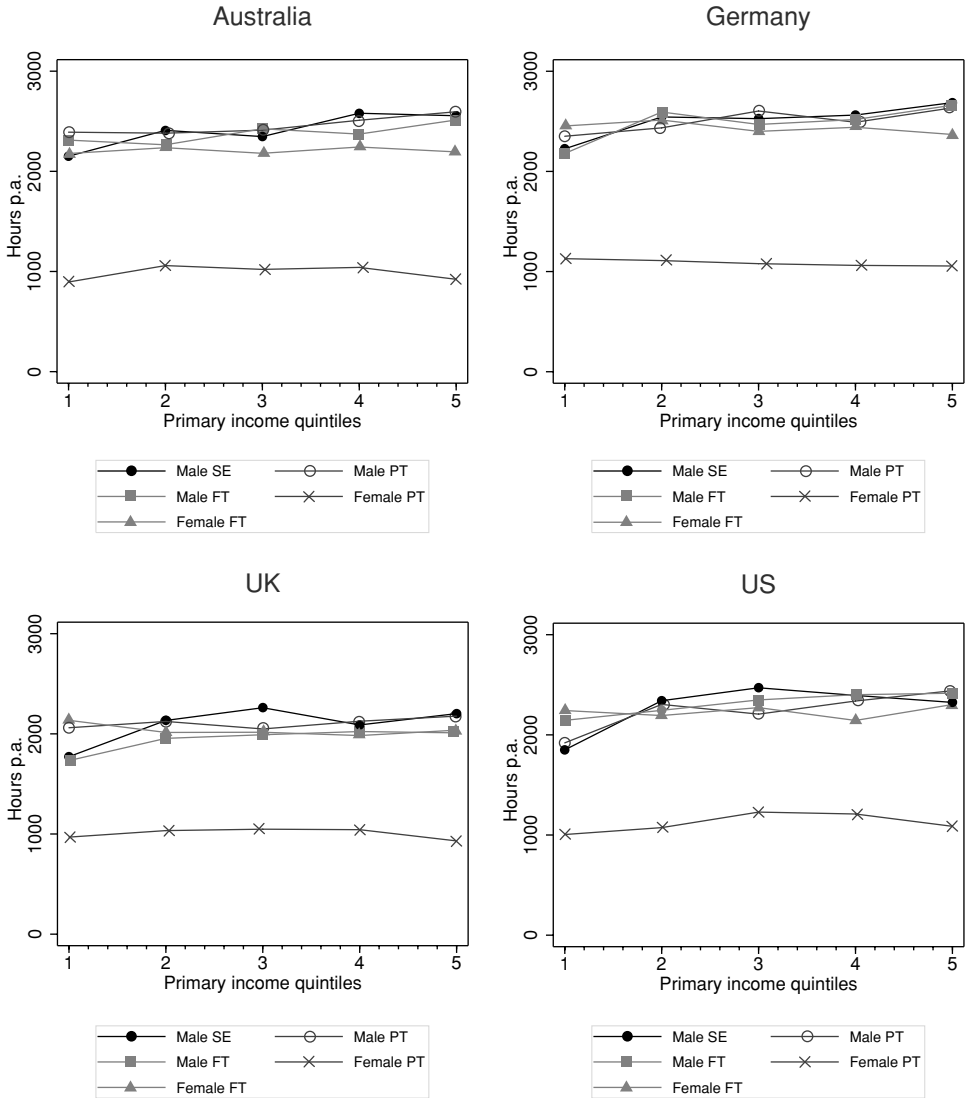


Figure 1.5 Labour supplies of household types, by primary income

Figure 1.5 plots the average annual market hours of primary and second income partners across quintiles of primary income, in each of the four countries. The profiles are computed for three household types: SE: single-earner household; PT: two-earner households with the second earner in part-time employment; and FT: two-earner households with the second earner in full-time employment.

The results show graphically the tendency of primary earners of all three household types to work close to the same average hours in each quintile. There is also relatively little variation in average primary hours as primary income rises, especially across

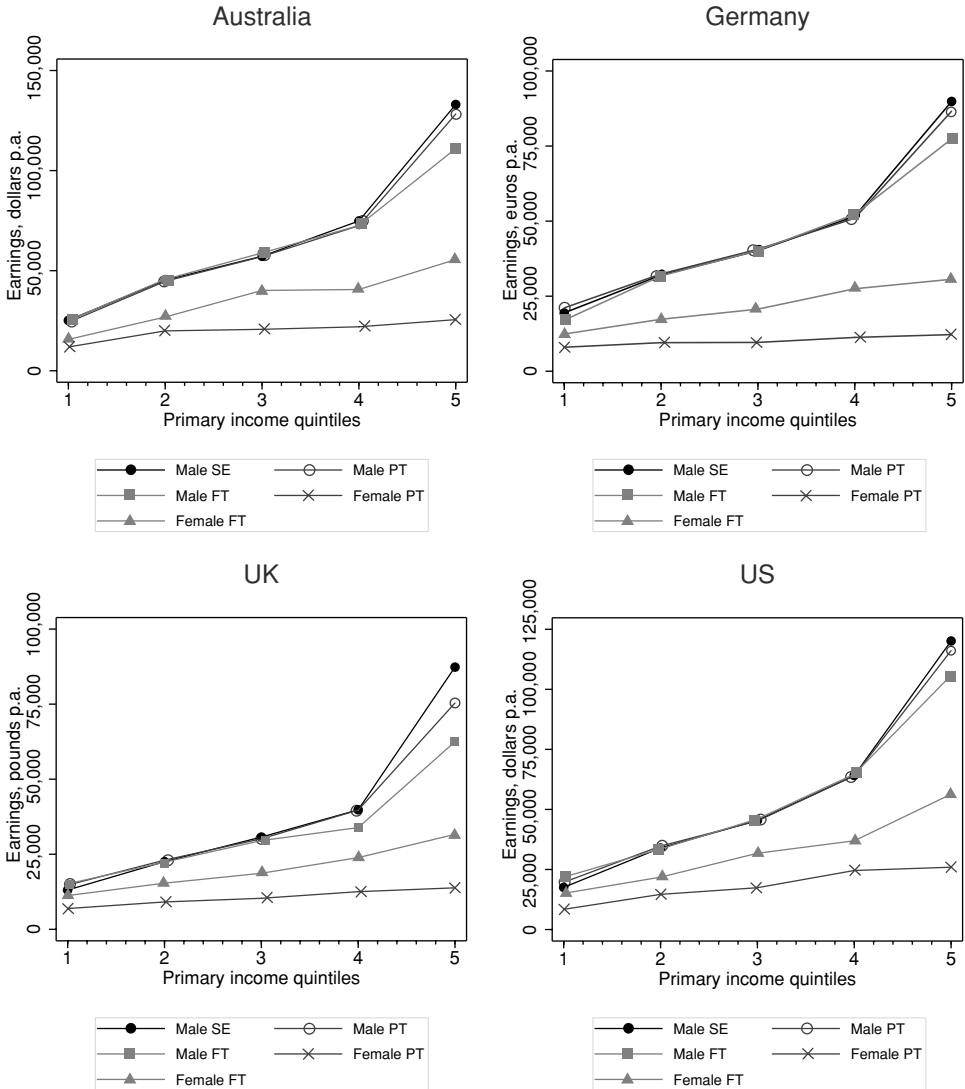


Figure 1.6 Earnings of household types, by primary income

quintiles 2 to 5. This reflects the fact that the vast majority of primary earners are male and almost all work full-time. Note that the second earner in the FT household works close to the same average hours as the primary earner. Thus, on average, the FT household works around twice the market hours of the SE household. The second earner in the PT household works around half the hours of partners in full-time employment, and so the PT household works around one and half times the hours of the SE household.

Figure 1.6 plots the resulting earnings profiles for each country. As foreshadowed, a crucial feature of the profiles of primary earnings is the relatively flat segment across

Table 1.6 *UK and US: household rankings by primary and household income*

Quintile	1	2	3	4	5	All
UK						
Primary income £p.a.	13,544	21,443	28,353	36,995	73,229	35,169
SE %	38.2	21.9	20.0	19.9	28.5	25.4
FT %	21.5	34.6	31.0	30.7	24.6	28.1
Household income £p.a.	17,156	29,212	38,194	49,932	90,691	45,760
SE %	63.9	25.1	14.2	11.0	19.6	25.4
FT %	10.9	23.8	30.7	38.3	38.7	28.1
US						
Primary income \$p.a.	19,793	34,586	46,528	66,047	133,355	61,887
SE %	34.3	20.0	17.8	18.5	28.7	22.8
FT %	22.8	46.4	43.9	41.2	25.5	38.3
Household income \$p.a.	25,241	49,584	67,843	93,525	170,730	88,309
SE %	51.3	23.9	18.9	12.9	16.2	22.8
FT %	11.1	35.6	41.3	46.8	40.5	38.3

quintiles 2 to 4 in all four countries. This means that the position of a family in a ranking defined on household income will be very sensitive to the earnings, and therefore to the labour supply, of the second earner because it will take only a small increase in earnings to shift a family from a low percentile of household income to a significantly higher point in the distribution.

To illustrate this, table 1.6 compares the distribution of SE and FT households by primary income and by household income in the UK and US. In both countries, the two household types tend to be well represented in each quintile of primary income, but with a tendency for the SE household to be over-represented in quintiles 1 and 5, and for the FT households to be more strongly concentrated in quintiles 2 to 4. Switching to a ranking defined on household income results in a dramatic change in the ordering of both household types in each country. The SE household becomes far more strongly represented in the lower two quintiles, and the FT household in the upper two quintiles.

In effect, a household income ranking places two-earner households working long hours for relatively low wages towards the upper end of the distribution. For example, in the case of the UK, a single-earner household with a primary income of, say £30,000, which is close to average weekly earnings for 2005–6,²⁵ would move from quintile 2 to quintile 5 of the household income ranking if the second partner switched from working at home to working in the market for the same income.²⁶ If the household has a preschool child, much of the net-of-tax second income might be spent on child care. Clearly, in the context of a static or ‘within period’ analysis, such a household could

²⁵ See OECD (2006).

²⁶ The upper threshold of household income for quintile 2 is £33,800 p.a., and the lower threshold for quintile 5 is £58,551 p.a.

not be considered to have the same standard of living as another in which only one parent needs to work full-time to earn £60,000 p.a., while the other works full-time at home. To argue to the contrary it is necessary to assume that home child care makes little to no contribution to family welfare, or that the mother who works in the market is so productive in the short time she has at home that this makes her family better off. This is obviously implausible.

The fundamental deficiency of the household income ranking is that it is defined on an income variable that omits home production appropriately weighted by price. It is not only an unsatisfactory measure of family welfare for the purpose of policy evaluation, it is also an unreliable basis for constructing measures of inequality.²⁷

1.4 Conclusions

In this chapter, we have presented the data that underlie and motivate the approach we adopt to modelling the household in this book, and have also given an indication of the kinds of policy problems that we will use our models to analyse. To summarise, these data on female labour supply and the allocation of time to household production, particularly child care, highlight three key observations which, taken together, have significant economic implications that are not captured by standard labour supply models:

- *There is a significant division of labour within the household once children arrive.* Although male and female employment rates appear to be converging, there are still many married women who specialise in household production after children arrive, as a substitute for market work. The fall in female labour supply with the arrival of children does not reflect the preference-determined substitution of leisure for market work. In fact, the allocation of time to leisure actually falls for both parents. Specialisation in production roles within the household implies exchange. The household therefore needs to be modelled as a small economy.
- *The fall in female labour supply, and the resulting division of labour within the household, are driven primarily by the choice between market and home child care.* The importance of incorporating household production in models of labour supply therefore arises from the empirical significance of the allocation of time to child care at home, as a substitute for working in the market and buying it in.
- *There is a high degree of heterogeneity in the division of labour within the household.* For a significant proportion of couples, both partners work full-time, and in others there is complete specialisation with respect to market and domestic work. In families where both parents work full-time, earnings must be partly allocated to buying in child care and related services, unless the parents have access to some form of extended family arrangement. The household in which one partner specialises in home production pays for child care through the opportunity costs of her time, and

²⁷ For an analysis in the context of the literature on measuring inequality, see Apps and Savage (1989).

that partner's time input to home production is remunerated implicitly through intra-household exchange of her output for her consumption of market goods funded by the earnings of her spouse.

In the next chapter, we outline briefly the standard model of the single individual decision-taker and summarise its main results. We then present and discuss some formal models of time allocation involving household production, showing how the single-person household model can be extended to this case. In chapters 3 and 4 we survey the literature on theoretical and empirical models of multi-person households, and show how household production can be integrated into them, arguing as we do so that this is a necessary step to take. In chapter 5 we first survey quite extensively the literature on the main existing approaches to the life cycle decision problem, and then present a 'family life cycle' model, which refines the approach to the life cycle outlined in this chapter and, we argue, gives a good explanation of the data. In chapters 6 to 9 we apply our models to a central problem in public economics, the analysis of optimal income tax systems and tax reform. Here we try to show that our approach yields simple and tractable models that give new insights into the problem of the taxation of couples.

2 Time allocation and household production

2.1 Introduction

In this chapter we introduce the main ideas underlying the theory of household production. To make the discussion self-contained, we first give a very brief account of the standard model of the consumer to establish its main assumptions and results. We then introduce household production by setting out Becker's (1965) model, and consider the powerful critique of this model made by Pollak and Wachter (1975). We present our own assessment of this critique, and conclude by applying some variants of the household production model to the question of the valuation of household output, incorporating elements of the Pollak and Wachter critique in the models we use.

2.2 The standard consumer model

We refer to the *standard consumer model* as the basic model of consumer demand and labour supply, which has provided the analytical framework for most of labour economics and public finance theory. It treats the household as a single individual who, in the simplest formulation, chooses a goods vector $x = [x_1, \dots, x_n] \in \mathbf{R}_+^n$ so as to maximise a strictly quasiconcave and increasing ordinal utility function $u(x)$. The assumptions underlying this utility function, which are often taken to be the expression of what is meant in economics by rational behaviour, are clearly concerned with *individual* as opposed to *group* decision-taking. Nevertheless, as far as the pure logic of the model is concerned, if the decision-taking process of a group, such as a family or multi-person household, is consistent with these assumptions, then the results of the model can be applied. The extent to which this is the case is an important area of controversy that will be discussed in the next two chapters. Here we focus the discussion on the treatment of the household's time allocation decisions.

There is an exogenously given price vector $p = (p_1, \dots, p_n) \in \mathbf{R}_+^n$ and total expenditure is then $e = px$. The simplest versions of the model treat income as exogenous, and although this can be a useful simplification, it is clearly a severe restriction of the model's applicability. It is relaxed by introducing an exogenously given *initial endowment vector* $\bar{x} = [\bar{x}_1, \dots, \bar{x}_n] \in \mathbf{R}_+^n$, where $\bar{x}_i \geq 0$ represents an amount of good

$i = 1, \dots, n$ owned initially by the individual. Choice of a consumption quantity $x_i \geq \bar{x}_i$ now is equivalent to the decision to buy ($x_i > \bar{x}_i$) or sell ($x_i < \bar{x}_i$) the good on the market. Given the price vector, we can define the individual's wealth¹ as $p\bar{x} + \mu$, where $\mu \geq 0$ is thought of as some generalised income claim, for example arising out of the ownership of assets or a government transfer. The budget constraint restricts expenditure not to exceed wealth:

$$px \leq p\bar{x} + \mu. \quad (2.1)$$

Alternatively, if we define $\hat{x}_i = x_i - \bar{x}_i$ as the net purchase (if positive) or sale (if negative) of good $i = 1, \dots, n$, in short, the *net demand* for good i , and $\hat{x}$ as the corresponding vector of net demands, we can write the budget constraint as

$$p\hat{x} \leq \mu. \quad (2.2)$$

In the case where $\mu = 0$, this says that the value of what is bought cannot exceed the value of what is sold, at all possible price vectors.

The model is specialised to analyse labour supply choices by designating one endowment, say $\bar{x}_1$, as the individual's total amount of time available, for example 24 hours a day. We then define $x_1 \in [0, \bar{x}_1]$ as leisure consumption, measured in time units, and $l \equiv -(\bar{x}_1 - x_1) \geq 0$, the sale of leisure time, as labour supply. The individual's time is often assumed to be the only income-generating good, in which case we set $\bar{x}_2 = \dots = \bar{x}_n = 0$. If the relative prices of these $n - 1$ goods are assumed to remain constant with respect to each other, we can use the composite commodity theorem of Hicks to replace them with a single consumption good $\tilde{x}$, with price $\tilde{p}$, and the wage rate w is the price of leisure, with wl wage income. μ is then called *non-wage income*. We replace the underlying utility function $u(x)$ by the equivalent function $U(\tilde{x}, l)$, where $U_l < 0$. Given the properties of $u(x)$, $U(\tilde{x}, l)$ is also a strictly quasiconcave, ordinal utility function, strictly increasing in $\tilde{x}$ and decreasing in l .

The consumer then chooses her labour supply and consumption by solving the problem

$$\max_{l, \tilde{x}} U(\tilde{x}, l) \quad \text{s.t.} \quad \tilde{p}\tilde{x} \leq wl + \mu. \quad (2.3)$$

This has a well-defined solution expressed in terms of the consumption demand function

$$\tilde{x} = D(\tilde{p}, w, \mu) \quad (2.4)$$

and labour supply function

$$l = S(\tilde{p}, w, \mu). \quad (2.5)$$

The assumptions of the theory, in particular the exogeneity of prices and linearity of the budget constraint, as well as the properties of the utility function, imply a set of testable restrictions on the consumption demand and labour supply functions:

¹ If this is a single-period or timeless model, wealth and income are synonymous and can be used interchangeably.

- Homogeneity of degree zero in the price, wage rate and non-wage income. Multiplying $\tilde{p}$, w and μ by the scalar $k > 0$ leaves consumption demand and labour supply unchanged, because the budget constraint is unchanged.
- The adding-up property, $\tilde{p}D(\tilde{p}, w, \mu) - wS(w, \mu) = \mu$. The assumption that the utility function is strictly increasing in consumption, i.e. exhibits non-satiation, guarantees that the solution satisfies the budget constraint as an equality.
- Slutsky equations for labour supply and consumption:

$$\frac{\partial l}{\partial w} = s_{lw} + l \frac{\partial l}{\partial \mu} \begin{matrix} \geq \\ \leq \end{matrix} 0 \quad (2.6)$$

$$\frac{\partial l}{\partial \tilde{p}} = s_{l\tilde{p}} - \tilde{x} \frac{\partial l}{\partial \mu} \begin{matrix} \geq \\ \leq \end{matrix} 0 \quad (2.7)$$

$$\frac{\partial \tilde{x}}{\partial w} = s_{\tilde{x}w} + l \frac{\partial \tilde{x}}{\partial \mu} \begin{matrix} \geq \\ \leq \end{matrix} 0 \quad (2.8)$$

$$\frac{\partial \tilde{x}}{\partial \tilde{p}} = s_{\tilde{x}\tilde{p}} - \tilde{x} \frac{\partial \tilde{x}}{\partial \mu} \begin{matrix} \geq \\ \leq \end{matrix} 0 \quad (2.9)$$

where s_{lw} , $s_{l\tilde{p}}$, $s_{\tilde{x}w}$, $s_{\tilde{x}\tilde{p}}$ are the compensated demand derivatives or pure substitution effects, with $s_{lw} \geq 0$, $s_{\tilde{x}\tilde{p}} \leq 0$. Note that the assumptions of the theory in general place no restrictions on the signs of the partial derivatives of (uncompensated) demands with respect to prices and income. We may have either leisure or consumption as an inferior good (in which case the other must be normal), and if leisure is a normal good, implying $\partial l / \partial \mu < 0$, labour supply may vary inversely with the wage, at least over some range.

- Symmetric, negative semidefinite Slutsky matrix:²

$$S = \begin{bmatrix} -s_{lw} & s_{l\tilde{p}} \\ s_{\tilde{x}w} & s_{\tilde{x}\tilde{p}} \end{bmatrix} \quad (2.10)$$

with $s_{\tilde{x}w} = s_{lw}$ (symmetry) and $-s_{lw}, s_{\tilde{x}\tilde{p}} \leq 0$, $-s_{lw}s_{\tilde{x}\tilde{p}} - s_{\tilde{x}w}s_{l\tilde{p}} \geq 0$, (negative semidefiniteness).

The key aspect of this labour supply model is the idea that there are just two uses of time: paid market work and the direct consumption of one's own time as 'leisure'. The disutility of work is simply the forgone utility of leisure. Any variation in the allocation of time between work and leisure across consumers facing the same (net of tax) wage rates and non-wage incomes can only be explained by a difference in preferences. This is obviously a strong simplification, which for purposes of deriving a basic model of labour supply has proved useful. It is not, however, useful for all purposes, and can in some contexts be misleading. For example, if a consumer is supplying no time to

² The negative of the compensated labour supply derivative with respect to the wage, $-s_{lw}$, is the compensated demand derivative of leisure with respect to the wage.

market work, she must in this model be consuming all her time endowment as leisure. Moreover, empirical measures of the difference between total time available and the time spent in market work do not in general correspond to the time devoted to leisure consumption, in the sense used here, because of the existence of household production. For many interesting and important issues in economics, it is therefore essential to consider alternative models of the uses and allocation of time.

2.3 The Becker model

The division of time into ‘work’ and ‘leisure’ can be modified and refined to permit a deeper analysis of the allocation of time itself. Becker’s model is motivated by the desire to analyse such questions as:

- What is the relationship between the growth in wages and incomes and the allocation of expenditures across goods with differing time intensities; for example, how does growth in wages affect demands for labour-saving appliances and convenience foods?
- What is the significance of household production and how does change in its technology affect market labour supply?
- How do changes in working hours affect the pattern of demand for goods?
- How does an increase in taxation of goods affect the supply of labour?

Time is after all the ultimate scarce resource, and its allocation is an interesting and important subject, to which the simple dichotomy of ‘work’ and ‘leisure’ fails to do justice.

Becker in fact dispenses with leisure altogether. He argues that the variables that determine utility directly are ‘commodities’ such as warmth, security, good health, entertainment, satisfaction of hunger, and so on, and that these are generated by the individual combining market goods with her own time in a process of production within the household.³ Sleep, for example, which might be thought of as the ultimate form of ‘leisure’, is really a production activity which combines the individual’s time with market goods such as a bed, pillows and a duvet to produce a commodity that could be called ‘rest’.

By placing strong linearity and separability conditions⁴ on the consumption technology – the processes by which these commodities are produced – Becker is able to formulate the consumer’s optimisation problem in a simple and tractable way, readily giving answers to the sorts of questions just posed. We now briefly set out the formal model.

Let $y = [y_1, \dots, y_m] \in \mathbf{R}_+^m$ denote the vector of utility-generating commodities. The consumer has a standard utility function $u(y)$ with the usual properties. She combines

³ Other important early contributions to the theory of household production were by Gorman (1956) and Lancaster (1966). For good general discussions of these see Deaton and Muellbauer (1980) and Gronau (1986). Here we discuss Becker’s approach because of its focus on time allocation rather than on modelling product quality.

⁴ Though these are intended primarily ‘for simplicity’, and non-linear models are also considered; see, for example, Becker and Michael (1973).

a market goods vector $x = [x_1, \dots, x_n] \in \mathbf{R}_+^n$ with her own time to produce these commodities, according to a Leontief type of linear production technology that we can describe as follows. The $(n + 1) \times m$ *input requirements matrix* $\mathbf{A}$ is written as

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdot & \cdot & a_{1m} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \cdot & \cdot & a_{nm} \\ \tau_1 & \cdot & \cdot & \tau_k \end{bmatrix} \quad (2.11)$$

where the component $a_{ij} \geq 0, i = 1, \dots, n, j = 1, \dots, m$, gives the amount of market good i required to produce one unit of commodity j , and $\tau_j \geq 0$ gives the amount of the individual's time required to produce one unit of y_j , and is therefore referred to as the time intensity of production of commodity y_j . These coefficients are assumed to be exogenously given constants independent of the amount of any y_j . It follows that there are constant returns to scale and no joint production.⁵ Then we have the household technology constraint

$$[x, t] = \mathbf{A}y \quad (2.12)$$

where $t = \sum_{j=1}^m \tau_j y_j$ is the total time spent in household production. This gives the $(n + 1)$ vector of demands for market goods and for the total required time input corresponding to any desired commodity vector. Further constraints are given by the budget and time constraints respectively:

$$px - wl \leq \mu \quad (2.13)$$

$$t + l = 1 \quad (2.14)$$

where p is again the price vector of market goods, w the wage rate, l the market labour supply, μ the non-wage income, and total time available is normalised to 1. 'Leisure', the direct consumption of one's own time, is now replaced by the time spent in household production. It follows that the disutility of market work is the forgone utility from the commodities that could have been produced with the time supplied to the market. The consumer's 'value of time' at the margin is given by the value of the shadow price – the Lagrange multiplier – associated with the time constraint (2.14) at the optimal solution to the consumer's problem.

The model can then be solved by maximising $u(y)$ subject to the constraints (2.12), (2.13), (2.14). However, it is more insightful to adopt an indirect approach. Define the *full prices*

$$\pi_j = \sum_{i=1}^n p_i a_{ij} + w \tau_j \quad j = 1, \dots, m. \quad (2.15)$$

⁵ The classic examples of joint production are: raising a sheep produces both mutton and wool; rearing a cow produces milk, meat and leather. In the household, an individual's time could produce more than one good: for example, cooking dinner whilst looking after a child.

These full prices π_j have a simple intuitive interpretation: they represent the cost of the amounts of the market goods required to produce one unit of the domestic commodity j , $\sum_{i=1}^n p_i a_{ij}$, together with $w\tau_j$, the cost of the time required, valued at the market wage rate, the opportunity cost of time used in household production.⁶ Because of the linearity assumption, *these prices are independent of the amounts of the y_j produced*. They are in effect constant marginal production costs. Given knowledge of the matrix $\mathbf{A}$, the market wage rate w and price vector p , the vector of full prices $\pi = (\pi_1, \dots, \pi_m)$ can be readily computed. Clearly, any one full price depends on the relative intensities of use of the market goods and labour, in conjunction with the goods prices and the wage rate:

$$\frac{\partial \pi_j}{\partial p_i} = a_{ij}; \quad \frac{\partial \pi_j}{\partial w} = \tau_j. \quad (2.16)$$

We can collapse all three constraints⁷ into the single *full income constraint*:

$$\sum_{j=1}^m \pi_j y_j \leq w + \mu \quad (2.17)$$

where $w + \mu$ is the consumer's full income or wealth, consisting of the value of the time endowment at the market wage plus non-wage income. The consumer then has the maximisation problem

$$\max_{y \geq 0} u(y) \quad \text{s.t.} \quad \pi y \leq w + \mu. \quad (2.18)$$

Note that the structure of this problem is exactly the same as that of the standard consumer model, with the commodities y_j taking the place of the market goods, and the vector of full prices π taking the place of the market goods price vector p . The first-order conditions consist of a standard set of equalities of marginal rates of substitution between commodities to ratios of the corresponding full prices. Thus demands will depend on the relative intensities of use of market goods and time in the production of each good. The solution of the problem yields:

- commodity demands $y_j^*(\pi, w, \mu)$, with the standard properties;
- *derived demands* for market goods

$$x_i^*(\pi, w, \mu) = \sum_{j=1}^k a_{ij} y_j^*(\pi, w, \mu) \quad i = 1, \dots, n; \quad (2.19)$$

- time allocations to household production

$$t_j^*(\pi, w, \mu) = \tau_j y_j^*(\pi, w, \mu) \quad j = 1, \dots, m; \quad (2.20)$$

⁶ This assumes that there are no constraints on substituting between market and household work, e.g. because of unemployment or rationing on the labour market, and that the consumer is not at a corner solution, using all her time for household production and supplying no time to the market.

⁷ Just substitute for l from (2.14) into (2.13), and then for x and t from (2.12), and use the definition of the π_j in (2.15).

- market labour supply

$$l^*(\pi, w, \mu) = 1 - \sum_{j=1}^m t_j^*(\pi, w, \mu) = 1 - \sum_{j=1}^m \tau_j y_j^*(\pi, w, \mu). \quad (2.21)$$

Time is divided between market work and producing domestic commodities, while leisure, as a separate category of time use, disappears. Market labour supply must, by virtue of the budget constraint, generate sufficient income to cover the cost of the market goods used in household production.

Of particular importance is the fact that, under the given assumptions on the household production technology, the full prices:

- are independent of preferences and the scale of domestic consumption and production;
- are identical across consumers facing the same goods prices, wage rates and household technology;
- can be estimated from the technology parameters, market prices and wage rates.

2.4 The Pollak and Wachter critique

As Pollak and Wachter pointed out in their critique of Becker's household production model, the linearity assumptions are stronger than necessary to give the first of the properties listed at the end of the previous section, which they argue is crucial to the 'usefulness of the model'. It is sufficient:

- first, that there is no joint production, in which case the output of each commodity can be produced according to its own independent production function, with the standard properties of differentiability and strict quasiconcavity,

$$y_j = h^j(x_{1j}, \dots, x_{nj}, t_j) \quad j = 1, \dots, m. \quad (2.22)$$

Each input quantity x_{ij} , t_j must enter only in production of y_j – there is no joint production;

- second, that each of these production functions has constant returns to scale, i.e. is linearly homogeneous.

We can in that case think of the consumer as a firm solving the standard optimal production problem

$$\min_{x_{ij} t_j} C_j = \sum_{i=1}^n p_i x_{ij} + w t_j \quad (2.23)$$

subject to the production function constraints in (2.22), to obtain the cost functions $C_j(p, w, y_j)$. Then under constant returns to scale these cost functions can be written as

$$C_j(p, w, y_j) = c_j(p, w) y_j \quad (2.24)$$

with the *unit cost functions* $c_j(p, w)$ independent of scale. The derivatives

$$\frac{\partial c_j}{\partial p_i} = x_{ij}^0(p, w) \quad (2.25)$$

give the *unit demands*, the amount of each good x_i that is optimal, at the given price vector and wage rate, to produce one unit of y_j , and likewise

$$\frac{\partial c_j}{\partial w} = t_j^0(p, w) \quad (2.26)$$

gives the unit demand for time spent in producing one unit of y_j . It follows that the demand functions for goods and time can be written as

$$x_i(p, w, y) = \sum_{j=1}^m x_{ij}(p, w, y_j) = \sum_{j=1}^m x_{ij}^0(p, w) y_j \quad i = 1, \dots, n \quad (2.27)$$

$$t(p, w, y) = \sum_{j=1}^m t_j(p, w, y_j) = \sum_{j=1}^m t_j^0(p, w) y_j \quad (2.28)$$

and the labour supply function is

$$l(p, w, y) = 1 - t(p, w, y). \quad (2.29)$$

Finally, the implicit price of each commodity is then simply the unit cost $\pi_j = c_j(p, w)$.

The assumption of linear homogeneity, as opposed to linearity, means that the input requirement coefficients, previously the a_{ij} and τ_j in the linear model, are now functions of the goods prices and the wage. When the latter are constant we have a linear technology, but in the analysis of the effects of price changes on demands we have to take into account that the input requirement coefficients vary with prices.

It is not even necessary to make the assumption of linear homogeneity to be able to define the implicit prices. As we know from standard production theory, whatever the specific nature of returns to scale, provided the production functions are differentiable and quasiconcave we can always obtain the cost functions $C_j(p, w, y_j)$, and then equate the implicit prices to the marginal costs, $\pi_j(p, w, y_j) = \partial C_j(p, w, y_j) / \partial y_j$. Using the identity

$$\sum_{j=1}^m C_j(p, w, y_j) = \sum_{j=1}^m p_i x_{ij}(p, w, y_j) + w t(p, w, y) \quad (2.30)$$

and the time constraint, we can model the consumer as solving the problem

$$\max_{y \geq 0} u(y) \quad \text{s.t.} \quad \sum_{j=1}^m C_j(p, w, y_j) \leq w + \mu, \quad (2.31)$$

the first order conditions for which can be written as

$$u_j - \lambda \pi_j(p, w, y_j) = 0 \quad j = 1, \dots, m \quad (2.32)$$

$$\sum_{j=1}^m C_j(p, w, y_j) = w + \mu. \quad (2.33)$$

Thus again in equilibrium marginal rates of substitution between any pair of commodities are equal to the ratio of their full prices.

However, under non-constant returns to scale, the full prices $\pi_j(\cdot)$ can only be determined at the solution to the problem. The budget constraint in (2.31) is in general non-linear, and its slope at the optimal solution, which determines the values of the π_j , cannot be specified a priori. In other words, marginal costs of the commodities are not constant and so will be determined by the desired outputs of, or demands for, the goods.⁸ The commodity prices are endogenous and in general vary with the scale of consumption. So, for example, two consumers may have the same household production technology and face the same wage rate and goods prices, but if their non-wage incomes differed, they would have differing implicit prices of the commodities in the presence of non-constant returns to scale, given that their commodity demands vary with income.

Even if there are constant returns to scale, the presence of joint production, under which a given input produces more than one output, also makes it impossible to define implicit prices independently of output levels, essentially because the production functions and cost minimisation problems can no longer be specified independently.

For example, suppose we have two household goods y_j which are produced jointly under a Leontief technology with the individual's time t , though each requires a different market good, x_j . Write the production functions as

$$y_j = \min \left(\frac{x_j}{a_j}, \frac{t}{\tau_j} \right) \quad j = 1, 2 \quad (2.34)$$

so that the input requirement functions are

$$x_j = a_j y_j \quad (2.35)$$

$$t = \max(\tau_1 y_1, \tau_2 y_2). \quad (2.36)$$

Then, if $\tau_1 y_1 < \tau_2 y_2$, the marginal cost of good 1 is $\pi_1 = p_1 a_1$, since its output can be expanded using only x_1 ; while if $\tau_1 y_1 \geq \tau_2 y_2$, the marginal cost of good 1 is (for output increases) $\pi_1 = p_1 a_1 + w \tau_1$, since both time and x_1 are needed to expand output. Similarly for y_2 . Thus marginal costs or implicit prices cannot be defined independently of outputs. This is not an artefact of the example. Pollak and Wachter prove generally that, even when there are constant returns to scale, relative commodity prices are

⁸ There is the further problem that if there are *increasing* returns to scale, the feasible set of commodity bundles defined by the constraint may not be convex, and so the first-order conditions (2.32), (2.33) are not sufficient for the optimum and may not be necessary – there may be multiple local optima or corner solutions.

independent of the amounts of commodities produced if and only if there is no joint production.⁹

Pollak and Wachter see this as a serious limitation to the ‘usefulness’ of Becker’s approach, since they regard non-constant returns to scale, as well as joint production, as very likely to be present in household production. In particular, they stress the possibility that the consumer may derive utility or disutility directly from the time spent on household tasks, which is a form of jointness in production – an individual’s time yields utility directly as well as producing a commodity which in turn produces utility. In the presence of non-constant returns to scale and joint production, the implicit prices can no longer be regarded as determined exogenously by technology, goods prices and the wage rate, but also depend on the consumer’s preferences.¹⁰

They propose therefore an alternative approach. Since each y_j is a function of goods and time use, simply substitute into the utility function to obtain utility as a function of goods and time use, $U(x_{11}, \dots, x_{ij}, \dots, x_{nj}, t_1, \dots, t_j)$. Then solve the problem of maximising this function subject to the linear budget constraint

$$\sum_i \sum_j p_i x_{ij} + w \sum_j t_j \leq w + \mu \quad (2.37)$$

to yield demand functions $x_{ij}(p, w, \mu)$, $t_j(p, w, \mu)$. The parameters of these demand functions will of course be a mixture of technology and preference parameters, which in general cannot be separately identified in empirical estimation, but at least the exogeneity and linearity of the budget constraint are preserved.

One comment on the approach proposed by Pollak and Wachter is that, in its emphasis on exogeneity and linearity of the budget constraint, it ignores the fact that budget constraints defined on market prices and net-of-tax income are in reality non-linear. Goods are often sold on non-linear tariffs, with fixed charges, bulk discounts and so on, so that average price per unit falls with quantity bought. Income support programmes and income tax structures typically induce non-linear budget constraints in the consumption-labour supply space.¹¹ Moreover, the prices consumers pay may well be endogenous, for example if consumers with unequal search costs search through a distribution of prices, or if the ability to take advantage of bulk discounts depends on income. There is an element of what could perhaps be called selective realism in the Pollak–Wachter position, with the problems of the household production approach being contrasted with an assumed ideal world of exogenous linear budget constraints. The latter is no more realistic than a world with constant returns to scale and no jointness in household production. Empirical demand analysis, whether it recognises it or not, is in fact beset by problems arising from the non-linearity of budget constraints.¹²

⁹ See Pollak and Wachter (1975), p. 203.

¹⁰ It could of course be pointed out that the assumption in the standard model, that the disutility of time spent at work is simply the utility of the leisure consumption forgone, is also quite special. One could argue that market work may also generate utility or disutility directly, for example through ‘on the job consumption’, ‘job satisfaction’ and so on.

¹¹ See, for example, chapter 6 for further discussion.

¹² For a thorough exploration of the econometric problems arising from this see Pudney (1989).

The ‘usefulness’ of a particular model formulation depends on the purposes of the analysis. It is clear from their discussion that Pollak and Wachter have in mind usefulness for the purposes of empirical demand analysis. Our concerns in this book are rather different. As will have been clear from our discussion in chapter 1, for purposes of analysing public policy we regard it as essential to incorporate domestic production into models of the household.

In particular, we think it important to keep the preferences of the household quite separate from its technology and budget constraints, rather than confounding them, as the suggested alternative approach does. The household’s technology and budget constraints determine its utility possibilities – its utility possibility frontier in the standard welfare economics sense – and the instruments of economic policy determine the household’s standard of living through its effects on this. Taxation and redistribution policy is very much concerned with redistributing income from households with higher to households with lower utility possibilities, treating their preferences as identical. Moreover, we regard the fact that the implicit prices of household goods are endogenous and vary across households as central to the analysis, although, for reasons of modelling strategy – keeping things simple – we model this as being due mainly to differences in human and physical capital and the wage and price vectors households face rather than to non-constant returns and joint production. We therefore concur with Gronau’s conclusion:¹³

The theory of home production, rather than serving as a blueprint for empirical research, is an analytical tool. The distinction between consumption and production is essential to the analysis of work at home (as distinct from consumption time). It consequently proves important for the analysis of labor supply (in particular that of married women), and the measurement of home output.

On the other hand, we would go further and argue that the theory of household production also does have an important empirical purpose and usefulness. As we pointed out in the previous chapter, the existence of household production raises important issues for public policies such as taxation and income redistribution, and so it is necessary for economists to take a greater interest in collecting data on household production and in estimating models which take it into account.¹⁴ One of the purposes of the theory is to inform us about the data we need to collect.¹⁵

2.5 Market substitutes, implicit prices and the value of household production

An interesting aspect of the work of Gronau is that, while further developing Becker’s model of household production both theoretically and empirically, he made the useful step of redefining the domestic outputs simply as goods in the usual sense, rather than

¹³ See Gronau (1986). This survey is full of insights and can still be read with great profit.

¹⁴ Gronau is in fact one of the pioneers in doing this.

¹⁵ The importance of measuring the output of home child care with varying market inputs is now evident in the extensive literature that evaluates alternative forms of care in terms of cognitive development and other outcomes. See, for example, the recent study by Bernal and Keane (2006).

the abstract ‘commodities’ of Becker’s model. Thus we can think of child care, meal preparation, laundry, shopping, house cleaning, gardening and so on as the production of goods and services quite comparable to the goods and services produced on markets. Indeed, in some papers¹⁶ Gronau went further, assuming that these household goods and services had perfect market substitutes. This is a tempting thing to do, because in that case the prices of the household goods are no longer endogenous to the household, but are determined exogenously on competitive markets and are observable, and this enormously simplifies empirical application of the models. These then become very similar to the models of household farm production that are found in the development economics literature.¹⁷ The basic model is that of a small open economy determining its production equilibrium by equating domestic marginal costs with world prices, and exporting or importing the difference between its production and consumption. The optimal production decision is separable from the optimal consumption decision as a result of the assumption of exogenously given market prices.

Thus, given the market price p of the output of the single domestic good y , the household finds its production optimum by maximising its net income or profit from domestic production

$$\pi = py - wt - qb \quad (2.38)$$

where t is time spent in household production, w is the wage, b is a bought-in market input and q is its price. Given the household’s production function for the good, $f(t, b)$, the production equilibrium (y^*, t^*, b^*) is found from the conditions

$$pf_t(t^*, b^*) = w \quad (2.39)$$

$$pf_b(t^*, b^*) = q \quad (2.40)$$

$$y^* = f(t^*, b^*). \quad (2.41)$$

Given the resulting net income π^* the household finds its consumption equilibrium by solving the problem

$$\max_{x, y, y_c, z} u(x, y_c, z) \quad \text{s.t.} \quad x + py_c + wz \leq w + \pi^* \quad (2.42)$$

where x is a market consumption good with price normalised at 1, y_c is its consumption of the household good, and z is leisure. Given its demand and supply functions $y_c(p, w, \pi)$, $y(p, w, q)$, it buys or sells the difference $y_c(p, w, \pi) - y(p, w, q)$ on the market. In the farm household models of the development literature, the household is taken to be a net seller, $y_c(p, w, \pi) < y(p, w, q)$, while in household models, it is assumed to be a net buyer, $y_c(p, w, \pi) > y(p, w, q)$.

Note that this assumes not only that there is a competitive market for y , but also that there is one for the household’s labour, and that it does in fact sell its labour on

¹⁶ See, for example, Gronau (1977) and (1980).

¹⁷ See for example Singh *et al.* (1986), for thorough discussions of both theoretical and empirical models of this kind.

this market, so that the wage is the opportunity cost of household time. If this is not the case, then we are no longer in the world of exogenously given market prices, and the nice separation between production and consumption decisions breaks down. The household's problem then becomes

$$\max_{x, y, y_c, z} u(x, y_c, z) \quad (2.43)$$

$$\text{s.t. } y = f(t, b) \quad (2.44)$$

$$x + p(y_c - y) + qb \leq 0 \quad (2.45)$$

$$t + z = 1. \quad (2.46)$$

In this case, the household must produce more of the domestic good than it consumes if it is to buy any amounts of the market goods, and it will have an endogenously determined shadow wage or price of time $\omega^* = pf_t(t^*, b^*) = pu_z/u_{y_c}$, derived as the ratio of the values of the Lagrange multipliers on its time and budget constraints at the optimum. This is of relevance to a significant proportion of households in developed countries, where one member specialises in household production and supplies no time to the labour market, even though one exists. This must imply that the shadow wage of that individual is greater than or equal to her market wage.¹⁸

The value of household production

A problem that has generated a large macroeconomic literature is that of measuring the value of household production, since it is recognised to be probably the largest single omission from the value of GDP.¹⁹ One approach to this assumes that household goods have perfect market substitutes and uses the prices of the latter to value the estimated amounts of domestic outputs, to give the value of household production as py^* in this simple model. The fact that the resulting estimates are unbelievably large should be taken as a warning signal about the reliability of this 'perfect substitutes' assumption.

Gronau identifies two main reasons for the inappropriateness of this assumption, which are not necessarily mutually exclusive. The first is simply that although many domestic outputs have some kinds of market substitute, these are not perfect, in the sense that the marginal rate of substitution between the market and domestic goods is a constant at all levels of consumption (linear indifference curves). The second is that, as Pollak and Wachter pointed out, the individual derives (positive or negative) utility from the time spent in performing the household task – think of recreational shopping, gardening, 'quality time' with children, the boredom of doing laundry and washing dishes.

We can use the simple model of this section to show how household production should be valued in the first of these two cases. Let production and consumption of the

¹⁸ This is further discussed in chapter 3.

¹⁹ As it is sometimes expressed: if a man marries his housekeeper national income goes down. Estimates put the omitted value at around 35 per cent of GDP in the developed economies.

household good now be identical, $y_c \equiv y$. Then taking x as the market substitute for y , the model applies exactly. We define the implicit price of the domestic good, p , as its marginal cost in household production, $p^* = w/f_t(t^*, b^*) = q/f_b(t^*, b^*)$. This again gives the value of household production as p^*y^* , where, however, p^* is endogenous to the household, depending as it does on the household's production function, and so not directly observable.

A number of authors²⁰ have made attempts to estimate the parameters of the household production functions to derive these marginal cost prices, but the basic problem of course is lack of data, particularly on outputs, but also to some extent also on inputs. Note, however, that where we are interested only in measuring the value of household production, we may not need these directly. Consider the household's budget constraint in (2.45). From it we have that

$$p^*y^* = w(1 - z^*) + \pi^* - x^* \quad (2.47)$$

where

$$\pi^* = p^*y^* - wt^* - qb^*. \quad (2.48)$$

The first equation tells us that, given that we can observe the values of the household's leisure and market consumptions, we just need an estimate of the implicit net income it makes on its household production. The second equation tells us that this, however, involves us in a circularity, since this net income depends on the value we are trying to measure. Suppose, however, that assuming constant returns to scale in household production would not involve too large an error. Then $\pi^* \approx 0$, and we have not one but two ways of measuring p^*y^* : either as the value of total time not spent in pure leisure (available from time-use studies) *minus* the value of market consumption, or as the value of time spent in household production (also available from time-use studies) *plus* the cost of bought-in market inputs into that activity.²¹

Note that this latter measure tells us that the second common method of valuing household production in macroeconomic studies, simply valuing time spent in household production at the market wage, not only implicitly assumes constant returns to scale, but also understates the true value because it omits the value of bought-in inputs.

Finally, assume that $\pi^* \neq 0$, and let ρ^* denote the ratio of the average to the marginal product of t at the household equilibrium. Then we can write

$$\pi^* = -[(1 - \rho^*)wt^* + qb^*] \quad (2.49)$$

²⁰ See in particular, for the earlier literature, Gronau (1980), Graham and Green (1984) and Nelson (1988). For more recent contributions to this topic see Crossley and Yu (2004), Deaton and Paxson (1998), Gronau and Hammerness (2006), Kalenkoski *et al.* (2005) and Li and Vernon (2003).

²¹ The closeness of these two separate estimates would give a measure of the appropriateness of the constant returns to scale assumption.

and substituting this into (2.47) gives (recalling that $w(1 - z^* - t^*) = x^* + qb^*$, market labour income equals total expenditure on market goods)

$$p^* y^* = \rho^* w t^*. \quad (2.50)$$

Thus some reasonable estimate of ρ^* will give the value of household production from the value of time spent in doing it.

To model the second case, assume that time spent respectively in household production *and* market labour supply l enter the utility function directly, in addition to leisure. As we suggested earlier, there is just as much justification for regarding time spent in market work as yielding positive or negative utility in and of itself. We can write the household's problem now as

$$\max u(x, y, z, t, l) \quad \text{s.t.} \quad x + qb \leq wl \quad (2.51)$$

$$l + z + t = 1 \quad (2.52)$$

$$y = f(t, b). \quad (2.53)$$

Clearly, even assuming constant returns to scale, we have again lost the nice separation between production and consumption decisions. In particular we obtain as the conditions determining time allocations

$$p^* f_t = w + \frac{(u_t^* - u_l^*)}{\lambda^*} \quad (2.54)$$

$$p^* f_b = q \quad (2.55)$$

where λ^* is the household's marginal utility of income, and p^* is again the implicit price of the household good (= marginal rate of substitution between x and y) at the household equilibrium. Thus the extent to which this model deviates from the results of the simpler model depends on the difference $(u_t^* - u_l^*)$ between the marginal utilities of time spent on household and market work at the equilibrium.

We emphasise that these are *marginal* utilities: the argument that some time spent on household or market work yields positive utility is not relevant; what matters are the marginal minutes of time spent in each of these activities. Using the simpler model, and ignoring the presence of t and l in the utility function, is then equivalent to assuming that these two marginal utilities are just about equal – *at the margin*, household and market work are equally disagreeable, as much as we may enjoy playing with the kids and chatting to colleagues over a business lunch. In the rest of this book we maintain this assumption, actually in the belief that it is not unreasonable.

2.6 Conclusions

In the discussion of household production in this chapter we have deliberately been ambiguous about what the household actually is, and this ambiguity is fully consistent with the approach in the literature. In fact, the utility functions used in the literature are assumed, as here, to have the properties of individual utility functions, though in informal discussion and illustration reference is usually made to multi-person households

or families. In our view, we cannot ignore the importance of household production in modelling households, and its relevance to many of the central issues of public policy. However, we think that this importance and relevance are only fully brought out when household production is incorporated into multi-person household models. In the next two chapters, we go on to consider respectively the theoretical and empirical aspects of such models in some depth.

3 Household models: theory

3.1 Introduction

This chapter surveys the literature on theoretical models¹ of the household, paying particular attention to some of the earlier contributions, and using them to place the current state of the theory in perspective. The first and most basic distinction we make is that between models which analyse the decisions of a single consumer/worker, and those that begin with the view of a household as consisting of more than one individual. We reserve the term ‘household models’ for the latter, and call the former ‘individual models’. We define a household as a set of two or more individuals who live together and are involved in joint (pairwise or group) decision-taking in respect of their allocations of time and money. Models of the household in this sense, also commonly called ‘family models’, are the subject of this chapter.

A second distinction concerns the purpose the models are meant to serve. Here we can contrast the approaches of the two early pioneers in this field, Paul A. Samuelson and Gary S. Becker. In his classic paper on social indifference curves,² Samuelson noted that his proof that these indifference curves could not exist, in the sense that they constitute a preference map with the same properties as that in the individual model, presented major problems for the theory of consumer demand, since observed household demands must in general be aggregates of demands of the individuals in the household. As Samuelson puts it:

Who after all is the consumer in the theory of consumer’s (not consumers’) behavior? . . . In most of the cultures actually studied by modern economists the fundamental unit on the demand side is clearly the ‘family’, and this consists of a single individual in but a fraction of cases . . . If community indifference curves do not exist, how can we expect family demand functions observed in the market place to obey the consistency axiom of revealed preference or any other regularity conditions? Why shouldn’t the pluralistic decisions of the family group result in price-quantity situations in which the point A was at one time selected even though B was cheaper, while at another time B was selected in preference to A? Could not this contradiction occur even though nobody had changed his or her preferences?

¹ Chapter 4 discusses the more recent empirical literature. ² Samuelson (1956).

Then, how can a household be modelled in such a way that its preference-ordering possesses the properties of those in the individual model? The answers given by Samuelson and others to the general questions raised by the nature of household decision-taking is the main theme of this chapter. We would phrase the question motivating Samuelson's approach in general terms as: what becomes of the standard propositions of economics, whether theoretical or empirical, positive or normative, when we realise that the basic decision-taking unit is a household and not an individual?

Becker's approach, on the other hand, was concerned not with reformulating the existing body of economic propositions, but with extending the domain of economic analysis to areas that before him had not generally been considered amenable to or appropriate for it. He was concerned to show that economic methodology could give important new insights into the workings of the household as an economic institution, in matters such as marriage, divorce, fertility, bequests, human capital formation, role allocation, and so on. Here, we will not review this important body of literature,³ but simply note that his results were mainly achieved in the context of a model that adopted the simplest possible solution to Samuelson's problem, by assuming that the household's preferences are those of an individual, the 'head of the household'. This is a much more drastic and simplistic step than that proposed by Samuelson himself, even though in the later literature the two are often, quite incorrectly, lumped together.

A third distinction we make in this introduction is that between models that do and those that do not incorporate household production, the use of market goods and household members' time to produce goods and services for consumption within the household. It is perfectly possible to introduce household production into the individual consumer model, as we saw in the previous chapter. It is also possible to formulate household models without household production, where, as in the individual model, each worker/consumer divides her time between just two uses, market labour supply and 'leisure', the direct consumption of one's own time. Throughout his discussion Samuelson, in order 'to eliminate nonessential complications', assumes a two-person pure exchange economy, so that his household consists of individuals with given endowments of consumption goods.⁴ While keeping the formal difficulties to a minimum, it seems to us to be essential to either of the basic purposes of household models – the reformulation of the standard results of economics, and the extension of its domain of application – to incorporate household production. We will make this argument in detail at various points in the rest of this chapter.

We organise this overview of the theoretical literature on household models by classifying them into three types, according to the principle by which the resource

³ For an extensive survey and discussion of Becker's work in general see Pollak (2003). Bergstrom (1997) also gives a wide-ranging survey of the broad field of household economics. Jacob Mincer was also an influential early contributor to the field; see, for example, Mincer (1963).

⁴ He obviously saw it as unproblematic to extend the model to include production. If there are joint costs and increasing returns to scale, however, the analysis is not quite so unproblematic, and some authors have seen these as central aspects of household production. See in particular the critique of Becker's household production model by Pollak and Wachter (1975), discussed in the previous chapter.

allocation decisions of the household are assumed to be solved. We distinguish between *cooperative* models, *non-cooperative* models, and *equilibrium* models.

Under the first heading we group Samuelson (1956), Becker (1981), the bargaining models of Alesina *et al.* (2007), Chen and Woolley (2001), Lundberg and Pollak (1993), McElroy and Horney (1981), Manser and Brown (1980), and Ott (1992), as well as the non-bargaining models of Apps and Rees (1988), (1997b), (2002), Basu (2006), Browning and Chiappori (1998) and Chiappori (1988), (1992),⁵ and regard them as particular formulations of Samuelson's consensual approach. We will try to distinguish the contribution each type of model makes to the development of this approach (regardless of whether or not the authors concerned saw themselves as in fact working in the Samuelson tradition).

Non-cooperative models, such as those of Leuthold (1968), Ashworth and Ulph (1981), and Konrad and Lommerud (1995), (2000), characterise the household allocation as the Nash equilibrium of a non-cooperative game. In the literature we often find the assertion that a key difference between cooperative and non-cooperative models is that the equilibrium allocations in non-cooperative models are not assumed to be Pareto-efficient, while those in cooperative models are assumed to be. This could, however, be misleading.

To see this, it is useful to recall the distinction between the *Pareto property* of social welfare functions and the *first- or second-best Pareto efficiency* of equilibrium resource allocations. The cooperative models all have, explicitly or implicitly, maximands that we will call household welfare functions, and these all have the Pareto property – an increase in the utility of any one individual, *ceteris paribus*, is always a good thing. Non-cooperative models do not possess maximands in this sense, i.e. they are not based on the idea that the household as a group solves a well-defined maximisation problem, even though the individuals within it maximise their own utilities. Nevertheless, models adopting the cooperative approach need not result in first-best Pareto-efficient equilibria, while non-cooperative models can result in first-best Pareto-efficient outcomes. The efficiency of the realised resource allocation depends on the constraints that are imposed on the allocations that are feasible, i.e. on the entire structure of the model.

For example, in a two-period Nash bargaining model the inability to make binding commitments between periods may lead to second-best outcomes,⁶ even though the Nash bargaining model assumes the Pareto property in the function to be maximised. On the other hand, in an infinitely repeated non-cooperative game, first-best Pareto-efficient allocations are often in the set of subgame perfect Nash equilibria. In other words, it is neither necessary nor sufficient that a model be cooperative for it to produce first-best Pareto-efficient allocations, and so it is a source of confusion to define the

⁵ Vermeulen (2002) surveys the theoretical and empirical work on the model presented in these papers by Chiappori and Browning and Chiappori, which we shall refer to as the 'collective model'. Sometimes in the literature this term seems to be applied to all models of the household which are *not* observationally equivalent to the individual model, including therefore bargaining models, but we think it less likely to lead to confusion if we define the term more narrowly.

⁶ See, for example, Ott (1992) and Lundberg and Pollak (2003).

differences between these classes of model in terms of the Pareto efficiency of their outcomes. It should also be noted that some of the cooperative models use aspects of the results of the non-cooperative models, for example in defining threat points in the cooperative Nash bargaining game, or in drawing on the results for private-contribution public goods games.⁷ We discuss these issues further in sections 3.2, 3.3 and especially 3.4 below.

Finally, we have models such as the Walrasian exchange model of Apps (1981), (1982), the marriage market models of Becker (1973), (1974), and the marriage labour market model of Grossbard (1976), (1984), (2003). Here, unlike both cooperative and non-cooperative models, the household members do not act as if they were in a situation of strategic interdependence, but simply take individually optimal decisions subject to the constraints presented by competitive markets. Then, while household members do not behave cooperatively in the decision-making sense, competitive market forces determine equilibrium solutions to their resource allocation problems that are typically Pareto-efficient (though possibly very unequal). An important insight that these models provide, in common with the bargaining models, is that conditions on markets external to the household have an important influence on the equilibrium allocation of resources within it.

3.2 Cooperative models

The common element of the models reviewed in this section is that the objective functions they postulate for the household, used to solve the household's resource allocation problem, are based on Samuelson's idea of a consensual or cooperative household agreement. The models differ in the attention they pay to the question of *how* the household reaches this agreement, and the effect this has on the form of its objective function. Furthermore, they assume that the household members pool their incomes in deriving the household budget constraint, in the sense that it is feasible to make any lump sum income transfers between individuals that may be required to reach the household optimum. This does not mean that all the models imply the so-called 'pooling hypothesis', which suggests the inappropriateness of this terminology and the need for an alternative, as we argue in more detail below.

3.2.1 Samuelson's theorem

To solve the problem of how to derive transitive household preference orderings that allow the derivation of well-behaved household demand functions for goods, Samuelson proposed that the household be assumed to reach a consensus that allows the application of a type of Bergson–Samuelson social welfare function (SWF), which in this context we will call the household welfare function (HWF). If we denote the utility functions of the h household members by $u^i(x^i)$, $i = 1, \dots, h$, with $x^i \in \mathbf{R}_+^n$ their consumption

⁷ See Warr (1983) and Bergstrom *et al.* (1987) for the theory of these games.

vectors, we can define the HWF as $H(u^1(x^1), \dots, u^h(x^h))$. Note that the individual utility functions are ‘selfish’, in the sense that they represent the individual preference orderings over consumption bundles.⁸ Aspects of household relationships such as love and caring can best be thought of as entering into the determination of the HWF, i.e. as part of whatever process determines the household consensus.

To understand what is implied by the translation of this concept from general equilibrium welfare economics to the analysis of the household, it is useful to spend a little more time on the meaning of the SWF than Samuelson did in his paper.⁹ The central idea is to abstract from the issue of exactly how the social group, in this case the household, arrives at its complete, transitive preference-ordering over levels of well-being of its members, as well as leaving open the choice of a specific function to represent this preference-ordering. Whatever this process may be, it is assumed that the outcome can be represented as a function possessing a minimum of three basic properties:

- *the Pareto property*: $\partial H / \partial u^i \equiv H_i > 0$, the function is strictly increasing in the well-being of each individual;
- *quasiconcavity*: given two utility profiles $[u^i(\cdot)]$ and $[\hat{u}^i(\cdot)]$ such that $H(u^1(\cdot), \dots, u^h(\cdot)) \leq H(\hat{u}^1(\cdot), \dots, \hat{u}^h(\cdot))$, and a third utility profile $[\bar{u}^i(\cdot)] = \lambda[u^i(\cdot)] + (1 - \lambda)[\hat{u}^i(\cdot)]$, for $\lambda \in [0, 1]$, we have that $H(\bar{u}^1(\cdot), \dots, \bar{u}^h(\cdot)) \geq H(u^1(\cdot), \dots, u^h(\cdot))$. Loosely, household indifference surfaces in the space of individual utilities are linear or strictly convex to the origin;
- *differentiability* to any required order at all points in the domain of the function.

The first of these properties seems reasonable for any household. As far as quasiconcavity is concerned, the welfare economics literature takes as limiting cases the weighted utilitarian case, $H = \sum \alpha_i u^i(x^i)$, $\alpha_i \geq 0$, $\sum \alpha_i = 1$, and the Rawlsian $H = \min_{u^i} [u^i(x^i)]$. The first displays no inequality aversion, in the sense that all profiles that give the same (weighted) sum of utilities are equally good, whatever they may imply for inequality of utility among the household members. The Rawlsian HWF possesses infinite inequality aversion, in that the household would be seeking to maximise the utility of its worst-off member. All strictly quasiconcave functions between these extremes are possible household welfare functions, with varying degrees of inequality aversion. The third property is a useful technical assumption.¹⁰

The key idea is that the function provides an analytical device that allows the economist to derive the implications for resource allocation of any specific set of ethical or value judgements that would determine its precise form, or indeed of all

⁸ Though they may be extended to incorporate whatever externalities and public goods may be present within the household. These extensions have been extensively studied in the welfare economics literature.

⁹ For fuller discussion of the concept, in the context that preoccupied theoretical and applied welfare economists at the time, see Bergson (1938) and Samuelson (1947). Needless to say, there has been considerable further development since then. For good surveys see Boadway and Bruce (1984), Jehle and Reny (2001) and Mas-Colell *et al.* (1995).

¹⁰ Note that the limiting case of the Rawlsian HWF possesses only the weak Pareto property $H_i \geq 0$ and is not everywhere differentiable, and so strictly speaking is excluded by these assumptions. It can, however, be approximated arbitrarily closely by a function that does possess these properties, so we continue to regard it as a possible HWF.

sets of value judgements that possess these three properties. Its widespread use in public economics, international economics and applied welfare economics generally over the past half-century confirms its usefulness.

As we argue below, the assumption that the household chooses its resource allocation by a process of Nash bargaining is equivalent to choosing a specific form for the HWF. The assumption about the process by which the household consensus is reached is sufficient to place a specific structure on the HWF. The key innovation that resulted from this, yielding a significant extension of Samuelson's approach, was, in effect, to introduce exogenous variables such as prices, wages and non-wage income, as well as other 'extrahousehold environmental parameters'¹¹ (EEPs) into the HWF as conditioning variables that determine the household's ordering over the utility profiles of its members.

Suppose now that the household faces a price vector $p \in \mathbf{R}_{++}^n$ and the household members have exogenously given individual non-wage incomes μ_i , with $\mu = \sum_i \mu_i$. If the individuals pool their incomes, we can write the household's budget constraint as $px \leq \mu$, and Samuelson formulated the household's resource allocation problem as that of maximising $H(\cdot)$ subject to this constraint. We summarise his discussion of this problem by the following theorem.¹²

Theorem (Samuelson, 1956): If there exists a HWF $H(\cdot)$ with the three properties given above, the following statements hold:

(i) *Decentralisation*: if $\hat{x}^i$ are solutions to the problem $\max_{x^i} H(\cdot)$ s.t. $px \leq \mu$, there exist functions $s_i(p, \mu)$, with $\sum_i s_i(p, \mu) = \mu$, such that $\hat{x}^i$ are also solutions to the problems $\max_{x^i} u^i(x^i)$ s.t. $px^i \leq s_i(p, \mu)$, $i = 1, \dots, h$. Samuelson termed the functions $s_i(p, \mu)$ the *sharing rule*. Conversely, existence of this sharing rule implies the existence of a HWF.

(ii) *Aggregation*: let $v^i(p, s_i)$ be the indirect utility functions derived as the value functions of the problems $\max_{x^i} u^i(x^i)$ s.t. $px^i \leq s_i$ for any s_i . Then the $s_i(p, \mu)$ are the solutions to the problem¹³ $\max_{s_i} H(v^1(p, s_1), \dots, v^h(p, s_h))$ s.t. $\sum_i s_i \leq \mu$. Furthermore let $x = \sum_i x^i$ be the household's aggregate demand vector. Then the value function in this latter problem, say $v(p, \mu)$, possesses all the properties of an indirect utility function, with in particular $\nabla_p v(p, \mu) = -v_\mu x$ (Roy's identity). We can call v_μ the household's marginal utility of income.

Intuitively, we can interpret the HWF as formally identical to an individual utility function with weak separability in consumptions, with the quantity of each physical

¹¹ The terminology introduced in McElroy (1990).

¹² We state this theorem here in its modern form. This is meant to summarise Samuelson's lengthy discussion of sharing rules as well as his formal theorem. For proofs, see Mas-Colell *et al.*, (1995), ch. 4. The usual context for this theorem is the economy as a whole, but of course it immediately applies to the household as a small economy. The key insight is '[The] problems of home economics are, abstractly conceived, exactly of the same logical character as the general problem of government and social welfare' (Samuelson, 1956, p. 10).

¹³ We are free to choose utility functions which make $H(\cdot)$ here a quasiconcave function of the s_i , so that the second-order conditions will be satisfied.

good consumed by each household member viewed as a separate good. Then the standard results on two-stage budgeting¹⁴ apply: the household can be modelled as first allocating its total expenditure among its members, and then within each expenditure category, i.e. for each individual, allocating the given expenditure optimally over goods.

Note that the sharing rule functions necessarily contain as arguments prices and total non-wage income, because these are the exogenous variables in the problem that generates them. Intuitively, shifts in the household's budget constraint and corresponding utility possibility set can be expected to change the distribution of income among household members *given* its unchanged preferences over their relative well-being. Note, however, that they do not depend on the *individual* non-wage incomes: any reallocation of these that leaves the total unchanged always also leaves the share functions unchanged. This, as we shall see below, is generally in contrast to the results of bargaining models.

The aggregation aspect of the theorem is also important. This implies that the aggregate household demands for goods, the vector $x(p, \mu)$, can be modelled *as if* they were the demands of a single fictitious consumer with the indirect utility function $v(p, \mu)$. This is a complete description of preferences, and an expenditure function and a direct utility function $u(x)$, which we refer to as the *household utility function* (HUF) for this fictitious individual, can be derived from it in the usual way. The assumptions on the HWF, combined with the form of its maximisation problem (including the pooled budget constraint), give a problem in the space of aggregate household consumption vectors, $\max_x u(x)$ s.t. $px \leq \mu$, that is identical in form to the problem of an individual consumer, and so the resulting aggregate demands will have all the regularity properties of the demand functions derived from the model of the individual consumer. We can derive aggregate demands by maximising a HUF subject to the aggregate budget constraint. We should, however, not lose sight of the fact that underlying these demands is an allocation of consumptions among household members that reflects the household's sharing rule or, equivalently, maximises the HWF.

Becker's approach, identifying the HWF with the individual utility function of the 'head of the household', is in one sense a special case of Samuelson's. On the other hand, the models of the household Becker formulates are much richer than Samuelson's general equilibrium pure exchange model and are concerned with the analysis of a far wider range of issues. Opinions can differ on the extent to which the significance of the insights Becker derives outweighs any reservations one may have about the ethical presuppositions represented by the form of his HWF, as well as about its descriptive realism.¹⁵

Samuelson also considered the approach to modelling household demand that has since become known as Gorman aggregation.¹⁶ This can be thought of as providing

¹⁴ See Gorman (1961) and Deaton and Muellbauer (1980).

¹⁵ We have had students who maintain that Becker's HWF is fully descriptive of the households they have grown up in. For a good overall assessment see Pollak (2003).

¹⁶ See Samuelson (1956), Gorman (1959) as well as Deaton and Muellbauer (1980).

a solution to the problem of deriving well-behaved aggregate demand functions for a group of consumers, for example a household, without having to specify explicitly a resource allocation process within the group, other than that individuals within it receive income shares, somehow determined. Suppose each household member i has a strictly positive demand for each good $j = 1, \dots, n$ that takes the functional form

$$x_{ij} = a_{ij}(p) + b_j(p)s_i \quad (3.1)$$

where s_i is i 's total expenditure on goods and $a_{ij}(p)$ and $b_j(p)$ are given functions. Then $\sum_i x_{ij}$, the aggregate household demand for each good j , is a function only of the price vector and aggregate household income $\mu = \sum_i s_i$. This is because, when all demands are strictly positive, a reallocation of this aggregate income among household members leaves aggregate demand unchanged, since the coefficients $b_j(p)$ are the same for all i .

Note, however, that if a good, for example one person's leisure, is always consumed in a positive amount by only that person, and in zero quantity by all others, then its total household demand is a function only of that individual's income, and not of aggregate household income. A pure redistribution of total income that changed one individual's income share would change the demand for her leisure.

The demands in (3.1) are generated by a utility maximisation process if and only if each household member has the Gorman polar form of indirect utility function

$$v^i(p, s_i) = \alpha_i(p) + \beta(p)s_i. \quad (3.2)$$

Thus one could argue that placing these very strong restrictions on preferences would allow an econometrician to work with aggregate household demand functions without having to construct a household model. Samuelson dismissed this approach as much too restrictive. Moreover, if demands are generated by individual utility maximisation, given some distribution of total expenditures, then the resulting household resource allocation must be Pareto-efficient, as a specific implication of the fact that all consumers face the same price ratios of all pairs of goods. So even assuming the Gorman polar form implies some restriction on the household decision process.

Samuelson clearly felt confident that he had given a satisfactory solution to the problem he had posed. Why was this solution apparently rejected by most of the subsequent contributors to the literature? One reason put forward¹⁷ was that the process of household formation cannot be analysed by a household utility function approach. This is not correct. Prior to formation of a household, the potential members must be able to predict the allocations they will obtain within the household and compare them to the next best alternative. They must do this by solving the HWF-maximisation problem. This is qualitatively no different from what is assumed in models that use Nash bargaining, i.e. a specific HWF, to analyse household formation or the marriage decision,¹⁸ and much more general. Another¹⁹ was that Samuelson's approach misses the elements of

¹⁷ See, for example, Nerlove (1974). ¹⁸ See, for example, Manser and Brown (1980).

¹⁹ See Manser and Brown (1980).

conflict as well as cooperation that characterise household decision-taking. This is also not correct. Samuelson's formulation is quite general as to the processes that generate the HWF, and these may well involve conflict. The HWF simply represents the outcome of its resolution.

McElroy and Horney (1981), in motivating their Nash bargaining approach, argue that Samuelson's solution implies that the outcomes of family decisions are 'empirically indistinguishable from those of constrained utility maximisation' in the individual consumer model. This is the key point, and we now turn to it.

3.2.2 Symmetry, anonymity and pooling

3.2.2.1 Symmetry The aggregation part of Samuelson's theorem says that the household can be analysed *as if* it generates its aggregate demands $x(p, \mu)$ by solving the problem $\max_x u(x)$ s.t. $px \leq \mu$, where $u(\cdot)$ has all the properties of an individual utility function. This implies that the aggregate household demands will have the same properties as those of individual demands in the standard consumer model, and in particular that the Slutsky matrix of compensated demand derivatives, where 'compensation' in this case must be defined as holding the value of the HWF constant, will be symmetric and negative semidefinite.

For later reference, it is useful to illustrate Samuelson's theorem in terms of a simple two-person labour supply model. Define the first two goods as the respective leisure consumptions of the two household members, with therefore the first two prices as their wage rates w_i , $i = 1, 2$. The remaining $n - 2$ goods are consumption goods, and let us assume that their prices remain constant relative to each other, so that we can apply Hicks's composite commodity theorem and write the utility functions as $u^i(x_i, z_i)$, where x is the composite consumption commodity and z is leisure. The pooled household budget constraint is $\sum_i (w_i z_i + x_i) \leq \sum_i (w_i + \mu_i)$, where we have normalised wages and prices so that the composite consumption commodity has a price of 1, and total time endowment of each household member is normalised at 1, so $(1 - z_i)$ is that individual's labour supply.²⁰

Given the sharing rule interpretation of the household equilibrium, each individual solves $\max_{x_i, z_i} u^i(x_i, z_i)$ s.t. $w_i z_i + x_i \leq s_i$ for given s_i to yield the indirect utility function $v^i(w_i, s_i)$ with, from Roy's identity,

$$\frac{\partial v^i}{\partial w_i} = -\frac{\partial v^i}{\partial s_i} z_i. \quad (3.3)$$

The solution to the household's distribution problem

$$\max_{s_i} H(v^1(w_1, s_1), v^2(w_2, s_2)) \quad \text{s.t.} \quad \sum_i s_i \leq \sum_i (w_i + \mu_i)$$

²⁰ Given the assumption that each i has a positive market labour supply. If one of the household members, for example a child, *necessarily* has zero market labour supplies, we can simply set the corresponding $z_i = 1$. To deal with cases in which an individual may choose between zero and positive labour supply, we simply have to extend the formulation of the problem to include the possibility of corner solutions. See subsection 3.2.3.1.

has first order conditions

$$H_i \frac{\partial v^i}{\partial s_i} = \lambda \quad i = 1, 2. \quad (3.4)$$

The value function of this problem $v(w_1, w_2, \mu_1, \mu_2)$ has, by the Envelope theorem, the derivatives

$$\frac{\partial v}{\partial w_i} = H_i \frac{\partial v^i}{\partial w_i} + \lambda \quad i = 1, 2 \quad (3.5)$$

$$\frac{\partial v}{\partial \mu_i} = \lambda. \quad (3.6)$$

Then substitution gives

$$\frac{\partial v}{\partial w_i} = \lambda(1 - z_i) \quad i = 1, 2 \quad (3.7)$$

as should be the case if $v(\cdot)$ is an indirect utility function for the fictitious aggregate household member. The direct utility function corresponding to this indirect utility function can be written as $u(x, z_1, z_2)$ with $x = \sum_i x_i$.

Now, since $u(x, z_1, z_2)$ has the properties of an individual utility function, maximising it subject to the budget constraint $\sum_i w_i z_i + x \leq \sum_i (w_i + \mu_i)$ yields leisure demand functions $z_i(w_1, w_2, \mu)$ to which the standard results of the individual labour supply model apply. In particular, we have symmetry of the compensated leisure demands for some given constant utility level²¹ u^0

$$\left[\frac{\partial z_1}{\partial w_2} \right]_{u=u^0} = \left[\frac{\partial z_2}{\partial w_1} \right]_{u=u^0}. \quad (3.8)$$

These are, in this model, the negatives of the compensated labour supply derivatives.

In the literature it is claimed that the symmetry of compensated labour supply derivatives of husbands and wives is rejected by the data, which in turn implies rejection of this implication of Samuelson's theorem. As we argue at some length in chapter 4, because they are based on a false interpretation of the data, by assuming that all time not spent in market work is leisure, and do not take account of household production, these studies do not in fact provide a satisfactory empirical test of Samuelson's model. Nevertheless, we doubt that symmetry would hold in an empirical analysis that did correctly specify the model.

3.2.2.2 Anonymity and pooling The formulation of the budget constraint, with aggregate household expenditure constrained by aggregate household income, rests, as we pointed out earlier, on the assumption that members of the household in effect pool their incomes, in the sense that they are prepared implicitly or explicitly

²¹ That is, some given value of the HWF.

to make whatever lump sum transfers of income between themselves are necessary to achieve the household optimum. The individual shares in the total value of consumption are derived from the household optimisation, rather than being fixed a priori by some sharing rule *not* derived in this way.²² In the interests of clarity, we believe that the term ‘income pooling’ should be reserved for this kind of formulation of the budget constraint in a household model.

As we just saw, in Samuelson’s model consumption and leisure demands, labour supplies and the household sharing rule are functions only of aggregate household income. This implies that the effect on these demands, supplies and income shares of an increase in an individual’s non-wage income does not depend on who that individual is – an additional \$1 of income has the same effect regardless of to whom it accrues. This has come to be known in the literature as the ‘pooling hypothesis’, since it appears to be the result of the assumption that income is pooled to obtain the budget constraint.

This terminology is unfortunate, because there are several models, including the bargaining models that we look at in the next section, which have household budget constraints that assume income pooling, but in which the effects of a change in individual income do depend on the identity of that individual, at least for some types of income change. Moreover, in non-cooperative models of household public good provision, where income pooling is not assumed, if both individuals supply positive amounts of the public good at the Nash equilibrium, then a redistribution of income between the two individuals has no effect, and only total income matters. In other words, income pooling as such is neither necessary nor sufficient for the ‘pooling hypothesis’ to be a result of a model, and whether it is or is not depends crucially on the rest of the structure of the model, as well as on what type of income change is being considered. It seems to us therefore that it would be much better to call this effect the ‘anonymity hypothesis’, since it is really saying that the identity of the income recipient does not matter. It is this result of a model, and not income pooling per se, that would be rejected when the data seem to show that the effects of an income change depend on precisely whose income has in fact changed. Samuelson’s model does imply the anonymity hypothesis. There is now quite a substantial body of empirical work that concludes that this hypothesis is rejected by the data.²³

This discussion suggests that Samuelson’s formulation was *too* successful in solving the problem of household decision-taking – it produced a model with identical results to the individual model, as McElroy and Horney pointed out. For them and other proponents of bargaining models, the answer was to base the model on the *process* by which the household consensus is achieved. Before considering the Nash bargaining approach in some depth, we look at three generalisations of Samuelson’s model: the incorporation of household production; the inclusion of children; and the generalisation of the household welfare function.

²² What Samuelson called a ‘shibboleth sharing rule’.

²³ For the early contributions see Schultz (1990) and Thomas (1990). The study by Lundberg *et al.* (1997) has also been very influential.

3.2.3 Generalising Samuelson

3.2.3.1 Household production The book by Apps (1981), and the papers by Apps (1982), and Apps and Rees (1988), (1997b), (1999), (2002), fall within the Samuelson tradition, in that they were concerned with extending existing results, primarily in labour supply and public economics rather than demand theory in general, to household models. They extend Samuelson's approach in that they include household production as an essential component of the models. Indeed, for the issues with which these papers were concerned, it was seen to be essential not only to model the multi-person household, but also to incorporate household production. There was no point in doing one without the other. This of course is also the approach of this book.

We first present two specific models, which extend the approach to household production discussed in chapter 2 to the case of two-person households. We use them to analyse some aspects of female labour supply. We then go on to generalise these models, and in doing so extend them to include children and child care.

Model 1

We assume that each household member produces a separate domestic good with output proportional to his or her time input. We allow male and female domestic productivities in general to vary across households, but it is sometimes useful to assume that male productivity is the same across households and only that of the female varies. This variation could be due to either or both differences in human capital and differences in household physical capital, which, since it is assumed exogenous, does not appear explicitly.²⁴

In a household, each partner produces a good which is partly consumed by her or himself and partly traded for the other good. The production functions are

$$y_j = k_j t_j \quad j = f, m \quad (3.9)$$

where t_j is the time spent in producing the household good y_j by individual $j = f, m$.

The individuals have the strictly quasiconcave and increasing utility functions

$$u_i(x_i, y_{if}, y_{im}) \quad i = f, m \quad (3.10)$$

where

$$\sum_{i=f,m} y_{ij} = y_j \quad j = f, m \quad (3.11)$$

and y_{ij} is the amount of i 's consumption of the domestic good produced by j . There is just one market consumption good, x , with price normalised at 1. Note that there are no pure leisure and no bought-in market inputs to household production in this model. These assumptions, together with the linearity of the technology, allow considerable

²⁴ The cost of the fixed household capital could be regarded as just being subtracted from non-wage income, which could therefore be negative.

simplification of the comparative statics analysis, which in many contexts compensates for the loss in generality.

The household's budget constraint is given by

$$\sum_{i=f,m} x_i = \sum_{i=f,m} (w_i l_i + \mu_i) \quad (3.12)$$

where μ_i is i 's non-wage income, l_i is labour supply and w_i is the wage rate. The price of the market consumption good is normalised at 1.

Given the individual time constraints

$$l_i + t_i = 1 \quad i = f, m \quad (3.13)$$

where total available time is normalised to 1, if we define the implicit prices of the domestic goods as

$$p_j = \frac{w_j}{k_j} \quad j = f, m. \quad (3.14)$$

We can replace the separate constraints (3.9), (3.12) and (3.13) with the single *full-income budget constraint*

$$\sum_{i=f,m} \left(x_i + \sum_{j=f,m} p_j y_{ij} \right) = \sum_{i=f,m} (w_i + \mu_i) \equiv X. \quad (3.15)$$

Note also that the price of at least one of the domestic goods varies across household types, with more productive households having the lower price. These prices have a straightforward interpretation: they are the values, in terms of consumption, of the time required to produce one unit of the respective household goods, i.e. they are the marginal opportunity costs²⁵ of the domestic outputs. All households have the same full income X , so high productivity households are clearly better off, because they face a lower implicit price of at least one domestic good, and so can enjoy higher total consumption at any given full income.

Adopting Samuelson's household welfare function approach, we can regard the household as generating its consumption allocation by first choosing income shares $s_i = s_i(w_f, w_m, \mu_f, \mu_m)$, such that

$$\sum_{i=f,m} s_i = X. \quad (3.16)$$

This simply reflects the distributional choices of the household members. Then each solves the problem

$$\max u_i(x_i, y_{if}, y_{im}) \quad \text{s.t.} \quad x_i + \sum_{j=f,m} p_j y_{ij} = s_i \quad i = f, m \quad (3.17)$$

²⁵ Note that the opportunity cost of time spent in household production is the market wage, which assumes strictly positive labour supply and no restrictions on the demand side of the labour market. For the case of a corner solution with zero market labour supply see below.

to yield demand functions $x_i(p_f, p_m, s_i)$, $y_{ij}(p_f, p_m, s_i)$, together with the marginal utilities of individual income $\lambda_i(p_f, p_m, s_i)$, the shadow prices of the individual budget constraints.

The problem in (3.17) is a standard consumer problem and we can in the usual way define the indirect utility functions

$$v_i = v_i(p_f, p_m, s_i) \quad i = f, m. \quad (3.18)$$

Then applying Roy's identity gives the partial derivatives

$$\frac{\partial v_i}{\partial p_j} = -\lambda_i y_{ij} \quad i, j = f, m \quad (3.19)$$

From the solution to (3.17) we can also derive the market labour supplies as

$$l_j(p_f, p_m, s_f) = 1 - \frac{1}{k_j} \sum_{i=f,m} y_{ij}(p_f, p_m, s_i) \quad j = f, m. \quad (3.20)$$

Thus an individual's labour supply is the difference between total time available and the amount of time required to meet total household demand for his or her domestic output. When non-market labour time is spent in household production rather than pure leisure, the demand for it is the aggregate of the individual demands of all household members, and not just of the individual concerned.

Model 2

In this model both household members' time is used, together with a bought-in market good, to produce a single, generalised household good. We rewrite the utility functions as

$$u(x_i, y_i, z_i) \quad i = f, m \quad (3.21)$$

where x and y are market and household consumption goods respectively, and z , measured in units of time, is leisure. The time constraints are now

$$l_i + t_i + z_i = 1 \quad i = f, m. \quad (3.22)$$

The production function of the household good is concave and strictly increasing and given by

$$y = g(b, t_f, t_m; k) \quad (3.23)$$

where b is the bought-in market good and k is now thought of as a parameter indicating general household productivity,²⁶ with $\partial y / \partial k > 0$. The budget constraint now

²⁶ This can again be thought of as being determined by the household's given stock of physical capital and/or the human capital of the household members. The important point is that it may vary across households. Exactly how variations in k affect the marginal products of the three inputs can be left open.

becomes

$$\sum_{i=f,m} x_i + qb = \sum_{i=f,m} (w_i l_i + \mu_i) \quad (3.24)$$

with q the price of the market input²⁷ and w_i the wage rate of individual i .

We could again solve the household's resource allocation problem by maximising an HWF subject to the time and production function constraints (3.22) and (3.23), and the budget constraint (3.24). However, it turns out to be more useful to exploit the idea of the household as a small economy, with a production sector and a consumption sector.²⁸

Turning first to the production sector, we can obtain the implicit price of the household good, p , by solving the problem of minimising production cost subject to an output constraint:

$$\min_{t_i, b} C = \sum_{i=f,m} w_i t_i + qb \quad \text{s.t.} \quad y = g(b, t_f, t_m; k). \quad (3.25)$$

This yields demand functions for domestic time use $t_i(q, w_f, w_m, y; k)$ and a cost function $C(q, w_f, w_m, y; k)$. We set

$$p = \frac{\partial C(q, w_f, w_m, y; k)}{\partial y} \quad (3.26)$$

i.e. the implicit price of the household good equals its marginal cost. Thus in general p is not only specific to the household, but also variable with respect to output y .

However, if the production function is linearly homogeneous, we can write the cost function as

$$C = c(q, w_f, w_m; k)y \quad (3.27)$$

where $c(\cdot)$ is a unit cost function, and we set $p = c(q, w_f, w_m; k)$. Thus the price of the domestic good is still household-specific but now invariant to output. Note that this price depends on the wage rates of both household members, and

$$\frac{\partial c}{\partial w_i} = \hat{t}_i \quad (3.28)$$

where $\hat{t}_i$ is the amount of i 's time required to produce one unit of the household good.

Define

$$\pi = pg(b, t_f, t_m; k) - \sum_{i=f,m} w_i t_i - qb. \quad (3.29)$$

Given that the t_i, b are chosen to solve the minimisation problem (3.25), π is the implicit profit from household production (in the case of constant returns to scale $\pi = 0$). Then

²⁷ This could, but need not, be the same across households. In some applications of this model, where, for example, the market input could be bought-in child care, it would be of interest to allow q to vary across households.

²⁸ The following derivations are perfectly standard but serve to introduce the notation.

household full income is

$$X = \pi + \sum_{i=f,m} (w_i + \mu_i). \quad (3.30)$$

X determines the total value of the resources available to the household for expenditure on the market and household consumption goods and leisure. Thus we can write the full income budget constraint as

$$\sum_{i=f,m} x_i + C(y) + \sum_{i=f,m} w_i z_i \leq X \quad (3.31)$$

or, in the special case of constant returns to scale

$$\sum_{i=f,m} (x_i + p y_i + w_i z_i) \leq X. \quad (3.32)$$

Samuelson's theorem says that there exist functions $s_i(p, w_f, w_m, \mu)$, with $\sum_i s_i = X$, such that the individual supplies and demands solve the problem

$$\max_{x_i, y_i, z_i} u_i(x_i, y_i, z_i) \quad \text{s.t.} \quad x_i + p y_i + w_i z_i \leq s_i \quad i = f, m \quad (3.33)$$

yielding individual demand functions and an indirect utility function

$$x_i(p, w_i, s_i) \quad y_i(p, w_i, s_i) \quad z_i(p, w_i, s_i) \quad v_i(p, w_i, s_i) \quad i = f, m \quad (3.34)$$

with all the usual properties.

Using the time constraints, labour supplies in this model are given by

$$l_i = 1 - t_i(q, w_f, w_m, y; k) - z_i(p, w_i, s_i) \quad i = f, m \quad (3.35)$$

where, in the demand functions for t_i , $y = \sum_{i=f,m} y_i(p, w_i, s_i)$.

In later chapters we apply these models to the analysis of consumption and saving over the life cycle, taxation and income redistribution policy. Here, we use them to discuss heterogeneity across households, the variation in household utility possibilities and the determinants of labour supplies.

Across-household heterogeneity and utility possibilities

A major concern in much of public economics, in particular in the analysis of taxation and income redistribution, is the comparison of welfare across households. Issues of equity or income distribution involve evaluating the gains and losses to different households caused by specific policies. In practice, a frequently used indicator of household well-being is money income. The higher a household's income from market labour supply, the better off the household members are assumed to be. There are two well-recognised objections to this in the public economics literature. The first is that one individual's 'capacity for satisfaction' from a given money income may differ from that of another, so that utility differences may not reflect income differences. This is, however, regarded as something of a nuisance, and is typically circumvented in analytical work by assuming that individuals have identical utility functions. A more important

objection is that individuals may differ in their innate productivity or capacity to generate labour income, so that a given money income could be generated by differing levels of effort, implying therefore differing levels of utility of the individuals, after the disutility of effort is taken into account. As we shall see in chapters 6–9, this is a major consideration in the formulation of models of optimal income taxation and tax reform. The aim is to try to tax individuals' innate productivity endowments, rather than simply their money incomes.

A serious limitation of the existing discussion is, however, its identification of the household with the individual, and its implicit assumption that time is divided between market work and leisure. If two individuals have the same utility function, the same money income²⁹ and the same amount of leisure, identified as non-working time, they would in this model have the same utility level. However, real public policy concerns multi-person households with significant amounts of domestic production. What is the relationship between money income and individual welfare in that case?

We concern ourselves here with the issue of the *utility possibilities* of the household: the set of possible utility levels available to household members for all possible allocations of its resources. If we assume identical utility functions across households, then we can structure the discussion entirely in terms of the full income budget constraint of each household in model 2, since this determines the household's utility possibilities. For convenience we reproduce this here

$$\sum_{i=f,m} (x_i + py_i + w_i z_i) \leq X \quad (3.36)$$

where, in the special case of constant returns,

$$p = c(q, w_f, w_m; k) \quad (3.37)$$

where k is productivity in domestic production.

First, we can say that, other things being equal, the higher the household's full income X on the right-hand side of (3.36), the higher the household's budget constraint and therefore its utility possibilities. Thus if household 1 has higher non-wage income than household 2 then, other things equal, it is better off. If the domestic good is a normal good, a higher non-wage income is consistent with f in 1 having a lower labour supply than f in 2. Thus female labour supply and household labour market income could be lower in household 1 (holding wage rates constant between the two households), but it has the higher utility possibilities.

Now consider the left-hand side of (3.36). Other things (i.e. full incomes) equal, the household is better off the lower the price of the household good, and therefore, again other things equal, the higher the domestic productivity and the lower the price of the bought-in market input. Given wage rates constant across households, the high-productivity household 2 will be better off than 1. If there is a positive association between f 's productivity in household production and her market labour supply, then

²⁹ It is assumed throughout that people face the same prices for market consumption goods.

household 2 will also have the higher female labour supply and market income. On the other hand, if the association between productivity in household production and f 's labour supply is negative, this would make household 2's market income lower than 1's. In this case household utility possibilities are negatively associated with household market income, and the latter is a poor indicator of household welfare. Unfortunately, the theoretical analysis carried out below shows that both cases are possible. Thus we have to conclude that in order to compare welfare across households, and in particular to assess whether total household market income is a suitable welfare indicator, much more needs to be known about the empirical relationship between female labour supply and domestic productivity.

Assuming still that preferences are identical across households, in a model *without* household production, differences in labour supplies across households would be explained by differences in wage rates, in non-wage incomes, and in consumption goods prices. The latter are, however, also usually assumed identical across households. The higher the wage rates and/or non-wage incomes a household faces, the higher will be its budget constraint, and therefore its utility possibilities.³⁰ Redistributive taxation seeks to shift income from households with higher to those with lower utility possibilities. It typically does this in practice by redistributing income from higher-income to lower-income households. Thus it rests on the assumption that market income is a good indicator of the utility possibilities of the household, which is not a problematic assumption in a model without household production.

To make this discussion more precise, we take the expression for labour supply in (3.35) in model 2, and consider its derivative with respect to k . This is

$$\frac{\partial l_i}{\partial k} = -\frac{\partial t_i}{\partial k} - \left(\frac{\partial t_i}{\partial y} \frac{\partial y}{\partial p} + \frac{\partial z_i}{\partial p} \right) \frac{\partial p}{\partial k}. \quad (3.38)$$

Thus we can say:

- if increased domestic productivity reduces the time i spends in household production, this will, other things being equal (especially the output of the household good), increase her market labour supply and household income;
- since increased domestic productivity will reduce the implicit price of the household good, market labour supply will increase, other things being equal, if the term in brackets in (3.38) is positive;
- the first part of this term is, however, negative if a fall in the price of the household good increases demand for it (sufficient for this is that it is a normal good), since it corresponds to the increase in i 's time required to meet the increased demand for the household good induced by the price fall;

³⁰ As long as the members of the household are net sellers of their labour on the market, their utility levels increase with their wage rates despite the fact that this increases the implicit prices of their household good. Consider, for example, the effect of an increase in oil prices on the GDP of an oil-exporting country: this must be positive, even if it is also a large oil consumer.

- the second part of the term will also be negative if leisure time and the domestic good are complements, while it will be positive, and therefore tending to offset the first part, if they are substitutes.

Thus in general it is an open question whether an increase in domestic productivity is associated with an increase or a decrease in female labour supply.³¹ It is then an empirical question as to whether this effect reverses the positive association between labour income and utility possibilities, which the above equation suggests is perfectly plausible, or reinforces it. To accept that labour income *is* a good indicator of household utility possibilities is to assume that the second is the case, without, however, any empirical evidence to support this.

The effects of wage changes on an individual's (uncompensated) labour supply

We focus on the effects of changes in wage rates on female labour supplies and, for variety, adopt model 1. Thus differentiating through (3.20) with respect to w_f gives

$$\frac{\partial l_f}{\partial w_f} = -\frac{1}{k_f} \left\{ \sum_i \frac{\partial y_{if}}{\partial p_f} \frac{\partial p_f}{\partial w_f} + \sum_i \frac{\partial y_{if}}{\partial s_i} \frac{\partial s_i}{\partial w_f} \right\}. \quad (3.39)$$

We see in this equation how productivity, preferences and household income distribution interact to determine female labour supply. The inverse productivity parameter $1/k_f$ determines how a given change in the demand for the household good produced by f translates into a change in demand for her time. The higher the productivity, the smaller the change in labour supply given the demand change. The productivity coefficient also determines the sensitivity of the implicit price of the domestic good to a change in the wage rate, since

$$\frac{\partial p_f}{\partial w_f} = \frac{1}{k_f}. \quad (3.40)$$

Thus, the larger this productivity, again the smaller the change in labour supply, because the smaller the effect of the wage rate on the implicit price or opportunity cost of the domestic good.

Preferences determine both the price and income derivatives of demands, $\sum_i \partial y_{if}/\partial p_f$ and $\sum_i \partial y_{if}/\partial s_i$. These are the counterparts of the derivatives of leisure demands with respect to wage rates in the standard consumer model. Here, they will depend on the degree of substitutability between domestic and market consumption goods.

Note that the assumption of separability between leisure and consumption, which is often made in the economics literature, is harder to make intuitively plausible when translated into the terms of the present model, since it is hard to believe that a change in the price of the market good would not affect the demand for the domestic good,

³¹ This ambiguity was fully recognised in early work on female labour supply; see in particular Gronau (1973).

and conversely, and so time allocation would also necessarily change. Note also that the basic restriction from the standard consumer model holds also in this model – if the household good is a normal good, these demand derivatives are respectively negative and positive.

The last term is the distributional effect. A change in the female wage rate increases household full income, but the effect on female labour supply works through the household sharing rule, since the changes in individual income shares in conjunction with the individual income demand derivatives determine the overall change in demand for the household good produced by f , and therefore the change in her market labour supply.

There is also a simple representation in terms of elasticities. We can write (3.39) in elasticity terms as

$$\varepsilon_{ff} = \frac{t_f}{l_f} \left(\eta_{ff} - \sum_{i=f,m} v_{if} \omega_{if} \right) \quad (3.41)$$

where:

$$\varepsilon_{ff} = \frac{w_f}{l_f} \frac{\partial l_f}{\partial w_f} \quad (3.42)$$

is the labour supply elasticity;

$$\eta_{ff} = - \sum_i \frac{\partial y_{if}}{\partial p_f} \frac{p_f}{y_f} \quad (3.43)$$

is the price elasticity of demand for f 's household good;

$$v_{if} = \frac{\partial y_{if}}{\partial s_i} \frac{s_i}{y_f} \quad (3.44)$$

is i 's income elasticity of demand for f 's household good; and

$$\omega_{if} = \frac{\partial s_i}{\partial w_f} \frac{w_f}{s_i} \quad (3.45)$$

is the elasticity of i 's income share with respect to a change in f 's wage.

We can apply these results to the question of how we might explain a leading stylised fact about male and female labour supplies: *female labour supply elasticities with respect to their wage rate are much higher than those of males*. This is consistent, *ceteris paribus*, with:

- women having a higher ratio of time spent on household work to time spent on market work;
- the price elasticity of demand for the household goods they produce is higher;
- the proportional effect of a change in the household income distribution is lower.

Consider now the effect of a change in the male wage rate on female labour supply. Again differentiating through (3.20) gives

$$\frac{\partial l_f}{\partial w_m} = -\frac{1}{k_f} \left\{ \sum_i \frac{\partial y_{if}}{\partial p_m} \frac{\partial p_m}{\partial w_m} + \sum_i \frac{\partial y_{if}}{\partial s_i} \frac{\partial s_i}{\partial w_m} \right\}. \quad (3.46)$$

There is again the mix of productivity, preference and distributional effects. A change in the male wage rate changes the implicit price of the male household good and therefore the demand for the female household good, in a way that will depend on whether they are substitutes or complements. Both productivities are relevant: hers determines how the change in demand for her good translates into a change in labour supply, his determines the effect of the change in his wage on the price of his household good. Finally, the change in income shares determined by his wage rate change determines, in conjunction with the income demand derivatives, the induced effect on demand for her good.

We can again express this in elasticity form, as

$$\varepsilon_{fm} = \frac{t_f}{l_f} \left(\eta_{fm} - \sum_{i=f,m} v_{if} \omega_{im} \right) \quad (3.47)$$

where:

$$\varepsilon_{fm} = \frac{w_m}{l_f} \frac{\partial l_f}{\partial w_m} \quad (3.48)$$

is the cross-wage labour supply elasticity;

$$\eta_{fm} = - \sum_i \frac{\partial y_{if}}{\partial p_m} \frac{p_m}{y_f} \quad (3.49)$$

is the cross-price elasticity of demand; and

$$\omega_{im} = \frac{\partial s_i}{\partial w_m} \frac{w_m}{s_i} \quad (3.50)$$

is the elasticity of i 's income share with respect to a change in m 's wage.

Analogous expressions can be derived from model 2. Note that we obtain a much richer theory of the determinants of female labour supply than that it is simply dependent on her preferences for leisure v. consumption.

Female labour market participation

The preceding analysis assumed that the household was at an interior solution, with the optimal values of all its choice variables, in particular both market labour supplies l_i , strictly positive. However, in many households the secondary earner, usually female, supplies no market labour, and we now want to examine this corner solution. To do this we take model 2, and simplify the notation by omitting non-wage income. Since we shall not be carrying out comparative statics, we formulate the household problem

simply as that of finding a Pareto-efficient allocation³²

$$\max u(x_f, y_f, z_f) \quad \text{s.t.} \quad u(x_m, y_m, z_m) = u_m^0 \quad (3.51)$$

$$(\lambda) \quad \sum_{i=f,m} x_i + qb = \sum_{i=f,m} w_i l_i \quad (3.52)$$

$$(\rho) \quad \sum_{i=f,m} y_i = g(b, t_f, t_m; k) \quad (3.53)$$

$$(\tau_i) \quad l_i + t_i + z_i = 1 \quad i = f, m \quad (3.54)$$

$$l_f \geq 0 \quad (3.55)$$

where we impose the non-negativity condition only on l_f because only there might it be binding, by assumption. The Lagrange multipliers are shown in front of the corresponding constraints. The first-order conditions include

$$\frac{\partial u}{\partial x_f} = \lambda \quad (3.56)$$

$$\frac{\partial u}{\partial y_f} = \rho \quad (3.57)$$

$$\frac{\partial u}{\partial z_f} = \tau_f \quad (3.58)$$

$$\rho \frac{\partial g}{\partial t_f} = \tau_f \quad (3.59)$$

$$\lambda w_f \leq \tau_f \quad l_f \geq 0 \quad (\lambda w_f - \tau_f) l_f = 0. \quad (3.60)$$

Take the case where $l_f = 0$ at the optimum – f supplies no time to the market. In a general sense, the reason for this is that preferences, productivities, wage rates, and the distribution of utilities within the household are all such that the total demand for the household good, together with her leisure demand, lead to a total demand for her time that exhausts her time constraint.³³ It is interesting, however, to consider the conditions at the margin. Note that we can define the implicit price of the household good as

$$p = \frac{\rho}{\lambda}. \quad (3.61)$$

Since ρ has the dimension of u_f/y , while λ has the dimension of u_f/x , p has the dimension of x/y , or units of the numeraire x per unit of y , as it should be. Then using (3.59) and (3.60) gives

$$w_f \leq p \frac{\partial g}{\partial t_f}. \quad (3.62)$$

³² The resulting equilibrium conditions will be common to all HWFs with the Pareto property. However, as we explain at some length below, to carry out a comparative statics analysis requires an HWF, since the assumption of Pareto efficiency alone is insufficient for this purpose.

³³ Note that in any model in which her time is divided only between market work and leisure, this corner solution would imply that she consumes all her time as leisure, and finances her consumption of goods by a transfer from her husband! Quite a privileged, if parasitic, existence.

In words, at a corner solution, when f supplies no time to the labour market, the value of her marginal product in household production is at least as great as her market wage. The intuition is clear: diverting a unit of her time at the margin from household to market production earns more of the market good, but this does not compensate for the loss of the household good. It follows that the market wage places a lower bound on, but may considerably understate, the value of her time in household production. Moreover, since f 's time is allocated between household production and leisure, we have that

$$p \frac{\partial g}{\partial t_f} = \frac{\tau_f}{\lambda} = \frac{\partial u}{\partial z_f} \bigg/ \frac{\partial u}{\partial x_f} \geq w_f. \quad (3.63)$$

Thus the marginal value of leisure time may well exceed the wage rate, because the opportunity cost of leisure is the value of domestic output forgone. Finally, it is worth emphasising that the values of these shadow prices depend on the entire set of first-order conditions, including those relating to m 's choice variables, and the constraints, in particular that on the household utility distribution.

3.2.3.2 Children and child care We would argue that in formulating models that explicitly include children, we should treat them as individuals with their own utility functions, rather than as 'public goods' consumed by their parents. This approach seems to be far more consistent with the basic idea underlying the formulation of household models. In the spirit of Samuelson, we leave open the question of the extent to which children participate in the decision process of the household.

It can be argued that, in the absence of children, household production takes on far less significance as an economic activity than when they are present. The arrival of children generates a very large demand for child care, as well as for other forms of household production, and it is this that makes inclusion of household production in the model absolutely essential. This in turn has important implications for female labour supply and household income, and indeed the entire economics of the household.³⁴

As in the two-adult household considered in the previous sections, the household with children can be modelled as if it were a small economy, and the standard results of general equilibrium theory can be applied in a straightforward way. The household buys a vector x of market goods which it consumes directly, a vector b of market goods which it uses as intermediate goods in domestic production, including possibly child care, a vector g of household public goods, and it produces a vector y of household private goods. We identify domestic child care outputs explicitly as c_k , $k = 1, \dots, K$ where $K \geq 1$ is the number of children in the household and k denotes an individual child. The household technology is assumed subject to non-increasing returns and no

³⁴ See the discussion of the life cycle in chapter 5 below, where this view is supported by the data.

joint production, and is given by the twice-differentiable, concave production functions

$$y_j = h^j(t_f^j, t_m^j, b^j) \quad j = 1, \dots, n \quad (3.64)$$

$$c_k = c^k(t_f^k, t_m^k, b^k) \quad k = 1, \dots, K \quad (3.65)$$

where t_i^j is i 's time input into production of household good j and t_i^k is the time i spends on care of child k . $i = f, m$ are the two adults and only they supply time to household production.³⁵ We have of course $b = \sum_j b^j + \sum_k b^k$.

The utility functions of adults are defined on their consumptions of the vector of market goods, x^i , the vector of household private goods, y^i , pure leisure (a scalar), z_i , and the vector of household public goods, g :

$$u_i = u^i(x^i, y^i, z_i, g) \quad i = f, m. \quad (3.66)$$

The utility functions of children are defined on market goods, x^k , household private goods, y^k , child care outputs, c_k , and household public goods, g :

$$u_k = u^k(x^k, y^k, c_k, g) \quad k = 1, \dots, K. \quad (3.67)$$

All utility functions are strictly quasiconcave and increasing, and at least twice continuously differentiable.

Each adult divides his or her time between general household production, child care, market labour supply l_i and pure leisure. Thus we have the time constraints:

$$t_i + l_i + z_i = 1 \quad i = f, m \quad (3.68)$$

where $t_i = \sum_j t_i^j + \sum_k t_i^k$ is the total time i spends in household goods production and child care. It is assumed that all of a child's time is pure leisure.³⁶

To complete the model we specify the household budget constraint. Corresponding to the vector x is the given price vector p , to the vector b the given price vector q , and to the vector g the given price vector r . Some market goods may of course be both final consumption and intermediate goods, in which case the corresponding elements of p and q will be identical. Given the market wage rates w_i and non-wage incomes μ_i , the budget constraint is

$$px + qb + rg \leq \sum_{i=f,m} (w_i l_i + \mu_i) \quad (3.69)$$

which, eliminating the time constraints, can be equivalently written as

$$px + qb + rg + \sum_{i=f,m} w_i (t_i + z_i) \leq \sum_{i=f,m} (w_i + \mu_i). \quad (3.70)$$

³⁵ If domestic help, including child care, is bought on the market then this is one of the market goods with the corresponding wage rate as its price. The essential assumption is that children do not contribute to domestic production.

³⁶ In both developing and developed countries children as usually defined, i.e. persons under the age of 16, may in fact supply both market and household labour. This can be handled formally by increasing the number of 'adults' in the household.

The household acts as if it maximises a household welfare function:

$$H = H(u_f, u_m, u_1, \dots, u_K)$$

subject to the constraints on the domestic technology and expenditure given in (3.64), (3.65) and (3.70). Given our assumptions, the first-order conditions for this problem are both necessary and sufficient for an optimum and are thoroughly familiar from standard general equilibrium theory.³⁷ Assuming all choice variables are strictly positive at the optimum we have:

- marginal rates of substitution between any pair of goods are equalised across household members;
- inputs are allocated so as to equalise their marginal value products in all uses;
- where both adults supply market labour, i 's market wage measures the marginal value product of time i spends on any household production or child care activity and also the marginal value of leisure;³⁸
- the Samuelson conditions characterise the consumption of household public goods, in particular the optimal consumption of such a good equates its price to a weighted sum of its marginal utilities to the household members, where the weights are the values of the partial derivatives H_i^* , H_k^* at the optimal solution;³⁹
- the amount of care given to child k equates the marginal cost of this care to its weighted marginal value product, where the weight is again the equilibrium value H_k^* .

We can solve for the vectors of demands, outputs of household goods and child care, and labour supplies as functions of prices, wage rates and non-wage incomes. All these functions are in principle observable, though existing datasets provide only very incomplete information that does not allow them to be estimated at the level of generality of this model. Apps and Rees (2002), discussed in section 4.3.2 of chapter 4, provide an example of the sorts of assumptions that have to be made to estimate this model with existing datasets.

An objection to this approach might be: what is the point of having a model that cannot be estimated? Our reply would be that model-building should not be constrained by what happen to be the currently available datasets.⁴⁰ Model-building unconstrained by data availability should lead to an awareness of the kinds of data that need to be collected. It also helps us to see the way in which empirical models make assumptions that amount to making up the missing data.

³⁷ As long as we are only concerned with characterising the household equilibrium, we could just as well have assumed the household simply finds a Pareto-efficient equilibrium, rather than maximising an HWF. Comparative statics analysis, however, requires an HWF.

³⁸ Where one adult supplies no market labour, the marginal value product of his/her time in household production is equal to the marginal value of leisure and is equal to or greater than the market wage.

³⁹ That is, an individual's weight in the household resource allocation, as measured by this derivative, determines the extent to which his or her preferences for the household public good influence its total output.

⁴⁰ For example, macroeconomics would not exist if that principle had been followed. National income accounting came after Keynes.

This model helps to clarify exactly what we might mean by ‘the costs of children’. The cost of a child is simply the value of the bundle of goods – market, household and public – plus the cost of child care, that the child receives at the equilibrium household allocation, i.e. it is the value of the resources the household chooses to allocate to her at its equilibrium.

On the given assumptions, there can be associated with the equilibrium allocation a set of implicit prices for the domestically produced goods, including child care, given by their marginal costs at the equilibrium. Let the vector of implicit prices of household private goods be denoted by π^* and let τ_k^* denote the equilibrium household price of child care for the k ’th child. Then the *full consumption cost of child k* is given by

$$C_k^* = px^{k*} + \pi^* y^{k*} + \tau_k^* c_k^* + r[g^* - g_{-k}^*] \quad (3.71)$$

that is, by the cost of the bundle of market, public and domestic private goods and child care that the child consumes at the equilibrium household allocation (denoted by asterisks). Here $r[g^* - g_{-k}^*]$ is the cost associated with the increased demand for household public goods that child k imposes, with g_{-k}^* denoting the vector of household public goods that would optimally be provided in the absence of child k .⁴¹ Our contention then is that it is this measure of the cost of a child that should ideally be estimated in child cost studies. More usually, such studies construct ‘adult equivalence scales’ for children based either upon minimal physiological needs, or estimates of the increase in income required to hold parental utility constant, given their expenditure on consumption goods for the children, i.e. $\sum_k px^{k*}$.

3.2.3.3 Generalising the household welfare function In the next section of this chapter we discuss the Nash bargaining approach in some detail and argue there that the key general idea it contains is that the household’s preference-ordering over the utility profiles of its members depends on their wage rates and non-wage incomes. It is the absence of these in Samuelson’s HWF that leads it to give the anonymity and symmetry results. Returning, for simplicity of notation, to the two-person household model, and writing the HWF introduced in the previous section now⁴² as

$$H = H(u^1(x_1, z_1), u^2(x_2, z_2); w_1, \mu_1, w_2, \mu_2) \quad (3.72)$$

allows us to take over the idea that the household preference ordering depends on these exogenous variables, without constraining us to accept the specific rationale – their

⁴¹ For example, if a shared good such as housing space is a household public good, then its consumption is likely to increase with the number of children. If not, this term is just zero.

⁴² We could also adopt McElroy’s suggestion of including a vector of EEPs that influence this preference-ordering though not the household budget constraint. Examples of such variables typically cited are: indicators of gender discrimination in labour markets; the ratio of marriageable men to women, the so-called sex ratio; taxes and transfers that change with the individual’s marital status; and the nature of divorce laws. These variables will often be specific to the precise context of the analysis. For example, analysis of household decisions in a developing country may include dowries and bride prices among these variables.

embodiment in formal threat points – that makes their presence natural in the Nash bargaining model. We call this the *generalised household welfare function* (GHWF).

We then define the *generalised Samuelson model* by the problem⁴³

$$\max H(u^1(x_1, z_1), u^2(x_2, z_2); w_1, \mu_1, w_2, \mu_2) \quad (3.73)$$

$$\text{s.t.} \quad \sum_i (w_i z_i + x_i) \leq \sum_i (w_i + \mu_i) + t \quad (3.74)$$

where t is thought of as a transfer to the household that is not assigned to any one individual, and does not affect the household's preferences over utility profiles.

It is perhaps pointless to speculate on whether Samuelson would have approved of this step at the time he wrote his paper on social indifference curves. On the one hand, he clearly conceived of social welfare functions in very general terms indeed. Thus in his discussion of social welfare functions in the context of the economy as a whole⁴⁴ he writes:

Without inquiring into its origins, we take as a starting point for our discussion a function of all the economic magnitudes of a system which is supposed to characterize some ethical belief . . . Any possible opinion is permissible . . . We only require that the belief is such as to admit of an unequivocal answer as to whether one configuration of the economic system is “better” or “worse” than any other, or “indifferent”, and that these relationships are transitive . . . numerous individuals find it of interest to specialize the form of [the SWF] . . . For one thing, prices are not usually included in the welfare function itself, except very indirectly through the effects of different prices and wages upon the quantities of consumption, work etc.

Applied to the household, we could certainly accept that the opinion that the household's preferences over utility profiles of its members should depend on their wages and non-wage incomes is a permissible ethical belief, in the sense that we may well expect household members to agree that it should be the case.⁴⁵ Then we can model their objective function in a way that corresponds to this belief. Samuelson's exclusion of prices from the *social* welfare function is no doubt due to the fact that, in the case of the economy as a whole, these are endogenous variables. Indeed, in applications of the social welfare function, for example in optimal tax theory, it is usual to write it as a function of individual utilities alone. However, in the case of the household, wages and non-wage incomes are exogenous, and so an objection on grounds of endogeneity

⁴³ In the general notation of the consumption good model introduced initially, the problem is

$$\begin{aligned} \max_{x^i} & H(u^1(x^1), \dots, u^h(x^h); p, \mu_1, \dots, \mu_h) \\ \text{s.t.} & \quad p \sum_i x^i \leq \sum_i \mu_i + t. \end{aligned}$$

⁴⁴ Samuelson (1947) pp. 221, 222.

⁴⁵ Possibly because they define the values of outside options or ‘threat points’, but other reasons are also possible. For example, they may be taken as an indicator of the contribution an individual makes to the household budget, and therefore of his or her entitlement to a share in household full income.

would not apply. We regard the GHWF as the appropriate extension of the idea of a social welfare function to the case of the household.

We assume the GHWF to possess the Pareto property and quasiconcavity with respect to the utility profile $(u^1(\cdot), u^2(\cdot))$ for any given vector $[w_1, \mu_1, w_2, \mu_2]$, as well as differentiability. It is natural to place the further *generic* restrictions, which hold everywhere except possibly at special points in the domain:

$$\mathbf{A1} : H_{w_i}, H_{\mu_i} \neq 0 \quad i = 1, 2.$$

It is not possible to be more specific about the signs of these derivatives, as we will show later by examining the Nash bargaining and collective models as special cases of the GHWF approach. More structure is introduced by assuming:

$$\mathbf{A2} : H_{iw_i}, H_{i\mu_i} \geq 0 \geq H_{iw_j}, H_{i\mu_j}, i, j = 1, 2 \quad i \neq j,$$

where H_{iw_i} ($H_{i\mu_i}$) and H_{iw_j} ($H_{i\mu_j}$) cannot both be zero.

In words: at the margin, the weight the household places on an individual's welfare is always non-decreasing in her own wage and non-wage income, always non-increasing in those of the other individual, and at least one of these derivatives with respect to the wage (resp. to the non-wage income) is non-zero. This guarantees that all of these variables have a non-trivial effect on the equilibrium resource allocation, since they change the household's marginal rates of substitution between individual utilities at all points. Specifically, the conditions imply

$$\frac{\partial}{\partial w_1} \left[\frac{H_1}{H_2} \right], \quad \frac{\partial}{\partial \mu_1} \left[\frac{H_1}{H_2} \right] > 0 \quad (3.75)$$

$$\frac{\partial}{\partial w_2} \left[\frac{H_1}{H_2} \right], \quad \frac{\partial}{\partial \mu_2} \left[\frac{H_1}{H_2} \right] < 0 \quad (3.76)$$

where H_1/H_2 is the household's marginal rate of substitution between 1 and 2's utility in the (u^1, u^2) -plane. The household's indifference curves in this plane at any point steepen with increases in w_1 and μ_1 and flatten with increases in w_2 and μ_2 . These restrictions imply that an increase in individual i 's wage rate or non-wage income always increases i 's *relative* weight in the GHWF.

The GHWF and labour supplies

The restrictions on household labour supply functions implied by Samuelson's HWF are identical to those in a standard individual model. This changes dramatically when we move to a GHWF with the above general properties. It turns out that, as we should expect from Pollak's (1977) analysis of price-dependent preferences,⁴⁶ the Slutsky matrix in the household's labour supply problem cannot be shown to be symmetric and negative semidefinite. In defining the compensated derivatives, 'compensation' of changes in wage rates and non-wage incomes must be understood as holding household

⁴⁶ This saw the reasons for including prices in an individual utility function as either Veblenesque 'snob effects' or the use of prices as indicators of quality.

welfare, and not individual utilities, constant. Here we summarise and provide intuition for the main results, details of which are given in the appendix to this chapter.

Non-anonymity The effect of a change in non-wage income will depend on the identity of the recipient of that income, and a pure income redistribution within the household will change labour supplies.

The simplest way to show this is to note that Samuelson's decentralisation proposition still applies, in that there is an identity between the labour supplies derived by solving the problem in (3.73), (3.74), and those derived by first solving the household's income distribution problem

$$\max_{s_i} H(v^1(w_1, s_1), v^2(w_2, s_2); w_1, \mu_1, w_2, \mu_2) \quad (3.77)$$

$$\text{s.t.} \quad \sum_i s_i \leq \sum_i (w_i + \mu_i) + t \quad (3.78)$$

for the sharing rule functions $s_i(w_1, \mu_1, w_2, \mu_2, t)$, and then solving the individual utility maximisation problems

$$\max_{x_i, z_i} u^i(x_i, z_i) \quad \text{s.t.} \quad x_i + w_i z_i \leq s_i(w_1, \mu_1, w_2, \mu_2, t) \quad (3.79)$$

yielding leisure demands $z_i(w_i, s_i(w_1, \mu_1, w_2, \mu_2, t))$. It then follows that

$$\frac{\partial z_i}{\partial \mu_i} = \frac{\partial z_i}{\partial s_i} \frac{\partial s_i}{\partial \mu_i}; \quad \frac{\partial z_i}{\partial \mu_j} = \frac{\partial z_i}{\partial s_i} \frac{\partial s_i}{\partial \mu_j} \quad i, j = 1, 2 \quad i \neq j \quad (3.80)$$

and standard comparative statics analysis shows that $\partial s_i / \partial \mu_i \neq \partial s_i / \partial \mu_j$ in general (see part B of the appendix).

(Non-)symmetry of compensated labour supply derivatives We now derive necessary and sufficient conditions under which the compensated derivatives of i 's labour supply with respect to j 's wage will be equal, $i, j = 1, 2 \quad i \neq j$, and also show that these conditions do not in general hold in this model. To do so we take the dual of the problem in (3.73), (3.74):

$$\min_{x_i, z_i} t = \sum_i (x_i - w_i(1 - z_i) - \mu_i) \quad (3.81)$$

$$\text{s.t.} \quad H(u^1(x_1, z_1), u^2(x_2, z_2); w_1, \mu_1, w_2, \mu_2) \geq H. \quad (3.82)$$

The solution to this problem yields an expenditure function $t(w_1, \mu_1, w_2, \mu_2, H)$, giving the unassigned income required to reach any given household welfare level. Applying the Envelope theorem gives

$$\frac{\partial t}{\partial w_i} = -(1 - z_i) - \beta H_{w_i} \quad i = 1, 2 \quad (3.83)$$

where β is a Lagrange multiplier. Then by Young's theorem

$$\frac{\partial^2 t}{\partial w_j \partial w_i} = \frac{\partial z_i}{\partial w_j} - \beta H_{w_i w_j} - H_{w_i} \frac{\partial \beta}{\partial w_j} \quad (3.84)$$

$$= \frac{\partial z_j}{\partial w_i} - \beta H_{w_j w_i} - H_{w_j} \frac{\partial \beta}{\partial w_i} \quad (3.85)$$

$$= \frac{\partial^2 t}{\partial w_i \partial w_j} \quad i, j = 1, 2 \quad i \neq j. \quad (3.86)$$

Thus we have that the necessary and sufficient conditions for the compensated leisure demand derivatives to be equal, i.e.

$$\frac{\partial z_1}{\partial w_2} - \frac{\partial z_2}{\partial w_1} = H_{w_1} \frac{\partial \beta}{\partial w_2} - H_{w_2} \frac{\partial \beta}{\partial w_1} = 0 \quad (3.87)$$

are:

$$(i) \quad H_{w_i} \equiv 0 \quad i = 1, 2 \quad (3.88)$$

or

$$(ii) \quad H_{w_i} \neq 0 \quad \text{and} \quad \frac{H_{w_j}}{H_{w_i}} = \frac{\partial \beta / \partial w_j}{\partial \beta / \partial w_i} \quad i, j = 1, 2 \quad i \neq j. \quad (3.89)$$

(i) is of course ruled out by assumption, while we show in part A of the appendix that (ii) also does not in general hold.

Own-wage effects on compensated leisure demands/labour supplies We can show in the usual way that the expenditure function $t(\cdot)$ has the usual property of concavity in (w_1, μ_1, w_2, μ_2) , and so, using (3.83) we have

$$\frac{\partial^2 t}{\partial w_i^2} = \frac{\partial z_i}{\partial w_i} - \beta H_{w_i w_i} - H_{w_i} \frac{\partial \beta}{\partial w_i} \leq 0. \quad (3.90)$$

However, the sign of $(\beta H_{w_i w_i} + H_{w_i} \partial \beta / \partial w_i)$ is in general indeterminate, and so we cannot conclude that $\partial z_i / \partial w_i < 0$. The intuition is as follows. When w_i increases, i 's leisure becomes more expensive for the household, and so the usual substitution effect would work towards reducing it, but at the same time, her marginal weight in the GHWF increases, implying that she will be given a higher income share. This then increases her demand for leisure (assuming it is normal) and so could increase the household's demand for her leisure overall, even as a compensated effect. Clearly, this latter effect disappears if $H_{w_i} \equiv 0$, as in Samuelson's HWF model.

More positively, we can place some testable restrictions on the compensated derivatives of the sharing rule, the functions $s_i(\cdot)$. We do this by using the problem in (3.81), (3.82) to derive a type of Slutsky equation for these functions. By carrying out the standard comparative statics analysis for this problem we can show:

A compensated increase in i 's non-wage income always increases her share and reduces j 's. Thus we have:

$$\frac{\partial s_i}{\partial \mu_i} = \frac{v_s^i H_{i\mu_i} - v_s^j H_{j\mu_i}}{\Delta} + \frac{\partial s_i}{\partial t} \quad i, j = 1, 2 \quad i \neq j \quad (3.91)$$

$$\frac{\partial s_j}{\partial \mu_i} = \frac{v_s^j H_{j\mu_i} - v_s^i H_{i\mu_i}}{\Delta} + \frac{\partial s_j}{\partial t} \quad i, j = 1, 2 \quad i \neq j \quad (3.92)$$

where $\Delta > 0$ is the bordered Hessian of the Lagrange function. From **A2** we have that the first terms in these two equations are respectively positive and negative. These are the compensated effects in question, since the second term in each expression is an income effect, equivalent to the effect of a change in non-assigned income t on the respective income shares. In general these income effects may be positive or negative. However, call a household *fair* if these income effects are both always positive. Then we have:

In a fair household, an uncompensated increase in i 's non-wage income always increases her share and may increase or reduce j 's The intuition for the latter ambiguity is of course that while the income effect of the increase in μ_i tends to increase s_j , the increase in the marginal weight the household places on i 's utility tends to reduce it. Note that in Samuelson's HWF model both compensated terms are zero and only the income effect remains.

However, the effect of a compensated change in i 's wage cannot be signed unambiguously:

$$\frac{\partial s_i}{\partial w_i} = \frac{v_s^i H_{iw_i} - v_s^j H_{jw_i}}{D} + \frac{H_i v_{sw}^i}{D} + \frac{\partial s_i}{\partial t} \quad i, j = 1, 2 \quad i \neq j \quad (3.93)$$

$$\frac{\partial s_j}{\partial w_i} = \frac{v_s^j H_{jw_i} - v_s^i H_{iw_i}}{D} - \frac{H_i v_{sw}^i}{D} + \frac{\partial s_j}{\partial t} \quad i, j = 1, 2 \quad i \neq j. \quad (3.94)$$

The ambiguity is created by the term $H_i v_{sw}^i$, which also arises in Samuelson's HWF model and has an interesting intuitive explanation. Holding the s_i and the marginal household weights on individual utilities, H_i , constant, an increase say in 1's wage could cause a reduction in the marginal household utility derived from an additional dollar to 1, thus changing the marginal rate of substitution between s_1 and s_2 along a household indifference curve, and leading *ceteris paribus* to an increase in s_2 and reduction in s_1 . Note that Samuelson's problem in (3.2) already has preferences that are effectively price-dependent in the (s_1, s_2) -plane.

In this subsection we have set out a general framework for cooperative household models, based on Samuelson's idea of a household welfare function, but extending it to incorporate the key insight from the Nash bargaining models – the idea that the household's preference-ordering over the utility profiles of its members depends on their wage rates (or prices more generally) and non-wage incomes. Applying what appear to be reasonable general restrictions on the effects of changes in these exogenous variables

allows a straightforward derivation of the general implications of cooperative models, for the case of household labour supplies. By placing the existing specific cooperative models in this broader context, we hope to have suggested the possibility of other approaches to modelling household preferences that depart from the special structure of these specific models.

3.2.4 *Nash bargaining and the collective model*

3.2.4.1 *Nash bargaining* The two papers that introduced Nash bargaining models into the household economics literature illustrate nicely the two main concerns of this literature pointed out earlier in the Introduction. The paper by Manser and Brown (1980) is concerned with analysing the marriage decision, and takes as its point of departure the work by Becker on the theory of marriage.⁴⁷ The paper stresses the importance of modelling the household as consisting of individuals with their own utility functions, and of retaining the elements of conflict as well as cooperation that are involved in the household resource allocation decision. Though the paper is correct in arguing that Becker's formulation of the household objective function, as the utility function of the 'head of the household', suppresses these aspects of household decision-taking, they are wrong in dismissing Samuelson's approach as if it also did this. The HWF retains the separate utility functions of individuals, is quite general about the types of considerations that determine the trade-offs between their utilities – by no means ruling out conflict – and, in the equivalent concept of the sharing rule, provides a flexible and convenient way of analysing the intra-household income distribution.

The second paper, by McElroy and Horney (1981), is on the other hand concerned with the issue Samuelson saw as central: what can be said about the properties of household demands when we recognise that the household consists of more than one individual? The paper shows that by modelling the household resource allocation process as a cooperative Nash bargaining game, well-defined demand functions can be derived on which testable restrictions can be placed, differing in important ways from those derived from the standard demand theory (and therefore from Samuelson's model as set out in section 3.2.1). The paper provides a detailed analysis of the structure of the household demand functions implied by the Nash bargaining hypothesis.

The two papers differ in the details of the formulation of their models. Manser and Brown emphasise the existence of household public goods and love and caring in explaining why marriage generates a surplus of utility over that which is obtained when the two individuals live separately. The existence of this surplus explains why marriage (or more generally formation of a household) takes place, while the analytical problem, solved by Nash bargaining, is to determine how this surplus is shared. In McElroy and Horney's paper, goods may also be private (i.e. consumption by one individual reduces the amount available for the other), love plays no explicit role (though the utility of

⁴⁷ See Becker (1973), (1974).

each individual depends also on the consumptions of the other), and the existence of a surplus from marriage is simply assumed.

The existence of a surplus to be bargained over is one necessary component of a Nash bargaining model; the specification of the ‘threat points’, the utilities that will be achieved by the respective parties if no agreement is reached, is the other. Both papers take the maximum (expected) utility each individual could achieve outside the household in question as the threat point. Since McElroy and Horney are concerned with an ongoing household, this has come to be called a ‘divorce threat point’, though in the case of Manser and Brown, since they are analysing conditions under which a not-yet-existing household will be formed, this is a misnomer. Their threat point is the next best expected utility an individual can obtain on the marriage market.

We formalise the bargaining approach in a way that emphasises its application to labour supply decisions, although this was not in fact the model in either of the two original papers. Let $v^i(w_i, \mu_i)$ denote the highest expected utility that individual $i = 1, 2$ could obtain outside the household in question. These are exogenously given reservation constraints. They are obtained by solving the problem $\max u^i(x_i, z_i)$ s.t. $w_i z_i + x_i \leq w_i + \mu_i$, where z_i is again i ’s leisure consumption and x_i consumption of a household public good⁴⁸ with price normalised at 1. The two-person household is then assumed to solve the problem

$$\max H^N = [u^1(x, z_1) - v^1(w_1, \mu_1)][u^2(x, z_2) - v^2(w_2, \mu_2)] \quad (3.95)$$

$$\text{s.t. } \sum_i w_i z_i + x \leq \sum_i (w_i + \mu_i). \quad (3.96)$$

Note that in this model, as in Samuelson’s, the budget constraint implies income pooling: the possibility of unrestricted lump sum transfers between the individuals. Leisure and consumption demands of the household are now $z_i(w_1, \mu_1, w_2, \mu_2)$, $x(w_1, \mu_1, w_2, \mu_2)$. The household surplus is generated because with the same leisure consumptions each member can have more of the household public good than when living separately. The demand functions have the ‘adding-up’ and zero degree homogeneity⁴⁹ properties of standard Marshallian demand functions, but the properties of anonymity and symmetry of compensated leisure demands discussed in the previous section no longer hold in general, as McElroy and Horney show.

Note that there is a specific form of GHWF in this problem, for, after all, what is the Nash product H^N , the maximand in the Nash bargaining problem, but a representation of the household’s preference-ordering over the individual utilities? As with the Stone–Geary individual utility function, it is a type of Cobb–Douglas function (defined on

⁴⁸ Taking consumption x as a household public rather than private good is the simplest way, in the absence of household production, to ensure a utility surplus from formation of the household. If x here were a private good, it would not be possible for both individuals to be better off in the joint household. Our own preference, however, is to follow Becker and show how such a surplus can be generated by improvements in the production set arising from joint rather than individual household production.

⁴⁹ The indirect utility functions $v^i(\cdot)$ are invariant to equiproportionate changes in all prices and incomes, and these also leave the budget constraint unaffected.

utility pairs rather than goods) on a restricted coordinate space⁵⁰ with origin at (v^1, v^2) rather than $(0, 0)$. It is easy to show⁵¹ that it possesses the general properties of a GHWF mentioned in the previous section, though not for all types of redistribution of non-wage incomes.

Lundberg and Pollak (1993) pointed out that the model has a problem, in that it does predict anonymity with respect to individual non-wage income changes that leave threat points unchanged. Thus suppose, in the event of divorce, individual 1 would receive a state transfer t , whereas, while married, individual 2 receives this transfer. Now consider a policy change that switches payment of t from 2 to 1 also while they are married. This leads to no change in the threat points, and also no change in total income, the right-hand side of the budget constraint, so there would be no effect on the Nash bargaining equilibrium. Against this, Lundberg and Pollak (1993) argued that it might be expected that such a change would increase 1's influence over the household resource allocation, while Lundberg *et al.* (1997) later presented evidence to show that a similar policy change in the UK did in fact cause a change in household consumption patterns in the expected direction. This argument motivated their formulation of the 'separate spheres bargaining model', to which we now turn.

3.2.4.2 Nash bargaining with non-cooperative threat points The idea underlying the separate spheres bargaining model is to base the threat points in the Nash bargaining game on separate individual budget constraints for the partners in the household, in a way that ensures that anonymity will not hold, even under the kind of transfer policy change just discussed.

The first step is to introduce the idea, originally proposed by Ulph (1988), and Woolley (1988), of basing the threat points not on the utilities the partners would receive if they divorced, but rather on the utilities they achieve in a non-cooperative Nash equilibrium within the existing household. There are in any case good arguments against post-divorce utilities as threat points – divorce may simply be too drastic and costly a fall-back point in the event of failure to agree on division of the utility surplus generated by the two-person household, and so may not be a credible threat. But a Nash bargaining model must have threat points, and so an alternative is to base them on a non-cooperative equilibrium within an ongoing household.

The next step is to characterise these threat points. How would people behave if, while remaining within the household, they do not cooperate to achieve the full benefits that are feasible, because they cannot agree on how to share them? Suppose now that there are two household public goods, x_1 and x_2 , so that the individual utility functions are $u^i(x_1, x_2, z_i)$, $i = 1, 2$. In a cooperative equilibrium, the individuals will jointly agree on how much of these goods to buy and consume.⁵² In a non-cooperative game with

⁵⁰ See also Lundberg and Pollak (1993), p. 995. ⁵¹ See the appendix.

⁵² Note that there is no household production in this model, as is the case with all the bargaining models considered so far. Ott (1992) appears to have been the first to introduce household production into the Nash bargaining model.

Nash equilibrium as the solution concept (sometimes called Cournot–Nash equilibrium to distinguish it more sharply from the Nash *bargaining* equilibrium), the partners will independently choose how much of these to buy to maximise their own individual utilities, taking as given the amount bought by the other. As in standard oligopoly models, this defines a set of reaction functions, the solution of which determines the Nash equilibrium.

The papers on voluntary-contribution public goods games by Warr (1983) and Bergstrom *et al.* (1987) show that two types of Nash equilibrium are possible in such games. Both players may choose to contribute positive amounts of each public good, in which case a key result is that a small transfer of income between them will leave the supplies of the public goods and the individual utilities of the players unaffected. In that case, therefore, a Nash bargaining model with its threat points defined by this non-cooperative equilibrium would, as pointed out earlier, actually predict anonymity with respect to the kinds of policy change Lundberg and Pollak are concerned with, for exactly the same reason as before: such a transfer changes neither the threat points nor the budget constraint.

Alternatively, there may be a non-cooperative equilibrium at a corner solution, where each of the players supplies nothing of one of the public goods, free-riding on the expenditure made by the other. In this type of equilibrium, a small transfer between the players does affect the equilibrium, increasing or decreasing the supply of a public good according to whether the transfer is away from or towards the free-rider for that good.

As the final step Lundberg and Pollak argue that households are characterised by a division of responsibilities based on ‘socially recognised and sanctioned gender roles’, under which each partner specialises in buying just one of the public goods, so that we *necessarily* have the corner solution case in the non-cooperative equilibrium. In this equilibrium, assuming partner $i = 1, 2$ buys only the one public good x_i , each finds his or her equilibrium consumptions x_i^*, z_i^* by solving the problem

$$\max_{x_i, z_i} u^i(x_i, x_j^*, z_i) \quad \text{s.t.} \quad w_i z_i + p_i x_i \leq w_i + \mu_i \quad i, j = 1, 2, \quad i \neq j \quad (3.97)$$

where x_j^* is the equilibrium supply of the other’s public good, and p_i is the price of x_i . Note that in this non-cooperative game there is no income pooling: the partners are constrained in their choices by their own incomes. Then the threat points in the associated Nash bargaining game are given by the indirect utilities

$$v_i^* = u^i(x_i^*, x_j^*, z_i^*) = v^i(w_1, w_2, p_1, p_2, \mu_1, \mu_2). \quad (3.98)$$

In general, each threat point is a function of both partners’ wage rates and non-wage incomes, since the non-cooperative Nash equilibrium values x_i^*, z_i^* are determined by all these variables. The cooperative Nash bargaining solution is then found by solving

$$\max_{x_1, x_2, z_1, z_2} N^S = [u^1(x_1, x_2, z_1) - v_1^*][u^2(x_1, x_2, z_2) - v_2^*] \quad (3.99)$$

$$\text{s.t.} \quad \sum_i (w_i z_i + p_i x_i) \leq \sum_i (w_i + \mu_i). \quad (3.100)$$

Thus changes in either one of the μ_i affect both the threat points by changing the Nash equilibrium of the non-cooperative game, and anonymity will not hold, even though the budget constraint in the cooperative game assumes pooling. Note that the solution of the Nash bargaining game is Pareto-efficient, since the GHWF N^S possesses the Pareto property, and unrestricted transfers are permitted by the pooled budget constraint. On the other hand, the threat points are not Pareto-efficient, because in that equilibrium each partner takes no account of the utility the other derives from his or her supply of the public good when choosing how much of that good to buy.

The exclusion of household production in this model is what in fact makes it necessary to insist on a corner solution for the non-cooperative equilibrium, and this implies assigning a very powerful role to ‘socially recognised and sanctioned gender roles’. Can these, however, really be so powerful that an individual cannot *buy* both of the public goods on the market, for example increasing his or her consumption of the other’s public good if not enough of it is being provided in the non-cooperative equilibrium? This of course undoes the model, since in equilibrium both contributions to each public good will be positive. Examples of gender specialisation in household provision that come to mind, for example with females supplying child care and males house maintenance, relate to specialisation in household production activities rather than expenditures. Even here it is hard to believe that only complete specialisation is feasible or permissible.

Konrad and Lommerud (1995) show that the required non-neutrality of pure transfers between the individuals in a non-cooperative Nash equilibrium can be far more plausibly rationalised in a model where individuals differ in their productivities in household production, so that specialisation is based on comparative advantage rather than socially sanctioned gender roles.⁵³ Thus they show that it is not necessary to have a corner solution in the non-cooperative game in order to obtain non-anonymity with respect to transfers.

Furthermore, it is of course possible that one of the spouses supplies no time to the labour market, which after all is the case in a large proportion of households. In that case, the non-cooperative equilibrium in the separate spheres model involves a zero supply of her public good (absent non-wage income) which strengthens the argument against a corner solution in this equilibrium. Moreover, in the final Nash bargaining equilibrium she can only buy this public good with a transfer from her spouse. Thus, although we agree that non-cooperation rather than divorce on the whole provides a more plausible rationalisation of threat points, and succeeds in extending the scope of the non-anonymity result, the model certainly gains in plausibility from being based on separate spheres in household production rather than simply in expenditures.

The model of Nash bargaining with non-cooperative Nash equilibrium threat points in Chen and Woolley (2001) does not make the separate spheres assumption, and allows the possibility that both partners buy positive amounts of the public good in the threat point equilibrium. However, in this model the overall Nash equilibrium solution is still

⁵³ Though the reasons for these differences in productivities might be related to conventional modes of upbringing.

always sensitive to the non-wage incomes⁵⁴ of the partners, essentially because the bargaining is assumed to be over transfers of income, rather than over the allocations of expenditures to private and public consumption goods. The partners take their Nash bargained equilibrium income shares and then individually maximise their own utility in choosing their own consumption and the amount of the household public good they buy. In terms of Samuelson's theorem, bargaining is over the sharing rule. The result of this is that, although their Nash bargaining GHWF possesses the Pareto property, the final household consumption allocation is *not* Pareto-efficient, and this is perhaps the most striking result of the model. The reason is that in spending their incomes individually rationally, they choose their expenditures on the household public good non-cooperatively.

The question then of course arises: why do the partners not perceive that they could both be better off if they bargained directly over expenditure allocations, rather than subjecting themselves to the two-stage process of a cooperative income allocation and then a non-cooperative expenditure allocation? Second-best allocations always arise out of some additional constraint on the first-best allocation process. In this case, this is the constraint that bargaining must be over income shares and not consumption choices. It is hard to see, however, why rational households would subject themselves to such a constraint, when it leads to Pareto-inferior outcomes. They are after all free to choose what they bargain about.

In conclusion: bargaining models impose a very specific and not especially convenient structure on the GHWF. Amongst the proponents of the approach, there seems to be no consensus on a satisfactory specification of the threat points, while the results of the models are sensitive to these. Moreover, the 'ethical basis' for Nash bargaining, defined by its four axioms – the Pareto property, invariance to positive linear transformations of the utility functions, independence of irrelevant alternatives, and symmetry⁵⁵ – are not as a group particularly compelling as a representation of family decision-taking, however intuitively appealing the idea is that some kind of explicit or implicit negotiation takes place within households. For example, there is no place for interpersonal utility comparisons in Nash bargaining, yet casual observation suggests that these are extremely prevalent in families. The requirement that just two individuals determine the household allocation is restrictive, and no justification is given for assuming that the bargaining game is cooperative, i.e. that binding commitments are possible. We believe that these qualifications leave a lot of room for alternative approaches to the formal modelling of household decision-taking.

3.2.4.3 The collective model The history of the collective model begins with something of a puzzle. In introducing the model, Chiappori (1988) makes the apparently unequivocal claim:

⁵⁴ There is no labour supply in their model, all income is in any case exogenous.

⁵⁵ This says that in a symmetric game both players should receive the same utilities in the bargaining solution. A game is symmetric if in the (u_1, u_2) -coordinate plane the feasible set of utility pairs is symmetric about the 45° line and the threat point utilities lie on this line.

[I]t tries to derive falsifiable conditions upon household behavior from a “collective rationality” concept; however, instead of referring to some definite bargaining concept, it only makes a very weak and general assumption – namely that the household always reaches Pareto-efficient agreements. The question which is investigated through the paper is thus the following: does Pareto efficiency alone imply restrictions upon observable household behavior?

Similarly, Chiappori (1992) states:

Indeed, it is only assumed that agents are either egoistic or “caring” in the Beckerian sense and that internal decision processes are cooperative, in the sense that they systematically lead to Pareto-efficient outcomes.

Now it is straightforward to characterize a given household equilibrium allocation by Pareto efficiency conditions alone, and, by the second theorem of welfare economics, to interpret this as being generated as if the household first shared its aggregate income among its members, who then maximise their individual utilities. However, the first point made by Samuelson (1956) was precisely to insist that Pareto efficiency alone is not enough to enable restrictions to be placed on the results of a comparative statics analysis of this equilibrium.⁵⁶ Simply knowing that, following some change in prices or income, the new equilibrium is Pareto-efficient does not allow us to say where this will be in the Pareto-efficient set in relation to the previous equilibrium. Restrictions on demand functions expressing the results of this comparative statics analysis therefore cannot be derived from Pareto efficiency alone.

The puzzle is resolved later on in Chiappori (1988) and (1992), where it becomes clear that he is actually assuming the existence of a household sharing rule or, equivalently, of *some* GHWF. This becomes most explicit in Browning and Chiappori (1998), where the ‘household utility function’ or, in the present terminology, the GHWF, is explicitly written⁵⁷ as

$$H^C = \alpha(w_1, w_2, \mu)u^1(x_1, z_1) + [1 - \alpha(w_1, w_2, \mu)]u^2(x_2, z_2). \quad (3.101)$$

Thus the collective model is characterised by a weighted utilitarian household welfare function, the weights varying with the wage rates and total household income. The phrase ‘assuming Pareto efficiency alone’ can be interpreted as being intended to emphasise that the model does not adopt a specific Nash bargaining formulation, but abstracts, in the spirit of Samuelson, from explicit consideration of the process by which the household agreement is reached. It cannot mean, however, that no HWF is being assumed.

The well-known property of a weighted utilitarian SWF is that it possesses no inequality aversion: all utility pairs that yield the same total are equally good. The fact that the model developed as a reaction against Nash bargaining may explain why the choice of a

⁵⁶ This was also noted in Apps and Rees (1988).

⁵⁷ In the notation of the model we are presently using. In Chiappori (1988) and (1992) the subject of the analysis was labour supplies, and so the GHWF would be written as here. In Browning and Chiappori (1998) the analysis was of consumption expenditures and the function is written as: $\alpha(p, \mu)u^1(x^1, x^2, X) + [1 - \alpha(p, \mu)]u^2(x^1, x^2, X)$ where p is a price vector, x^i are consumption vectors and X is a household public good.

GHWF with zero inequality aversion is not discussed in the literature on the collective model, even though it implies a substantive restriction on the distributional preferences of the household.⁵⁸ It follows the Nash bargaining models in generalising the HWF to contain as arguments prices and income, and possibly also EEPs, though it does not distinguish between the individual non-wage incomes.⁵⁹

The central issue⁶⁰ raised in Chiappori (1988) and (1992) is that of whether the household's sharing rule $s_i(w_1, w_2, \mu)$ can be identified empirically by using labour supply functions of individual household members estimated on typically available datasets. The general answer is that it cannot, but there are special cases in which its *partial derivatives* can be derived from estimated labour supply functions. Thus in Chiappori (1988) and (1992) the household consists of only two individuals, each of whom supplies labour to the market, and there are neither children nor household production. Chiappori shows that in this model it is possible to identify the partial derivatives of the sharing rule, essentially because the only way in which the household non-wage income and the other partner's wage rate enter the labour supply function of each individual is through this sharing rule. Estimates of the coefficients of the partner's wage and of household non-wage income from empirical labour supply functions then provide the basis for identification of the partial derivatives of the sharing rule. Necessary for this identification, however, is that we have as many labour supply functions as individuals with shares in the household income. If there are (non-working) children who of course receive shares in household consumption, there are not enough parameter estimates and the model is underidentified.

Chiappori (1992) places some emphasis on the importance of identification of the partial derivatives of the sharing rule for the analysis of public policy, for example income taxation. In any optimal tax or tax reform analysis based on a household model, the derivatives of individual income shares with respect to net-of-tax wage rates appear prominently in the conditions determining tax rates.⁶¹ However, knowledge of these derivatives alone is not in general sufficient to allow solution of the conditions for the tax rates. Knowing by how much an individual's income share would change does not help us to decide on the desirability of such a change in the absence of information about the *levels* of the income shares. Marginal social utilities of income depend on levels of income as well as the income share derivatives. For example, the fact that 55 cents of a one dollar increase in income would flow to 1 and 45 cents to 2 cannot be judged to be a good or bad thing unless we know how well off relative to each other 1 and 2 are in the first place. Moreover, if, because of data limitations, it really is the case that only aggregate household non-wage income is observable, while income shares

⁵⁸ In the social choice literature there is extensive discussion of the arguments for and against utilitarianism as a basis for the social welfare function.

⁵⁹ Strictly speaking, therefore, the model does not make any predictions about anonymity with respect to individual non-wage income changes, since it does not identify these. Chiappori (1988) explains the choice of total non-wage income on the grounds that typical datasets do not provide information on individual non-wage incomes.

⁶⁰ The demand restrictions implied by the Browning–Chiappori model are discussed in the next chapter.

⁶¹ See chapter 7 for a discussion of this.

are functions of individual non-wage incomes, then we have no way of retrieving the derivatives of the sharing rule with respect to individual non-wage incomes.

However, as Apps and Rees (1997b) argue, and Chiappori (1997) acknowledges, the major limitation to the applicability of the result arises out of the fact that the model excludes household production. In the presence of household production, it is no longer sufficient to have empirical estimates of the labour supply functions, since these do not yield the leisure demand functions. Non-market time is divided between household production and leisure. It is therefore also necessary to have estimates of the individual demand functions for time spent in household production, together with data on domestic output. Even when these are available, further restrictions are required in order to identify the derivatives of the sharing rule along the lines suggested by Chiappori. One possibility is to assume constant returns to scale in household production. Apps and Rees show that this assumption is necessary to reduce the number of unknown partial derivatives to equal the number of equations assumed to be empirically estimated.⁶²

When data on domestic output are missing, an alternative assumption that has the effect of constructing the missing data is necessary. One possibility is the assumption that household goods have perfect market substitutes, as in the farm household production model mentioned in chapter 2 (though note that in the empirical work on this model farm output data are also collected). In this case the price of the good produced by the household is exogenously given and observable. Chiappori shows that in this case his earlier results on retrievability continue to hold.

In the more general case where the household goods do not have perfect market substitutes and output data are unavailable, even with the constant returns to scale assumption the sharing rule can be identified only up to an additive function of wages, i.e. not even its partial derivatives can be identified, as confirmed by Chiappori. The problem is that in this case the implicit (endogenous) price of the domestic good is identified only up to a multiplicative constant. In other words it could be anything (positive). This supports the basic contention of the Apps and Rees paper: empirical applications of the sharing rule approach simply have to have more comprehensive datasets.

3.3 Non-cooperative models

The distinction between cooperative and non-cooperative games hinges on the ability to make binding commitments to implement an agreed set of actions. Thus, in Nash bargaining, the parties can somehow commit themselves to implement whatever actions produce the utility pair given by the Nash bargaining solution. This possibility of making binding commitments, for example in the form of legally enforceable contracts, is

⁶² This is shown in proposition 4 of Apps and Rees (1997) and proposition 2 of Chiappori (1997). Somewhat puzzlingly, Chiappori claims that the latter 'generalises proposition 4 in Apps and Rees since we do not need to assume that the production function is linear' (p. 199). This must be based on a misreading, because proposition 4 in Apps and Rees clearly states that the condition is linear *homogeneity* of the production function.

exogenously given. In non-cooperative games no such exogenously given commitment possibilities exist, and although equilibria that are 'cooperative', in the sense of being Pareto-efficient, may be possible, they must be supported by the self-interest of the players, and so their existence is endogenous to the game being played.

It seems clear that, descriptively speaking, household decision-taking is better characterised as non-cooperative in this sense. There may be formal laws, social norms and customs that constrain individual actions within the household, but it seems fanciful to suggest that these amount to a mechanism for making complete, binding commitments. There are three possible responses to this observation.

The first is to try to rescue the Nash bargaining approach by appealing to some results in the theory of non-cooperative bargaining games.⁶³ Thus as Binmore *et al.* (1986) show, in an infinitely repeated non-cooperative bargaining game where two players alternately make offers at fixed intervals of time, as the length of this time interval goes to zero, the solution of the game converges to that of the corresponding cooperative Nash bargaining game. Though this is a beautiful result, it does not in our view provide a realistic basis for adopting the cooperative bargaining approach to household decision-taking.

The second response is to take the non-cooperative approach seriously and to model household decision-taking as a game in which the solution concept is Nash equilibrium or some refinement of it. Here there are two approaches in the literature. One, already discussed in the previous section, is to draw on the theory of voluntary-contribution public goods games, modelling household consumption as a public good. The main general limitation of this is that it gives too prominent a place to household public goods in the household's consumption decisions. The second approach, which is designed for the analysis of labour supply decisions, is to derive best-response or reaction functions in these variables and solve for the Nash equilibrium. This approach was introduced by Leuthold (1968), and further developed by Ashworth and Ulph (1981), though neither paper points out explicitly that it is finding a non-cooperative Nash equilibrium of the household. A comprehensive and insightful discussion of both the theory and econometrics of these kinds of models is provided by Kooreman and Kapteyn (1990).

Leuthold takes a household consisting of two individuals with utilities defined only on total household consumption of a market good and own leisure consumption, and with a pooled budget constraint. Each independently maximises his/her own utility function, assumed to have the Stone–Geary form, subject to this budget constraint, but with the labour supply of the other, which enters into this constraint via the wage income variable, taken as given. This is how the strategic interdependence between the players in this game comes about. The relationships that emerge from this maximisation are the individual reaction functions. The household equilibrium is then found by solving the simultaneous equations defined by these functions, giving the Nash equilibrium.

⁶³ The proponents of the Nash bargaining models discussed above saw no reason for such a justification. They took the ability to commit simply as given, and saw the bargaining approach as appropriately capturing the elements of conflict as well as cooperation in household decision-taking.

This solution is then the basis for the comparative statics analysis of the model and its empirical estimation.

Ashworth and Ulph allow utilities to depend also on the partner's leisure consumption, the case of 'caring preferences', and adopt a flexible functional form for the utility function, but given the data that are available in the typical datasets, including theirs, the cases of 'selfish' and 'caring' preferences are observationally equivalent. They explicitly test and reject the hypothesis that the parameters of the two utility functions are identical, which they interpret as being equivalent to rejection of the model that treats the household as if it were a single consumer. Neither paper explores the question of the allocation of total consumption between the two individuals.

Kapteyn and Kooreman argue that an important and questionable aspect of this Nash equilibrium solution is that it is not Pareto-efficient. Each individual could be made strictly better off by moving from this equilibrium to a point on the Pareto frontier, but such a move cannot be supported as an equilibrium in this one-shot non-cooperative game. Thus Kooreman and Kapteyn follow Ulph (1988) and Woolley (1988) in suggesting the equilibrium as the basis for threat points or reservation values, though whereas the latter two would embed them in a Nash bargaining game, Kooreman and Kapteyn impose them as constraints on the problem of maximising a Samuelsonian HWF. The difference between the two approaches is that whereas proposing Nash bargaining implies assuming a specific cooperative game without a justification for the possibility of binding agreements, Samuelson's approach leaves the process by which the HWF is derived entirely open. Kooreman and Kapteyn go on to show that, given the datasets typically available for estimation of these types of models, they will in general be under-identified, and illustrate with a specific example how more data can improve identification.

The third and most satisfactory response to the observation that household members are engaged in a non-cooperative game is to recognise that this is in fact a repeated game, rather than the one-shot game that has been modelled in the literature discussed so far. This opens up the possibility of application of a rich body of game-theoretic literature, which so far seems only to have been exploited by Lundberg and Pollak (1994) and Basu (2006).

The intuitive idea that rational households ought to be able to do better than the one-shot non-cooperative Nash equilibrium receives support in the theory of repeated non-cooperative games, where it is shown how 'cooperative', i.e. Pareto-efficient, equilibria can be supported by threats of punishment for deviation from them by any individual player. Since a potential deviant will weigh up the cost of future punishment against the immediate gain from deviation, the future must not be too heavily discounted for such threats to work, which we will assume to be the case in the following discussion. The simplest form of punishment following deviation from the agreed equilibrium would be reversion to the one-shot non-cooperative equilibrium for the rest of the game,⁶⁴ but

⁶⁴ See Friedman (1977) who has extensively analysed such strategies. Note that these games, known as supergames, take the form of infinitely repeated plays of the same one-shot game.

more sophisticated ‘carrot and stick’ strategies,⁶⁵ in which punishments harsher than Nash reversion are imposed for a limited period, followed by return to the cooperative equilibrium, can support Pareto-efficient outcomes as subgame perfect equilibria.

There are three important conceptual issues in the analysis of such games, each of which is relevant to the application of these ideas to models of the household. First is the issue of the finiteness of the repeated game. Since the lives of household members are finite, attempting to apply the theory developed for infinitely repeated games to the household runs into the paradox of backward induction. In the last period of the game there is no possibility of supporting cooperative behaviour by the threat of future punishment for a deviation, and so the only equilibrium is the one-shot Nash equilibrium. In the second-to-last period, everyone knows the equilibrium that will be played in the last period and so no threat can support cooperation in this period – again only the one-shot Nash equilibrium can be sustained. This argument can then be applied period by period right back to the first. Thus ‘cooperative’ equilibria are not possible.

There are, however, a number of cases in which this problem can be circumvented. One important case is that in which there is uncertainty about the terminal period of the game: for almost all periods, there is some probability that there will be a future period. Then, cooperation may still be sustained by the threat of punishment in a possible future. Also important is the case in which there is some probability that any given household member is simply of the cooperative type,⁶⁶ because, say, she has internalised social and cultural norms that would make one play the game cooperatively rather than selfishly (individually rationally). If all players are of this type then there is no problem, but even if one is not, she may find it individually rational to behave for at least part of the game *as if* she were, in order to maintain a reputation for being a cooperative type. In both these types of games, we would expect to see the breakdown of cooperative behaviour in the later stages of the game. A third example is the case in which there are multiple Nash equilibria in the one-shot game, which seems perfectly possible for household models in general.⁶⁷ Then the backward induction chain could be broken, for example by supporting cooperation in the next-to-last period by the threat of playing the worst possible Nash equilibrium instead of the best possible Nash equilibrium in the last period. This would work if the discounted value of the difference in pay-offs between the two equilibria exceeds the immediate gain from deviation.

The second issue concerns renegotiation proofness.⁶⁸ Punishment for deviation could be costly to the punisher, and, given that a deviation has occurred, it might appear rational to ‘kiss and make up’, ‘forgive and forget’, and not carry out the self-lacerating punishment. But anticipation of this *ex ante* would then make the cooperative solution unsustainable. Thus attention would have to be restricted to cooperative equilibria that can be supported by punishments that would credibly be carried out, for example

⁶⁵ See Abreu (1986). ⁶⁶ For analysis of such games see Kreps *et al.* (1982).

⁶⁷ Though models using non-cooperative equilibria, such as Chen and Woolley (2001), Leuthold (1968), and Lundberg and Pollak (1993) usually exclude this possibility by using specific functional forms for the utility functions. For the game theory in this case see Benoit and Krishna (1985).

⁶⁸ See, for example, Farrell and Maskin (1989).

because they give the punisher, for the duration of the punishment, a higher pay-off than she obtains at the cooperative equilibrium that is to be supported.

The third issue concerns completeness of information. Information in a game is complete if each player knows at each point in time the actions chosen by all players at all previous points in time. Suppose, however, that the previous choice of action by the other players is not observable by any one player, and, because of some kind of underlying uncertainty, cannot be inferred from the outcomes of the game.⁶⁹ When a particular set of pay-offs is realised, there is a positive probability that it was generated by actions that deviated from the cooperative agreement. The optimal punishment strategies in this case take the form of setting a critical level of pay-off and carrying out the punishment if the actual pay-off deviates from this, where this critical level reflects a choice of probability that it was not due to chance.

As a simple example: suppose that if he studies appropriately hard for a maths exam, your son should obtain an A with probability 0.5, a B with probability 0.4, and a C with probability 0.1. You cannot observe how hard he studies, so promise him a ticket to the next game of his favourite football team if he gets an A, nothing if he gets a B, and that he is grounded for a week and has to work on his maths if he gets a C. You know that if you do this, he will certainly study hard. The slightly paradoxical thing about this is that if he does get a C, you have to carry out the punishment, even though you know that he was simply unlucky and did work hard. If you take him to the football game anyway, since you yourself get pleasure from doing that, he has no incentive to work hard in future, and if he anticipates your doing this, he has no incentive to study hard now.

The theory of finitely repeated non-cooperative games offers a rich set of ideas for application to the economics of the household. It suggests that under some circumstances a household may achieve Pareto-efficient outcomes, though this cannot be taken for granted, and in some contexts it may be interesting and important to analyse this explicitly within the framework of a properly formulated non-cooperative game. We pursue this point in the last section of this chapter.

Finally, it should be noted that, typically, the set of equilibria that may be supported by threat strategies in a repeated game can be quite large, so that there is still a problem of equilibrium selection. One proposal might be that the equilibrium could be selected by Nash bargaining over the set of utility pairs sustainable as equilibria by threats of punishment for deviation, but for reasons given earlier we regard this as too special and constricting. Lundberg and Pollak (1994) see this multiplicity as allowing a role to be played by social and cultural norms, custom and tradition. An obvious way to formalise this is by the GHWF. We can explore the implications of any specific set of assumptions about these norms and customs, as well as those holding for all sets of assumptions that result in a GHWF with the properties set out earlier, by maximising

⁶⁹ For analysis of models of this type in the context of oligopoly see Abreu *et al.* (1986), Green and Porter (1984), Rees (1985) and Rotemberg and Saloner (1986).

this over a set of utility possibilities that can be regarded as supportable as equilibria in an appropriately formulated finitely repeated non-cooperative game.

3.4 Equilibrium models

This class of models sees a household as embedded in a competitive market system, rather than in an explicit cooperative or non-cooperative game. The household's resource allocation is determined by the equilibrium of the market system. An early forerunner was Becker's model of the marriage market which, in its very simplest form, saw the sex ratio – the ratio of men to women in the marriageable age group – as determining entirely who gets the rents from forming a household. The gender in excess demand obtained all the rents, that in excess supply none.⁷⁰ Here, we look briefly at the models in Apps (1981), (1982) and Grossbard (1984).

Apps used a two-sector general equilibrium model to analyse the incidence of income taxation in an economy in which differences in market wage rates associated with non-economic characteristics, in particular gender, are due not to innate differences in productivity,⁷¹ as in the standard optimal tax models, but rather to labour market discrimination, which 'crowds' women into low-wage occupations. In this model, men supply labour in a high-wage sector to produce a market good, and women work in a low-wage sector to produce a second market good, as well as working within the household to produce a domestic good. The implicit price a man pays for this domestic good cannot be less than the opportunity cost of his wife's time, valued at the wage women receive in the low-wage market sector, since women are assumed free to move between market and household work. Women are not, however, allowed access to the high-wage sector. There is a perfectly competitive marriage market so that, in equilibrium, no woman could demand more than the going market price for the household good. There are no transfers or income pooling within the household: each spouse consumes goods to a value given by the income from his or her own labour supply plus individual non-wage income. Goods prices and men's and women's wage rates are determined in a competitive general equilibrium by equalities of demands and supplies of all goods and labour.

Anonymity does not hold in this model – an increment of income to a man has in general different effects to those of an income increment to a woman. Moreover, symmetry of compensated labour supply derivatives with respect to each other's wage rates also does not hold. For example, an increase in the male net wage (caused, say, by a reduction in the tax rate) affects the labour supply of a woman not through a pooled budget constraint, but rather through a set of comparative statics effects on the price of the market good produced by men and the demands for the household and market goods produced by women.

⁷⁰ See Becker (1973), (1974) and, for a survey of more recent matching models, Bergstrom (1997).

⁷¹ At least, not initially, though in the long run the effect of labour market discrimination is to create differences in the accumulation of human capital and therefore in individual productivities. These effects are modelled endogenously.

Although it does not involve bargaining or threat points, the model highlights the important idea that the utility level partners can achieve within the household depends crucially on their outside opportunities, in this case conditions on their respective labour markets. Labour market discrimination that reduces the market wage of women will also reduce what they can obtain within the household. It embodies a form of the ‘separate spheres’ idea of Lundberg and Pollak (1993), but applied to household production rather than purchases of household public goods.

Grossbard (1984) also presents a general equilibrium model which determines within-household allocations, though it is somewhat more general in having markets for both male and female household labour as well as market work. It does not make the assumption that there is labour market discrimination against women, but instead assumes that the sexes are treated symmetrically. It focuses on defining an implicit wage for time spent in household production,⁷² allowing the partners, however, to make lump sum redistributions to compensate, if need be, for low implicit wages, and in this respect therefore intersects with the cooperative models discussed earlier. It does not model formally how these transfers are determined, but implicitly assumes some kind of bargaining process. It allows partners to specialise in different household goods, therefore also anticipating the ‘separate spheres’ idea, and again the structure of the model implies that anonymity and symmetry with respect to wealth and price changes do not hold. The main aim of the model is to extend the theory of labour markets, to show how variables that influence within-household labour markets, primarily the sex ratio and compensating marriage differentials, also have explanatory power in determining outcomes on labour markets external to the household.

The boundary between these ‘equilibrium models’ and the cooperative models discussed earlier is not so clear-cut as this attempt at classification may perhaps suggest, but is to a large extent more a matter of emphasis. For example, the bargaining models can incorporate conditions on labour markets and factors such as gender discrimination and the sex ratio either into threat points or EEPs. Many insights and qualitative results are shared by the two sets of models. The equilibrium models, however, place more weight on the price mechanism and the idea of the interrelationship between markets internal and external to the household, the cooperative models more on solving a centralised household decision problem.

3.5 The Pareto property and Pareto (in)efficiency

Cooperative household models see the household as acting as if it maximises a GHWF. The Pareto property, that the GHWF is strictly increasing in the utilities of its members, is one of the general properties of this function. However, whether the household actually achieves a first best Pareto-efficient allocation is a separate issue, relating to the entire structure of the model being analysed. In this section we consider two examples, one empirical, one theoretical, of the application of this point.

⁷² Following her earlier work on household labour market models, see Grossbard (1976).

In a careful and thorough empirical study, Christopher Udry (1996) showed convincingly that, in a large sample of West African households, household resource allocations were not Pareto-efficient. Since the cooperative theories of the household surveyed earlier in this chapter can be said to ‘assume Pareto efficiency’, this was interpreted as powerful evidence against these theories. Should we therefore accept Udry’s conclusion that ‘this [evidence] implies that the conventional pooling model of household resource allocation is false and both cooperative bargaining models and the more general model of efficient household allocations are inadequate for describing the allocation of resources across productive activities in households’?

This is exactly where the distinction between the Pareto property of a GHWF and the first- or second-best Pareto efficiency of a specific achieved resource allocation becomes important. Udry analysed a rich dataset giving information on landholdings, input quantities and output yields for a large sample of households in the West African country of Burkina Faso. The plots of land farmed by a given household can be divided into those controlled by the men in the household and those controlled by the women. By comparing inputs and yields of plots planted to the same crop by the same household in a given year,⁷³ he was able to show that the gender of the person controlling the land exerts a significant influence on the yield: men’s plots yield significantly higher outputs than women’s plots. Pareto efficiency would therefore require a reallocation of inputs across plots.

The key point however, as Udry emphasises, is that it is not possible for this reallocation to be brought about by contracts, i.e. binding agreements. If, for example, women’s land were to be worked by men, it would come to be seen as land under the men’s control. To retain their rights over the land, women are constrained to work it themselves. This constraint, arising out of an absence of contractual possibilities, creates a second-best situation.

The initial allocation of land appears to be determined by the marriage market, and control over land appears to be what determines the individual’s influence over the household resource allocation – for example, her bargaining power if we are working within the framework of a bargaining model. That is, the amount of land under one’s control is an extra-household environmental parameter (EEP) in the terminology of McElroy (1990). Thus, the reason a woman would not let a man take over the working of her land is that it would consequently worsen her position within the household. Even though it would increase the total value of household output, and so, given her initial power within the household, could increase her immediate consumption, in the future this power would decline and the subsequent worsening of her consumption share deters her from making the transfer of land in the first place.

This does not, however, prove that a model which assumes a GHWF with the Pareto property is refuted, though it does caution against assuming that this implies that the household resource allocation will be (first best) Pareto-efficient. What is necessary in

⁷³ The study controls very carefully for variables such as soil fertility, crop type and differences in production technology that might account for these differences.

this case is to impose the second-best constraint on the maximisation of the GHWF, expressing the fact that land cannot be reallocated in order to maximise returns, but rather is fixed by the original allocation. The model will then predict a second-best – ‘Pareto-inefficient’ – allocation, exactly as was observed. Stronger evidence than that produced by Udry is therefore necessary to show that the households were not behaving in accordance with the theoretical model as far as the formulation of the household’s objective function is concerned.

We develop this point further within a simple theoretical framework, which again shows that achieved resource allocations may be Pareto-inefficient even though the household seeks to maximise a GHWF with the Pareto property. This inefficiency again arises because we assume that the household members are unable to make binding commitments over time, although within any given time period they are able to do so. The model tries to capture the following situation. A newly formed household sees its future in two phases. In the first, it will have a high demand for household production. If one of the partners specialises in this, as is usually the case, she will reduce her market labour supply and correspondingly accumulate less work-related human capital. In the second phase, there is a much lower demand for household production, but the partner who previously specialised in it will face a lower market wage rate. If the couple can commit in the first period to consumption levels in the second, they can achieve a Pareto-efficient allocation which takes into account that their joint income will be lower in the second period the more household production they have in the first. The weight each receives in the GHWF is determined by their wage rates in the *first* period, because that is when they negotiate the allocation.

However, if they cannot make a binding commitment to consumption levels in the second phase, they must recognise that any first period agreement would be renegotiated at that time in the light of the then-prevailing wage rates. Thus, we have to impose as an *additional* constraint on their choice of allocations in the first period the restriction on the possible allocations they will be able to negotiate in the second. This creates the second-best Pareto inefficiency.

We take a two-period model in which only one partner, f , carries out household production, in the first period only. The key assumption is that, because human capital acquisition is work-related, her market wage in the second period is a decreasing function of the amount of household production she carries out in the first period, since this displaces time spent in market work in that period. We model the household’s choices first on the assumption that it is able to commit in the first period to individual consumption levels in the second, and show that we obtain a Pareto-efficient equilibrium. We then show the inefficiency that results when commitment is not possible.⁷⁴

⁷⁴ This type of problem has been thoroughly analysed in general terms in the ‘transactions cost’ literature, associated primarily with Coase, Grossman, Hart and Williamson. See Hart (1995) and Williamson (1989) for comprehensive accounts of this literature. The problem arises because complete contracts cannot be written or enforced. Marriage seems a particularly striking case of an incomplete contract. Pollak (1985) appears to have been the first to introduce the ideas from this literature into household economics. The model in Apps (1981), (1982) has the idea of market productivity decreasing with specialisation in household production. The present

3.5.1 The commitment case

Assume that m supplies one unit of market labour inelastically and that his wage is constant over time at w_m , while f 's wage is w_{f1} in period 1 and $w_{f2} = \omega(y)$, in period 2, with $\omega'(y) < 0$. Production of the household good y is carried out in the first period only. We assume that one unit of f 's time produces one unit of y , and so the marginal opportunity cost or implicit price of y is w_{f1} . When the household can commit to future consumption values, it solves the problem

$$\max H(u_{f1}(x_{f1}, y_f) + u_{f2}(x_{f2}), u_{m1}(x_{m1}, y_m) + u_{m2}(x_{m2}), w_{f1}, w_m) \quad (3.102)$$

$$\text{s.t.} \quad \sum_i (x_{i1} + w_{f1}y_i) \leq w_m + w_{f1} \quad (3.103)$$

$$\sum_i x_{i2} \leq w_m + \omega(y) \quad (3.104)$$

where x is a market consumption good with price normalised at 1 and $\sum_i y_i = y$. We assume time-separable utility with no utility discounting. To concentrate on essentials, we also assume no capital market. Since the household chooses its allocation at time 1, the relevant argument in the GHWF, which determines the marginal weight f 's utility receives, is her first-period wage w_{f1} . Then the first-order conditions with respect to consumptions can be written as:

$$\frac{\partial u_{f1}/\partial y_f}{\partial u_{f1}/\partial x_{f1}} = w_{f1} - \delta\omega' = \frac{\partial u_{m1}/\partial y_m}{\partial u_{m1}/\partial x_{m1}} \quad (3.105)$$

$$\frac{\partial u_{f1}/\partial y_f}{u'_{f2}} = \frac{w_{f1}}{\delta} - \omega' = \frac{\partial u_{m1}/\partial y_m}{u'_{m2}} \quad (3.106)$$

implying of course Pareto efficiency. Here $\delta = \lambda_2/\lambda_1$ is a discount factor, where λ_1 and λ_2 are the marginal utilities of household income in periods 1 and 2 respectively.

The household takes full account of the fact that part of the cost of household production in the first period is a lower wage for f in the second, arising from the loss of her human capital, and so expresses the implicit relative price of the domestic good in period 1 as $w_{f1} - \delta\omega'$ (recall $\omega'(y) < 0$). The term $-\delta\omega'$ acts as a tax on current consumption of the domestic good. It arises because the household's income in period 2 will be lower, the higher f 's domestic output is in period 1.

3.5.2 The non-commitment case

The model is as before, but now both partners realise that any prior agreement on consumptions in period 2 will be renegotiated in the light of wage rates prevailing at that time. Then the time 2 allocation will be chosen as the solution to the following

model is a simpler version of a Nash bargaining formulation of the non-commitment problem developed by Ott (1992) ch. 6. See also Lundberg and Pollak (2003) for a similar type of Pareto inefficiency. The simplicity of the present model is perhaps an argument for the GHWF approach.

problem

$$\max H(u_{f2}(x_{f2}), u_{m2}(x_{m2}), \omega(y), w_m) \quad (3.107)$$

$$\text{s.t. } \sum_i x_{i2} \leq w_m + \omega(y). \quad (3.108)$$

That is, the weight given to f in the second period will reflect her wage in that period. It follows that the optimal solutions to this problem are functions $x_{i2}^*[\omega(y)]$. Moreover, given the assumptions on the GHWF set out in section 3.2.3.2, straightforward comparative statics yields

$$\frac{dx_{f2}^*}{dy} \neq \frac{dx_{m2}^*}{dy} \quad (3.109)$$

as we would expect. A change in y changes each period 2 consumption not only because of an income effect, but also because the relative weights on the individuals' utilities change in m 's favour.

At time period 1 the household solves the problem

$$\max H(u_{f1}(x_{f1}, y_f) + u_{f2}\{x_{f2}^*[\omega(y)]\}, u_{m1}(x_{m1}, y_m) + u_{m2}\{x_{m2}^*[\omega(y)]\}, w_{f1}, w_m) \quad (3.110)$$

$$\text{s.t. } \sum_i (x_{i1} + w_{f1}y_i) \leq w_m + w_{f1}. \quad (3.111)$$

From the first-order conditions for this problem we obtain

$$\frac{\partial u_{i1}/\partial y_i}{\partial u_{i1}/\partial x_{i1}} = w_{f1} - \frac{\lambda_2}{\lambda_1} \frac{dx_{i2}^*}{dy} \quad i = f, m. \quad (3.112)$$

Then, since the second terms on the right-hand side are unequal for each i , the first-period allocation will not be Pareto-efficient. In choosing their consumptions, they take into account the effects of f 's loss of human capital on their individual consumptions and utilities in the household equilibrium in period 2 and, since these are different, their marginal rates of substitution between the two goods in period 1 will differ. They cannot correct this because of the inability to make binding commitments to the allocation in period 2.

Appendix to chapter 3

Part A

The problem is

$$\min_{x_i, z_i} t = \sum_i (x_i - w_i(1 - z_i) - \mu_i) \quad (3.113)$$

$$\text{s.t. } H(u^1(x_1, z_1), u^2(x_2, z_2); w_1, \mu_1, w_2, \mu_2) \geq H, \quad (3.114)$$

first-order conditions of which are

$$1 - \beta H_i u_{x_i}^i = 0 \quad i = 1, 2 \quad (3.115)$$

$$w_i - \beta H_i u_{z_i}^i = 0 \quad i = 1, 2 \quad (3.116)$$

together with the constraint as an equality, with β as a Lagrange multiplier. Standard comparative statics analysis then gives

$$H_{w_1} \frac{\partial \beta}{\partial w_2} - H_{w_2} \frac{\partial \beta}{\partial w_1} = A \frac{D_{15} + w_1 D_{25}}{H_1 D} + B \frac{D_{35} + w_2 D_{45}}{H_2 D} + \frac{H_{w_1} D_{45} - H_{w_2} D_{25}}{D} \quad (3.117)$$

where

$$A = H_{w_2} H_{1w_1} - H_{w_1} H_{1w_2} \quad (3.118)$$

$$B = H_{w_2} H_{2w_1} - H_{w_1} H_{2w_2} \quad (3.119)$$

D is the bordered Hessian of the Lagrangean, and D_{i5} , $i = 1, \dots, 4$ is the corresponding cofactor of D . Then assumption A2 implies that A and B cannot be zero, and inspection of D shows that the expression in (3.117) cannot be zero.

Part B

The problem

$$\max_{s_i} H(v^1(w_1, s_1), v^2(w_2, s_2); w_1, \mu_1, w_2, \mu_2) \quad (3.120)$$

$$\text{s.t.} \quad \sum_i s_i \leq \sum_i (w_i + \mu_i) + t \quad (3.121)$$

has first-order conditions

$$H_i v_s^i - \lambda = 0 \quad i = 1, 2 \quad (3.122)$$

together with the constraint as an equality, with λ a Lagrange multiplier. Standard comparative statics analysis then gives

$$\frac{\partial s_i}{\partial \mu_i} = \frac{H_{i\mu_i} v_s^i - H_{j\mu_i} v_s^j}{\Delta} + \frac{\partial s_i}{\partial t} \quad i, j = 1, 2 \quad i \neq j \quad (3.123)$$

$$\frac{\partial s_i}{\partial \mu_j} = \frac{H_{i\mu_j} v_s^i - H_{j\mu_j} v_s^j}{\Delta} + \frac{\partial s_i}{\partial t} \quad i, j = 1, 2 \quad i \neq j \quad (3.124)$$

where $\Delta > 0$ is the bordered Hessian of the Lagrange function. Since assumption A2 implies $H_{i\mu_i} \neq H_{i\mu_j}$, these two are not equal.

Part C

Here we show how the forms of the Nash bargaining and collective models, viewed as special cases of the GHWF approach, correspond to assumptions A1 and A2 in subsection 3.2.3.3.

A1

Nash bargaining model

$$H_{w_i}^N = -\lambda_i^0(1 - z_i^0)(u^j - v_j^0) \quad i, j = 1, 2 \quad i \neq j \quad (3.125)$$

where z_i^0 is the uncompensated leisure demand of individual i at her threat point, and λ_i^0 is her marginal utility of income at that point. This is strictly negative at all points of interest, given a positive ‘marriage surplus’, but zero at the points $u^j = v_j^0$ and $z_i^0 = 1$. For μ_i we have

$$H_{\mu_i}^N = -\lambda_i^0(u^j - v_j^0) \quad i, j = 1, 2 \quad i \neq j \quad (3.126)$$

which again is strictly negative at all points of interest, but zero at $u^j = v_j^0$. Note that this formalises an important point made by Lundberg and Pollak (1993), which led to the formulation of their ‘separate spheres’ bargaining model: a pure redistribution of incomes between household members has no effect in the Nash bargaining model if it leaves the threat points unchanged, i.e. if $\lambda_i^0 = 0$ for such a redistribution.

Collective model

$$H_{w_i}^C = \alpha_{w_i}(u^1 - u^2) \gtrless 0. \quad (3.127)$$

Here $\alpha_{w_1} > 0$, $\alpha_{w_2} < 0$, and the sign of the derivative $H_{w_i}^C$ depends on the relationship between the two utility values, with zeroes at all allocations such that $u^1 = u^2$. Likewise

$$H_{\mu}^C = \alpha_{\mu}(u^1 - u^2) \gtrless 0 \quad (3.128)$$

A2

Nash bargaining model

$$H_i^N = u^j - v_j^0; \quad H_{iw_i}^N = 0; \quad H_{iw_j}^N = -\frac{\partial v_j^0}{\partial w_j} < 0 \quad (3.129)$$

$$H_{i\mu_i}^N = 0; \quad H_{i\mu_j}^N = -\lambda_j^0 < 0 \quad i, j = 1, 2 \quad i \neq j \quad (3.130)$$

Collective model

$$H_{1w_1}^C = \alpha_{w_1} > 0; \quad H_{1w_2}^C = \alpha_{w_2} < 0 \quad H_{2w_i}^C = -H_{1w_i}^C \quad i = 1, 2. \quad (3.131)$$

$$-H_{2\mu}^C = H_{1\mu}^C = \alpha_{\mu} \gtrless 0. \quad (3.132)$$

4 Empirical household models

4.1 Introduction

In this chapter we survey the more recent literature that attempts to take household models to the data. The first section sets the scene, by reviewing the household utility function approach, or variants of it, to modelling the labour supply decisions of couples. We show what is involved in estimating the model on datasets with missing information on the allocation of non-market time between domestic production and leisure. The following sections review the empirical studies on multi-person household models. We show what goes wrong when household production is assumed not to exist. Our basic contention here is that, as shown in chapter 1, the significance of time use for household production is an empirical fact, not a theoretical hypothesis.

We place a great deal of emphasis on the fact that available datasets, even time-use datasets, are inadequate to allow reliable estimation of structural parameters, and that significant advances in this area require more comprehensive datasets. In the meantime, we need to understand exactly what kinds of data we need that we do not have, and the ways in which assumptions implicit in commonly made empirical specifications effectively construct data. The implications of missing information on pure leisure; time allocations to domestic work, including child care; domestic outputs; and individual consumptions of market and domestic goods will be discussed in some detail.

4.2 The household utility function model

4.2.1 *The theoretical underpinning*

The empirical literature on the labour supplies and consumption demands of two-adult households has tended, until quite recently, to focus on the estimation of various specifications of the household utility function (HUF) model.¹ Most studies assume

¹ For surveys, see Blundell and MaCurdy (1999), Heckman (1993) and Killingsworth and Heckman (1986), and for a comprehensive review of earlier work, see Killingsworth (1983).

that the underlying theoretical model has the general form

$$\max_{x, z_i} u = u(x, z_f, z_m) \quad (4.1)$$

$$\text{s.t.} \quad \sum_i w_i z_i + x \leq \sum_i (w_i + \mu_i) \equiv X \quad (4.2)$$

where x is total household consumption of a Hicksian composite commodity and z_i is i 's pure leisure, $i = f, m$.² The μ_i and w_i are i 's non-wage income and wage rate respectively; X is full income. The individual time constraints are

$$l_i + z_i = 1, \quad i = f, m \quad (4.3)$$

where l_i is market labour supply and total time available is normalised to 1. The resulting demand functions are $x(w_f, w_m, X)$, $z_i(w_f, w_m, X)$, $i = f, m$.

As we made clear in the previous chapter, the underlying process that generates the household's demands is the maximisation of a household welfare function (HWF), and $u(\cdot)$ can be thought of as belonging to a fictitious individual who chooses the aggregate consumption and leisure demands of the household. Thus these demands satisfy symmetry of the Slutsky matrix and anonymity with respect to non-wage income changes. The literature is then concerned with the estimation of empirical specifications of these demand functions. Since this model predicts symmetry of the compensated leisure demand or labour supply functions, rejection of this symmetry by the data is interpreted to imply rejection of this model.

However, in this empirical work, the data identified with the variables z_i do not actually measure pure leisure demands. The convention has been to estimate the model on household survey data containing information on market labour supplies, wage rates and non-labour incomes at the level of the individual and on market consumptions at the level of the household.³ While it was always clear that data on individual consumptions of the market good were missing, it was less clearly understood that data on pure leisures were also missing in these datasets. Only information on *total non-market time use*, obtained by subtracting hours of market labour supply from the total time constraint, was available.⁴

We know from time-use data that non-market time includes time spent on household production for consumption by other family members, as well as time used wholly for *own* consumption, i.e. pure leisure. Empirical work based on these datasets therefore can only provide estimates of the aggregate household demands for the time input

² See, for example, the discussion of the 'standard family labor supply model' in Blundell and MaCurdy (1999), p. 1658. The survey presents the model for a household consisting of two working-age individuals, with children and any other dependants included in a vector of household attributes X_t , in period t . As the survey states: 'Families are assumed to maximise joint utility over consumption, C_t , and the leisure of each family member, L_{1t} and L_{2t} . For such a household utility may be written $U(C_t, L_{1t}, L_{2t}, X_t)$.'

³ Kooreman and Kapteyn (1987) is a notable exception. The authors use time-use survey data to estimate a model of the allocation of time within the household.

⁴ This is widely recognised in studies that estimate models using data for developing countries. See, for example, Schultz (1990).

of each adult into non-market activities that provide domestic consumption for all members, *plus* the individual demand for pure leisure. Thus the empirical model is an aggregate family demand system for all goods, including domestically produced goods, that appear in the household utility function. It becomes a model of individual leisure demands only if it is estimated on time-use data for pure leisure.

We can formalise this by setting up a household model with a Samuelson HWF of the form

$$H = H(u^f(x_f, y_{ff}, y_{fm}, z_f), u^m(x_m, y_{mf}, y_{mm}, z_m)) \quad (4.4)$$

where y_{ij} is the amount of a domestic good consumed by individual i that is produced by individual j . Thus the individual utility functions are defined on four goods: market consumption, pure leisure, consumption derived from i 's time in household production, and consumption derived from j 's time in household production. From this we can write the HUF as $u(x, y_f, y_m, z_f, z_m)$, i.e. as a function of aggregate market and household goods and individual leisures.

For simplicity, take the linear household technology introduced in the previous chapter:

$$\sum_{i=f,m} y_{ij} = y_j = k_j t_j \quad j = f, m \quad k_j > 0 \quad (4.5)$$

so that the time constraints become

$$l_j + t_j + z_j = 1 \quad j = f, m.$$

The implicit prices (marginal opportunity costs) of the domestic goods are then

$$p_j = \frac{w_j}{k_j} \quad j = f, m \quad (4.6)$$

and so the household full income budget constraint becomes

$$x + \sum_{j=f,m} (p_j y_j + w_j z_j) = \sum_{j=f,m} (w_j + \mu_j) \equiv X. \quad (4.7)$$

Maximising household utility subject to this budget constraint gives aggregate Marshallian household demands

$$x(p_f, p_m, w_f, w_m, X), \quad y_j(p_f, p_m, w_f, w_m, X), \quad z_j(p_f, p_m, w_f, w_m, X) \\ j = f, m \quad (4.8)$$

with the properties of symmetry and anonymity. Now we have that

$$\tau_j \equiv 1 - l_j = z_j(p_f, p_m, w_f, w_m, X) + \frac{1}{k_j} y_j(p_f, p_m, w_f, w_m, X) \quad (4.9)$$

$$= \tau_j(w_f, w_m, X) \quad j = f, m \quad (4.10)$$

where τ_j is total non-market time, erroneously interpreted simply as leisure. Then an empirical model that estimates the function $\tau_j(w_f, w_m, X)$ with non-market time as the dependent variable can be interpreted as estimating a type of reduced form of

this Samuelson model, extended to include household production. It is clearly not the leisure demand function. The parameters of these empirical demand functions $\tau_j(\cdot)$ will be combinations of preference and household technology parameters, which, however, cannot in general be identified from the parameter estimates. Given that these estimated parameters are not parameters of a true leisure demand function, we cannot in general expect that they will satisfy restrictions such as symmetry of compensated demands.

Clearly, the data that would be required to estimate and test this model satisfactorily would be measures of both inputs and outputs in household production, so that estimated productivities could be used, in conjunction with wage rates, to obtain estimates of the prices of domestic goods. These could then be used, together with estimates of market prices and non-wage incomes, to estimate demand functions for household and market goods and true leisure demands. This approach is further discussed in section 4.3.

The studies in the labour supply literature that have attempted to estimate the preceding version of the HUF model have tended to concentrate on the econometric issues associated with corner solutions and missing data on the market wage rates of non-participating married women,⁵ and with the difficulties that arise when the tax-benefit system leads to budget sets that are not convex.⁶

The concern with corner solutions is indeed justified, given the proportion of married women of prime working age who do not participate. The question is, how to model them. Heckman (1980) describes the problem as follows:

The important point to note is that traditional estimates of the coefficients of labor supply functions of working women confound two effects: movement along a given labor supply function for working women, and movement across taste distributions. Thus, for example, presence of an additional child under six has a dramatically negative effect on hours of work for a working woman (–925 hours reduction in supply). But working women with an additional child have a greater average taste for market work, since only the most work-prone women remain after the ‘imposition’ of a child.

The data presented in chapter 1 support this estimate of the effect of the presence of a young child in the household. However, the interpretation of unexplained heterogeneity in female labour supply in terms of ‘taste’ variation – being ‘work-prone’ or not – or unobserved ‘fixed effects’ is unsatisfactory, and represents a major limitation of the household utility function approach. As we have emphasised in earlier chapters, non-participation represents the household’s choice that the mother will use her own time to care for a young child, rather than to work in the market and buy in child care. This choice is driven by a number of factors, already listed in earlier chapters, among which differences in productivity in household production could be important.⁷ To

⁵ See, for example, Hausman and Ruud (1984), Kooreman and Kapteyn (1986) and Ransom (1987).

⁶ There is a large literature that uses simplified versions of the HUF model as the theoretical framework for modelling tax reform because of the difficulties encountered with non-convex budget sets. For computational simplicity, a number of more recent studies estimate family labour supply models with discrete hours choices, following the approach of van Soest (1995).

⁷ This is acknowledged in Heckman (1980): ‘there may be self selection by the individuals being investigated. One observes market wages for certain women because their productivity in the market exceeds their productivity in the home. (Note that this does not imply that [as a group] the more productive women are the ones that work.)’

attribute heterogeneity simply to preferences is to sweep under the carpet, so to speak, an interesting and important set of economic issues.

In one sense, one might argue that all the emphasis on ‘corner solutions’ is misplaced, or at least that it directs attention to the wrong kind of problem. After all, if we observe that a woman, or a man for that matter, works in a bank rather than an insurance company, we do not see this as implying a corner solution in the supply of labour to the insurance sector. Why should the decision to work in the household rather than the market be any different? A key aspect of the difference is the difficulty of obtaining data on the economic factors that motivate the decision to work in the household, among them the marginal value product in household work, which is due in part to a long history of lack of interest in collecting these data.

Considerable progress can in fact be made in estimating household models by using information on domestic work, child care and leisure activities available in time-use data, as we show in the next section. These data also provide an insight into why there are ‘corner solutions’. As the figures in chapter 1 show, non-participants, the vast majority of whom are female, work much longer hours in household production than participants, especially on child care. The data suggest that corner solutions reflect the high degree of substitution of home production for market work with the arrival of the first child, in response to the very high demand for child care generated by that event. In other words, corner solutions, at least in the earlier phases of the life cycle, are about child care and the different choices that can be made in providing that care.

4.2.2 *Models with fixed labour supplies*

A large number of studies attempt to avoid the difficulties encountered in modelling jointly determined labour supplies, especially in the context of non-linear and non-convex budget sets,⁸ by estimating simplified versions of the HUF in which the labour supply of one partner is assumed to be fixed. Here we present a model in which male labour supply is fixed and only female labour supply is variable. The approach, labelled by Killingsworth (1983) the ‘male chauvinist’ model, has been justified with reference to the strong evidence that there is little variation in the labour supply of prime aged males, as for example in Blundell *et al.* (1988), p. 25:

We consider the effects of tax reform only on the labour supply of married women because our research on household labour supply [see, for example, Blundell and Walker (1986) and Blundell and Meghir (1985)] indicates that the male hours of work is not a dimension of household behaviour that exhibits significant sensitivity to economic variables. This finding is not unusual and indeed is a main conclusion of the exhaustive survey of the literature in Pencavel (1987).

We first set out the approach and then show the implications of taking account of household production. Under the assumption of fixed male labour supply, the

⁸ See, for example, Hausman (1981), Blomquist (1983), Arrufat and Zabalza (1986), MaCurdy *et al.* (1990), and Arellano and Meghir (1992).

household's budget constraint is given by

$$x + w_f z_f = w_f + v \quad (4.11)$$

where $v = \mu + w_m \bar{l}_m$ is now exogenous income. The empirical application of the model implies the estimation of household demands, $x(w_f, v)$, $z_f(w_f, v)$, which are the solutions to the problem

$$\max u(x, z_f) \quad \text{s.t.} \quad x + w_f z_f = w_f + v.$$

This then yields the female labour supply function

$$l_f(w_f, v) = 1 - z_f(w_f, v). \quad (4.12)$$

Blundell *et al.* (1988) specify functional forms for female labour supply that satisfy aggregation, with the error term of each interpreted as the 'stochastic component of preferences' (p. 27). The labour supply model of Arrufat and Zabalza (1986) is also an example of a model of this kind, with optimisation errors in addition to errors which are said to capture preference heterogeneity.

We now take the model of the previous subsection, and consider what happens to it when we fix male labour supply. That model is based on the problem of maximising the HUF $u(x, y_f, y_m, z_f, z_m)$. In order to reduce the problem to that of choosing household consumption and the labour supply of f alone, we assume that the time m spends in household production as well as pure leisure is fixed. Thus the household's problem becomes:

$$\max_{x, y_f, z_f} u(x, y_f, \bar{y}_m, z_f, \bar{z}_m) \quad \text{s.t.} \quad x + (p_f y_f + w_f z_f) = w_f + v. \quad (4.13)$$

The resulting demand functions are

$$x(p_f, w_f, v), \quad y_f(p_f, w_f, v), \quad z_f(p_f, w_f, v) \quad (4.14)$$

and again we have

$$\tau_f \equiv 1 - l_f = z_f(p_f, w_f, v) + \frac{1}{k_f} y_f(p_f, w_f, v) \quad (4.15)$$

$$= \tau_f(w_f, v). \quad (4.16)$$

Alternatively, we can assume that empirical work of this kind assumes that the pure leisure is fixed and productivities do not vary across households. In that case, by suitable choices of units, the k s can all be set to 1 and the price of y_j is just the wage rate w_j . The household's problem then becomes:

$$\max_{x, y_f, z_f} u(x, y_f, \bar{y}_m) \quad \text{s.t.} \quad x + w_f y_f = v. \quad (4.17)$$

The resulting demand functions are

$$x(w_f, v), \quad z_f(w_f, v), \quad y_f(w_f, v) \quad (4.18)$$

and again we have

$$\tau_f \equiv 1 - l_f = y_f(w_f, v) \quad (4.19)$$

$$= \tau_f(w_f, v). \quad (4.20)$$

This suggests that estimating the function $l_f(w_f, v)$ involves the omission of the variable p_f , or equivalently household productivity.

4.3 Estimation of models on time-use data

Surprisingly, there has been relatively little discussion of the limitations of the treatment of household production in the HUF model, and of the problems involved in estimation of the model on datasets with missing information on the allocation of time to leisure and household production. Instead, emphasis has been on using the same limited datasets to estimate the partial derivatives of sharing rules, and to test whether symmetry and anonymity of the HUF model are rejected empirically. In this section we draw upon Apps and Rees (1996), (2002) to discuss the issues involved in trying to estimate models which are, from the theoretical point of view, satisfactorily specified, but which face the problem of missing data on key variables. We also try to make clear how misleading the results on estimates of behavioural parameters and sharing rules can be, when empirical models go only part of the way in recognising the real nature of households. Replacing the household utility function approach with a model representing the household as consisting of two individuals with their own utility functions does not in itself represent a major improvement if it ignores household production and child care. It may actually be a step backwards, if it provides estimates so inaccurate that they are worse than having no estimates at all. For example, we show below that estimating a collective model while ignoring the existence of household production suggests that wives have more than a 30 per cent higher value of consumption and 60 per cent higher amount of leisure time than their husbands. Consider the implications of using that kind of estimate to design tax policy that takes into account the within-household income distribution! Moreover, in order to minimise the problems that not going the whole way in modelling the household creates, empirical studies often restrict the data on which they are estimated, for example to households consisting only of fully employed couples with no children present. Not only does this raise issues of sample selection bias, but also severely limits the applicability of the results. We now go on to discuss these points in some detail.

4.3.1 *Comparison of the ‘exchange’ and ‘transfer’ models of the household*

Apps and Rees (1996) construct a dataset by merging data on hours of market work, earnings, non-labour incomes, demographic characteristics and levels of education with one giving detailed data on non-market time use.⁹ The study estimated two models, one

⁹ For a further discussion of data sources and methodology used to merge time-use data with household income and expenditure survey data, see section 5.4.1.

called the 'exchange model', which is essentially a version of model 2 presented earlier in chapter 3, and the other a 'transfer model', the collective model of Chiappori (1988) and (1992). These names reflect the fact that the former sees economic activity within the household in terms of specialisation and exchange, while the latter sees at least part, and possibly the whole, of the consumption of one partner as being financed by a transfer from the other. One aim of the paper was to show the differences in results obtained when estimating a model which explicitly takes account of household production, as compared to one that assumes this does not exist, and that all non-market time is pure leisure. The time-use data allowed this to be done. A second aim, however, was to clarify the limitations even of this quite comprehensive dataset, and to show what assumptions had to be made to be able to estimate a satisfactory household model in the absence of all the required data. Finally, a third aim was to show the importance of female labour supply heterogeneity, by comparing the estimation results for 'traditional households', in which wives specialise in household work, with 'non-traditional households', where wives have a significant labour force attachment.¹⁰

The production and demand systems of the exchange model are estimated simultaneously. On the production side we assume constant returns to scale and specify a translog household production function to estimate the parameters of the unit cost function for domestic output.¹¹ We estimate demand functions for the market consumption good, the household good and pure leisure for the Deaton and Muellbauer (1980) 'Almost Ideal' (AI) demand system specification of preferences.

Since data on individual market and domestic consumptions are missing, we cannot compute full income shares. In the absence of these data we assume pure exchange,¹² so that each individual's full income can be computed from the data on wage rates and non-labour incomes available at the level of the individual. This implies zero intra-household transfers, which is clearly just one possible assumption among many. The point was to show that some such assumption to compensate for missing data is necessarily always made by any model which estimates household demands, and at least this procedure has the virtue of making this explicit. Moreover, to assume otherwise raises the question as to why we would expect a transfer between partners on average, if we accept that there are no relevant innate productivity differences on the basis of gender. With missing data on the individual consumptions of the domestic good, we were also limited to estimating an aggregate share equation for the domestic good,¹³ with unidentified preference parameters assumed to be the same for both partners.

The transfer model was estimated not by nesting it in the exchange model, but by proceeding as a researcher would if he believed that this model was a correct specification of household behaviour and, in particular, really believed that household production did

¹⁰ More than 500 hours per year.

¹¹ Notable pioneering papers in the area of household cost function estimation are Gronau (1980) and Graham and Green (1984).

¹² This is the Apps (1981) model.

¹³ Given the parameter restrictions required for adding up, only the domestic good and leisure shares need to be estimated.

not exist. Thus each individual's time input into household production is added to his or her pure leisure to obtain the dependent variable in the 'leisure' share equations. This of course represents a misspecification of the leisure variable, which arises in any model that takes 'total non-market time' as pure leisure, as we discussed earlier. This leads also to a misspecification of the individual income variable. Because the transfer model treats time in household production as pure leisure, one individual in the household must receive, and the other pay, a transfer equal to the difference between the value of his or her market income and consumption. In the case of a traditional household in which the female partner has no market income, if she also has no non-labour income, the transfer from the male partner is the full value of her market consumption. This must be how someone who believes in the transfer model would specify individual incomes. Estimation of a 'sharing rule' based on the collective model is essentially doing the same thing, but without making this explicit.

As we might expect, the parameter estimates for independent variables common to the two models were very different. Of particular interest from the point of view of behaviour relationships that enter policy formulation, for example taxation, were the results on labour supply elasticities. The exchange model yielded (uncompensated) elasticities of 0.1803 for females and 0.0106 for males, while the transfer model yielded 0.2347 for females and 0.1136 for males. The paper argues that misspecification of the leisure and income variables and omission of the implicit price of household production accounts for the higher estimates from the transfer model. As we might also expect, the transfer model generates very high values for the estimates of the transfers made by the primary to the secondary earner, with on average husbands 'transferring' around one-third of their income to their wives. This resulted in a mean full consumption of \$49,569 for wives and \$36,557 for husbands. It also implies that wives had much higher levels of leisure consumption, with a mean of 5,257 hours p.a. as opposed to 3,665 p.a. for their husbands. These results of course follow directly from ignoring the existence of household production.

As the paper makes clear, all results are dependent on the assumptions made to compensate for the absence of data on the outputs of domestic production and the allocation of consumption among individuals in the household. The former are required for estimating productivities and the implicit prices of household goods, the latter for estimating individual shares in full income. We continue to maintain that significant progress will be made in estimating the parameters of behavioural relationships of individuals in households, of sufficient reliability and robustness for use in policy formulation, only when such data are made available.

4.3.2 *Estimating the costs of children*

The need for a measure of the costs children create for a household arises in a wide range of policy contexts. Horizontal equity in taxation requires that the tax levied on a given income should take account of how many children that income has to support. Poverty alleviation requires that the transfer a household receives will depend on the number of children it contains.

The method of estimating child costs that finds least favour with economists is that of estimating the costs of a child's physiologically determined 'basic needs', for food, clothing, heat, shelter and so on. Indeed as both Browning (1992) and Nelson (1993) show, this is the most common approach to equivalence scale calculation in the US and many other countries. Methodologically, economists dislike this approach because of its prescriptive, non-choice-based nature and its implicit identification of 'welfare' with a narrow set of measures of physical well-being. More substantively, it finds disfavour with those who stress that poverty and deprivation are relative concepts. To tie income support to the cost of achieving some minimum subsistence level is to define poverty in absolute physiological terms. On the other hand, a definition of poverty in relative terms requires estimation of the amounts households actually spend on their children. The approach discussed here takes the view that the cost of a child is an outcome of the choice of an intra-family distribution of consumption, that is, of how much consumption of all goods, including child care and other domestic outputs, a household chooses to allocate to its children at an equilibrium.

At a basic level there is the question of why children cost their parents anything in the first place. The standard analysis of intertemporal choice in a world of complete information and perfect markets shows how an individual can use the capital market to determine an entire lifetime consumption stream given an endowed lifetime income stream. Why should that individual not be a child, or parent acting as the child's agent? This would imply that parents need not incur child costs.

The problem is of course essentially one of contract incompleteness. There are considerable uncertainties surrounding the future income stream of a newly born child, and it is prohibitively costly to write a contract, which would probably be non-enforceable, specifying an action for every possible contingency. There is also an agency problem – what would there be to stop a parent mortgaging the future income of the child to increase his or her own consumption? It is then this incompleteness in capital markets which creates the need for intergenerational transfers to cover child costs.¹⁴

But then the impossibility of a contract by which a parent would be compensated for these costs from the future income of the child¹⁵ creates the possibility that the costs that would be incurred are not optimal from the child's point of view. This is not only to say that altruism may not be sufficient to achieve optimality for the child, but that the parent's own wealth constraint and preferences define the consumption and investment possibilities, especially in human capital formation, for the child. It is then possible to base the case for public intervention in support of child costs on the existence of these kinds of market incompleteness.¹⁶ Moreover, as we argue in chapter 5, capital

¹⁴ See also Becker (1991).

¹⁵ This is of course the problem that motivated Samuelson's formulation of the overlapping generations model (see Samuelson 1958). Cigno (1993) proposes a solution to this problem, in the context of a three-generation overlapping-generations model, using the idea of a self-enforcing family constitution.

¹⁶ This implies that there is an efficiency argument, as well as a distributional argument, for transfers to families and children. For an interesting development of this approach to the analysis of transfers to families, see Cigno *et al.* (2000).

market imperfections in the more usual sense also inhibit the smoothing of consumption over time, resulting in significant falls in parental full consumption on the arrival of children.

The modelling approach to the problem of estimating child costs so far adopted in much of the literature is so narrow and in its empirical applications itself so ad hoc that it is not at all clear that any improvement over the physiologically based estimation has been made. The approach is to postulate a utility function, indirect utility function or expenditure function, in which children appear as parameters, and to define child costs as the amount of money which would have to be paid to hold utility constant as the 'child parameters' vary. That is, we have a compensating variation measure of child costs, which can be put in terms of an equivalence scale by dividing it by the income the childless household or individual would require to achieve the same utility level.

As Nelson (1993) makes clear, this formulation implies a basic redefinition of the question that child cost measurement is designed to answer, and one which puts in question the policy relevance of the results. The policy concern is with the material welfare of each household member including, or especially, that of the children, and not simply the utility level of the adults. The main aim of policy is not simply to compensate adults for the amount of their consumption they choose to divert to their children. A further important policy issue is that of ensuring that transfers aimed at benefiting children actually reach them and do not just result in increased adult consumption.¹⁷ Under the 'adult utility function' approach this is not even an issue, and the approach certainly does not provide an analytical framework within which it can be discussed.¹⁸

At the conceptual level, as against the level of policy relevance, at least as damning is the Pollak and Wales (1979) critique. This establishes first that the required compensating variations or equivalence scale index numbers cannot be estimated from the kinds of expenditure survey data usually used for the purpose; and second that if the true compensating variations are indeed estimated, the resulting 'child costs' would in fact be negative and so would again be valueless for the kinds of purposes to which public policy wants to put them. If adults have children it must be because their utilities are higher with than without them, which implies a negative compensating variation. Despite this critique, a number of models have been applied to obtain estimates of equivalence scales.¹⁹ Results vary widely, with estimates of the 'cost of a child' ranging from 12 per cent to 100 per cent of an adult's consumption.²⁰

¹⁷ For example, in the UK the system of tax allowances for dependent children was replaced by a system of direct cash payments on the grounds that the resulting transfer was more likely to be received by mothers and therefore more likely to result in increased consumption by children.

¹⁸ In other words there may well be a dissonance between the welfare weights individuals, here children, receive in a 'household welfare function' and those they receive in a 'social welfare function'. For further discussion of this see chapter 7.

¹⁹ Surveys of the main approaches – originating respectively with Engels, Rothbarth, Prais and Houthakker, Barten, and Gorman – are given in Browning (1992) and Nelson (1993). These also make clear the implausibility of the assumptions underlying each of the approaches.

²⁰ See table 2 in Browning (1992).

As well as discarding the approach based on holding parental utility constant, it is also necessary to take into account the fact that a major part of the costs of a child is the parental time that has to be diverted to it from other uses. Although this has long been recognised in the theoretical literature,²¹ it has largely been neglected in empirical work on child costs.²² This work is based exclusively on consumption expenditure data, even though there is every reason to expect the expenditure of parental time to be even more important quantitatively than that on the child's consumption of market goods. Here we define the cost of a child in terms of its *full consumption*, that is, the sum of consumption of market goods, domestically produced goods, and parental time in the form of child care.

The model underlying the empirical work is that presented in chapter 3. All household members, including children, are treated as individuals with their own utility functions, on which the household welfare function is defined. Adults choose allocations of their time over market labour supply, domestic production, child care and pure leisure. They also choose an allocation of market income over consumption goods and bought-in inputs (including possibly child care) into the household production process.

As in the work discussed in the previous subsection, we distinguish between 'traditional households', in which the secondary earner, usually female, works hardly at all outside the home, and 'non-traditional households', in which the secondary earner's market labour supply is substantial, approaching that of the primary earner in the upper part of the distribution. The data presented in chapter 1 show that in the former type of household the consumption allocation of children consists of a much larger proportion of maternal child care and domestic production, and a much smaller proportion of market goods, than in the latter type of household. An approach which focuses on market consumption, as does the traditional equivalence scale literature, would therefore regard child costs as much lower in the traditional household, which clearly is absurd. This absurdity does not become obvious in the literature because it averages across household types, thus biasing estimates of true child costs downwards. Here we maintain the distinction between the household types, not only because it gives a more accurate estimate of child costs, but also because it is interesting to observe the substitution between a child's consumption of domestic child care and household goods, on the one hand, and market goods, on the other, as we move between household types.

The procedure used in the empirical work can be summarised as follows. We merge time-use survey data on the allocation of time to child care and domestic work with information on wage rates and non-labour incomes, for a sample of two-child families.

²¹ See in particular Cigno (1996).

²² Notable exceptions to this are Gustafsson and Kjulin (1994) and Colombino (2000). The former use Swedish time-use data and net wage rates to estimate the time costs of child care and other child-induced housework, distinguishing between mother's and father's inputs. The latter constructs a model of the household resource allocation and estimates child costs in terms of the consumption allocation, including parental time, that the child receives at this allocation. It does not, however, consider the differential impact of child costs on parents.

This allows the straightforward computation of domestic child care costs and their variations between the parents and across household types. Data on adult consumptions of pure leisure are similarly available, but the individual adult allocations of market and domestic consumption goods have to be estimated. This is done by: assuming that adult preferences over own-consumption bundles are the same in households with and without children;²³ assuming pure exchange, and therefore no lump sum transfers, between the adults, as discussed in the previous subsection; and then estimating the adult preference parameters on data for households without children.

The production and demand systems are estimated simultaneously, as in the exchange model, for the same functional forms: the translog for the constant returns to scale production system and the AI demand system specification of preferences. We also specify a sharing rule with restrictions that allow us to estimate the size of transfers between each parent and the children. This allows estimation of the value of consumption of the market good, domestic good and parental child care allocated to each child.

Using data on the observed changes in individual leisure consumptions when the household acquires children, it is possible to impute corresponding changes in the allocations of income and consumption separately to the adults. The data show that female leisure consumption falls by significantly more than that of the male when children are added to the household. This is interpreted to mean that the imputed income of the female in the household equilibrium also falls significantly relative to that of the male, so that a larger transfer to the children is made by females than by males. In effect, pure leisures are used as assignable goods in identifying changes in individual parental full income shares following the arrival of children.

In broad summary, the main results are as follows. In households²⁴ with the traditional market/household division of labour, the overall cost of both children's consumption of market goods is estimated to be in the range of 23 to 34 per cent of the household total. These estimates increase to around 40 to 47 per cent in non-traditional households (in which the female partner works an average of 1,508 hours per year). When the costs of parental time devoted to child care and domestically produced goods are added in, the value of both children's consumption allocation is estimated to be around 51 to 56 per cent of total consumption in the first kind of household, and around 49 to 54 per cent in the second. Because the definition of child costs in this study is much broader than that in most of the literature, which considers only the child's share in consumption of market goods, these estimates of child costs are also substantially higher than those typically reported.

The approach also makes it possible to distinguish between the transfers to children made by each adult in the household, differentiated as to the extent of the female

²³ As in Gronau (1991). Thus a parent's personal preference-ordering over bundles containing beer and nappies is assumed not to change after the arrival of a child, but a household's expenditure pattern over these bundles certainly will, because the parents assign a weight in the household utility function to the utility of the child.

²⁴ The sample consists of households with two children and both parents present.

partner's participation in the labour market. Taking the male's consumption as the benchmark, when attention is restricted to market goods, the estimates of a child's consumption in a traditional household are 24 to 40 per cent of that of the adult male. In non-traditional households the estimates rise to 53 to 69 per cent. One explanation for this is that market purchases of 'bought-in' child care as well as transfers to the child from the employed mother are both higher. When the values of domestic production and parental child care are added into the child's consumption, the figures rise to 82 to 98 per cent of the total consumption of the adult male in traditional households, and 78 to 91 per cent in non-traditional households. These results underline the need to incorporate the time costs of child care and domestic work which, as the time-use data show, are largely those of the female partner.

In the absence of data on domestic outputs and individual consumptions of market and household goods, alternative sets of assumptions to those made in this study could clearly be made, and these would lead to different results. The point of the study was not to provide the definitive measure of child costs, but rather to suggest how child costs should be defined, and to set out a modelling methodology which can allow the costs so defined to be estimated. The variables that appear in the theoretical model²⁵ are all in principle observable. In this discussion we have tried to make clear the importance of recognising the assumptions that have to be made in the light of the limitations of the available data, to identify the key variables on which data are missing, and to stress the need for data on these missing variables. Further progress will be made only by confronting the problem that any set of assumptions is to some extent arbitrary, and that estimates of child costs (or equivalence scales) using existing datasets must be based, implicitly or, as here, explicitly, on such a set of assumptions.

4.4 Empirical work on the collective model

Here we discuss five of the leading contributions to the empirical literature on the collective model.²⁶ We do not believe that this sample selection poses a serious problem of bias.

Of the papers that apply the collective model using expenditure survey data, we focus here on Browning *et al.* (1994) and Browning and Chiappori (1998). In both papers, the data relate to samples of households in which both adults work full-time and there are no children present, as well as to households consisting of single persons working full-time. Effectively, therefore, labour supplies are taken as exogenously fixed,²⁷ so that variation in commodity demands across households is taken to be unrelated to

²⁵ As set out in chapter 3.

²⁶ For a comprehensive survey of the pre-2002 literature see Vermeulen (2002).

²⁷ This is done by including in the sample only households where both partners work 'full-time'. But 'full-time' work nevertheless involves a not insignificant variance in hours worked, especially for men, as the data presented in chapter 1 show. Note that the study by Browning and Meghir (1991) shows convincingly that the assumption of separability between labour supplies and consumption demands is decisively rejected empirically, and so this assumption could not be used to achieve effective exogeneity of labour supplies.

wage rates as such, which also do not enter the household sharing rule. The latter depends only on a price vector for consumption goods as well as total household expenditure and a vector of EEPs. In Browning *et al.*, however, there is a single cross-section of households, which, it is assumed, all face identical prices, and so only total expenditure and selected EEPs enter as independent variables in the estimations. No explicit theoretical rationale is given for why commodity prices rather than wage rates should play such an important role in determining the household income distribution, but the choice of whether to base the analysis on goods rather than labour markets makes some such assumption inevitable, given the nature of available datasets. If of course wages do enter the sharing rule, as in the earlier theoretical papers, since they can be expected to vary across households, this could be an important source of omitted variable bias.

In Browning *et al.* the aim of the paper is both to test the unitary and collective models and, more ambitiously, to estimate the partial derivatives of the household sharing rule. It is first shown that, in this setting, the partial derivatives of the sharing rule with respect to total expenditure and an EEP, say μ and z respectively, can be identified from empirical estimates of the demand functions for an assignable good, x . This is a good whose individual consumptions x^i , $i = f, m$, can be identified in the data and, just as importantly, for which it can be assumed that x^i generates utility only for spouse i , and not for the other. That is, it must be possible to write i 's demand for x^i as²⁸

$$x^i = D^i(s^i(z, \mu)) \quad (4.21)$$

where $s^i(z, \mu)$ is i 's share in total expenditure, i.e. the sharing rule, with $\sum_{i=f,m} s^i(z, \mu) = \mu$. On the usual assumptions underlying the collective model, in particular that this sharing rule exists,²⁹ it is straightforward to show that, given that the condition in (4.28) is satisfied, the four partial derivatives s_z^i, s_μ^i can be identified from empirical estimates of the functions $D^i(s^i(z, \mu))$. Thus differentiating these demand functions gives the system

$$D_z^i = D_{s^i}^i s_z^i \quad i = f, m \quad (4.22)$$

$$D_\mu^i = D_{s^i}^i s_\mu^i \quad i = f, m \quad (4.23)$$

from which we obtain the two equations

$$\alpha_i = \frac{s_z^i}{s_\mu^i} \quad i = f, m \quad (4.24)$$

where $\alpha_i \equiv D_z^i / D_\mu^i$ can be identified from the parameter estimates of the empirical demand system.

²⁸ Note that prices can be suppressed in these demand functions since they are assumed constant across households. There would be a problem, however, if the demand functions contained wage rates, which do vary across households.

²⁹ Equivalent to assuming some GHWF, though in Browning *et al.* (1994) it is not specified explicitly.

Note furthermore that we have the two identities

$$\sum_{i=f,m} s_z^i = 0 \quad (4.25)$$

$$\sum_{i=f,m} s_\mu^i = 1. \quad (4.26)$$

Thus we have the linear equation system

$$\begin{bmatrix} 1 & 0 & -\alpha_f & 0 \\ 0 & 1 & 0 & -\alpha_m \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} s_z^f \\ s_z^m \\ s_\mu^f \\ s_\mu^m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad (4.27)$$

which can be solved uniquely for the partial derivatives of the sharing rule if and only if the determinant of the left-hand-side matrix, which is equal to $\alpha_f - \alpha_m$, is non-zero. Thus we have the necessary and sufficient condition for identifiability

$$D_z^f / D_\mu^f \neq D_z^m / D_\mu^m. \quad (4.28)$$

In the empirical work, the two goods chosen for demand estimation are men's and women's clothing. As the paper recognises, this requires that each spouse does not have preferences towards or participate in the decision about the clothing the other wears, so that the demands are functions only of own income shares, and notes that '[m]any readers will be thoroughly skeptical about this'. Moreover, the issue of sample selection bias created by estimating the demands on a sample of fully employed couples with no children present in the household is explicitly put to one side.

The main results of the paper are that the collective model (in this quite general form) is not rejected for couples,³⁰ while, on the sharing rule, the paper concludes that '[t]he only factors that seem to affect sharing within the household are the differences in ages and incomes of the members and the . . . total expenditure', which are statistically highly significant and have fairly substantial effects. Age differences and individual incomes are taken as EEPs.

There is, however, an important conceptual issue arising out of taking individual incomes as EEPs, which does not appear to have been perceived by the authors of this paper, perhaps because labour supply is essentially regarded as exogenous. If labour supply is in fact endogenous, then so is individual income, and so at least one of the EEPs in the model is not actually exogenous. The paper by Basu (2006) shows that in the case where individual incomes enter as conditioning variables in the GHWF, the model becomes quite different from the collective model.³¹ The reason is quite obvious: if an individual knows that her share in household income will, other things

³⁰ Since there is no price variation in the data, symmetry of the Slutsky matrix cannot be tested. This is the main purpose of Browning and Chiappori (1998), discussed below.

³¹ Even though in fact Basu takes the weighted utilitarian form of GHWF as the basis for his analysis.

being equal, increase with her individual labour income, which is endogenous, then this will obviously affect her choice of labour supply.³² Thus the sharing rule becomes a matter of strategic interaction within the household, and, as Basu shows, analysis of the resulting game becomes quite intricate.

The paper by Browning and Chiappori (1998) derives and tests restrictions on observable demand functions again for a collective model in which labour supplies are assumed to be exogenous. No attempt is made to identify parameters of a sharing rule. The dataset in this paper has data on prices which vary across households both spatially and over time, and this allows the derivation and testing of restrictions on compensated price effects which provide a sharp test of the unitary v. collective models.

Individual utilities are defined on the vectors of individual consumptions x^i , $i = f, m$, and a vector of household public goods X . The household is explicitly assumed to maximise a weighted utilitarian GHWF of the form

$$H = \alpha(p, \mu)u^f(x^f, x^m, X) + [1 - \alpha(p, \mu)]u^m(x^f, x^m, X)$$

subject to the pooled budget constraint $p(x^f + x^m + X) \leq \mu$. Thus the model allows for externalities between consumptions of the household members as well as for ‘non-egoistic’ preferences. The weighting function $\alpha(\cdot) \in [0, 1]$ is assumed to be differentiable and linearly homogeneous in the price vector p and exogenous total expenditure μ , though no restrictions are placed on the signs of its derivatives. The function is interpreted as a measure of the ‘power’ f has in the consumption decisions of the household, and it is left as an empirical matter to determine how changes in particular prices and total expenditure would affect this power.³³

The solution of the household’s problem gives Marshallian demand functions for total household consumptions (in general only aggregate household consumptions are empirically observable), which can be written as $D_j(p, \mu, \alpha(p, \mu))$, $j = 1, \dots, n$. Note that the partial derivatives $\partial D_j / \partial p_k$, $\partial D_j / \partial \mu$, $j, k = 1, \dots, n$, *since they correspond to constant α* , will have all the standard properties of Marshallian demands in the individual demand model: for constant α , H has all the properties of an individual utility function. Thus, *for constant α* we can define the $n \times n$ matrix $\Sigma = [\sigma_{jk}]$, where

$$\sigma_{jk} = \frac{\partial D_j}{\partial p_k} + D_j \frac{\partial D_j}{\partial \mu}$$

and, by analogy with the Slutsky matrix in the standard consumer case, Σ is symmetric and negative semidefinite. We can also define the n -vector of derivatives $[\partial D_j / \partial \alpha]$, giving the effects on household demands of a change in f ’s ‘power’ within the household.

³² For example, do women go out to work because it will give them more ‘say’ within the household? This, Basu’s hypothesis, is actually very plausible.

³³ Note that axiom 3 in the paper, which introduces this GHWF, implies axiom 2, which states that the household seeks to achieve a Pareto-efficient allocation, and so makes the latter axiom redundant.

Empirically, however, we cannot observe the demand functions $D_j(\cdot)$ but only the reduced form functions $\hat{D}_j(p, \mu)$, where we have the identities

$$\hat{D}_j(p, \mu) = D_j(p, \mu, \alpha(p, \mu)) \quad j = 1, \dots, n$$

implying for any j, k

$$\begin{aligned} d\hat{D}_j &= \frac{\partial \hat{D}_j}{\partial p_k} dp_k + \frac{\partial \hat{D}_j}{\partial \mu} d\mu \\ &= \frac{\partial D_j}{\partial p_k} dp_k + \frac{\partial D_j}{\partial \mu} d\mu + \frac{\partial D_j}{\partial \alpha} \left(\frac{\partial \alpha}{\partial p_k} dp_k + \frac{\partial \alpha}{\partial \mu} d\mu \right) \\ &= dD_j \quad j, k = 1, \dots, n. \end{aligned}$$

Now we know that we can compensate for a sufficiently small price change dp_k , holding utility constant, if we change expenditure by $d\mu = \hat{D}_k dp_k = D_k dp_k$. Substituting into the above equation therefore gives

$$\frac{\partial \hat{D}_j}{\partial p_k} + \hat{D}_k \frac{\partial \hat{D}_j}{\partial \mu} = \sigma_{jk} + \frac{\partial D_j}{\partial \alpha} \left(\frac{\partial \alpha}{\partial p_k} + D_k \frac{\partial \alpha}{\partial \mu} \right) \quad j, k = 1, \dots, n$$

The values on the left-hand side of this equation can be calculated from empirical estimates of the system of demand functions $\hat{D}_j(p, \mu)$. The right-hand side places a testable restriction on the $n \times n$ matrix of values, which we denote by S , that we thereby obtain. Browning and Chiappori call this the *pseudo-Slutsky matrix*. If we denote the term $(\partial \alpha / \partial p_k + D_k \partial \alpha / \partial \mu)$ by v_k , we can write this equation as

$$S = \sum + \begin{bmatrix} v_1 \frac{\partial D_1}{\partial \alpha} & v_2 \frac{\partial D_1}{\partial \alpha} & \cdot & \cdot & v_n \frac{\partial D_1}{\partial \alpha} \\ v_1 \frac{\partial D_2}{\partial \alpha} & v_2 \frac{\partial D_2}{\partial \alpha} & \cdot & \cdot & v_n \frac{\partial D_2}{\partial \alpha} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ v_1 \frac{\partial D_n}{\partial \alpha} & v_2 \frac{\partial D_n}{\partial \alpha} & \cdot & \cdot & v_n \frac{\partial D_n}{\partial \alpha} \end{bmatrix}$$

It is clear that the second matrix on the right-hand side has rank 1, since any two columns or rows are linearly dependent.³⁴ Thus Browning and Chiappori define the testable restriction for the collective model as the condition that the matrix S be the sum of a symmetric negative semidefinite matrix and a matrix of rank 1. If the household's behaviour were equivalent to the standard consumer model, on the other hand, this latter matrix would not exist, and S would be symmetric and negative semidefinite.

The intuition for this result is straightforward. If the problem is solved with the weight α held constant, we would obtain the standard Slutsky matrix, since this is a problem with a non-price-dependent HWF, and Samuelson's theorem then applies. To the compensated demand effect for any one price change must, however, then be

³⁴ For example, just multiply the second column by $-v_1/v_2$ and add it to the first.

added the compensated effect of that price change on demand via its effect on the distributional weight α and the corresponding household distribution of income. In other words, in this model, the effects of a price change are decomposable into the standard effects on individual demands resulting from a change in the budget constraint, and the effects on demands resulting from the change in the household's preferences over utility profiles.

Browning and Chiappori report the results of estimating this model on demand data for single-person households and for two-person households without children. All individuals in both types of households are in full-time employment. As we have argued elsewhere, for such households domestic production is unlikely to be very important. Their main finding was that whereas the Slutsky symmetry restriction was not rejected for singles, it was rejected for couples, while the theoretical restriction on the matrix S was not rejected for couples. These are intuitively very satisfying results.

We now turn to the three papers that are concerned with labour supply models. The paper by Fortin and Lacroix (1997) attempts to test the 'unitary' and 'collective' models and estimate parameters of the sharing rule by estimating labour supply functions. A very carefully written theoretical section of the paper makes clear that by a 'unitary model' they mean one based on a household utility function derived from the type of HWF in (4.4), while a 'collective model' is one based on a GHWF with wage rates and non-wage income as conditioning variables. They do not make the mistake, common in this literature, of claiming that the collective model 'assumes nothing more than Pareto efficiency'. Since they do not place the restriction that the GHWF must take the weighted utilitarian form, the results for their 'collective model' are also consistent with models with any other specific form of GHWF.

They exclude household production from their model, and use a sample containing only households in which both spouses have a significant market labour supply³⁵ and there is no more than one child present. This means that they could have, more or less, a case in which Chiappori's model would claim that the partial derivatives of the sharing rule can be identified – ideally there should be no children present, but this would have made their sample too small. The results for these sharing rule parameter estimates are quite poor, with standard errors very large relative to parameter values, and only one parameter (out of eight) significantly different from zero at the 10 per cent level. We would argue that this is not surprising, given that the exclusion of household production implies that the model is basically misspecified.³⁶

The implications of unitary and collective models respectively can be expressed in terms of restrictions on parameter values in the estimated labour supply equations, and the results reject anonymity ('pooling') and symmetry almost entirely. The restrictions

³⁵ They point out that Chiappori's formulation assumes that both partners have positive and continuously varying labour supplies, which is why they exclude households with a non-participating partner. They acknowledge the issue of sample selection bias and do their best to deal with it.

³⁶ In the previous section we showed how estimating a model that omits household production, when this is in fact a significant use of the non-market time of household members, leads to problems of misspecification and omitted variables, resulting in biased and inconsistent parameter estimates. In particular, labour supply elasticity estimates were much larger than those derived from a model that takes household production into account.

of the collective model are not in general rejected, except in the important case of young couples with a pre-school child, where they are rejected even at the 10 per cent level. Since the data show that household production is particularly important in households with pre-school children (see, for example, the data presented in chapter 1), we would again interpret this as rejecting the hypothesis, implicit in their model, that household production can be ignored. Overall, we would conclude that even when an empirical analysis is as carefully conducted as this one, the omission of household production is an important source of specification error.

The papers by Donni (2007) and Blundell *et al.* (2007) are both concerned with extending Chiappori's original model to the case in which there is some kind of corner solution in one individual's labour supply. Donni's study uses data from the French labour market, a characteristic feature of which is that, due to regulation of working hours, male workers are taken to be rationed at the maximum number of hours they are allowed to work. Only female labour supply varies continuously up to the rationing limit. Donni makes an interesting extension to Chiappori's theoretical results to show that in this case, where the male labour supply function cannot be used to identify the partial derivatives of the sharing rule, using the wife's labour supply function together with one consumption demand equation gives an alternative procedure.

As in the case of Fortin and Lacroix, Donni uses a dataset of couples with at most one child, since taking couples with no children present would have led to an unacceptably small sample size. Significance levels of the estimated structural coefficients are very low, with only 19 out of 49 structural parameters significant at the 10 per cent level, and none of the parameters of the estimated sharing rule significant.³⁷ On the other hand, the effect of her wage on the wife's labour supply is significant and large, with a labour supply elasticity of 2.7, which the author notes is very large compared to the majority of estimates obtained from other studies, but explicable in terms of the particular nature of the French labour market.

In Blundell *et al.* the corner solution is, so to speak, at the opposite end of the interval from that considered by Donni: men are regarded as either working not at all or full-time, while women's working hours are a continuous variable. The object is to extend the collective model to the case in which males make a discrete choice between zero hours and full-time working. As Donni points out, in commenting in his own paper on Blundell *et al.*, it is rather a strong assumption to make that the small proportion of men not working all do so by choice (as is required by the Chiappori framework) and are not in some way rationed on the labour market. The sample on which estimation is based consists only of UK households with no children present, in which the female is aged at least 35, and both male and female are under 60, and this accounts for less than 20 per cent of all couples. As our time-use data presented in chapter 1 showed, these are couples for which there is considerable female labour supply heterogeneity, but where time spent in household production is, despite the absence of children, still substantial. The authors summarise their findings³⁸ as:

³⁷ Table 4, part (a). ³⁸ Blundell *et al.* (2007), p. 442.

The data are not at odds with the collective model, but the unitary model is not rejected either. Once we estimate the collective model we find the female wage elasticity is about 0.30, which is very close to the figure that earlier UK studies have found . . . Moreover, we find that the level of male consumption is sensitive to male wages, employment and other income. He gets to consume 81% of the increase in his earnings and 73% of an increase in the female wage and other income . . . The effect of the female wage on the sharing rule is very badly determined preventing any conclusions about its effect on her consumption. Generally some of the results have proved fragile mainly because of the relatively low number of observations when the male does not work.

4.5 Conclusions

Our basic argument in this chapter has been that satisfactorily specified models of household labour supplies face the difficulty that conventional datasets do not provide the information necessary to apply these models empirically. In particular, data on domestic outputs and inputs and individual consumptions of market and household goods are missing. Models of labour market behaviour that, implicitly, react to this problem by adopting specifications that assume that households allocate no time to household production provide results of limited or doubtful validity. In particular, the overall impression that we gain from studies attempting to estimate parameters of household sharing rules is that, despite the fact that samples are selected in a way most favourable to the possibility of identifying the sharing rule, and that they are carried out by extremely competent econometricians, the results are very weak and certainly do not in our view constitute a reliable basis for the formulation of public policy. Progress will only be made when we have more comprehensive datasets.

5 Labour supply, consumption and saving over the life cycle

5.1 Introduction

As we discussed in chapter 1, for many purposes in the analysis of public policy it is necessary to consider households over their entire life cycle. In this chapter, we first survey the existing literature on life cycle models, examining how extensions to the basic Fisherian model of intertemporal consumption choice have been developed to try to explain why consumption tends to track income to a much greater extent than is predicted by this model. We then go on to widen the scope of the discussion by developing a model of the ‘family life cycle’, which seeks to integrate the analysis of consumption and saving with that of time allocation in multi-person households. We present a simple formal model and then go on to combine time-use data with data on consumption and saving to present an empirical picture of the life cycle that differs considerably from that on which the models in section 5.2 are based. It emphasises the effect of fertility decisions on female labour supply as a major determinant of changes in household income, consumption and saving over time.

5.2 Life cycle models

The basic microeconomic foundations for models of consumer saving and consumption over the life cycle are provided by the neo-classical model of intertemporal choice formulated by Irving Fisher.¹ This can be viewed as a version of the standard initial endowments consumer model introduced in chapter 2. The goods vector, $x = [x_1, x_2, \dots, x_T] \in \mathbf{R}_+^T$ is defined as a time stream of Hicksian composite consumption goods, with x_t , $t = 1, 2, \dots, T$ consumption at date t and T the last date in the consumer’s lifetime. The endowments vector is $\bar{x} = [\bar{x}_1, \bar{x}_2, \dots, \bar{x}_T] \in \mathbf{R}_+^T$, taken for the moment to be exogenous. The consumer has access to a capital market, on which she is able to exchange a claim to consumption at any date t for one at any other date s .

¹ See Fisher (1906), (1907), (1930). For an excellent exposition of the purely consumption aspects of Fisher’s work, with extensions to macroeconomic applications such as the the Life Cycle Hypothesis and the Permanent Income Hypothesis, see Deaton and Muellbauer (1980). Deaton (1992) and Browning and Lusardi (1996) give valuable surveys of more recent developments as well as good overviews of the empirical work.

The usual way of describing the institutional set-up that allows this is to say that at any time t the consumer may buy or issue bonds that pay one unit of the consumption good at time $t + 1$, with r_t the one-period (real) interest rate on such a bond. b_t denotes the number of bonds the consumer issues ($b_t > 0$, she borrows) or buys ($b_t < 0$, she lends), and the capital market is perfect, in the sense that r_t is the same for borrowing and lending and invariant to the amount borrowed or lent, and to the identity of the consumer.

The consumer knows with certainty her endowment vector and future interest rates. Then her single-period budget constraint is given by

$$x_t \leq \bar{x}_t + b_t - (1 + r_{t-1})b_{t-1} \quad t = 1, 2, \dots, T \quad (5.1)$$

with, if we assume no inheritances and bequests

$$b_0 = b_T = 0. \quad (5.2)$$

It is a common, though not essential, simplification to assume that the interest rate is constant over time, $r_t = r$, all t . Then, defining the *market discount factor*

$$\delta^t = (1 + r)^{-t} \quad (5.3)$$

the sequence of single-period budget constraints can be collapsed into the single wealth constraint

$$\sum_{t=1}^T \delta^{t-1} x_t \leq \sum_{t=1}^T \delta^{t-1} \bar{x}_t \equiv W \quad (5.4)$$

where W is the value of the consumer's wealth. Defining $p_t = \delta^{t-1}$ as the price of a unit of consumption at date t in terms of consumption at date 1 (the numeraire), and introducing the utility function $u(x)$, with all the usual properties, the consumer's optimisation problem is

$$\max_x u(x) \quad \text{s.t.} \quad \sum_{t=1}^T p_t x_t \leq W. \quad (5.5)$$

Since this is a completely standard problem, all the usual results hold (see chapter 2). Interest in the model therefore focuses on the interpretations we can make about the consumer's behaviour over time.

The first point to note, which is an assumption rather than a result of the model, is that the consumer is planning in terms of her entire lifetime: she chooses a lifetime consumption stream in the light of her lifetime income endowment stream. Given the information she is assumed to possess, both about her future endowments and about future prices or interest rates, and the possibilities for exchange provided by the perfect capital market, anything else would hardly seem rational. Hence the term 'life' in the expression 'life cycle model'.

A more substantive point is that in the Marshallian consumption demand functions derived from this model, $x_t(p, W)$, where $p = (1, p_2, \dots, p_T) \in \mathbf{R}_+^T$, all endowment

vectors $\bar{x}$ that have the same discounted value W , i.e. that imply the same wealth, imply the same optimal consumption time stream. This is where the model starts to bite in terms of the life *cycle*.

One way of defining what the life cycle might mean in this model is in terms of the form taken by the endowment vector $\bar{x}$. If we think of date 1 as the beginning of the consumer's working life, and date $R \leq T$ as the date of her retirement, it could be plausible to think of successive values of $\bar{x}_t$ as at first increasing, reaching a maximum in 'middle age', falling slowly thereafter until R , and then falling sharply to zero. Although we are for the moment begging questions about the individual's labour supply, and the role this may have to play in the model, this seems to be a plausible picture of an average person's 'life cycle' earnings.

The interesting point, however, is that, in this model, this picture is quite irrelevant. Any other pattern of endowed incomes that yields the same wealth will give rise to the same consumption time stream. In other words, the consumer is able to decouple her lifetime consumption time stream from her endowed income time stream by use of the capital market. This is the single most important, and contentious, result of this model.

We can illustrate this point more sharply if we make a common assumption about the structure of the consumer's intertemporal preferences. Suppose they take the additively separable form

$$u(x) = \sum_{t=1}^T \rho^{t-1} v(x_t) \quad v' > 0, v'' < 0 \quad (5.6)$$

where $\rho \leq 1$ is the so-called felicity discount factor, and is defined as $\rho = (1 + i)^{-1}$, where $i \geq 0$ is the felicity interest rate, and has the interpretation that it is the subjective interest rate the consumer has to be paid to induce her to postpone a unit of utility for one period. The consumer's (strictly concave) utility function of consumption in every period is the same, but she discounts later utilities exponentially relative to earlier ones.

Then the first-order conditions for the maximisation problem imply

$$\frac{v'(x_{t+1}^*)}{v'(x_t^*)} = \frac{\delta}{\rho} \quad t = 1, 2, \dots, T-1. \quad (5.7)$$

Given strict concavity of the utility function, we have the results

$$\rho = \delta \Leftrightarrow x_{t+1}^* = x_t^* \quad t = 1, 2, \dots, T-1 \quad (5.8)$$

$$\rho < \delta \Leftrightarrow x_{t+1}^* < x_t^* \quad t = 1, 2, \dots, T-1 \quad (5.9)$$

$$\rho > \delta \Leftrightarrow x_{t+1}^* > x_t^* \quad t = 1, 2, \dots, T-1. \quad (5.10)$$

Thus the consumption path will either be perfectly flat, monotonically falling or monotonically rising, depending only on the relative values of the two exogenous parameters, and regardless of the specific shape of the endowed income stream. Note that this result also depends on the assumed constancy of the market interest rate.

For the intuition on these results, we express them in terms of the relationship between the market interest rate, the reward for postponing consumption, and the felicity interest rate, the required reward for postponing utility, a measure of ‘impatience’

$$i = r \Leftrightarrow x_{t+1}^* = x_t^* \quad t = 1, 2, \dots, T-1 \quad (5.11)$$

$$i > r \Leftrightarrow x_{t+1}^* < x_t^* \quad t = 1, 2, \dots, T-1 \quad (5.12)$$

$$i < r \Leftrightarrow x_{t+1}^* > x_t^* \quad t = 1, 2, \dots, T-1. \quad (5.13)$$

If the market insufficiently rewards the consumer for postponing utility, she will consume more of her wealth sooner, if it rewards her more than sufficiently, more of it later, and when reward just offsets required return she maintains constant consumption.

Finally, a useful characterisation of the optimal path can be given if we denote ρ/δ by $\hat{\rho} \geq 1$ and call it the ‘net felicity discount factor’. Then the first-order conditions yield

$$\hat{\rho}^{t-1} v'(x_t^*) = \lambda \quad t = 1, 2, \dots, T \quad (5.14)$$

where λ is the consumer’s marginal utility of wealth, constant over time. Then she chooses her consumption time stream to maintain constancy of the marginal utility of consumption discounted at the net felicity discount factor.

Figure 5.1 gives a picture of these possible optimal consumption time streams over the life cycle, in relation to a stylised graph of the endowed income stream, labelled $\bar{x}\bar{x}'$. This is the sort of picture predicted by the standard model of intertemporal consumption choice.

A problem,² often also referred to as the ‘excess sensitivity puzzle’, is that when empirical data on consumption and income are compared, the consumption time stream has a similar profile to the income stream, and appears to track it over the earlier years, as shown by the broken curve CC in the figure. This suggested that the simplifications underlying the model need to be reconsidered, and their generalisations tested to see if they lead to modifications of the predicted profile which help it to fit the data.

The leading candidates for the model’s generalisation are:

- *endogenous income*: far from being exogenous, income will be generated by the consumer’s labour supply, choices of which should be analysed explicitly;
- *capital market imperfections*: even casual observation of real-world capital markets shows they are far from satisfying the requirements for perfection;
- *uncertainty*: clearly consumers do not possess such complete information about future interest rates and endowments as is assumed in this model;
- *demographic factors*: most households do not consist of single individuals, but rather, as we discussed extensively earlier in this book, of families, and the size and composition of the family changes systematically over time – indeed, the idea of the life

² First pointed out by Thurow (1969). Strictly speaking it is an empirical rejection of the model set out in (5.1)–(5.5).

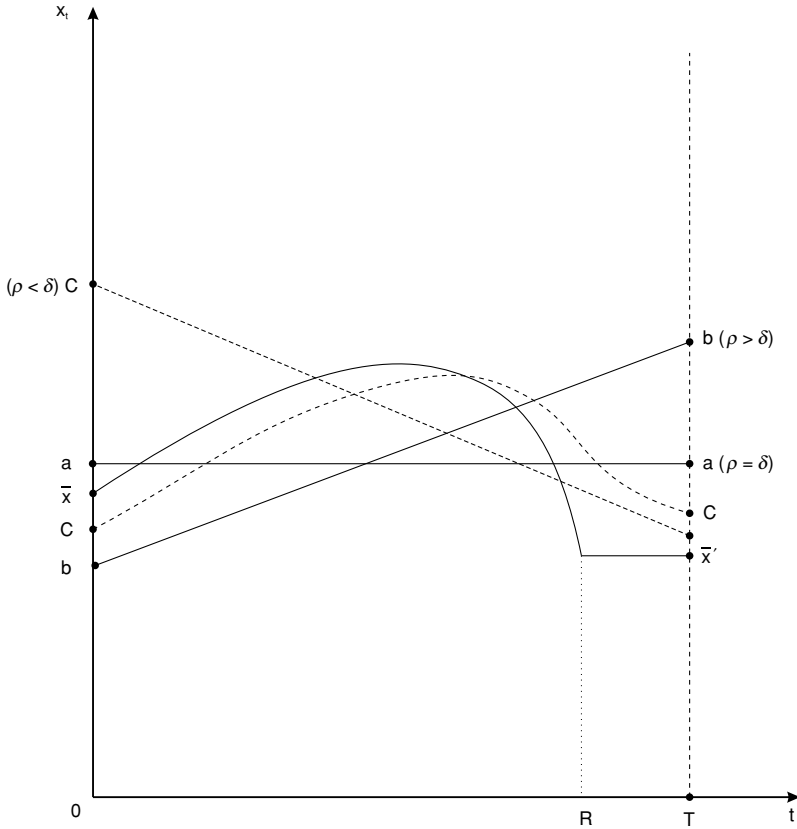


Figure 5.1 Life cycle income and consumption paths

cycle may be better expressed in terms of these changes than in terms of the time profile of incomes.

In the next four subsections, we consider one by one the extensions to the basic model that have been made to take account of these aspects of reality, with a focus on what they imply about the relationship between income and consumption over time.³

5.2.1 Endogenous labour supply

It is straightforward to extend the model to incorporate endogenous labour supply.⁴ Let $l_t \geq 0$ denote the consumer's labour supply at date t , μ_t her non-wage income (for example, a transfer from the state, but *not* income arising from the capital market) and w_t her wage rate, with all wage rates and non-wage incomes known with certainty.

³ This is done within the basic framework of lifetime optimisation just set out. It is of course possible to reject this framework itself, on the grounds that consumers are not so unboundedly rational. See Browning and Lusardi (1996) and the literature cited there for a discussion of 'behavioural' approaches to intertemporal decision-taking.

⁴ The original analysis of this, in a continuous time setting, was carried out by Heckman (1974).

Write the utility function now as $v(x_t, l_t)$, assumed strictly concave and with $v_1 > 0$, $v_2 < 0$. Then the maximisation problem is

$$\max_{x_t, l_t} \sum_{t=1}^T \rho^{t-1} v(x_t, l_t) \quad \text{s.t.} \quad \sum_{t=1}^T \delta^{t-1} x_t \leq \sum_{t=1}^T \delta^{t-1} (w_t l_t + \mu_t). \quad (5.15)$$

Now suppose that the consumption good and labour supply (or equivalently leisure) are separable, in the sense that variations in one have no effect on the marginal utility of the other, so that the utility function can be written as $u(x_t) - v(l_t)$. Then the first-order conditions for solution of the problem imply

$$\frac{v_1(x_{t+1}^*)}{v_1(x_t^*)} = \frac{\delta}{\rho} \quad t = 1, 2, \dots, T-1. \quad (5.16)$$

Thus in respect to the consumption time stream we see that it is characterised in exactly the same way as in the basic model. Although of course the precise solution for the optimum consumption time stream will be determined by the wage vector $w = (w_1, w_2, \dots, w_T)$ and non-wage income vector $\mu = (\mu_1, \mu_2, \dots, \mu_T)$, i.e. the demand functions are now $x_t(p, w, \mu)$, nevertheless because of the separability in the utility function the general shape of the time path of consumption is exactly as it was when income was exogenous, and is determined entirely by the relation between the exogenous parameters δ and ρ , in conjunction with the curvature of the utility function. The wage rate and labour supply determine income, but because of the separability of the utility function this is essentially exogenous to the determination of the time profile of consumption. On the other hand, this assumption is rather special, and we turn now to the more general case of non-separability.

In this case, imposing the constraint $l_t \geq 0$ so as to be able to endogenise the retirement decision, the first-order conditions give

$$\rho^{t-1} v_1(x_t^*, l_t^*) - \lambda \delta^{t-1} = 0 \quad (5.17)$$

$$\lambda \delta^{t-1} w_t - \rho^{t-1} v_2(x_t^*, l_t^*) \leq 0 \quad l_t^* \geq 0 \quad l_t^* [\lambda \delta^{t-1} w_t - \rho^{t-1} v_2(x_t^*, l_t^*)] = 0 \quad (5.18)$$

and the time paths of labour supply and consumption are jointly determined. In particular, we have

$$\frac{v_1(x_{t+1}^*, l_{t+1}^*)}{v_1(x_t^*, l_t^*)} = \frac{\delta}{\rho}. \quad (5.19)$$

This tells us first that in the retirement phase,⁵ when $l_t^* = l_{t+1}^* = 0$, the time path of consumption is again determined as in the basic model, but in the working phase, when $l_t^* > 0$, the time path of consumption will be influenced by the time path of labour supply, which in turn will be determined by the market interest rate, wage rates and non-wage incomes, i.e. we have $l_t(p, w, \mu)$.

⁵ Note that in this model the date of retirement R can be thought of as determined endogenously, as the time at which condition (5.18) is satisfied as an inequality at the point $l_R^* = 0$. Of course we need some assumption on wage rates to ensure that labour supply remains at zero thereafter.

We must now characterise the life cycle in terms of the exogenous wage rate. Suppose that the wage rate is increasing through the first phase of the working life, flattens out in middle age, and then falls. If labour supply is non-decreasing in the wage (and abstracting from the time path of μ_t), labour supply and wage income will follow a path with the same general shape. Taking the condition in (5.19), we can then ask: given the relation between l_{t+1}^* and l_t^* , what must be true about the form of the utility function to yield from this condition a similar qualitative relationship between the values of x_{t+1}^* and x_t^* ? This condition tells us what we have to assume to obtain the result that consumption tracks income over the life cycle, or at least its pre-retirement part.

To explore this question, write (5.19) as

$$v_1(x_{t+1}^*, l_{t+1}^*) = \frac{\delta}{\rho} v_1(x_t^*, l_t^*) \quad (5.20)$$

and assume that the changes in labour supply and consumption are sufficiently small that we can accept the Taylor series approximation

$$v_1(x_{t+1}^*, l_{t+1}^*) = v_1(x_t^*, l_t^*) + v_{11}(x_t^*, l_t^*)[x_{t+1}^* - x_t^*] + v_{12}(x_t^*, l_t^*)[l_{t+1}^* - l_t^*]. \quad (5.21)$$

Then inserting this into (5.20) and rearranging gives

$$[x_{t+1}^* - x_t^*] = \left(\frac{\delta}{\rho} - 1 \right) \frac{v_1}{v_{11}} - \frac{v_{12}}{v_{11}} [l_{t+1}^* - l_t^*] \quad (5.22)$$

Since $v_1 > 0$, $v_{11} < 0$, we have that, for the changes in labour supply and consumption to have the same sign:

- (i) if $\rho = \delta$ it is necessary and sufficient
- (ii) if $\rho > \delta$ and $[l_{t+1}^* - l_t^*] > 0$ it is sufficient though not necessary
- (iii) if $\rho < \delta$ and $[l_{t+1}^* - l_t^*] > 0$ it is necessary but not sufficient that $v_{12} > 0$.

The intuition is easy to see, particularly if we take the case of $\rho = \delta$, in which case, in the basic model, the optimal consumption path would be flat. As the wage rate rises over the life cycle, labour supply increases, as does income, and the marginal utility of consumption increases ($v_{12} > 0$), causing an increase in consumption demand. Thus income and consumption move together, caused by the movement in the wage. If $\rho > \delta$, consumption would be rising anyway and so this reinforces the effect of the increasing wage, while if $\rho < \delta$ consumption would be falling and this tends to offset that effect. If the wage and labour supply are falling, $[l_{t+1}^* - l_t^*] < 0$, and the words ‘necessary’ and ‘sufficient’ have to be interchanged in (ii) and (iii).

The consensus view in the life cycle literature seems to be, however, that changes in labour supply in response to changes in the wage are too small to generate the closeness of the observed correlation between income and consumption. Clearly, if $[l_{t+1}^* - l_t^*] \approx 0$, as the empirical evidence on *male* labour supply would suggest, then the role of labour supply changes in determining this correlation could be ignored. On the other hand, as we see in the later part of this chapter, female labour supply may

be subject to quite dramatic changes between periods, with possibly large effects on family income.

An interesting attempt to revive this line of argument was made by Baxter and Jermann (1999) who, drawing on the macroeconomic business cycle literature,⁶ emphasise the importance of household production. They rightly point out that the life cycle literature is framed entirely in terms of the consumption expenditure on market goods and the income derived from market labour supply. But market consumption is only one component of *full consumption*, defined as the sum of the values of consumption of market goods and goods produced within the household. It is quite possible that full consumption behaves in the way predicted by the standard model, being decoupled from income by use of the capital market, while its individual components do not. In particular, as the wage rate rises over the first part of the life cycle, the implicit price of household goods rises relative to that of market goods, causing a substitution towards market consumption. This will be reversed as the wage rate falls later in the cycle. Thus we observe income and market consumption moving closely together, even though movements in full consumption may not be associated with movements in income.

As we make clear later in this chapter, although it is certainly important to introduce household production and full consumption into the model, the mechanism suggested by Baxter and Jermann is unconvincing when we take account of two-person households. The time-use data show that most household production is undertaken by the female partner, whose wage rate does not show the same life cycle pattern as that of the male. Female wage rates tend to stay constant or fall over the life cycle. As we will show, taking account of the real nature of the household and incorporating time-use data into the empirical model leads to a quite fundamentally different view of the life cycle than that considered up until now. For the moment, however, we continue with the proposed extensions to the standard model.

5.2.2 *Imperfect capital markets*

The formal extension of Fisher's model to take account of capital market imperfections was made by J. Hirshleifer (1958). There is clearly a range of possibilities in modelling an imperfect capital market. As a minimum, we would assume that the interest rate on saving is below that on borrowing, each of which is, however, still invariant to the amount borrowed or lent. Hirshleifer dealt with this case, as well as that in which the lending interest rate falls and the borrowing rate rises as the amount respectively lent or borrowed rises. An extreme version of this would place an upper bound on borrowing set at zero, as, for example, in Deaton (1992). This is generally referred to as the 'liquidity constrained' case.

Here we consider first the following formulation, in which the liquidity constrained case is nested: a consumer faces the same, constant saving interest rate r_s , and a

⁶ See, for example, Benhabib *et al.* (1991), Greenwood and Hercowitz (1991), Rios-Rull (1993) and Rupert *et al.* (1995), (2000).

borrowing rate r_{bt} which is an increasing function of the amount borrowed, $b_t \geq 0$, such that

$$r_{bt} = r(b_t) \quad r'(\cdot) > 0 \quad r''(\cdot) \geq 0 \quad (5.23)$$

and

$$r(0) > r_s \quad (5.24)$$

for all t . Thus consumers can borrow, but at an increasing interest rate that is always higher than the lending rate. There is no capital rationing in the sense of an absolute upper bound on borrowing, but a consumer may well choose in equilibrium not to borrow, if $r(0)$ is 'too high'. In this case the consumer behaves exactly as if liquidity is constrained.⁷

Realistically, this borrowing function could vary across time and could also contain as arguments the consumer's income and/or assets, reflecting her default risk and ability to put up collateral for loans. However, on grounds of tractability we stay with this simple formulation. Its implication is that in equilibrium different consumers may face different borrowing rates at the margin, and these rates may vary across periods, depending on the consumer's borrowing in each period.

The utility function remains unchanged, and we again take endowed income as exogenous. We just have to reformulate the budget constraints. We now let $s_t \geq 0$ denote saving. We then have

$$x_t \leq \bar{x}_t - s_t + b_t + (1 + r_s)s_{t-1} - (1 + r_{b,t-1})b_{t-1} + \mu_t \quad t = 1, 2, \dots, T \quad (5.25)$$

$$s_T = s_0 = b_T = b_0 = 0 \quad (5.26)$$

$$b_t \geq 0, \quad s_t \geq 0 \quad t = 1, 2, \dots, T - 1. \quad (5.27)$$

It is no longer possible to collapse these one-period budget constraints into a single intertemporal wealth constraint, since the value of a consumer's wealth depends on the interest rate that is relevant for discounting in each period, which in turn depends on the optimal pattern of borrowing and lending. There is no one interest rate that defines a single discount factor that can be applied to the endowed incomes in each period to obtain a unique value of wealth. In periods in which the consumer saves, the marginal opportunity cost of consumption is $1 + r_s$, in periods when she borrows, the opportunity cost of consumption is the *marginal cost of borrowing* $m_t^*(b_t^*) \equiv 1 + r(b_t^*) + r'(b_t^*)b_t^*$, and in periods when she neither lends nor borrows, which may occur far more frequently than in the perfect capital market case, the marginal opportunity cost of consumption is a shadow price that lies between $1 + r_s$ and $1 + r(0)$. In equilibrium there will be a set of discount factors or relative prices at which we can compute the present value of the endowed income stream, but these prices and this wealth value are known only after we

⁷ It is counterfactual to assert that consumers do not or cannot borrow, but if interest rates are so high that they do not wish to, it is a reasonable simplification to model them as if they are quantitatively rationed, as, for example, in Deaton (1992).

have solved the consumer's problem. Nevertheless, it is straightforward to characterise this solution.

The consumer again maximises

$$u = \sum_{t=1}^T \rho^{t-1} v(x_t) \quad (5.28)$$

subject to the constraints (5.25) to (5.27).

Associating Lagrange multipliers $\lambda_t > 0$ with the constraints (5.25), the first-order (Kuhn-Tucker) conditions are

$$\rho^{t-1} v'(x_t^*) - \lambda_t^* = 0 \quad t = 1, 2, \dots, T \quad (5.29)$$

$$(1 + r_s)\lambda_{t+1}^* - \lambda_t^* \leq 0 \quad s_t^* \geq 0 \quad [(1 + r_s)\lambda_{t+1}^* - \lambda_t^*]s_t^* = 0 \quad t = 1, 2, \dots, T-1 \quad (5.30)$$

$$\lambda_t^* - m_t^* \lambda_{t+1}^* \leq 0 \quad b_t^* \geq 0 \quad [\lambda_t^* - m_t^* \lambda_{t+1}^*]b_t^* = 0 \quad t = 1, 2, \dots, T-1 \quad (5.31)$$

together with the constraints.

There are three mutually exclusive⁸ solution possibilities for each period:

1. $s_t^* > 0, b_t^* = 0$: the consumer saves in this period and

$$\rho^{t-1} v'(x_t^*) = (1 + r_s)\lambda_{t+1}^*; \quad (5.32)$$

2. $s_t^* = 0, b_t^* > 0$: the consumer borrows in this period and

$$\rho^{t-1} v'(x_t^*) = m_t^* \lambda_{t+1}^*; \quad (5.33)$$

3. $s_t^* = 0, b_t^* = 0$: the consumer neither saves nor borrows and

$$m_t^* \lambda_{t+1}^* \geq \rho^{t-1} v'(x_t^*) \geq (1 + r_s)\lambda_{t+1}^*. \quad (5.34)$$

Figure 5.2 illustrates the nature of the solution, for the case $T = 2$. Point γ is the initial endowment point, which lies at the kink in the budget constraint. The consumer may lend from there at a constant interest rate r_s to reach an equilibrium at a point such as α . Alternatively, she may borrow along the curve rightwards from γ to reach equilibrium at a point such as β , with m_t^* the slope of the curve at the optimal point. Clearly in this case borrowing will be less than if it were possible to borrow at a constant rate equal to r_s (as indicated by the broken line). Finally, she may remain at γ , in which case the indifference curve slope lies between $1 + r_s$ and $1 + r(0) = m_t^*(0)$.

Given an arbitrary pattern of initial endowments $\bar{x}_t$, there could be an equally arbitrary pattern of optimal saving and borrowing. However, to discuss the life cycle it is useful to think of the consumer as being in a phase of a run of years in which she continually saves, continually borrows, or continually does neither. Thus consider such phases in turn, and what must hold, given (5.29)–(5.31), in any two adjacent years within them:

⁸ It is straightforward to show that we cannot have both $b_t^* > 0$ and $s_t^* > 0$ at any t .

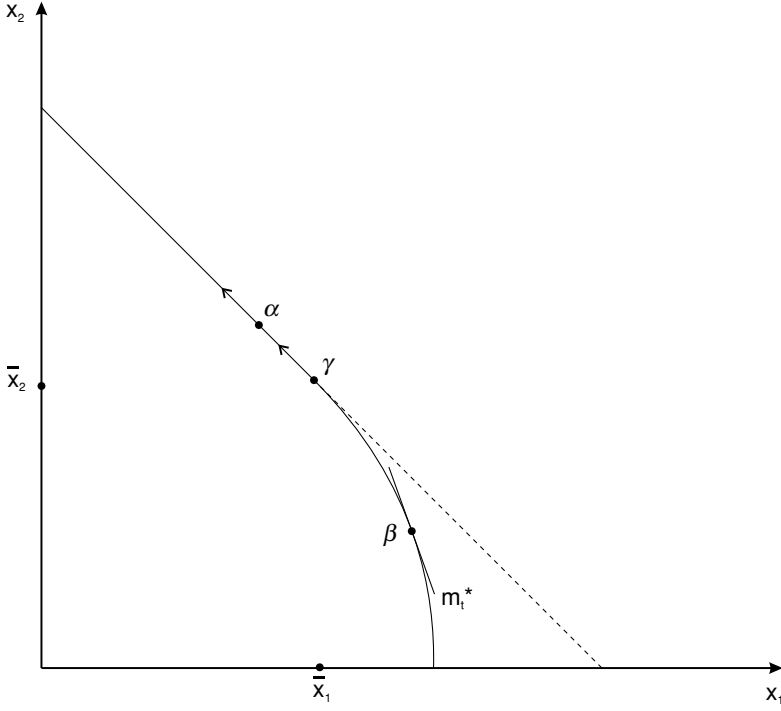


Figure 5.2 Possible equilibria in an imperfect capital market

1. *the saving phase*: the foc imply

$$\frac{v'(x_{t+1}^*)}{v'(x_t^*)} = \frac{\delta_s}{\rho} \quad (5.35)$$

where $\delta_s \equiv (1 + r_s)^{-1}$. Thus within this phase the previous ‘decoupling’ result continues to hold. Moreover, it might be argued that real interest rates on saving are sufficiently low that it is reasonable to take $\rho < \delta_s$, in which case the earlier results tell us that $x_{t+1}^* < x_t^*$, and consumption will tend to be falling over this phase.

2. *the borrowing phase*: in this case we have

$$\frac{v'(x_{t+1}^*)}{v'(x_t^*)} = \frac{\delta_{bt}}{\rho} \quad (5.36)$$

where $\delta_{bt} \equiv m_t^{*-1}$. Again there is decoupling, but now, since it is reasonable⁹ to assume $\rho > \delta_{bt}$, we have $x_{t+1}^* > x_t^*$ and consumption is rising over this phase.

3. *the no-borrowing no-lending phase*: in this case we simply have to look at the budget constraints:

$$x_t^* = \bar{x}_t + \mu_t. \quad (5.37)$$

Optimal consumption tracks income perfectly.

⁹ Think of borrowing at the margin on overdrafts, credit cards and other kinds of non-collateralised borrowing.

To relate these results to the kind of picture shown in figure 5.1, we then simply have to argue that in the earlier stages of the life cycle in which income first grows and then flattens out, the consumer is in a borrowing and/or no-borrowing no-lending phase, so that consumption will be tracking income, and then in the later stages, she is in a saving phase. This will again give a consumption path that tracks the path of endowed income. The basis for this story would be the argument that if consumers anticipate a sequence of increasing incomes they will tend to borrow, or of decreasing incomes they will tend to save. The imperfections in the capital market therefore act so as to translate the time profile of income into a similar time profile of consumption.

5.2.3 Uncertainty

There is an enormous literature on consumption and saving decisions under uncertainty. Here we consider only that part of it that is concerned with the life cycle, concentrating on recent developments in the theory of precautionary or ‘buffer stock’ saving.¹⁰ We use a simple model which, however, suffices to make the main points. Suppose, first, that a consumer lives for three periods, and has an endowed income in each period, which is a random variable drawn from the same probability distribution in each period. This distribution is

$$\pi_H : \bar{x}_H = \gamma_H \bar{x}_P \quad \gamma_H > 1 \quad (5.38)$$

$$\pi_P : \bar{x}_P \quad (5.39)$$

$$\pi_L : \bar{x}_L = \gamma_L \bar{x}_P \quad \gamma_L \in [0, 1) \quad (5.40)$$

where π_ω , $\omega \in \{H, P, L\}$ is the probability of state ω . The idea here is that $\bar{x}_P$ is normal or permanent income, γ_H is a ‘favourable shock’, for example a wage bonus or profit share, and γ_L is an ‘unfavourable shock’, best thought of as unemployment. Thus permanent income is income in the absence of random shocks. The shocks are not symmetric: γ_H need not be much above 1, while γ_L could be very low and possibly zero. Also, $\bar{x}_P$ will not be in general the mean of the distribution, although it can be thought of as the most probable value, with $\pi_P > \pi_H > \pi_L > 0$.

It is assumed that the capital market is perfect and the real interest rate r is certain. Thus, in the absence of the possibility of shocks, the consumer’s problem would simply be

$$\max_{x_t} \sum_{t=1}^3 \rho^{t-1} v(x_t) \quad \text{s.t.} \quad \sum_{t=1}^3 \delta^{t-1} x_t \leq \bar{x}_P \sum_{t=1}^3 \delta^{t-1}. \quad (5.41)$$

From the earlier analysis, therefore, we know that the slope of the time profile of consumption is completely determined by the ratio ρ/δ . It is usual in this literature to assume ‘impatience’, or $\rho/\delta < 1$, in which case we know that consumption would be

¹⁰ The best point of entry into this literature is Carroll (2001). See also Browning and Crossley (2001) for a critical view.

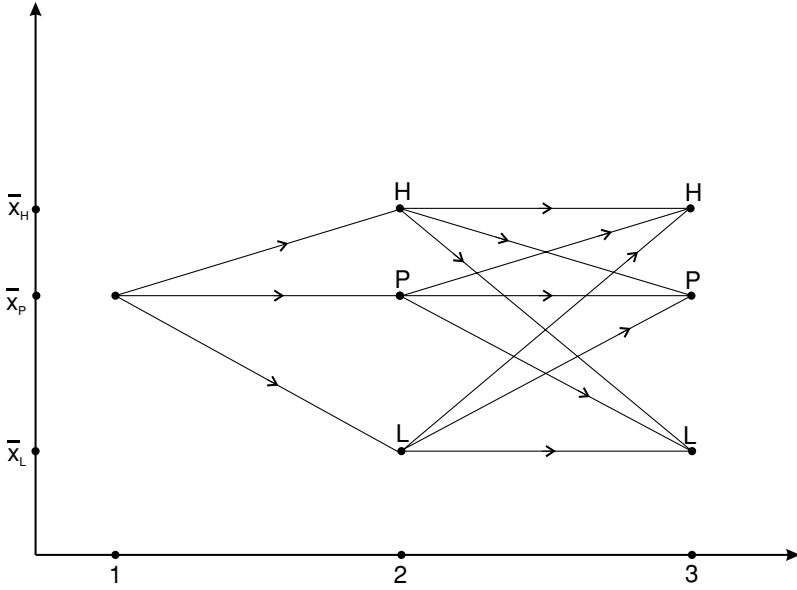


Figure 5.3 Income paths under uncertainty

falling over time, with the consumer borrowing initially to sustain consumption above permanent income, and then repaying later.

We assume initially that at date $t = 1$ the consumer has a certain endowed income¹¹ of $\bar{x}_P$. She has to choose in each period $t = 1, 2$ an amount s_t of saving ($s_t > 0$) or dissaving ($s_t < 0$). The distribution of endowments at $t = 2, 3$ is as shown in (5.38)–(5.40) and possible time paths are illustrated in figure 5.3.

We solve the consumer's problem recursively, beginning with date 2. Her expected utility function is

$$\bar{v} = v(x_1) + \rho \sum_{\omega \in \{H, P, L\}} \pi_{\omega} v(x_{2\omega}) + \rho^2 \sum_{\omega \in \{H, P, L\}} \pi_{\omega} v(x_{3\omega}) \quad (5.42)$$

and so, contingent on the state of the world $\bar{\omega} \in \{H, P, L\}$ at date 2, she solves the problem

$$\max_{s_{2\bar{\omega}}} \bar{v}_{2\bar{\omega}} = v(x_{2\bar{\omega}}) + \rho \sum_{\omega \in \{H, P, L\}} \pi_{\omega} v(x_{3\omega}) \quad (5.43)$$

given

$$x_{2\bar{\omega}} = \bar{x}_{\bar{\omega}} + (1 + r)s_1 - s_{2\bar{\omega}} \quad (5.44)$$

$$x_{3\omega} = \bar{x}_{\omega} + (1 + r)s_{2\bar{\omega}} \quad (5.45)$$

where of course in this problem s_1 is a given parameter.

¹¹ Nothing essential changes if we assume a different initial income.

It is simplest just to substitute into the maximand from these constraints and then derive the first-order condition

$$\frac{\partial \bar{v}_{2\bar{\omega}}}{\partial s_{2\bar{\omega}}} = -v'(\bar{x}_{\bar{\omega}} + (1+r)s_1 - s_{2\bar{\omega}}^*) + \rho(1+r) \sum_{\omega \in \{H, P, L\}} \pi_{\omega} v'(\bar{x}_{\omega} + (1+r)s_{2\bar{\omega}}^*) = 0 \quad (5.46)$$

and note also that the second-order condition, which is satisfied given strict concavity of $v(\cdot)$, is $\partial^2 \bar{v}_{2\bar{\omega}} / \partial s_{2\bar{\omega}}^2 < 0$.

Given the optimal values $s_{2\bar{\omega}}^*$, which are functions of s_1 , $s_{2\bar{\omega}}^*(s_1)$, we can define the value function $\bar{v}_{2\bar{\omega}}^*(s_1)$ giving the maximised value of the objective function in (5.43) as a function of s_1 . From the Envelope theorem we have that

$$\frac{d\bar{v}_{2\bar{\omega}}^*(s_1)}{ds_1} = (1+r)v'(\bar{x}_{\bar{\omega}} + (1+r)s_1 - s_{2\bar{\omega}}^*). \quad (5.47)$$

Moving to period 1, we then have the problem

$$\max_{s_1} \bar{v}(s_1) = v(\bar{x}_P - s_1) + \rho \sum_{\bar{\omega} \in \{H, P, L\}} \pi_{\bar{\omega}} \bar{v}_{2\bar{\omega}}^*(s_1) \quad (5.48)$$

with first-order condition

$$-v'(\bar{x}_P - s_1^*) + \rho(1+r) \sum_{\bar{\omega} \in \{H, P, L\}} \pi_{\bar{\omega}} v'(\bar{x}_{\bar{\omega}} + (1+r)s_1^* - s_{2\bar{\omega}}^*(s_1^*)) = 0. \quad (5.49)$$

Solving for s_1^* then yields the solutions for the $s_{2\bar{\omega}}^*$.

All this is of course perfectly standard stochastic dynamic programming. We now consider the specialisations of it that underlie the buffer stock saving model.

5.2.3.1 Prudence and precautionary saving The idea underlying the precautionary motive for saving, which goes back at least to Keynes, is that an increase in uncertainty should lead a risk-averse consumer to save more. This was formalised by Kimball (1990), who introduced and defined formally the term ‘prudence’ for this type of behaviour. Although supported both by intuition and some evidence, there is nothing in the model so far that necessarily implies that the consumer is prudent. To see this, we carry out the comparative statics analysis of optimal saving in state $\bar{\omega} \in \{H, P, L\}$ at date 2.

First we have to have a precise sense in which uncertainty can be said to increase. For this we use the idea of a mean-preserving spread (MPS), defined as follows. Mean endowed income at date 3 is

$$(\pi_H \gamma_H + \pi_P + \pi_L \gamma_L) \bar{x}_P. \quad (5.50)$$

So we can define an MPS by taking¹² $d\gamma_L < 0$ and

$$d\gamma_H = -\frac{\pi_L}{\pi_H} d\gamma_L > 0. \quad (5.51)$$

It is straightforward to show that every risk-averse consumer would prefer the initial distribution. Now taking the first-order condition (5.46) and applying the Implicit Function theorem gives

$$\frac{ds_{2\bar{\omega}}^*}{d\gamma_L} = -\frac{\rho(1+r)\bar{x}_P[\pi_L v''(\gamma_L \bar{x}_P + (1+r)s_{2\bar{\omega}}^*) - \pi_H v''(\gamma_H \bar{x}_P + (1+r)s_{2\bar{\omega}}^*)\pi_L/\pi_H]}{\partial^2 \bar{v}_{2\bar{\omega}}/\partial s_{2\bar{\omega}}^2}. \quad (5.52)$$

Given the second-order condition, this implies that the consumer is prudent, with $ds_{2\bar{\omega}}^*/d\gamma_L < 0$, if and only if

$$v''(\gamma_L \bar{x}_P + (1+r)s_{2\bar{\omega}}^*) - v''(\gamma_H \bar{x}_P + (1+r)s_{2\bar{\omega}}^*) < 0. \quad (5.53)$$

Since $\gamma_H > \gamma_L$, this requires that $v''(\cdot)$ be increasing in income, i.e. that $v'''(\cdot) > 0$, so that marginal utility $v'(\cdot)$ is a strictly convex function of income. This is the necessary and sufficient condition for the consumer to be prudent.

The intuition is shown with the help of figure 5.4. Given that the marginal utility curve is strictly convex, reducing the lower income and increasing the higher income in a mean-preserving way must increase expected marginal utility of income at date 3, the second term in condition (5.46). Thus marginal utility at date 2 must also be increased to satisfy the condition, and this can only be done by reducing consumption, i.e. increasing saving, at that date.

Requiring that a consumer be prudent has an immediate and quite dramatic implication: it means that one must reject the standard approach to consumption choice under uncertainty developed in the life cycle literature of the 1970s and 80s, because this assumed that if the consumer is risk-averse, she possesses a quadratic utility function, for which $v'''(\cdot) = 0$, and so she cannot be prudent. Indeed, the approach based on the assumption of prudent consumers typically uses the constant relative risk aversion (CRRA) utility function

$$v(x) = \frac{x^{1-\eta}}{1-\eta} \quad \eta > 1 \quad (5.54)$$

with quite important implications, as we show below.

The introduction of uncertainty means that the consumer's saving behaviour will result from the interplay of impatience and prudence. Whereas under certainty an impatient consumer will choose a declining consumption path, borrowing initially and then repaying, the introduction of significant uncertainty will cause her, if prudent, to

¹² For this we need $\gamma_L > 0$. If $\gamma_L = 0$ we could simply consider a mean-preserving reduction in risk, $d\gamma_L > 0$, and argue that this should reduce saving if the consumer is prudent.

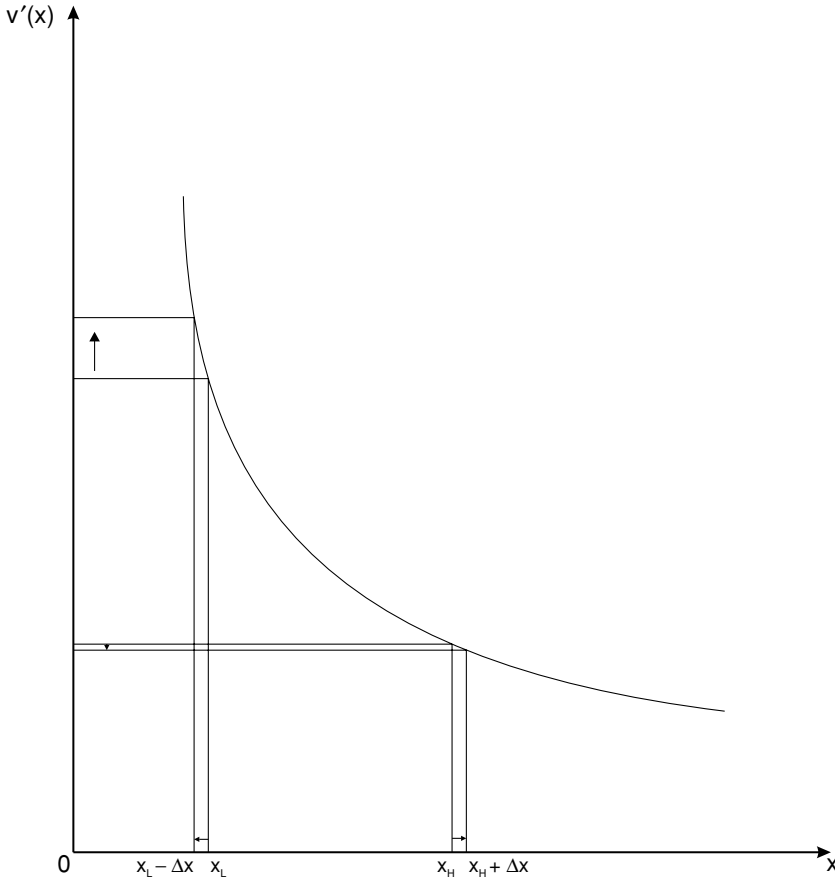


Figure 5.4 Prudence and the expected marginal utility of income

borrow less or, depending on the strength of her prudence and the degree of uncertainty, to save. The theory of buffer stock saving sets out to characterise the resulting consumption path and to show that it matches the data.

5.2.3.2 Prudence and liquidity constraints We saw earlier that in an imperfect capital market consumers may be deterred from borrowing by its high cost, or, in an extreme case, by an absolute constraint ruling out borrowing totally. As a result, an impatient consumer's consumption tracks income in a way that matches the data but not the predictions of the perfect market model. Proponents of the prudence-based approach argue that a similar result holds in this model, though it arises because consumers choose not to borrow to avoid the risk of having no future income, rather than being constrained from borrowing even though they would wish to.

To show what is involved here, take again the problem of choice of saving conditional on state $\bar{\omega} \in \{H, P, L\}$ at date 2, but impose the no-borrowing constraint $s_{2\bar{\omega}} \geq 0$. In

that case the first-order condition becomes

$$\frac{\partial \bar{v}_{2\bar{\omega}}}{\partial s_{2\bar{\omega}}} \leq 0 \quad s_{2\bar{\omega}}^* \geq 0 \quad s_{2\bar{\omega}}^* \frac{\partial \bar{v}_{2\bar{\omega}}}{\partial s_{2\bar{\omega}}} = 0. \quad (5.55)$$

Then the claim is that in the present model this constraint would never be binding, since $s_{2\bar{\omega}}^* > 0$ is always an interior optimum. This must imply in turn that at $s_{2\bar{\omega}} = 0$, we always have $\partial \bar{v}_{2\bar{\omega}} / \partial s_{2\bar{\omega}} > 0$, i.e. from (5.46) with $s_{2\bar{\omega}} = 0$

$$\frac{\rho}{\delta} \sum_{\omega \in \{H, P, L\}} \pi_{\omega} v'(\bar{x}_{\omega}) > v'(\bar{x}_{\bar{\omega}} + (1+r)s_1). \quad (5.56)$$

There is nothing in general that implies this inequality, particularly if the consumer is impatient and so $\rho/\delta < 1$.

However, in the simulation exercises that underlie analysis of this model the CRRA utility function (5.54) is typically assumed, under which $v'(x) = x^{-\eta}$, with a value for η of around 3, while a positive probability is placed on the state in which $\gamma_L = 0$. In that case, as long as $s_1 > 0$, (5.56) is satisfied for any $\bar{\omega} \in \{H, P, L\}$, since $\lim_{x \rightarrow 0} x^{-\eta} = \infty$. In other words, in the absence of saving, income in state L has an arbitrarily high marginal utility and so, as long as that state has positive probability, this inequality must be satisfied.¹³ The consumer always saves because of the risk of having no income during the next period.

The assumption that $\gamma_L = 0$ seems extreme and not particularly plausible. Unemployment benefit, social welfare payments, and incomes of other family members, could all be expected in reality to put a lower positive bound on income. In that case it is quite possible that even a prudent consumer would borrow. There would be an upper bound to this borrowing placed by the available income in state L , because the consumer will still never want to risk having no income, net of debt repayment, in that state.¹⁴ Since this positive borrowing distinguishes the present model from that with absolute borrowing constraints (though not with an imperfect capital market more generally) in a way consistent with the data, that ought to be considered an advantage of this model. People do borrow.

5.2.3.3 Buffer stock saving Suppose now that, rather than living for just three periods, the consumer faces an infinite time horizon, with the same endowed income distribution in each period.¹⁵ Thus at each period, the optimisation problem faced by the consumer is identical. For simplicity, therefore, we drop the t subscript. We now adopt Deaton's definition of *cash on hand* at any date as the sum of endowed labour income at that date and the previous period's saving plus the interest on that. This is the amount available to finance current consumption and saving, i.e. we have

¹³ That $s_1 > 0$, so that the right-hand-side marginal utility is finite, can be shown by applying the same argument to the first-period saving decision.

¹⁴ This also raises the issue of the nature of consumer bankruptcy laws, which in modern economies are rather more liberal than the Dickensian type of no-bankruptcy constraint apparently envisaged here.

¹⁵ The model can be extended to allow growth in permanent income with no essential change in results.

the identity

$$s = c - x \quad (5.57)$$

where c is cash on hand. We can reformulate the problem in (5.48) for any period in terms of the choice of consumption, rather than saving, just by substituting for s from this identity to obtain

$$\max_x v(x) + \rho \sum_{\omega \in \{H, P, L\}} \pi_\omega v_\omega^*(c - x) \quad (5.58)$$

where $v_\omega^*(\cdot)$ has the same interpretation as before, as the value of future utility along the optimal path, discounted to the next period. Note that c is a given parameter in this problem. Then using the first-order condition

$$v'(x^*) + \frac{\rho}{\delta} \sum_{\omega \in \{H, P, L\}} \pi_\omega v_\omega^{*'}(c - x^*) = 0 \quad (5.59)$$

we can define¹⁶ a *consumption function* $x^*(c)$, which gives the optimal consumption x^* as a function of the given cash on hand. Carroll and Kimball (1996) show that on the assumptions of the model, and for a wide class of utility functions, of which the CRRA form is one, this function is strictly concave, as illustrated by the curve labelled $x^*(c)$ in figure 5.5. Thus at any date, given the cash on hand c , optimal consumption $x^*(c)$ is given by the corresponding point on this curve.

Now cash on hand at any date t is not actually exogenous, but is determined by cash on hand and the consumption decision at the previous date $t - 1$, together with the income realisation x at date t . Thus define *permanent income cash in hand* as

$$c = \bar{x}_P + (1 + r)[c - x]. \quad (5.60)$$

This implicitly defines a level of consumption

$$x = \delta(\bar{x}_P + rc) \quad (5.61)$$

having the property that, given a sequence of income realisations equal to permanent income, cash on hand remains constant over this sequence.¹⁷ Consumption in this period must be restricted to the present value of permanent income plus the interest income on cash on hand. This function is drawn as the line $x(c)$ in figure 5.5.

Consider now the intersection point c^* in the figure. This defines the equilibrium buffer stock to which the consumer's saving tends. To see this, note that at any $c > c^*$, optimal consumption $x^*(c)$ lies above the line, implying that it is too high to maintain that level of cash on hand, and so c will tend to fall over time, while the converse argument applies for $c < c^*$. Once c^* is attained, the consumer will remain there,

¹⁶ Recall that the life cycle literature began with the investigation of the Keynesian consumption function by Modigliani and Brumberg (1954) and Friedman (1957).

¹⁷ Carroll and Kimball (1996) in fact define this in terms of the expectation of income $\sum_\omega \pi_\omega \bar{x}_\omega$, but in the present model it seems more appropriate to take permanent income.

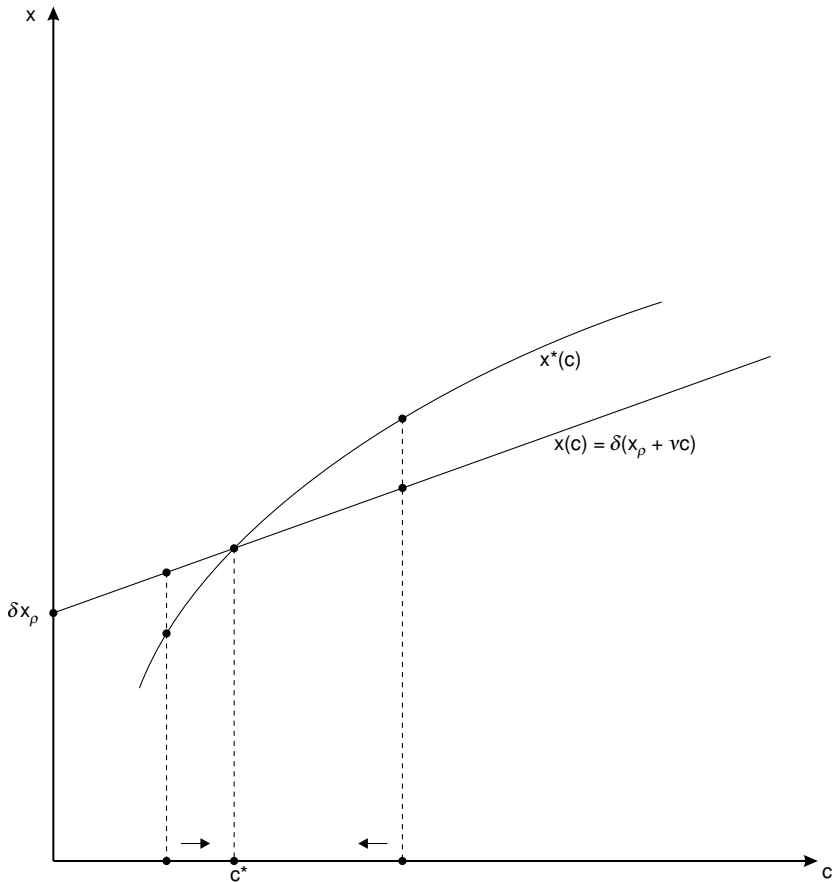


Figure 5.5 Equilibrium cash in hand and buffer stock saving

absorbing income shocks by increasing or reducing consumption. Thus c^* represents a stable equilibrium buffer stock of cash on hand.

All this of course is for a theoretical model with an infinite time horizon. The buffer stock saving model is related to the data by simulating a model incorporating impatience, a CRRA utility function implying therefore prudence, ‘plausible’ values of the preference parameters ρ and η , and ‘realistic’ values for the market interest rate and parameters of the probability distribution of endowed labour incomes over a typical consumer’s lifetime (including a small probability of zero income).¹⁸ This lifetime will of course be finite and will exhibit the typical pattern of initially growing permanent income, which flattens out and then declines post-middle age, until retirement. The simulation models produce consumption time paths that match the data, and in particular show consumption tracking income until middle age, and then falling more rapidly than income as savings build up pre-retirement.

¹⁸ For examples of these studies see, in particular, Carroll (1994), (1997), and Gourinchas and Parker (2002).

The story that ties together the theoretical and simulation models is that, in the earlier years, the consumer behaves essentially as if she were in an infinite horizon world. Since she begins with no wealth, she saves so as to build up a buffer stock of cash on hand, as suggested by figure 5.5, and this explains why she does not borrow. Consumption therefore grows with income, which effectively constrains it, even though she could borrow if she wished to. Later in life however – the model simulations place this within the age range 40–50 – the influence of the finiteness of the working life and the approaching retirement age becomes dominant, and buffer stock behaviour gives way to the more traditional life cycle behaviour of saving for old age. Thus its proponents claim that the buffer stock model is capable of giving a complete explanation of life cycle consumption and saving behaviour without imposing liquidity constraints or indeed, we might add, taking account of the endogeneity of labour income and life cycle demographics.

5.2.4 Demographics

The models we have considered so far essentially regard the household as a single consumer. In fact, as we have repeatedly emphasised, the typical household consists of a family, and a number of studies of life cycle consumption and saving have argued that taking this into account provides a very simple resolution of the ‘puzzle’ of the excess sensitivity of consumption to income. In the years in which income is growing, couples are starting families and raising children, who then eventually leave home, releasing parental income for saving for old age. It can be argued that this process, which is perhaps a better characterisation of what is commonly meant by the life cycle than the movement in wages and income, explains the tendency for consumption to grow with income. Indeed, at the simplest level, the study by Blundell *et al.* (1994) shows that when consumption is adjusted for the number of household members, with children being counted as 0.4 of an adult, then the per capita consumption time profile is virtually flat, consistent with the simple life cycle model. Other studies, by Attanasio and Browning (1995) and Attanasio and Weber (1995) also show that allowing for children’s consumption removes most of the hump in consumption. This is unexciting, but if true, it makes much of the theory considered so far in this chapter redundant, at least as far as explaining the excess sensitivity puzzle is concerned. In fact Browning and Ejrnaes (2002) argue that ‘if we take proper account of the numbers and ages of children then there is no need to introduce a precautionary motive’, and go on to demonstrate this with UK Family Expenditure Survey data. Other studies, by Attanasio *et al.* (1999) and Gourinchas and Parker (2002), however, using US data, find an explanatory role for both demographics and precautionary saving.

A weakness of the models analysing the effect of demographics on consumption over the life cycle is that they are not based on an explicit analysis of household decision-taking, but simply specify an individual utility function defined on consumption, in which demographic variables enter as conditioning variables which influence the marginal utility of consumption in some more or less arbitrarily specified

way. Moreover, the focus on ‘resolving the excess sensitivity puzzle’ is a very narrow one, and it would be interesting to extend the analysis both conceptually and in terms of data sources to enable a richer set of policy issues to be addressed. We now present a model which provides an explicit analysis of family decision-taking over the life cycle, and the remainder of this chapter goes on to apply this model empirically.¹⁹

5.3 A model of the ‘family life cycle’

Three out of the four approaches to life cycle decision-taking surveyed in the previous section are based essentially on the view of the household as consisting of a single worker/consumer. The attempt to make the case for the importance of ‘demographics’ simply takes the step of making individual preferences depend on demographic variables in an ad hoc way. In this section we emphasise the importance of viewing the life cycle in terms of how the household, as a family, moves through the various phases of its life. We begin by setting out a simple model, essentially the intertemporal extension of our earlier atemporal models, to give a theoretical framework, and then present data and empirical analysis that we believe support this way of viewing the household.

5.3.1 Consumption and time allocation over the life cycle

For the purposes of the theoretical model we divide the life cycle into four periods, assumed of equal length for simplicity.²⁰ In every period except the last a couple allocates its time between market work, household production and leisure. Period 1 is the pre-children phase, in period 2 children are present,²¹ in period 3 they have left the household, and in period 4 the couple has retired, and so the only uses of time are household work and leisure. We analyse first the within-period choices of expenditure and time allocations, and then the intertemporal allocation of expenditures.

Period 1 The individuals f and m jointly choose consumptions of the market and household goods x_{i1} , y_{i1} and leisure z_{i1} , $i = f, m$, to maximise an HWF²²

$$H_1[u_f(x_{f1}, y_{f1}, z_{f1}), u_m(x_{m1}, y_{m1}, z_{m1})] \quad (5.62)$$

¹⁹ This draws upon Apps and Rees (2001), (2003), (2005).

²⁰ In the empirical treatment below we refine this substantially, increasing the number of phases and allowing them to have variable length.

²¹ Note that we do not analyse the decision on whether and when to have children, or on how many to have, but take these as exogenous. We do not believe that taking account of the endogeneity of these decisions would change the conclusions we draw in the present analysis, but nevertheless this is an important area for future work.

²² Nothing substantive would change if we assumed a GHWF, but for simplicity we ignore issues of the within-household income distribution.

subject to the following constraints. First, household production is given by

$$\sum_{i=f,m} y_{i1} = y_1 = h(t_{f1}, t_{m1}) \quad (5.63)$$

with $h(\cdot)$ a concave, linearly homogeneous production function,²³ and t_{i1} are time inputs. The budget²⁴ and time constraints are

$$\sum_{i=f,m} x_{i1} = \sum_{i=f,m} w_{i1}l_{i1} + \mu_1 \quad (5.64)$$

$$t_{i1} + l_{i1} + z_{i1} = 1 \quad i = f, m. \quad (5.65)$$

In the usual way, we can find the implicit price of the household good, p_1 , as the marginal (= average) cost of its production, $\gamma(w_{f1}, w_{m1})$, and collapse the three constraints into the single full income constraint

$$\sum_{i=f,m} (x_{i1} + p_1 y_{i1} + w_{i1} z_{i1}) = X_1. \quad (5.66)$$

In a single period model, the income on the right-hand side of this constraint would be given by $\sum_{i=f,m} w_{i1} + \mu_1$, which we call full income, available for consumption of all goods, including the household good and leisure. However, we know that by use of the capital market, the value of full consumption, the left-hand side of (5.66), can deviate from full income in that period, and so the right-hand side of (5.66), X_1 , which we call *full expenditure*, is yet to be determined.

Solving for the optimal allocation, values of consumptions x_{i1}^* , y_{i1}^* , z_{i1}^* , and time allocations l_{i1}^* , t_{i1}^* and inserting the former into the household welfare function gives the period 1 value function $V_1(p_1, w_{f1}, w_{m1}, X_1)$. Life cycle considerations have their effect on period 1 decisions via the choice of X_1 . Given this, the within-period allocation is determined by wage rates and the price of the household good, as extensively discussed in earlier chapters.

Period 2 As discussed in chapter 4, we model children by including them as individuals with their own utility functions, which appear in the HWF. Without real loss of generality we can assume just one child with the utility function $u_k(x_{k2}, y_{k2}, c_k)$, where c_k is child care, produced by using parental time inputs t_{ik} and a bought-in market child care input b , with production function, also concave and linearly homogeneous, given by $c_k(t_{fk}, t_{mk}, b)$. As before, we can define prices for the household good, and child care, given by $p_2 = \gamma(w_{21}, w_{m2})$ and $q = \gamma_c(w_{f2}, w_{m2}, p_b)$, with $\gamma(\cdot)$ and $\gamma_c(\cdot)$

²³ We assume the production function is the same in each period but that could easily be generalised. Also the assumption of linear homogeneity is made just for notational simplicity. Only the case of increasing returns presents any analytical complications.

²⁴ Non-wage income μ can include fixed pension or superannuation contributions and so may be negative. Note we differentiate between long-term contractual saving in public or private pension or superannuation schemes, which are taken as fixed across all time periods, and market borrowing or saving, which are determined in each period. We take the former simply as exogenously given.

the respective unit cost functions and p_b the price of the bought-in child care input. The HWF now becomes

$$H_2[u_f(x_{f2}, y_{f2}, z_{f2}), u_m(x_{m2}, y_{m2}, z_{m2}), u_k(x_{k2}, y_{k2}, c_k)]. \quad (5.67)$$

Given the parental time constraints on the four possible uses of time

$$t_{i2} + t_{ik} + l_{i2} + z_{i2} = 1 \quad i = f, m \quad (5.68)$$

the full expenditure budget constraint is

$$\sum_{i=f,m,k} (x_{i2} + p_2 y_{i2}) + \sum_{i=f,m} w_{i2} z_{i2} + q c_k = X_2. \quad (5.69)$$

We assume here that children supply neither domestic nor market labour. We solve for the demands for market and domestic goods and child care, and for the parental time allocations, to derive the value function for this period, $V_2(p_2, w_{f2}, w_{m2}, q, X_2)$.

The effect of the presence of children in the household is to add to the demand for market and domestic consumption goods and to create demand for a new good, child care. All these demands will now depend inter alia on the price of child care which, in turn, except at a corner solution in which only parental time is used, will depend on the price of market child care as well as parental wage rates.²⁵ This model therefore suggests that *across-household* variations in consumptions, labour supplies and the mix of market and parental child care would be explained by variation in wage rates, the price of market child care, and productivities in producing domestic goods and child care.²⁶

Period 3 On the assumption that when the children leave home they place no further demands on parental time and income,²⁷ the period 3 problem is identical to that in period 1, and in the same way yields the value function $V_3(p_3, w_{f3}, w_{m3}, X_3)$.

Period 4 Since the individuals have retired, $l_{i4} = 0$ by definition, and so the model is as in period 3 except that the budget and time constraints are

$$\sum_{i=f,m} x_{i4} = X_4 \quad (5.70)$$

$$t_{i4} + z_{i4} = 1 \quad i = f, m. \quad (5.71)$$

²⁵ Where the optimal time allocation implies that one partner reduces market labour supply to zero, the corresponding market wage rate gives only a lower bound on the opportunity cost of that individual's time – see chapter 3 for further discussion of this point.

²⁶ A further factor that we regard as likely to be empirically important, but which is not captured in the present model, are the perceptions of the loss of work-related human capital and the expectations of the resulting future wage and employment possibilities, all of which are influenced by labour supply decisions in this period. Section 3.4 considered a simple model along these lines. For an empirical analysis, see Attanasio *et al.* (2003).

²⁷ We hear parents saying 'If only.' This is just a simplifying assumption. It means that we can drop children formally from the HWF.

The household's non-wage income in this period, μ_4 , consists entirely of transfer payments from pension or superannuation schemes (but not, of course, income from the returns to past saving on the capital market). The household maximises its HWF

$$H_4[u_f(x_{f4}, y_{f4}, z_{f4}), u_m(x_{m4}, y_{m4}, z_{m4})] \quad (5.72)$$

subject to the above three constraints together with the production function constraint. We derive the value function $V_4(X_4)$ as before. Note that no wage rates enter into this value function. All prices and wage rates in this retirement period are implicit or shadow prices, determined endogenously at the household optimum by the values of the Lagrange multipliers associated with the various constraints.

We no longer in general have a constant implicit price of the household good, even with the assumption of linear homogeneity of the household production function, since the marginal opportunity cost of an individual's time spent in household production is no longer the market wage, but rather an implicit wage rate, say ω_i , given by the marginal rate of substitution between consumption of the numeraire market good and leisure. This measures at the optimum the number of units of the market good the household is prepared to pay for an additional unit of i 's leisure. The implicit price of the household good is then its marginal cost, where the input of each individual's time is valued at his or her implicit wage ω_i . The cost of producing the household good, at these implicit wages, is minimised. At the equilibrium, each individual's marginal rate of substitution between market and domestic consumption is equated to this implicit price.

Note that now all non-leisure time can be devoted to household production, which could be rather cheap at the margin. Thus it would not be surprising to see a major substitution of household for market goods, and this could well be the explanation of the 'retirement income puzzle' discussed by Banks *et al.* (1998).²⁸ They show that the fall in (market) consumption on retirement is much larger than can be explained by the reduction in work-related expenditures – travel costs, clothing and so on – and suggest 'unanticipated shocks' upon retirement, for example learning that one's pension is smaller than expected, as a possible explanation. It seems more consistent both with the life cycle model and the data, however, to argue that *full* consumption remains relatively stable, and what we observe is a large substitution of household production for market consumption, due of course to the fact that on retirement all the household's working time is now devoted to domestic production. This explains why, in categories such as expenditure on buying food that is prepared and consumed at home, Banks *et al.* find a significant fall. There is simply a substitution of own time for market expenditure, a reversal of the 'switch to convenience foods' example, that was so much used in the early discussions of the Becker household production model.

This argument is further supported by a detailed study by Aguiar and Hurst (2005), who exploit a dataset giving the dollar value, the quantity and the quality of food

²⁸ See also Bernheim *et al.* (2001).

consumed by US households. They show that a 17 per cent fall in expenditure on food consumption by households on retirement was accompanied by a 53 per cent increase in the time retired households spent on food production – essentially in the time spent shopping for and preparing food. There was no evidence of a decline in the quantity and quality of food consumed, though there was a significant shift away from use of fast-food restaurants. This paper gives a fine example of how the use of more comprehensive datasets can give real insights into household behaviour.²⁹

Given the standard assumption of intertemporal separability of utility – in this case household welfare – we can now analyse the determination of the time stream of full expenditures, $X_1, \dots, X_4$. This obviously will depend upon what we assume about the capital market. We consider two possibilities.

Case 1: A perfect capital market Assume the household members agree on the intertemporal felicity discount factor³⁰ ρ , and that the market discount factor is δ . Then the household's problem is simply

$$\max_{X_t} \sum_{t=1}^4 \rho^{t-1} V_t(X_t) \quad \text{s.t.} \quad \sum_{t=1}^4 \delta^{t-1} X_t \leq \sum_{t=1}^3 \delta^{t-1} \left(\sum_{i=f,m} w_{it} + \mu_t \right) + \delta^3 \mu_4 \equiv W \quad (5.73)$$

where μ_4 includes all non-capital market income such as state pensions, superannuation payments and so on, and W denotes wealth.

We are already familiar with the nature of the solution to this problem. The full consumption stream can be decoupled from the full income stream and is chosen to satisfy the condition that discounted marginal utility of full expenditure on consumption of market and household goods and leisure, and on child care in period 2, is equalised across periods, i.e. we have

$$\left(\frac{\rho}{\delta} \right)^{t-1} V'_t(X_t^*) = \lambda^* \quad t = 1, \dots, 4 \quad (5.74)$$

with λ^* the marginal utility of household wealth. This alone does not imply any particular time profile for the consumption of the *market* good. This will depend on the within-period allocations.

Case 2: An imperfect capital market Here we adopt the model set out in subsection 5.2.2 earlier. The household can save at a given (low) interest rate r_s , or borrow at an interest rate $r_{bt} = r(b_t)$ in any period $t = 1, \dots, 4$, where the interest rate is an increasing function of borrowing b_t . A kink in the intertemporal budget constraint

²⁹ See also Hurst (2008). As the title of this paper suggests, the puzzle is easily resolved once better data on consumption and health expenditures of retirees are examined. For a bargaining-theoretic approach to this issue see Lundberg *et al.* (2003).

³⁰ They also have to agree on a time horizon, something which the present model assumes away. For an analysis in terms of a household model of some of the issues arising from differing longevities of husbands and wives, see Browning (2000).

is created by the assumption that $r(0) > r_s$ (see figure 5.2). A household may be in equilibrium by lending at rate r_s , borrowing at a marginal cost of $m_t^*(b_t^*) \equiv 1 + r(b_t^*) + r'(b_t^*)b_t^*$, or neither. A large difference $r(0) - r_s$ could lead to many households being in this latter position.³¹

Thus the household solves the problem

$$\max \sum_{t=1}^4 \rho^{t-1} V_t(X_t) \quad \text{s.t.} \quad (5.75)$$

$$X_t \leq \sum_{i=f,m} w_{it} + \mu_t - s_t + b_t + (1 + r_s)s_{t-1} - (1 + r_{b,t-1})b_{t-1} \quad (5.76)$$

$$b_t \geq 0, \quad s_t \geq 0 \quad t = 1, 2, 3 \quad (5.77)$$

$$X_4 \leq \mu_4 + (1 + r_s)s_3 - (1 + r_{b,3})b_3 \quad (5.78)$$

$$s_0 = b_0 = 0. \quad (5.79)$$

The key difference between the perfect and imperfect capital market models lies in the impact of changes in household per capita income in a given period on consumption in that period. In the case of a perfect capital market that impact is diffused over the entire lifetime, which dilutes its impact in the period in which it takes place. In an imperfect capital market, a change in income in a given period can have a large effect on optimal consumption in that period, the more so the greater the slope of the borrowing rate function.³²

The data we present below suggests the following picture for the solution to the problem presented in (5.75)–(5.79). In period 1, when the typical couple has two incomes and no children, saving is substantial. This is consistent with the precautionary saving of the buffer stock model, though, since in this family life cycle model there is no uncertainty, we would prefer to call this ‘anticipatory saving’. The couple anticipates that they are going to have children in the relatively near future, and that they need to smooth the resulting shock to household per capita income and consumption.

The data show that this shock is substantial. Not only is there an additional source of demand for consumption of market and domestic goods, but the new demand for child care has major implications for time allocation and the household’s market income.³³ The household has to decide on how it will meet these increased demands, especially for child care. We can think of it as choosing among different mixes of market and parental child care. One solution is for the secondary earner to drop out of the labour market altogether, another is for her to switch to part-time working, the third is for her to continue in full-time employment, using a lot of bought-in child care. As we showed

³¹ For these households, this is then observationally equivalent to Deaton’s liquidity constrained case, with an absolute no-borrowing constraint. We do, however, observe some households borrowing short-term in the form of overdrafts, credit card debt and charge accounts.

³² The extreme case is that where there is a no-borrowing constraint, since then the change in income translates exactly into a change in consumption when the constraint binds.

³³ An early empirical paper, Gronau (1980), also points out the strength of this effect, though it appears less sharply in his work because he defines the life cycle on age of the head of the household.

in chapter 1, roughly equal proportions of households choose each solution, though there is some variation across countries in this. The within-period model for period 2, given above, suggests that the determinants of this choice are wage rates, the price of bought-in child care, and relative productivities of parental and market child care. To this the intertemporal model adds capital market conditions: the interest rate on saving in period 1, since this determines the asset income in period 2; and the borrowing rate in period 2, since this determines the household's ability to smooth its consumption path in that period.

In those households that choose full-time parental child care, the secondary earner leaves the labour market, and a major effect of the move into phase 2 is a fall in household market income. At the other end of the spectrum, where the secondary earner continues full-time working, the costs take the form of payments for bought-in child care and substitution of market for domestic consumption goods. In either case, there is a reduction in the income available for saving, and so, as we would expect, saving drops dramatically in this period.

In the perfect capital market model, we would predict a substantial switch to borrowing, to smooth the drop in adult full consumption (including leisure) that the data show takes place in period 2. However, borrowing in period 2 remains relatively low for most households. This suggests that the imperfect capital market is likely to be a better model for predicting the data. Most households are at or close to the kink in period 2.

In period 3 household income recovers to some extent, as a significant proportion of secondary earners return to the labour market (though often only as part-time workers), and the end of the need to allocate time and money to child care increases adult consumption of market and household goods and allows increased saving in anticipation of retirement. However, saving does not return to its pre-children level. Moreover, even in this period there is considerable heterogeneity in saving behaviour, reflecting the heterogeneity in female labour supply: most saving is carried out by households in which the female has a significant market labour supply.

Note that the existing models of life cycle consumption and saving, reviewed in the previous section, do not capture these features of household behaviour. In particular the drop in saving between periods 1 and 2 is missed, even by those models that incorporate demographics, largely because the life cycle is defined in terms of age of 'head of household'. As we show below, this results in considerable averaging over households with and without children present. Likewise, the considerable heterogeneity in saving behaviour across households, reflecting the heterogeneity in female labour supply, is missed, largely because of the assumed separability of 'leisure', i.e. time allocation, and consumption and saving decisions. In the remainder of this chapter we present data and analysis that confirm and fill out this general picture.

5.4 Evidence on family life cycle profiles

To model the life cycle, we construct a dataset that extends the information base of studies currently available in the literature in the following respects:

- Time-use data are combined with income and expenditure survey data to obtain information on the household's time allocation and consumption of leisure, as well as its implicit expenditure on domestic output and child care.
- Data on government indirect benefits, as well as taxes and direct benefits, are used to show how government policy affects life cycle consumption.
- Data on the sources and cost of household borrowing, and on long-term saving in the forms of house purchase and superannuation payments, are used to identify the nature of the capital market imperfections faced by households.

We first describe these aspects of our data and the criteria used to define life cycle phases on the presence and ages of children.

5.4.1 *Time-use and household expenditure survey data*

We combine data on income and consumption with information on the allocation of time to activities outside the market by drawing on two surveys: the Australian Bureau of Statistics (ABS) 1998 Household Expenditure Survey (HES) and the ABS 1997 Time Use Survey (TUS).³⁴ The HES contains data collected by interview on household consumption expenditure, labour supply, earnings and non-labour incomes. The TUS provides detailed information collected by diary, for two diary days, on the allocation of time to labour market activities and nine major categories of non-market activities.³⁵ We compute hours of market work as the sum of time allocations to all employment categories, including work breaks, job search and associated travel.³⁶ The nine non-market activities are aggregated into two major categories, domestic work³⁷ and leisure.³⁸ In the domestic work category we also distinguish between child care and time allocated to other household services. Both surveys provide data on a common set of demographic, education and occupation variables.

The HES also provides detailed information on household debt, house price, mortgage and loan repayments, and contributions to mandatory retirement saving and life insurance. The information on loans is highly disaggregated by purpose, type of lender,

³⁴ Ideally, we would like to have access to panel data or, alternatively, to successive HES and TUS surveys. However, these are not available. Our analysis is, in effect, limited to the use of a single cross-section. We therefore do not take account of cohort effects. While we recognise that cohort effects can be important, it does not seem to us that they would alter the direction of our key results. For example, Attanasio *et al.* (2003) show that differences in behaviour of three successive cohorts of women in the US occur primarily in the younger, child-bearing age groups, while older age groups behave in a very similar way across all three cohorts. In using our data, we are essentially assuming that the younger age groups will indeed behave as the older groups when they reach that age, consistent with the evidence in Attanasio *et al.*

³⁵ The nine non-market activity episode classifications are: personal care, education, domestic activities, child care, purchasing goods and services, voluntary work and care, social and community interaction, recreation and passive leisure.

³⁶ By including all categories we obtain TUS data means that are very close to those for 'usual hours of work' obtained by questionnaire in the HES.

³⁷ Domestic work is computed as the sum of time allocations to the categories 'domestic activities', 'child care' and 'purchasing goods and services'.

³⁸ For each activity episode, information is recorded for a 'primary' and, if relevant, a 'secondary' activity. Where primary and secondary activities are reported, the weighting used is 0.6:0.4.

term and amount of loan. In addition, the HES includes estimates of indirect government taxes and benefits as well as detailed data on direct taxes and benefits.

We select samples of two-adult households from these datasets, excluding only those in which the female partner is aged from 40 to 44 years and there are no children present, on the assumption that these records may represent households that have decided not to have children. The TUS sample contains 1,927 records, and the HES 3,986 records. Using regression models estimated on the TUS data, we then merge information on time use with the income and consumption data for each record in the HES sample.³⁹

To ensure that the time constraint is satisfied for each record, we predict time-use ratios. We estimate as linear functions of observed variables the ratios of leisure to non-market time, $\tau_{z_i} = z_i / (a_i + z_i)$, and child care to domestic work time, $\tau_{\kappa_i} = \kappa_i / a_i$, $i = 1, 2$, where κ_i is child care time. The independent variables include dummy variables for age of youngest child, interaction variables that capture the effect of additional children at each age of youngest child, and the characteristics of the adults, including age and dummy variables for education and employment status. Non-market time is computed from the time constraint as $a_i + z_i = A - l_i$, where l_i is market labour supply, and the predicted value for τ_{z_i} is used to compute each individual's leisure time, z_i , as $(A - l_i) * (\tau_{z_i} + u_i)$, where u_i is a randomly selected element from the residuals of the regression estimated on the TUS data. Domestic time is computed from the time constraint as $a_i = A - l_i - z_i$.⁴⁰ Using the result for a_i and predicted value for τ_{κ_i} , the allocation of time to child care is computed by the same procedure.

5.4.2 The life cycle defined on age of head of household

When we organise our data by defining the life cycle in terms of the age of the head of the household,⁴¹ we obtain profiles of income and consumption that are consistent with the results of studies for countries with similar female employment rates. To illustrate this, table 5.1 and figure 5.6 present cross-sectional profiles of median net income,⁴² consumption expenditure and saving by age of the male partner as head of household. The profiles of income and consumption exhibit the familiar 'hump' shape and correlation, closely matching, for example, the profiles reported in Blundell *et al.* (1994, table 4.1) for the UK. Both income and consumption rise initially and then begin to decline after age 50, with positive saving up to age 60 and negative thereafter. The profiles appear to indicate buffer-stock behaviour when households are young (up to

³⁹ We take this approach because the HES sample has the advantage that it is much larger than the TUS sample and it also provides information on an annual or 'usual' weekly basis, and so the data for each record are less 'noisy' than the TUS data collected for two diary days.

⁴⁰ Diary data indicate that information on labour supply collected by questionnaire is reported with error. Thus, computing the value for a missing time-use variable from the time constraint may give rise to an endogeneity bias. Note, however, that the same objection must then apply to the estimation of standard labour supply models in the literature, where the convention is to use data for market hours only, and to compute missing non-market time (referred to as 'leisure') from the time constraint.

⁴¹ As, for example, in Attanasio and Browning (1995), Attanasio *et al.* (1999), Blundell *et al.* (1994), Deaton (1991) and Gourinchas and Parker (2002).

⁴² Computed as labour and non-labour income, net of taxes and including government transfers.

Table 5.1 *Median net income, consumption and saving, \$p.a., 1998, by age*

Age	Net income 1	Market consumption expenditure 2	Saving 3	Female partner % 4	Number of children 5	Predicted male wage 6	Household with children % 7	HES cell size 8
<30	41,704	36,941	3,016	70.8	0.73	15.61	43.6	389
30–34	42,796	37,639	3,900	62.7	1.41	16.90	73.1	465
35–39	44,876	42,163	1,612	71.9	1.80	17.79	85.8	551
40–44	47,320	43,554	2,652	72.3	1.72	18.03	82.8	519
45–49	52,390	47,508	4,030	74.4	1.31	17.71	70.8	486
50–54	50,684	44,114	4,628	72.8	0.67	16.97	39.8	483
55–59	39,312	38,618	1,040	60.1	0.31	15.36	21.7	341
60–64	25,376	31,460	−4,342	34.1	0.13	13.85	9.4	266
65–70	21,684	25,908	−2,184	8.7	0.10	12.44	1.3	239
70+	20,382	18,928	−884	6.6	0.01	10.45	0.1	379

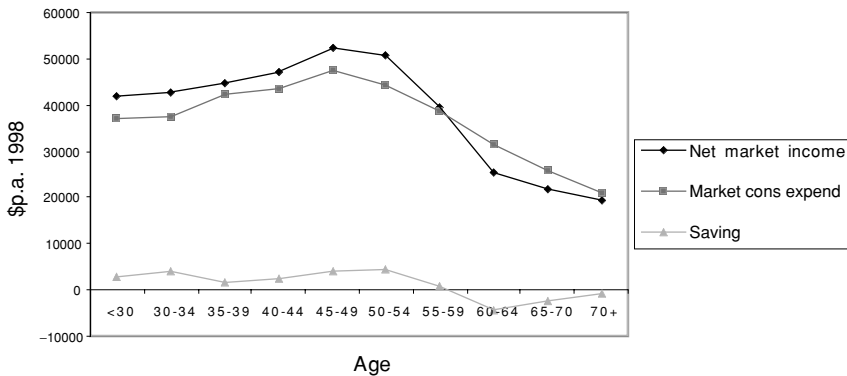


Figure 5.6 Median net income, consumption and saving, by age

age 40), followed by saving for retirement (from age 40 to 60). Table 5.1 also reports the profile of the predicted male wage,⁴³ which, as we would expect, also exhibits the usual hump shape.

A number of studies have already pointed out the role of female employment and demographic variables as determinants of this hump shape of income and consumption profiles.⁴⁴ In table 5.1 we can see that female participation drops during the child-bearing years and the number of children in the household peaks soon thereafter. Note, however, that although the female participation rate falls, household income continues to rise. Blundell *et al.* obtain the same result for the UK and comment specifically ‘what

⁴³ The male wage is predicted as the exponent of a linear function of a vector of characteristics, estimated on a sample of 2,217 employed males.

⁴⁴ For example, Blundell *et al.* (1994) and Attanasio and Weber (1995).

is particularly interesting here is that although female employment falls in the early years, household income does not’.

However, this association of rising household income with falling female employment is a consequence of the way in which the data are organised and gives a misleading picture of what is really at work here. When we partition the early age categories into couples who have not yet had children and families with very young children, we find that income rises only for the first group, precisely because female employment in this group remains very high. For the majority of families, as well as the average, the arrival of children causes a dramatic fall in female labour supply and, consequently, in household income. In other words, income rises with age in the early years only for those households in which both partners remain in full-time work – typically the young full-time two-earner couple without children. The shape of the income profile in figure 5.6 represents the median of *neither* household type. Rather, it is driven by the proportions of each household type in the early age categories, and also depends on the labour supply of mothers with young children.⁴⁵ Female participation rates fall in the second age category because, as we can see from table 5.1, the proportion of households with children rises from 44 per cent to 73 per cent from the first to the second age category, and most of these households have children aged under 4. In the third age category, 85.8 per cent of households have children and family size is larger, but the children are older and so female employment rises. The following three working-age categories represent households with older children, or with children who have left home, and therefore record higher female participation rates and higher incomes.

5.4.3 *The life cycle defined on children*

The first step in our approach is to define the life cycle not in terms of calendar years, but in terms of eight phases which seem to us to represent the key transitions in the life cycle of a typical household. Given the decision to have children, the life cycle evolves in a way which seems to be determined by them. This view of the life cycle leads to a representation of the data on labour supplies, consumption and saving for the average household which is as familiar to everyday experience as it is foreign to the economics literature on lifetime consumption decisions.⁴⁶ Before they have children, both household members have high labour supplies, high saving and plenty of leisure. The presence of pre-school children dramatically changes the pattern of labour supply, leading to large falls in female labour supply, saving and leisure. As the children grow up these changes are gradually reversed, with the state, through the public education system, taking over a large part of the burden of child minding and education,

⁴⁵ This of course varies across countries. Given the higher participation and full-time employment rates of young mothers in the US, for example, it is not surprising to find that studies using US data, such as that by Attanasio and Weber (1995), obtain a single humped profile of both female hours and household income, by age of head.

⁴⁶ This is not to say that the importance of ‘demographics’ has been ignored, as we hope our earlier discussion has made clear. Our contention is that the effects of having children on female labour supply choices, in the presence of an imperfect capital market, are much more significant than seems to be recognised in the literature, and that this significance is made clearer by the way in which we organise the data.

Table 5.2 *Median net income, market consumption and saving, \$p.a., 1998, by phase*

Life cycle phase	Net market income 1	Market consumption expenditure 2	Saving (1–2) 3	Number of children 4	HES cell size 5
1	52,104	40,422	7,800	–	449
2	39,676	38,765	1,040	1.56	508
3	42,120	40,393	1,872	2.30	518
4	45,292	43,417	2,652	2.09	510
5	55,120	53,680	988	1.71	518
6	49,764	42,382	6,136	–	501
7	25,740	30,817	–2,704	–	506
8	19,084	20,905	–1,352	–	507
All	40,664	38,249	1,508	–	4,017

allowing increases in the labour supply of the female as second earner. Consumption of market goods steadily increases and borrowing falls, with high saving levels again being achieved in the phase immediately after the children have left home.

The two samples are split into the eight phases according to criteria that capture the presence and age of children and the later transition of their parents from work to retirement, and that also give phases of nearly equal cell size. Phase 1 is limited to couples in which the female partner is aged under 40 years and there are no dependent children present. Phase 2 contains families with at least one child under 5 years and no older children present, or a child under 2 years but no child over 9 years. Phase 3 families have at least one child aged 5 to 9 years and may have a younger child aged 3 to 4 years or an older child. In phase 4 the children are predominantly in the 12 to 14 year age group. In phase 5 families have older dependent children still living at home. There are no children present in phase 6 to 8 households. Phase 6 is defined to include couples in which at least one partner is aged under 55 years or the male partner is under 60 years and not retired. Phase 7 represent couples in which the male partner is aged under 65 years or is between 55 and 60 and partially or fully retired. In phase 8 both partners are retired.

5.4.4 *The average household*

Table 5.2 reports life cycle profiles of median net household income,⁴⁷ expenditure on market goods and saving, in columns 1 to 3 respectively, using the HES sample. Column 4 gives the average number of dependent children in each phase and column 5 lists cell size. The excess sensitivity puzzle is confirmed by the profiles in columns 1 and 2, which show the strong tendency of household consumption to track net household

⁴⁷ Net household income includes all government direct (cash) benefits but not indirect benefits through, for example, the education and health systems.

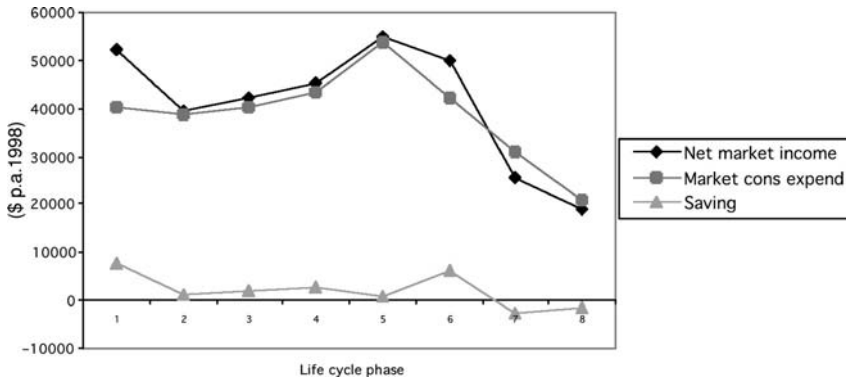


Figure 5.7 Median net income, consumption and saving, by phase

income, with the highest median consumption expenditure coinciding with the highest net income in phase 5. This is brought out clearly in figure 5.7. The figure shows that there is first a sharp fall in median income as the household moves from phase 1 to phase 2. This is then followed by the usual ‘hump’ shape in both net income and consumption from phases 2 to 8. Saving is at its highest in the pre-children phase, drops sharply in phase 2 with the arrival of children, and fails to rise to near its phase 1 level until phase 6 when the children have left home.

While it is clear that market income and consumption are strongly associated with the presence of children shown in column 4, data on these variables alone can give an entirely misleading picture of the true paths of income and consumption, and of the impact of demographic variation, because they exclude the household’s implicit income and expenditure on domestic production. Information on the household’s expenditure on leisure is also essential because, as the time-use data we now present show, it cannot be assumed to be separable from domestic work and labour supply.

Table 5.3 reports life cycle profiles of time allocations to market and domestic work. The table lists TUS weighted data means for male and female market hours, domestic hours and total hours of work, in columns 1 to 6 respectively. Comparing these hour profiles with those for net household income and consumption in table 5.2, it is immediately apparent that much of the variation in net household income across phases 1 to 6 reflects changes in female labour supply or, more specifically, the reallocation of time from market to domestic work by the second earner after the arrival of children. Across these phases there is relatively little variation in male market hours but large changes in the hours of the female partner, which are negatively related to domestic hours of work. This is shown graphically in figure 5.8.

The strong negative relationship between female market and domestic hours suggests that the two types of work become close substitutes after the arrival of children. The most dramatic substitution occurs in phase 2, reflecting the fact that young children generate a high demand for care. This can in general be provided at home or bought on the market, but the time-use data show that there is a very large domestic supply of child care.

Table 5.3 *Male and female hours of market and domestic work, p.a.**

Life cycle phase	Male hours of work			Female hours of work		
	Market 1	Domestic 2	Total 3	Market 4	Domestic 5	Total 6
1	2,286	614	2,900	1,789	1,012	2,801
2	2,271	1,514	3,785	681	3,416	4,097
3	2,266	1,504	3,770	684	3,163	3,847
4	2,259	1,266	3,525	976	2,541	3,517
5	2,244	928	3,172	1,057	1,987	3,044
6	2,220	842	3,062	1,099	1,679	2,778
7	1,080	1,180	2,260	414	1,769	2,183
8	0	1,458	1,458	0	1,750	1,750
All	1,784	1,167	2,951	811	2,169	2,980

* Weighted TUS data means

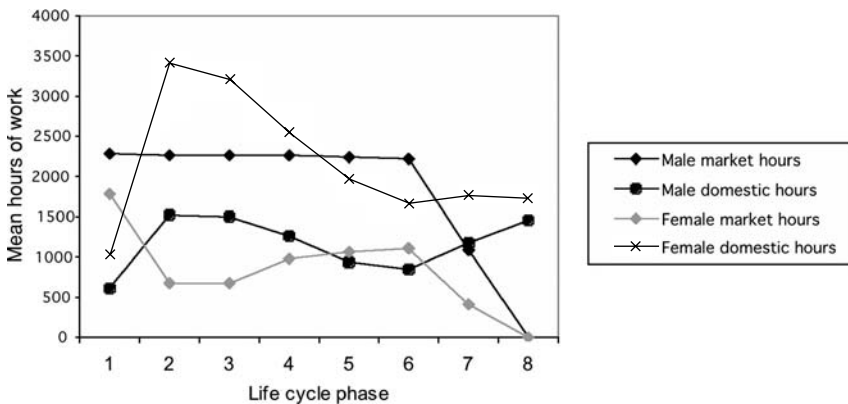


Figure 5.8 Male and female hours of market and domestic work

In phases 2 to 5, of the female partner's total domestic hours, she allocates, on average, 2,095, 1,715, 985 and 471 hours to child care, while the male partner allocates 870, 775, 465 and 179 hours, respectively. Even though family size is larger in phase 3 (see table 5.2), more time is allocated to child care in phase 2 because of the predominance of children aged 0–4 in that phase.⁴⁸ The data show that *total* hours of work rise, and therefore leisure falls, quite dramatically with the arrival of children, and this is then steadily reversed over successive phases of the life cycle.

When consumption expenditure includes the time cost of domestic production, its profile tends to track total hours of work, rather than net income. This is shown in

⁴⁸ The average number of children aged 0–4 in phase 2 is 1.43, and in phase 3 it is 0.75. There are no children of this age group present in subsequent phases.

Table 5.4 *Domestic and full consumption expenditure, \$p.a., 1998**

Life cycle phase	Domestic consumption expenditure 1	Domestic child care 2	Indirect government benefits 3	Government child care and education benefit 4	Taxes-benefits 5	Household full consumption expenditure 6	Adult full consumption expenditure 7
1	19,439	—	4,097	1,453	16,206	67,386	67,386
2	56,883	35,912	9,818	2,368	3,261	108,592	56,010
3	53,545	31,408	16,092	9,640	-4,076	113,646	52,659
4	43,437	17,292	18,841	13,132	-5,565	108,769	57,944
5	35,190	9,402	18,602	12,615	639	111,728	68,946
6	25,980	—	5,159	959	15,188	79,236	79,236
7	28,594	—	7,064	242	-1,924	71,583	71,583
8	28,722	—	12,174	32	-19,066	66,273	66,273
All	35,966	—	11,541	4,828	37	90,132	64,572

* Weighted means

table 5.4. Column 1 of the table presents a profile of domestic consumption expenditure, computed as the product of time allocated to domestic work and the net wage.⁴⁹ Column 2 of the table reports the average cost of parental time allocated to child care in these phases.

Studies reporting life cycle profiles of consumption typically omit indirect government benefits. These are important because they are large, averaging over \$11,500 per household in the sample; they vary quite dramatically across phases; and they tend to vary inversely with the household's cost of domestic child care. Column 3 of table 5.4 reports the profile of average indirect government benefits.⁵⁰ The very high levels in phases 3 to 5 are due to public spending on education and the child care it provides. Government indirect child care and education benefits are shown separately in column 4. Families with children at school or in tertiary education receive by far the larger share of support. In comparison, relatively little is spent on children in the pre-school phase.⁵¹ It is therefore not surprising to find that parents in phase 2 allocate a very large share of their resources to child care, and that they must work longer than parents in later phases to obtain a given level of consumption. The life cycle profile of household full

⁴⁹ On the assumption of constant returns to scale, expenditure on domestic consumption at the implicit price of domestic output is computed as the time input to household production evaluated at the net wage. To obtain the net wage we instrument for the gross wage and compute a tax rate from data on direct taxes and cash benefits. Data on earnings and hours are used to compute hourly earnings as the measure of the gross wage. For further details, see Apps and Rees (2002).

⁵⁰ The HES estimate of indirect government benefits covers non-cash benefits and services for education, health, housing and social security and welfare. For details of the calculation of these benefits, see ABS (2001).

⁵¹ Note that the data mean of \$9,818 for indirect government benefits in phase 2 includes medical costs for the birth of a child and post-natal care. Note also that part of the \$2,368 of government spending on child care and education in this phase goes to children of school age who are also present in some families with a child under 2.

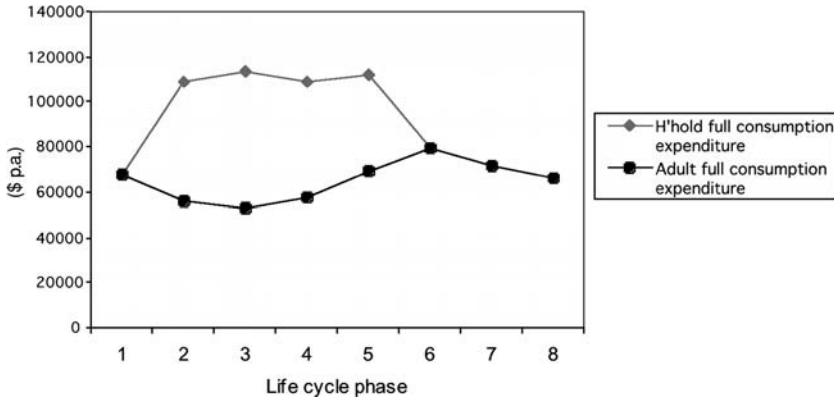


Figure 5.9 Household and adult full consumption

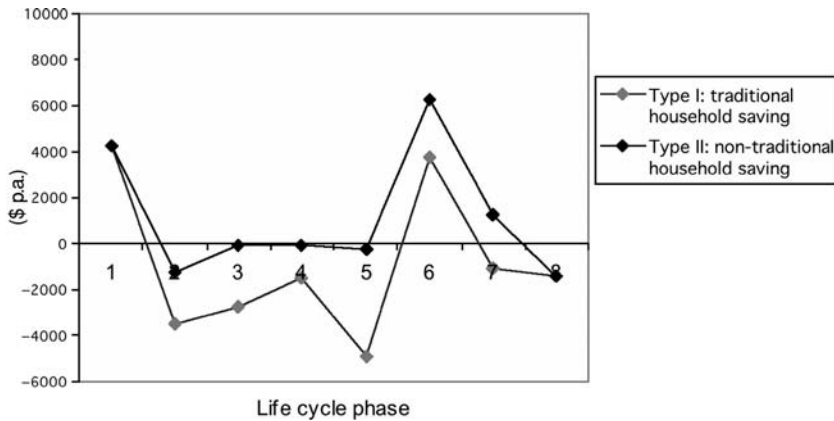


Figure 5.10 Median saving by household type

consumption, computed to include market and domestic consumption and government indirect benefits, is presented in column 5 and depicted in figure 5.9.

The final column of the table reports a profile of the parents' full consumption, obtained by subtracting child costs from household full consumption. Child costs are calculated as the sum of the parents' time costs of child care and government spending on child care and education, plus a share of market and domestic consumption (excluding child care and government education benefits) computed for an 'equivalence scale' that sets the cost of a child to 0.4 that of an adult.⁵² The results indicate that the full consumption cost of children is, on average, around half that of the adults in the household. This is consistent with the estimates of child costs reported in section 4.3.2.

The adult full consumption profile is depicted in figure 5.9. Its U-shape matches the adults' leisure profiles, suggesting that parents cut back on both consumption and

⁵² This ratio of child to adult consumption is used, for example, in Blundell *et al.* (1994) to deflate household market consumption expenditure.

leisure, instead of borrowing more, in order to support their children in the early child-rearing phases. The explanation for this that we suggest is that parents face higher interest rates in the early child-rearing phases, particularly in phase 2, together with a lack of access to good quality, affordable market child care.⁵³ Because there is very little government support for child care, and very high effective tax rates apply to the incomes of mothers who work,⁵⁴ the household's optimal choice is, first, to reallocate the mother's time from market to household work, since she generally faces a lower wage rate, and second, for both parents, but especially the mother, to work longer hours in total, and so reduce leisure, in phase 2.

In later years, the cost of children to parents is substantially reduced by public funding of education. In other words, when the child reaches school age the public education system takes over many of the child-minding activities that the household itself has to undertake for pre-school children. This allows the female partner to expand her market labour supply, while simultaneously reducing total hours of work once the pre-school phase is over. This effect is evident in phase 4 and is accentuated in phase 5. Household income, labour supply and market consumption expenditure all peak in phase 5, with teenaged and older children living at home, while saving is at its peak in the following phase, when the children have left home but market labour supply is still high. Thus, the profile of total hours of work, together with that of adult full consumption, is, we argue, to a significant extent an outcome of an imperfect capital market and variations in the public funding of the costs of children. Once the children have reached school age, access to public education allows parents to maintain family consumption without cutting back excessively on leisure.

This argument is supported by the data on saving and borrowing and on housing available in the HES. These show, on the one hand, how much families must save under a mandatory system of superannuation. They also have an overwhelming incentive to invest in owner-occupied housing if, ultimately, they are to buy housing over their life time at an affordable (and in fact very low) user cost. Table 5.5 lists contributions to superannuation and life insurance (column 1) and mortgage repayments of capital (column 2) by phase. When the sum of these is subtracted from saving, many households are found to be in the position of having to borrow in the short term to finance these forms of long-term contractual saving. The median of the amounts they must borrow in the short term in each phase, calculated as the difference between saving and the sum of mortgage repayments of capital and superannuation contributions, is shown in column 3.

⁵³ To appreciate the inefficiencies and consequent high cost of market child care, one need only consider the impact that government financial support, central planning and regulation has had on primary school care and education, and what would happen to female labour supply and school attendance if that sector were treated in the same way as child care.

⁵⁴ As we show in some detail in chapter 6, Australia now has a tax-transfer system that is, in effect, a system of joint taxation, due to recent reforms to family payments within the income tax system. Married mothers who work lose around half their earnings in taxes and reduced family payments, and so many cannot meet the cost of formal child care out of their net incomes. For a recent empirical analysis of the effects of child care costs and rationing of female labour supply decisions, see Del Boca and Vur (2007), and for a survey of the literature on the effects of policy on both participation and fertility, see Del Boca and Locelli (2008).

Table 5.5 *Long-term saving and housing debt*

Life cycle phase	Superannuation + life insurance policies 1*	Mortgage repayments (capital) 2*	Saving – (mortgage + superannuation) 3**	Housing debt % 4***
1	1,263	3,479	3,868	44.4
2	1,461	2,766	–2,833	39.1
3	1,508	2,702	–2,381	28.4
4	1,779	2,531	–1,296	24.6
5	2,812	2,316	–2,652	14.0
6	2,492	1,520	2,236	9.1
7	1,047	356	–3,315	2.4
8	123	62	–1,404	0.1
All	1,586	1,886	–1,092	18.5

*Weighted data means, \$p.a. 1998

**Median saving net of mortgage repayments of capital and superannuation contributions, \$p.a., 1998

***Debt to house price ratio

The imperative to save for full home purchase is indicated by the dramatic decline in debt to house price ratio from phases 1 to 8, shown in column 4, which is matched by a rise in the percentage of home ownership, from 59 per cent in phase 1 to 95 per cent in phase 8. It is straightforward to show that the user cost of owner-occupied housing, obtained by discounting repayments of capital and the initial equity at the time of purchase, becomes negative over time, due primarily to capital gain in an imperfect capital market and also to low transaction costs relative to private rental.⁵⁵

5.4.5 *Within-phase heterogeneity*

The underlying idea of the models considered earlier is that households choose lifetime paths of male and female labour supplies, saving and consumption of household and market goods, given wage rates (net of taxes), interest rates and productivities in household production. Differences in domestic productivities across households and prices of bought-in child care lead to differences in choices of these endogenous variables, for households facing the same wage and interest rates and capital market conditions. The data show that in fact these choices result in a very high degree of heterogeneity in respect of female labour supply and saving behaviour across households with the

⁵⁵ The data suggest that, under these conditions, home ownership is analogous to an annuity with a very high rate of return, especially if households minimise transaction costs by rarely moving over the life cycle. Its preferential tax treatment is also a contributing factor but cannot alone explain the differential between owning and renting over time if one assumes, implausibly, a perfect capital market.

Table 5.6 *Labour supply, income, saving and taxes, by household type, by age*

Life cycle phase	Male market hours 1*	Female market hours 2*	Household private income 3**	Male earnings 4**	Female earnings 5**	Saving 6**	Taxes net of cash benefits 7*
Panel A: Traditional households							
2	2,434	15	40,508	39,728	0	-3,536	1,304
3	2,451	99	41,600	37,492	0	-2,959	-4,139
4	2,485	365	46,956	37,648	3,129	-1,078	-4,803
5	2,447	414	61,204	40,040	7,124	-4,535	1,717
6	2,466	666	61,568	36,400	6,864	3,736	16,318
7	2,362	0	41,184	27,560	0	1,060	8,085
Panel B: Non-traditional households							
2	2,474	1,335	62,348	39,000	20,748	-1,092	11,076
3	2,551	1,610	65,624	41,600	21,788	410	5,339
4	2,511	1,911	68,744	37,544	25,584	249	3,898
5	2,557	2,135	86,320	41,964	30,420	-156	10,596
6	2,487	2,317	76,024	32,240	29,120	6,404	19,421
7	2,418	1,256	45,708	23,920	15,808	2,885	9,765

*Data means. **Medians, \$p.a. 1998.

same wage rates and demographics, which is concealed in the overall average figures considered above.

To give an indication of the empirical importance of this heterogeneity we construct life cycle profiles for two groups defined according to female labour supply as an indicator of domestic productivities. We are limited to this strategy for defining household types because of missing data on domestic output.

Ideally we would like to distinguish between those households in which female labour supply is zero or marginal throughout the life cycle, and those in which it is significant and relatively large over the entire life cycle. This categorisation requires panel data. Since we have access to cross-section data only, we present profiles for a sample of 'in-work' households partitioned into two groups of equal size within each phase according to the female partner's 'usual hours of work'. The sample is selected on the criterion that the male partner's 'usual hours of work' exceed 24 per week. The criterion yields a sample of 2,415 records from the HES dataset and cell sizes of 452, 445, 442, 438, 418 and 220 for phases 2 to 7, respectively.

Table 5.6 presents labour supply, income and saving profiles for the two groups across these phases. Columns 1 and 2 give data means for male and female market hours. The latter reveals the high degree of polarisation in female labour supply. Households tend to split into two distinct types: those in which the female partner is a non-participant or supplies relatively little market labour and those in which she is employed full-time or

works relatively long part-time hours.⁵⁶ We label the former, in panel A, ‘traditional’ and the latter, in panel B, ‘non-traditional’.

Columns 3 to 6 present median household private income, male and female earnings and household saving net of mortgage repayments and superannuation contributions. From the income profiles it can be seen that the difference between household private income and male earnings in each phase approximates female earnings, indicating that non-traditional households have higher incomes in each phase primarily because of the additional market hours and earnings of the female partner, and not because of higher wage rates. We also find that very little of the variation in female hours can be explained by wage rates or demographics.⁵⁷ From column 6 it can be seen that household saving also tracks female labour supply. Non-traditional households not only have higher household incomes, they also have much higher levels of saving or lower levels of borrowing, according to phase.⁵⁸ This is shown graphically in Figure 5.10.

The differences between household income and saving profiles of each type might lead one to expect corresponding differences in market consumption expenditure. However, this is not the case, due to large transfers from non-traditional to traditional households through the tax-transfer system. Column 7 of the table presents profiles of taxes net of benefits, including indirect taxes and government benefits, for each household type. In phase 3, for example, the traditional household receives, on average, a net benefit of \$4,139 whereas the non-traditional household has a net loss of \$5,339, a difference of over \$9,500. In phase 2 the gap is even larger, although in this phase both lose because of the low level of support for young children. These differences reflect the fact that, *ceteris paribus*, non-traditional households pay much more in income taxes, while receiving very little in family payments or compensation for the cost of child care. In the retirement phase, with higher levels of saving, they are less likely to be eligible for the income-tested age pension. In effect, non-traditional households save for their own retirement and contribute to financing pensions for traditional households, by working longer hours and paying higher taxes.

Because this outcome is achieved primarily by a system of family payments tightly targeted on the basis of second earner and family income, its impact is very high in terms of effective average tax rates on the majority of second earners, as shown in chapter 6. Under this type of regime, small differences in domestic productivities could be sufficient to give rise to the considerable heterogeneity in female labour supply decisions that we observe.

⁵⁶ Part-time employment status is defined as 1–35 hours of work per week and full-time as 35 hours of work or more per week.

⁵⁷ Traditional households in the ‘in-work’ samples have an average of 1.71, 2.42, 2.13 and 1.79 children in phases 2 to 5, respectively. The corresponding figures for non-traditional households are 1.47, 2.07, 1.97 and 1.73.

⁵⁸ Note that income generated by household production is inherently non-saveable, though to the extent that it substitutes for market consumption it may permit higher saving.

5.4.6 Empirical specification of a model

A number of simplifying assumptions are required to deal with the limitations of the data available for estimating a model designed to capture the above features of the family life cycle. The approach we take is based broadly on estimating a within-period production and demand system that is a highly simplified version of the model of child costs described in chapter 4. In fact, we make use of the findings of that model to compute the costs of children. The resulting children's share of full income within each household is then treated as a lump sum transfer from parents, and includes parental time costs as well as a share of a general domestic good and market consumption.

On the production side we specify constant returns to scale Cobb–Douglas (CD) functions and estimate the parameters simultaneously with those of an aggregate AI demand system for the adults. The within-period price of the domestic good, p_t at time t , is then computed for each record as a function of net wage rates and the production parameters specific to each record, together with a scaling factor, consistent with the CD form.⁵⁹ Given that we are estimating a two-adult demand system, aggregation restrictions require that family members face the same prices. This condition is not satisfied if the adults face different prices (wage rates) for leisure. To deal with this problem, we also specify the leisures as inputs to the production of a single leisure good, z , to obtain a single price, q_t , for leisure.

Suppressing the household type and phase subscripts, the indirect utility function takes the form

$$u(q_t, p_t, c_t) = (\ln c - \ln a(q_t, p_t))/b(q_t, p_t) \quad (5.80)$$

where c is full consumption and $a(q_t, p_t)$ and $b(q_t, p_t)$ are the price indexes of the AI demand system. We estimate the within-period model on data for the sample of households described in the preceding section, excluding those that report negative incomes or income primarily from sources other than wages and salaries. The sample contains 2,151 records. Full income is defined on the basis of a time constraint of 14 hours per day.⁶⁰

We specify the intertemporal demand system by introducing the phase subscripts and rewriting the above indirect utility function as

$$u_t = \hat{a}_j(q_t, p_t) + \frac{\ln c_t}{b_j(q_t, p_t)} \quad j = 1, \dots, 8 \quad (5.81)$$

with

$$\hat{a}_j(q_t, p_t) \equiv \frac{-\ln a(q_t, p_t)}{b_j(q_t, p_t)}. \quad (5.82)$$

⁵⁹ This approach has the potential for parameter bias due to endogeneity of time allocations. However, with missing data on domestic output there is inevitably a trade-off between this problem and parameter bias due to omitted domestic price variables.

⁶⁰ For further detail, see Apps and Rees (2003).

The solution to the household's problem yields the life cycle and across-household profiles of full consumption, and the estimated within-period demand and labour supply functions can then be used to derive profiles of market and domestic consumption, saving and second earner labour supplies, for the perfect and imperfect capital market models respectively.

Perfect capital market Given the functional form for indirect utility, the first-order conditions for this problem in the perfect capital market case are

$$\frac{\rho^t}{\delta^t b_j(q_t, p_t) c_t} = \lambda \quad t \in \phi_j \quad j = 1, \dots, 8 \quad (5.83)$$

$$\sum_{t=0}^T \delta^t c_t = W \equiv \sum_{t=0}^T \delta^t \left(A \sum_{i=1}^2 w_{it} + P_t \right) \quad (5.84)$$

where W is 'full wealth'. The important thing to note is that the marginal utility of consumption expenditure in each period depends on the prices of the domestic good, p_t , and leisure, q_t , and therefore on the wage rates and the domestic productivity. Thus the entire time profile of consumption, as well as its allocation within each period as between market and domestic consumption, depends on this productivity. The solution of the system is given very simply by

$$c_t = \alpha_t c_T \quad (5.85)$$

$$c_T = \frac{W}{\sum_{t=0}^{T-1} \delta^t \alpha_t + \delta^T} \quad (5.86)$$

$$\alpha_t \equiv \left(\frac{\rho}{\delta} \right)^{t-T} \frac{b_8(q_T, p_T)}{b_j(q_t, p_t)} \quad t \in \phi_j \quad j = 1, \dots, 7 \quad (5.87)$$

Imperfect capital market In principle, this problem could be fairly complicated to solve. However, from the data, we can establish that, at the margin, the average household is in equilibrium at the saving interest rate in phases 1, 6 and 7, and at borrowing interest rates in phases 2 to 5, where it can also be established from the data that the latter are higher than the former. Denoting the discount factors by $\delta(t, j)$, $j = 1, \dots, 8$, $t = 0, \dots, T$, we can use these to collapse the single-period budget constraints into a lifetime wealth constraint, which we write as

$$\sum_{t,j} c_t \delta(t, j) = W \quad (5.88)$$

where wealth W is computed from the full income data and the discount rates. The household again maximises the utility function in (5.80) subject to this wealth constraint, yielding the first-order conditions

$$\frac{\rho^t}{\delta(t, j) b_j(q_t, p_t) c_t} = \lambda \quad t \in \phi_j \quad j = 1, \dots, 8 \quad (5.89)$$

together with the wealth constraint. We then have to solve the equations

$$c_t = \hat{\alpha}_t c_T \quad (5.90)$$

$$c_T = \frac{W}{\sum_{t,j} \delta(t, j) \hat{\alpha}_t + \delta(T, 8)} \quad (5.91)$$

$$\hat{\alpha}_t \equiv \frac{\rho^{t-T} \delta(T, 7) b_8(q_T, p_T)}{\delta(t, j) b_j(q_t, p_t)} \quad t \in \phi_j \quad j = 1, \dots, 7 \quad (5.92)$$

for the optimal time path of consumption expenditure.

5.4.6.1 Results The models in section 5.2 showed how the optimal path of life cycle consumption c_t depends on the marginal utility of consumption in each phase, as a function of the discount factors $\delta(t, j)$, $t = 0, \dots, T$, and the price index $b_j(q_t, p_t)$. The price index, in turn, depends on the prices of the domestic good, p_t , and leisure, q_t , and therefore on wage rates and domestic productivity. As already noted, unless we have information on all price variables, including the discount rates, it is not possible to test alternative life cycle models without constructing data for those that are missing. Here, we have information that supports our hypothesis that the majority of households in the child-rearing phases, particularly those in the earlier phases, borrow short term facing an interest rate above their lending rate, and we select discount rates consistent with this hypothesis.

5.4.7 Intertemporal profiles of consumption

The criterion we apply as a ‘test’ of the alternative hypotheses concerning the capital market is how well the life cycle profiles of total and full consumption generated by the actual data on saving can be predicted using the estimated preference parameters and selected discount rates. The first step is the construction of a reference or baseline total adult consumption profile, which involves computing consumption expenditure in each period, c_t , $t \in \phi_j$, $j = 1, \dots, 8$. The calculation uses sample data means for time allocations, wage rates and the tax-transfer system in each phase, and median saving in phases 1 to 7, as reported in the preceding tables. Consumption in phase 8 is obtained by compounding up previous saving/borrowing at selected discount rates. We then compute reference full consumption and leisure profiles using the parameters of the within-phase production and demand system, incorporating error terms as specified, to ensure that we can predict the data.

We select a relatively low real lending rate of 0.5 per cent, which we consider plausible for the average lender facing relatively high transaction costs and for whom the real rate is often masked over time by inflation. As noted in section 5.3.2, compulsory superannuation payments and the capital component of mortgage repayments are treated as exogenous amounts subtracted from income in the preretirement phases (i.e. as taxes), compounded up and added back into P_t in phase 8. We compound up compulsory contributions to superannuation at an even lower real rate, of 0.1 per cent, to take account of the risk associated with choice of fund (which is now becoming increasingly

Table 5.7 *Reference consumption profiles, by phase*

Life cycle phase	Two-adult total consumption expenditure, c_t 1	Two-adult full consumption expenditure 2	Household full consumption expenditure 3
1	124,271	71,136	71,136
2	84,788	56,982	111,894
3	83,325	52,583	117,429
4	96,811	59,159	115,337
5	116,589	65,787	110,511
6	120,533	72,425	72,425
7	134,925	71,487	71,487
8	132,291	70,188	70,188

evident) in addition to transaction costs. In contrast, the capital component of mortgage repayments is compounded up at 1.0 per cent, to make an adjustment for the high rate of return to owner-occupied housing. Results based on these rates are presented in table 5.7.

Column 1 of the table reports the reference two-adult total consumption expenditure profile, computed as described. Column 2 gives the reference full consumption expenditure profile of the two adults, obtained by subtracting their leisure expenditure in each phase. Column 3 presents the reference household full consumption expenditure in each phase, computed as the sum of that of the adults and the transfer they make to the children. The costs of children computed for the reference case are held constant across the models.

To test the perfect capital market hypothesis we compare these profiles with those predicted using the same sample means for time allocation, wage rate, taxes, transfers and indirect government benefits, but not the data on saving. Instead we use the preference parameters, β_j , $j = z, y, x$, to obtain c_t . We then use the full set of parameters of the production and demand system to compute within-period full consumptions and leisure expenditures. Table 5.8 reports the results in the same format as in table 5.7.

The predictions for the same profiles under the imperfect capital market hypothesis are presented in table 5.9, again in the same format. The results are derived for real borrowing rates of 1.4 per cent, 1.6 per cent, 1.4 per cent and 1.1 per cent in phases 2 to 5, respectively. Though these rates may appear low in absolute terms, they are proportionately much higher than the lending interest rate. It is also important to keep in mind that they represent across-household averages, and that there is considerable heterogeneity across households in saving/borrowing behaviour within each phase, particularly in the early phases. Many households will in fact face either higher or lower rates.

The profiles of two-adult total consumption expenditure, c_t , reported in column 1 of each of the preceding three tables are compared graphically in figure 5.11. The

Table 5.8 *Perfect capital market consumption profiles*

Life cycle phase	Two-adult total consumption expenditure, c_t 1	Two-adult full consumption expenditure 2	Household full consumption expenditure 3
1	108,554	60,081	60,081
2	109,262	77,310	132,212
3	111,285	74,734	139,569
4	113,254	71,690	127,866
5	116,847	65,966	110,692
6	114,954	68,308	68,308
7	117,440	59,918	59,918
8	114,187	58,229	58,229

Table 5.9 *Imperfect capital market consumption profiles*

Life cycle phase	Two-adult total consumption expenditure, c_t 1	Two-adult full consumption expenditure 2	Household full consumption expenditure 3
1	123,280	65,118	65,118
2	84,495	56,745	111,655
3	82,108	51,646	116,492
4	95,322	58,043	114,216
5	115,723	65,179	109,900
6	130,555	79,905	79,905
7	133,334	61,918	61,918
8	129,687	68,443	68,443

reference profile is strongly U-shaped across phases in which dependent children are present, as we would expect from the evidence in section 5.2. The imperfect capital market model predictions match very closely those of the reference profile. In contrast, the perfect capital market model predicts that the household will smooth consumption expenditure on a per capita adult basis. The positive slope of the profile reflects the choice of a real interest rate greater than the time preference rate. The main departure of the imperfect capital market profile from the reference profile appears in phase 1. The model predicts a level of total adult consumption below that of the reference, but note that it is well above that predicted under the perfect capital market hypothesis.

The perfect capital market model also predicts relatively smooth profiles of adult full consumption and of leisure, which are clearly inconsistent with the data. Figures 5.12 and 5.13 show graphically the full consumption predictions of the two models for the household and for the two adults. Each figure includes the corresponding reference

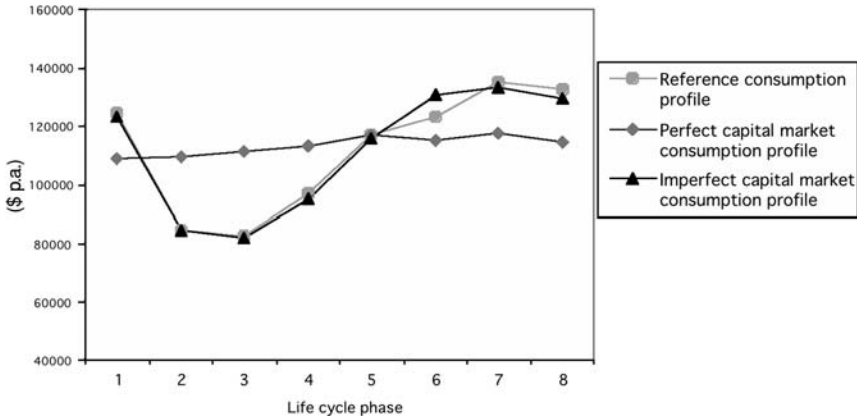


Figure 5.11 Two-adult total consumption profiles

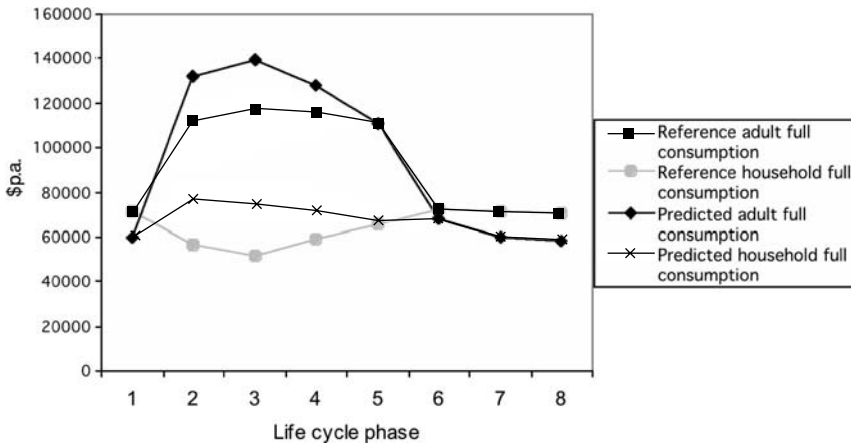


Figure 5.12 Perfect capital market model

profiles. Figure 5.13 illustrates some of the implications of the perfect capital market hypothesis. Because the model generates a relatively smooth profile of two-adult full consumption, adding in the costs of the children's full consumption gives a more strongly humped profile of household full consumption across phases 2 to 5 than indicated by the data. In other words, evidence of a more hump-shaped profile of household full consumption across the phases in which children are present is required in order to support the perfect capital market hypothesis.

Comparing Figures 5.12 and 5.13, it is immediately apparent that the imperfect capital market model predicts full consumption paths that are much closer to the reference profiles, although there are some differences. The most marked departure appears in phase 1, where the imperfect capital market model yields a lower level of full consumption, and a correspondingly lower level of expenditure on leisure.

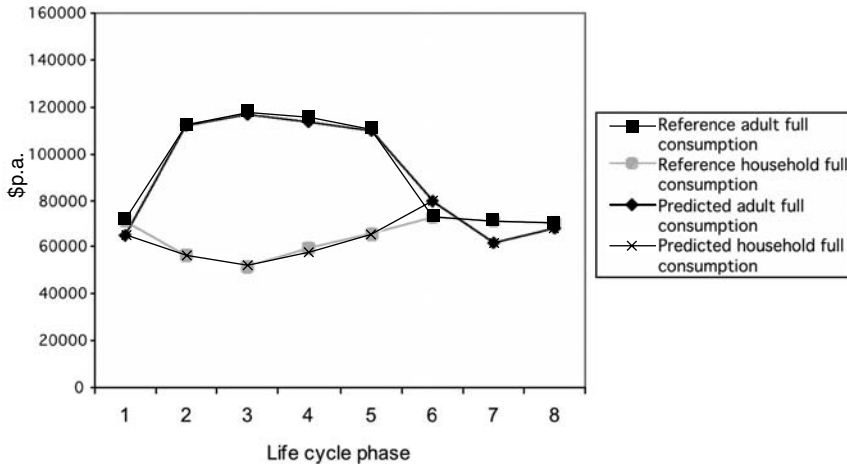


Figure 5.13 Imperfect capital market model

5.5 Conclusions

Our descriptive picture of a household's time allocation, income, saving and consumption over its life cycle, defined in terms not of calendar years but of key phases in the evolution of the family, helps resolve some of the 'puzzles' that have been noted in the existing literature, but suggests a new one: why, in the phase in which the household has pre-school children, are there such dramatic changes in time allocations, consumption and saving? The data suggest that the standard assumption of a perfect capital market is untenable, but so is the hypothesis that households are entirely liquidity-constrained and do not borrow in the short term.

The overall explanation we have proposed in this chapter goes as follows: in the pre-children phase when young couples are both working they save in anticipation of what will happen when they have children. When the children arrive, couples react differently in respect of the female's labour supply, in a way that depends on net-of-tax wage rates, the cost and availability of non-parental child care, productivities in household production, and longer-term labour market considerations such as the effects of reducing market work on human capital and future employment possibilities and wage rates. In any case, there will be a fall in income available for adult consumption and saving, because of some combination of loss of income from the second earner and costs of child care. Because of high borrowing rates in an imperfect capital market the couple do not smooth their consumption path by borrowing. They react instead by reducing market consumption, leisure and saving (other than through house purchase and long-term contractual saving). Once through this phase, with the school system taking over the function of child care, female labour supply, household income, saving and consumption start to increase, but there is still considerable heterogeneity across households that persists over the remainder of the lifecycle.

More generally, we are proposing an approach to life cycle saving and consumption behaviour which sees the endogenisation of the income process, driven by female labour supply choices, as central.⁶¹

Although we have simplified the analysis by ignoring uncertainty, we certainly would not want to claim that it is unimportant in reality, nor that a precautionary motive may not be operative in those phases in which households save. Indeed we would see that as an interesting direction in which to develop the model we have presented here.

Our results have strong implications for public policy, at a time when declining fertility is seen as the major cause of population ageing and consequential problems in sustaining social security programmes, such as pay-as-you-go pension systems. Greater support for households during the critical early childhood phase, for example by making available high-quality child care at low prices, could help overcome the problems presented by an imperfect capital market and reduce the opportunity costs of having children. Reducing the high effective tax burdens (documented in chapter 6) on working mothers, although increasing this opportunity cost, would result in increased female labour supply and so have a positive income effect on fertility, provided adequate child care is in place.⁶² The increased tax base resulting from higher female labour supply could make such policy measures virtually self-financing. Removing the disincentive effects of high incremental taxation on working wives would also reduce the considerable loss of human capital that takes place when many women substantially reduce their market labour supply after the arrival of children. Of all the various life cycle hypotheses, we would see the family life cycle hypothesis as having the most relevance to public policy.

⁶¹ Recall that three of the four approaches to the life cycle surveyed in section 5.2 took household income and labour supply as essentially exogenous.

⁶² This is argued more formally in some detail in Apps and Rees (2004).

6 Household taxation: introduction

6.1 Introduction

The seminal papers by Mirrlees (1971) and Sheshinski (1972) laid the foundations of the modern theory of income taxation for the cases of non-linear and linear taxes respectively. Their fundamental contribution was to formalise the idea of the trade-off between equity and efficiency in redistributive taxation, and to show how this determines, in conjunction with social preferences for redistribution, the structure of a tax system. This trade-off defines in effect a cost function for redistribution, which conditions the extent to which a ‘social planner’ would find it desirable to carry it out.

Their analytical framework, the behavioural model which defines the relationships between tax rates, labour supply, tax revenue and welfare, is that of the individual worker/consumer dividing his time between work and leisure, and using the income from labour supply to buy consumption goods. It is perhaps arguable that this model is a reasonable abstraction for the type of household that overwhelmingly characterised developed economies up to the late 1950s and early 60s, in which the male spouse specialised in market work and the female in household work. If the assumption is made that the within-household income distribution is made just as the ‘social planner’ would wish,¹ and the value of untaxed household production is constant across households and independent of the tax system, then the results of the Mirrlees/Sheshinski optimal tax analysis could be applied to the problem of taxation of these ‘traditional’ households, in which only one market labour supply is relevant. Even then, however, there are issues such as gender discrimination and the determinants of the distribution of productivities in market work, taken as exogenous in the optimal tax models, that could lead one to dispute this.² Be that as it may, the transformation of the nature of the household, extensively discussed earlier in chapter 1, makes it clear that the theory of taxation needs to be extended to deal with the issues raised by the taxation of couples, especially in the case of those with dependent children.³

¹ This is called ‘non-dissonance’ in Apps and Rees (1988) and is further discussed in chapter 7.

² See, for example, Apps (1981), (1982).

³ Not necessarily because of the presence of children as such, but because of the effects their arrival has on female labour supply, as we discussed in chapter 5.

The first part of this chapter provides a general, non-rigorous discussion of this subject, and sets out our classification of tax systems according to the structure of marginal rates and choice of tax unit. More analytical work on which the discussion is based is provided in the following chapters 7–9.

The second part of the chapter illustrates the types of tax systems encountered in practice. We compare the income tax systems that apply in Australia, Germany, the UK and the US, distinguishing between *formal* and *effective* marginal rate schedules. Each of these countries has a formal income tax system specified in terms of taxable income bands and a schedule of marginal rates. In all four, the formal marginal rate schedule applying to taxable income is progressive. In Germany and the US the tax base is joint income. In Australia and the UK, the base is individual incomes. We first compare the structure of average and marginal rates under the formal systems of each country. We then go on to identify the structure of effective rates resulting from combining the formal income tax with child benefits, tax credits and tax offsets. The analysis shows how these policy instruments can result in a structure of effective tax rates that is very different to that of the formal system, and how they can also be used to modify the tax base of a system ostensibly based on individual incomes in the direction of joint taxation.

6.2 Taxation systems

We begin by setting out a classification of possible tax systems. Given some measure of the income that is to be taxed, denoted by y , the first distinction we make concerns the relationship between this and the amount T of tax deducted, the tax function $T(y)$. We have three basic cases:

Linear taxation

The taxpayer receives a lump sum payment a and pays a single, constant marginal tax rate $t \in [0, 1]$, so that the tax function is

$$T = -a + ty. \quad (6.1)$$

We could think of this as consisting of a ‘universal benefit’ a , which is then ‘withdrawn’ as a function of income at a rate t . Sheshinski (1972) was concerned with the optimal values of the parameters (a, t) . Note that if and only if $a > 0$, this system has *average rate progressivity*, in the sense that the average tax rate T/y is increasing with y , i.e.

$$\frac{d}{dy} \left[\frac{T}{y} \right] = \frac{a}{y^2} > 0. \quad (6.2)$$

From the point of view of distributional fairness, this is often regarded as a desirable attribute of a tax system – higher income taxpayers pay a larger proportion of their income in tax. If $a = 0$ we have a *flat rate* or *proportional tax*, with the same marginal

and average rate at every income level, and therefore zero progressivity. If $a < 0$ and $t = 0$ we have a *poll tax*, which is regressive because it has a falling average rate, $a/y^2 < 0$.

Piecewise linear taxation

The taxpayer receives a lump sum payment $a \geq 0$, and there is a sequence of n *threshold income* values $y_1, y_2, \dots, y_n$, and $n + 1$ associated marginal tax rates $t_0, t_1, \dots, t_n$ such that the tax function is

$$T_0 = -a + t_0 y \quad y \in [0, y_1] \quad (6.3)$$

$$T_1 = -a + t_0 y_1 + t_1(y - y_1) \quad y \in (y_1, y_2] \quad (6.4)$$

$$T_2 = -a + t_0 y_1 + t_1(y_2 - y_1) + t_2(y - y_2) \quad y \in (y_2, y_3] \quad (6.5)$$

and so on until

$$T_n = -a + t_0 y_1 + t_1(y_2 - y_1) + \dots + t_n(y - y_n) \quad y \in (y_n, \infty). \quad (6.6)$$

The intervals $[0, y_1]$, $(y_{i-1}, y_i]$, $i = 2, \dots, n$ and (y_n, ∞) are the *tax brackets*.

Piecewise linear tax systems are the most frequently encountered in practice, typically with relatively few brackets. Typically, also, we have $a = 0$ and $t_0 = 0$, with the lowest bracket $[0, y_1]$ listed as a ‘personal exemption’, ‘tax-free allowance’, ‘standard deduction’, or something similar.

Formal tax systems also typically have the feature of increasing marginal rates or *marginal rate progressivity*, with $t_0 < t_1 < \dots < t_n$. In this case, the tax system is said to be *convex*.⁴ Obviously, marginal rate progressivity implies average rate progressivity, though, as we saw in the linear tax case, not conversely. If, for any tax rate t_i , $i = 0, \dots, n - 1$, we have $t_{i+1} < t_i$, we call this tax system *non-convex*.⁵

Although formal tax systems are almost universally convex, and so we might conclude that marginal rate progressivity is also universally regarded as a desirable attribute of a tax system, in practice there are often types of cash benefit system which run parallel to the formal tax system, and which create non-convexities in the *effective tax system*, i.e. the system which shows the net effect of the payment and withdrawal of such benefits on the actual tax rates that taxpayers face. We provide examples of this in section 6.4.

Nonlinear taxation

Here the tax function is some general continuous non-linear function $T(y)$. Mirrlees (1971) was concerned with the form this function should take.

⁴ Because, as we show below, it yields a convex budget set for the taxpayer.

⁵ Because the taxpayer's budget set will not be a convex set.

The formal German tax system, for example, is a non-linear function of the form

$$T_1(y) = 0 \quad y \in [0, y_1] \quad (6.7)$$

$$T_2(y) = a_2(y - y_1) + b_2(y - y_1)^2 + c_2 \quad y \in (y_1, y_2] \quad (6.8)$$

$$T_3(y) = a_3(y - y_2) + b_3(y - y_2)^2 + c_3 + T_2(y_2) \quad y \in (y_2, y_3] \quad (6.9)$$

$$T_4(y) = a_4(y - y_3) + b_4 + T_3(y_3) \quad y \in (y_3, \infty) \quad (6.10)$$

where the a s, b s and c s are given numbers. The quadratic functions in the second and third brackets are designed to give marginal tax rates, $T'_i(y)$, $i = 2, 3$, which increase linearly in taxable income. On the other hand, the marginal tax rate in the top bracket, a_4 , is constant. The parameters are chosen so that the marginal tax rate in the third bracket increases at a slower rate than that in the second bracket, so that there is a non-differentiability in the marginal rate function, a kink, at y_2 , as well as at y_3 . Thus, after an initial tax-free income range, the formal German tax system has marginal tax rates that increase in a piecewise linear way with taxable income, until they reach a constant top marginal rate. It is therefore a convex system, in the sense defined above. However, added to this formal tax system are taxes to support the churches and the reconstruction of eastern Germany, as well as a host of special allowances, e.g. for children, commuting expenses, night-time working, etc., so that ultimately the only way to work out the effective marginal tax rate structure is to compute it for typical cases, as we do in section 6.4 below.

In addition to this three-way classification of tax systems into linear, piecewise linear and non-linear, we have a classification according to how the individual incomes of household members are treated for tax purposes. Thus, given the individual taxable incomes⁶ y_f , y_m , the issue is the form of the tax function $T(y_f, y_m)$. Again, we distinguish three possible systems.

Joint taxation

Often also known as ‘income splitting’, under this system the taxable incomes of the individuals are summed, divided by 2, and then subjected to a given tax rate structure. In terms of our previous discussion we have $y = (y_f + y_m)/2$. Thus the tax function can be written as $T((y_f + y_m)/2)$. The immediate implication is that the marginal tax rates faced by the two individuals are equalised, regardless of the degree of inequality of those incomes: $\partial T/\partial y_f = \partial T/\partial y_m = T'/2$. A second implication is that the amounts of tax paid by a couple with the same total income is the same, regardless of how much is earned by each partner, so that a high-income single-earner household pays the same amount of tax as one containing two low-income earners with the same total income.

As income-splitting is sometimes described, ‘the first dollar earned by the wife is taxed at the same rate as the last dollar earned by the husband’. However, this ignores the existence of earned income tax thresholds. For example, in the US in 2007 each

⁶ That is, gross incomes adjusted for whatever tax allowances exist for *individual* incomes.

spouse had a ‘personal exemption’ of \$3,200, and so the marginal tax rate is zero with respect to earnings of either individual below this level. However, once this level is exceeded by the husband and reached by the wife the statement becomes true.

Given this definition of taxable household income, it can be combined with either linear, piecewise linear (as in the US) or non-linear (as in Germany) taxation.

Note that an important feature of a piecewise linear or non-linear joint taxation system is that the marginal tax rate applied to one income may change with the level of the other income. For a twice differentiable tax function, this can be concisely expressed as

$$\frac{\partial^2 T}{\partial y_i \partial y_j} \geq 0, \quad i, j = f, m, i \neq j, \quad (6.11)$$

with strict inequality holding at some income levels. In a piecewise linear system like the American one, the decision of a wife to go out to work at an income that exceeds her tax-free threshold will increase the marginal rate applied to the husband’s income if it pushes their joint income into a higher bracket (there is a discrete jump in his marginal tax rate). It will certainly do so at any point below the maximum marginal rate in the German system. The narrower the income brackets between which marginal tax rates vary, the more likely this is to happen, and of course in the continuous German system the ‘brackets’ relevant for marginal rates, below the top bracket, have a width of zero.

We call this important property of a joint taxation system *marginal rate interdependence*. Then it may be useful in certain contexts⁷ to define the *incremental tax burden* of one individual’s income at any given level of the other’s as

$$IB_i = \frac{T(y'_i, y_j^0) - T(y_i^0, y_j^0)}{y'_i - y_i^0} \quad i, j = f, m \quad i \neq j \quad (6.12)$$

where $y'_i - y_i^0 > 0$. This tells us the average rate at which the total tax burden on the household increases when the income of one individual increases and the other is held constant. In general of course this may differ from the average tax rate on the household

$$ATR = \frac{T(y_f, y_m)}{y_f + y_m}. \quad (6.13)$$

Individual taxation

Here the same tax function is applied separately to the incomes of the individuals in the household, so that the tax function (with a slight abuse of notation) takes the additively separable form $T(y_f, y_m) = T(y_f) + T(y_m)$. It follows immediately that we have *marginal rate independence*, since

$$\frac{dT'(y_i)}{dy_j} = 0 \quad i, j = f, m \quad i \neq j \quad (6.14)$$

⁷ For example, when considering the decision of a wife whether or not to take a job when her husband works full-time.

at all income levels. Likewise the incremental tax burden becomes simply $IB_i = [T(y'_i) - T(y_i^0)]/(y'_i - y_i^0)$ and for $y_i^0 = 0$ becomes identical with i 's average tax rate $T(y_i)/y_i$. This is a very important difference to the joint taxation system, where there is marginal rate interdependence. In the latter system, given the assumption that the husband, as primary earner, will certainly go out to work, the question of whether the wife will take a job depends at least in part on the change in the household's total tax bill that results, including any effect on his marginal tax rate of her increase in income, that is, it depends on the incremental tax burden with respect to her work decision. In an independent tax system this effect is absent. This is further discussed below, when we examine some actual tax systems.

The formal tax system in Australia has always been individual, while in the UK individual replaced joint taxation in 1990. Interestingly, the motivation for this does not seem to have been concern with the possible equity and efficiency effects of these systems in the economic sense, but with the more general idea that equality of treatment of women and men in all spheres of life naturally implied individual taxation. The woman's income should no longer be regarded as the property of the man.

As well as marginal rate independence, it also implies that in a piecewise linear system with marginal rate progressivity, the second earner in a household (typically, though not exclusively, female) will often be in a lower marginal tax rate bracket than the primary earner. In a linear tax system, on the other hand, there is no difference between individual and joint taxation, since marginal tax rates are necessarily equal in each case. Thus reductions in the marginal rate progressivity of an individual tax system can also be thought of as moving it in the direction of joint taxation, since marginal tax rates of primary and second earners will be brought closer together.

Note that, as we show in section 6.4, although the formal tax systems in the UK and Australia are individual, the fact that certain forms of transfer payments are withdrawn as a function of *household* income introduces elements of joint taxation into them, and it becomes a computational exercise to determine the extent to which the structures of effective marginal and average tax rates, as well as incremental tax burdens, differ significantly from those we would expect under a joint taxation system.

Selective taxation

We give this name to a tax system in which the incomes of individuals in the household are taxed separately according to different tax schedules. An example is provided by the theoretical analysis of Boskin and Sheshinski (1983),⁸ who consider the problem of finding the optimal values of the parameters (a, t_f, t_m) in the linear tax function

$$T(y_f, y_m) = -a + t_f y_f + t_m y_m \quad (6.15)$$

where m denotes the (male) primary earner and f the second earner. They argue not only that the marginal tax rates should be unequal, but that optimally $t_f < t_m$. We examine their analysis at some length in the next chapter.

⁸ See also the 'gender based taxation' of Alesina *et al.* (2007).

This tax function is an example of *separable selective taxation*, where the tax function takes the additively separable form $T_f(y_f) + T_m(y_m)$, and therefore possesses the property of marginal rate independence. A further example of this would be where the functions $T_f(y_f)$ and $T_m(y_m)$ are different and independent piecewise linear tax schedules. Most generally of course we simply have a tax function $T(y_f, y_m)$ in which we could allow any kind of marginal rate interdependence.

It may be thought that selective taxation exists only in the abstract world of optimal tax theory. However, real world examples exist. The effective tax system in Australia is a selective tax system, at least over a range. Every household with a child under 5 receives a cash transfer of \$3,584, which is withdrawn as a function of the *second earner's* income at a rate of 20 cents in the dollar, when this income exceeds \$4,380. This makes the second earner's tax schedule steeper than that of the primary earner over a given income range.

In principle then, combining the two classifications we could have nine possible formal tax systems, though this falls to eight when we note that individual and joint linear tax systems are identical. In practice, the overwhelming majority of formal tax systems are convex, piecewise linear, joint or individual systems, though effective tax systems can differ in important ways from the formal systems, especially in terms of convexity and the degree of jointness. The German system is a rare example of a continuous system, and the Australian effective tax system gives an example of selectivity, though certainly not of the kind that theorists such as Boskin and Sheshinski would have imagined.

6.3 The simple analytics of redistributive tax systems

First we compare the characteristics of linear, piecewise linear and continuous tax systems for a single taxable income y . We then extend the discussion to two-earner households and household production. We assume initially that we have an economy consisting of three consumers, with given market wage rates w_i , $i = 1, 2, 3$, and $w_3 > w_2 > w_1 > 0$. This wage, assumed to be a measure of an individual's productivity in market production, determines the consumer's *type*. Consumers have identical, quasilinear utility functions

$$u_i = x_i - v(l_i) \quad v', v'' > 0 \quad i = 1, 2, 3 \quad (6.16)$$

where l_i is labour supply. Their gross incomes are $y_i = w_i l_i$. It is then useful to write these utility functions as

$$u_i = x_i - v\left(\frac{y_i}{w_i}\right) = x_i - \psi_i(y_i) \quad i = 1, 2, 3. \quad (6.17)$$

Note three important properties of these utility functions, which play a useful role in the diagrammatic analysis to come:

- the marginal rate of substitution $\psi'_i(y_i) > 0$ is independent of x_i ; indifference curves are vertically parallel along any perpendicular to the y -axis in the (y, x) -plane;

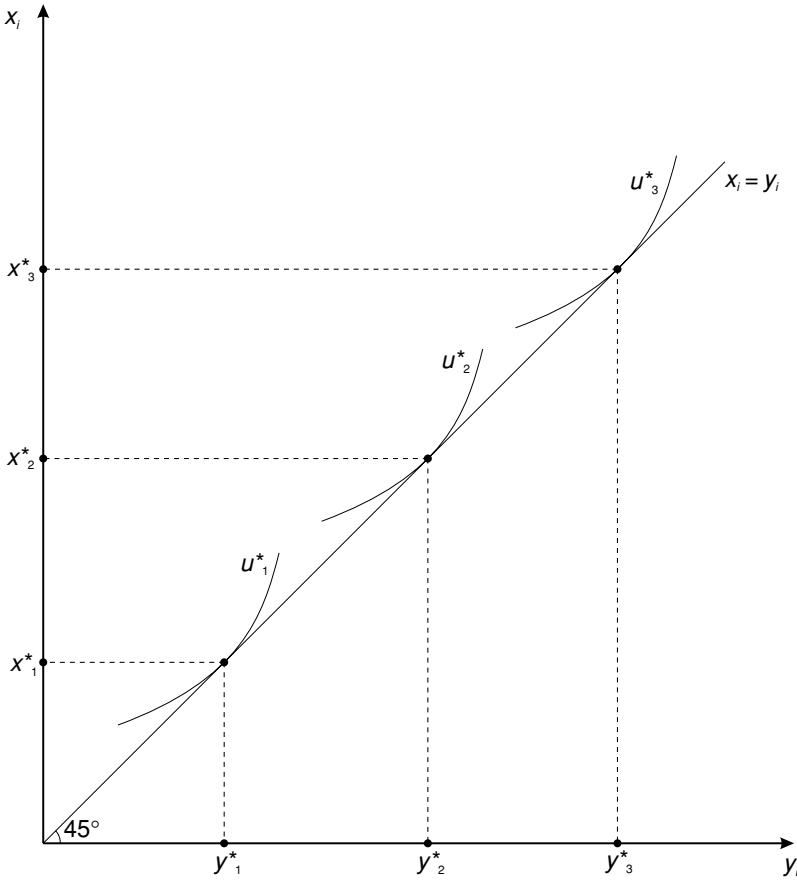


Figure 6.1 Equilibrium without taxation

- $\psi_1(y) > \psi_2(y) > \psi_3(y)$ at any point (y, x) . The higher the wage type, the less hard she has to work to achieve a given income;
- $\psi'_1(y) > \psi'_2(y) > \psi'_3(y)$ at any point (y, x) . The higher the wage type, the smaller the increase in effort required to achieve a given increase in income. This is a *single-crossing condition*.

Figure 6.1 shows an initial equilibrium of these consumers in the absence of taxation. Their budget constraints are simply $x_i = y_i$, and the first-order condition is $\psi'_i(y_i^*) = 1$. Using the last of the above properties it is easy to show that at the equilibrium $u_3^* > u_2^* > u_1^*$.

6.3.1 Linear taxation

Now suppose a planner wants to make the distribution of utilities more equal, and chooses a linear tax to do so. Figure 6.2 shows a possible equilibrium in this case. Everyone receives the same lump sum $\hat{a}$ and pays the same marginal tax rate $\hat{t}$, so that

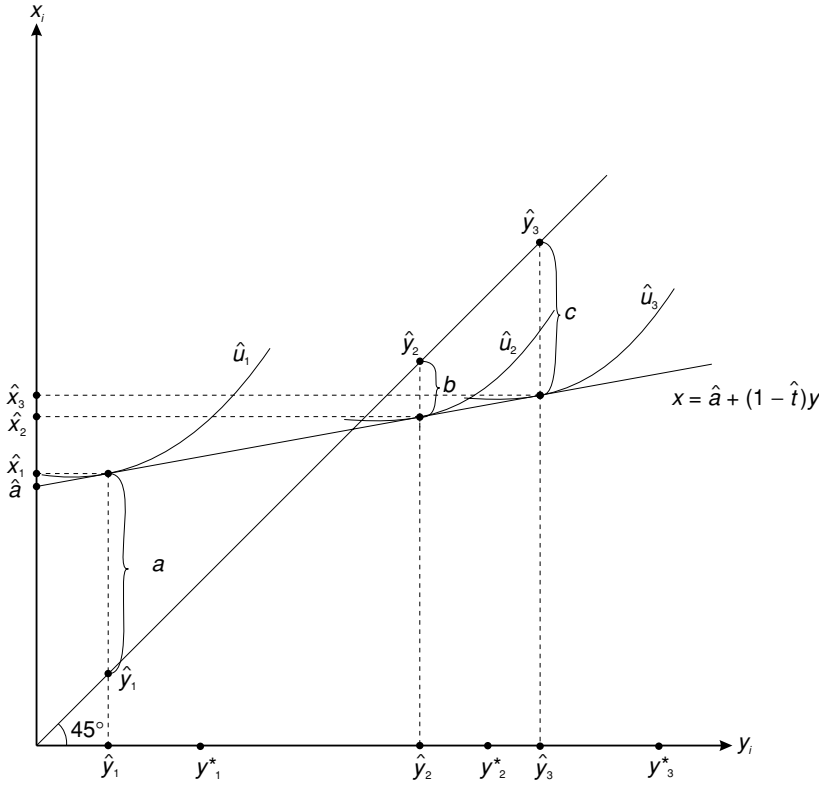


Figure 6.2 Equilibrium with linear taxation

their budget constraints become $x_i = \hat{a} + (1 - \hat{t})y_i$. Thus they are all in equilibrium at $\psi'_i(\hat{y}_i) = 1 - \hat{t}$. The lowest wage type is receiving a net transfer, the second wage type is paying a small net amount in tax,⁹ the highest wage type is paying a higher net amount. Since the planner's budget constraint must balance, aggregate consumption must equal aggregate income, so we have

$$\hat{x}_1 - \hat{y}_1 = \hat{y}_2 - \hat{x}_2 + \hat{y}_3 - \hat{x}_3 \quad (6.18)$$

as shown by the vertical distances a , b , c in the figure. We could alternatively express this constraint as

$$3\hat{a} = \hat{t} \sum_{i=1}^3 \hat{y}_i. \quad (6.19)$$

Clearly, raising a requires raising t – a vertical shift upwards in the budget constraint must also flatten it – and implies a greater redistribution from higher to lower wage

⁹ This is an arbitrary aspect of the example; it could go either way.

types, since the higher the wage, the greater the contribution to the total ‘revenue requirement’ $3a$ that a given type has to pay. How far the planner wants to go in this direction will depend on her ‘taste for redistribution’.¹⁰

However, there is a cost to redistribution, and hence a trade-off. We see that the reduction in the net wage caused by the tax reduces the labour supply and gross income of type i by $y_i^* - \hat{y}_i$, $i = 1, 2, 3$. These represent the disincentive effects of the taxation, and in the aggregate give the deadweight loss, measured in terms of gross income, $\sum_i (y_i^* - \hat{y}_i)$.¹¹ Clearly the higher is a , and therefore t , the flatter is the budget constraint and the greater are these deadweight losses. The cake shrinks as it is being redistributed. The optimal linear tax regime is found at the point at which the planner judges that the marginal social benefit from increased redistribution is just equal to the marginal cost in terms of reduced aggregate income.

6.3.2 Nonlinear taxation

James Mirrlees formulated the optimal taxation problem as one of asymmetric information, in the form now known as the adverse selection problem. If the planner were able to observe each consumer’s wage type, she could simply choose lump sum redistributive taxation, taking a fixed sum away from the higher wage types and giving it to the low wage type, without creating any distortions of labour supply or gross incomes. In the absence of deadweight losses, redistribution is costless and so she could achieve her preferred income distribution. However, if she cannot observe the wage type of an individual, this cannot work: each higher wage type will pretend to be a low wage type to get the transfer, and there will be no tax revenue to finance it.¹²

The solution is given by the idea of incentive compatibility: the planner designs a tax system that offers three alternative linear taxes,¹³ so structured that each wage type finds it optimal to choose the tax designed for her own type. Figure 6.3 shows the general form of the optimal solution, and compares it to the case of zero taxation shown in figure 6.1.

The optimal tax system has four key properties:

1. To achieve incentive compability, the wage type 2 must be indifferent¹⁴ between the pair $(\tilde{y}_1, \tilde{x}_1)$ and $(\tilde{y}_2, \tilde{x}_2)$, and the wage type 3 must be indifferent between $(\tilde{y}_2, \tilde{x}_2)$ and $(y_3^*, \tilde{x}_3)$. Clearly, type 1 strictly prefers $(\tilde{y}_1, \tilde{x}_1)$ to both of the other two allocations.

¹⁰ Note that here we are concerned with the case of *pure redistribution*. It could be that there is a government revenue requirement $G > 0$, say to finance a public good, so that taxation would be positive even in the absence of redistribution. However, as long as the planner has some concern for equity of the income distribution, the analysis would not change in any essential respect. If she does not, then a simple poll tax could be set, which would imply no deadweight losses.

¹¹ Note that the assumption of a quasilinear utility function implies that these changes do not depend on the height of the budget constraint but only on its slope. These changes are compensated changes; there are no income effects.

¹² This is shown more formally in chapter 8. ¹³ Strictly, two linear taxes and a lump sum tax – see below.

¹⁴ This reflects the assumption that when indifferent between lying about one’s wage type and telling the truth, one chooses the latter.

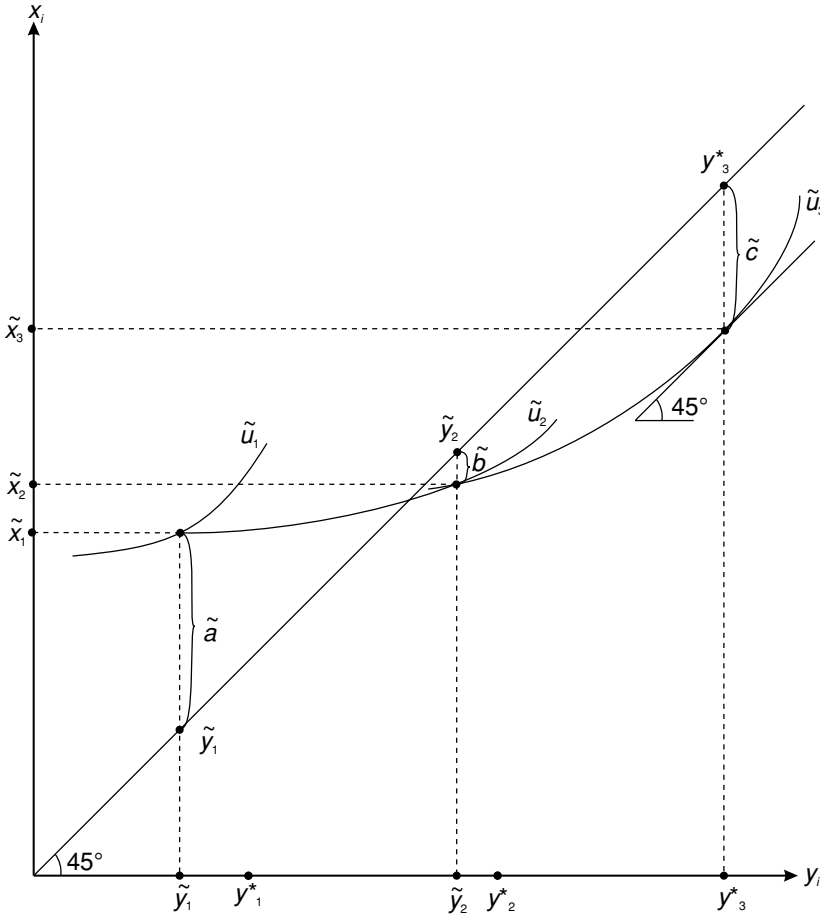


Figure 6.3 Equilibrium with non-linear taxation

2. The lowest wage type receives a net payment $\tilde{x}_1 - \tilde{y}_1 > 0$, while the highest type pays a tax $\tilde{x}_3 - y_3^* < 0$.¹⁵ Again the budget must balance, with $\tilde{x}_1 - \tilde{y}_1 = \tilde{y}_2 - \tilde{x}_2 + y_3^* - \tilde{x}_3$, as shown by the distances $\tilde{a}$, $\tilde{b}$ and $\tilde{c}$.
3. There are deadweight losses due to the distortions to the gross incomes of types 1 and 2, since $\tilde{y}_i < y_i^*$, $i = 1, 2$. Thus the total deadweight loss in gross income terms is $\sum_{i=1}^2 (y_i^* - \tilde{y}_i)$. As we show formally in chapter 8, it can never be optimal to leave the effort levels of the wage types below the top wage type undistorted at their levels in the absence of taxation, y_1^* and y_2^* . Essentially, distorting the effort supplies of the lower types to below their no-tax levels allows higher taxation of the higher

¹⁵ Again, it could go either way in the case of the middle type, depending on the parameters of the problem. Indeed, it is possible that the middle type would receive exactly the same pair as type 1. In the literature this is referred to as ‘bunching’.

wage types and more redistribution. The optimal tax structure then results from the trade-off between the cost of these distortions and the benefit of redistribution.

4. The last property of the solution is that the top wage type's gross income level is undistorted – it is kept at the first best level y_3^* . This is the famous 'no distortion at the top' result. It arises essentially because there is no gain, in terms of redistribution, to distorting the effort supply of the top type, because there is no higher type who could as a result be taxed more heavily. In other words, as long as we put type 3 somewhere on the indifference curve $\tilde{u}_3$ in figure 6.3, we achieve incentive compatibility, and so are free to choose the point on this curve at which to place her, and $(y_3^*, \tilde{x}_3)$ is the point which maximises the tax revenue we can obtain from her.¹⁶

Figure 6.3 shows the optimal solution in terms of gross income-consumption points (y_i, x_i) . It is, however, straightforward to represent these in terms of a tax system, as figure 6.4 shows. The planner offers a menu of the three taxes:

1. A linear tax with parameters $(\tilde{a}_1, \tilde{t}_1)$, for $y \in [0, \tilde{y}_1]$,¹⁷ and with $\tilde{t}_1 = 1 - \psi'_1(\tilde{y}_1)$. This gives the budget constraint $x = \tilde{a}_1 + (1 - \tilde{t}_1)y$, thus inducing type 1 to choose $(\tilde{y}_1, \tilde{x}_1)$.
2. A linear tax with parameters $(\tilde{a}_2, \tilde{t}_2)$, for $y \in [0, \tilde{y}_2]$,¹⁸ and with $\tilde{t}_2 = 1 - \psi'_2(\tilde{y}_2)$. This gives the budget constraint $x = \tilde{a}_2 + (1 - \tilde{t}_2)y$, thus inducing type 2 to choose $(\tilde{y}_2, \tilde{x}_2)$.
3. The lump sum tax $\tilde{T}_3$.

In general we have $\tilde{t}_1 \geq \tilde{t}_2$, and since in any case the marginal tax rate for the highest income group is zero, we certainly do not have marginal rate progressivity. These features of an optimal tax system in Mirrlees's model might be found puzzling in the context of real-world tax systems, but are perfectly rationalisable in terms of the model. The fact that a marginal rate may be small does not imply that someone may not be paying a lot of tax. The purpose of the *marginal* tax rate is to induce the individual to choose the appropriate gross income-consumption point by setting the slope of her budget constraint. The height of this constraint, and therefore the net amount of tax paid or received, is determined by the parameters $\tilde{a}_1$, $\tilde{a}_2$ and $\tilde{T}_3$ respectively. Moreover this 'no distortion at the top' result may be given undue prominence in this model with a small discrete number of types. In a model with a continuum of types, those near the top could be paying high marginal rates, and the type at the very top of the continuum is of measure zero.

¹⁶ Note that the 'no distortion at the top result' takes the form that optimal y_3 remains at the first best level y_3^* only because we have assumed quasilinear utility functions. If we took more general utility functions, the result would be that the first best *first-order condition*, that the slope of 3's indifference curve equals 1, would continue to hold, but the level of y_3 would be distorted. See chapter 8 for a fuller discussion of this point.

¹⁷ Note that the upper bound on income that may be earned under this tax must be set at $\tilde{y}_1$, otherwise higher wage types could take it, choose their optimal y -values, and be better off than at the planner's allocation.

¹⁸ See the previous footnote.

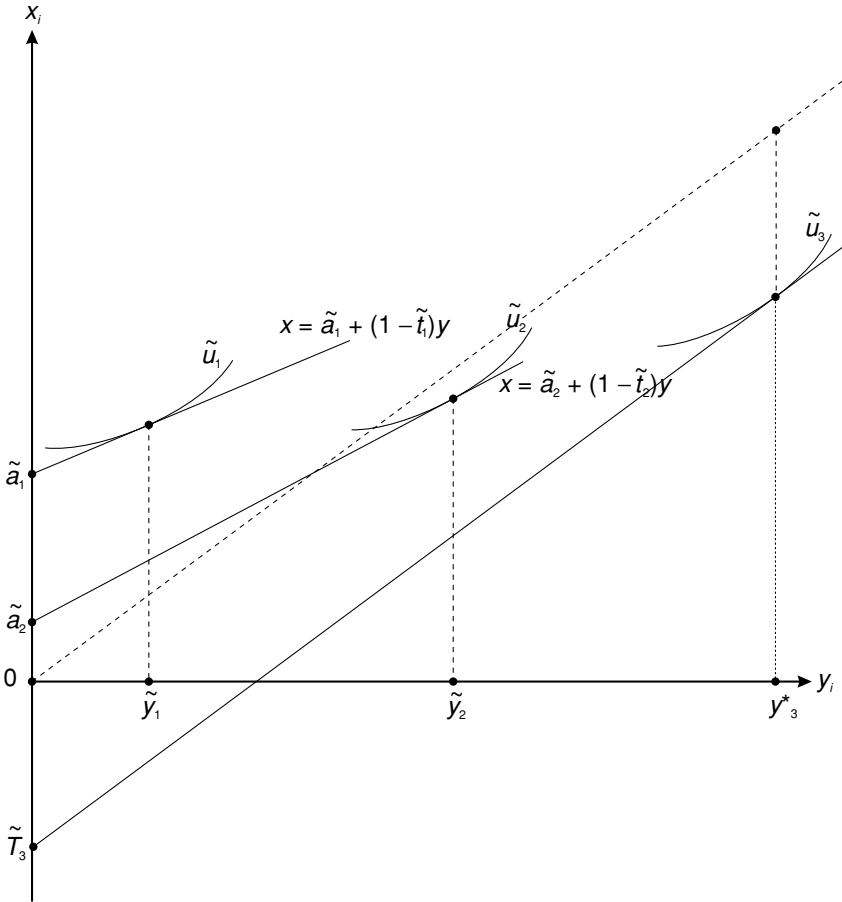


Figure 6.4 Implementing the optimum with taxes

6.3.3 Piecewise linear taxation

Linear taxation is simple to implement but rather tightly constrained in the possibilities for redistribution that it makes available. Since all individuals are offered the same budget constraint, i.e. are ‘pooled’, it does not have to deal with the problem of adverse selection. On the other hand, optimal non-linear taxation requires a great deal of information for its implementation: the optimal tax rates are sensitive to assumptions about the utility functions of the various types, the proportions of individuals who are of each type,¹⁹ and the distributional preferences of the planner, i.e. the social welfare function.²⁰ Piecewise linear taxation can be thought of as retaining the relative simplicity

¹⁹ Here we have simply assumed that these are equal. For the generalisation see chapter 8.

²⁰ One may also question whether empirically observed wage rates are a true measure of innate productivity endowments, since they are determined by endogenous decisions on human capital acquisition.

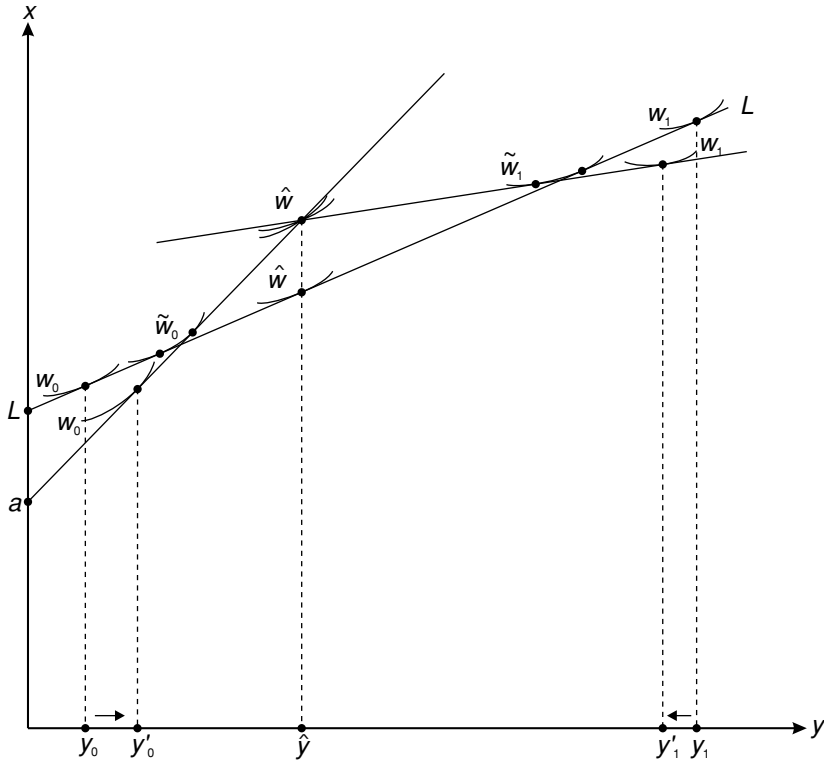


Figure 6.5 Effects of a convex piecewise linear tax

and practicality of linear taxation while expanding the possibilities for redistribution by allowing more than one marginal tax rate.

In examining optimal piecewise linear taxation we find it useful to relax the assumption of only three wage types and assume instead there is a continuum of types, with the wage defined on the interval $[w_0, w_1]$. On the other hand, we consider a piecewise linear tax system with just two tax brackets, since this suffices to make all the main points. We assume also that preferences are just as defined previously in (6.17), so that at any point in the (y, x) -plane there is a continuum of indifference curves with slopes $\psi_y(y, w)$, decreasing with $w \in [w_0, w_1]$. In figure 6.5 we first explore the question of why a convex piecewise linear tax system could improve upon an optimal linear tax.

In the figure, let the line LL denote the optimal linear tax system. Consumers will be distributed continuously at tangency points along this line, with y_0 the gross income of type w_0 , and y_1 that of type w_1 . Now suppose we replace this by the two-bracket piecewise linear tax system,

$$T_1 = -a + t_1 y \quad y \in [0, \hat{y}] \quad (6.20)$$

$$T_2 = -a + t_1 \hat{y} + t_2(y - \hat{y}) \quad y \in (\hat{y}, \infty) \quad (6.21)$$

with $t_2 > t_1$, (marginal rate progressivity). This implies the budget constraint

$$x = a + (1 - t_1)y \quad y \in [0, \hat{y}] \quad (6.22)$$

$$x = a + (t_2 - t_1)\hat{y} + (1 - t_2)y \quad y \in (\hat{y}, \infty) \quad (6.23)$$

Given this budget constraint, consumers will be continuously distributed at tangency points along it, with y'_0 denoting the gross income of type w_0 , and y'_1 that of type w_1 . However, there will now be ‘bunching’ at the kink in the budget constraint at $\hat{y}$. There will be a sub-interval of wage types, $[w', w'']$ choosing gross income $\hat{y}$, corresponding to the values of w for which the inequality

$$1 - t_1 \geq \psi_y(\hat{y}, w) \geq 1 - t_2 \quad (6.24)$$

holds. Note that the types in the interior of this sub-interval, (w', w'') , for which the strict inequalities in (6.24) hold, are not at tangency points, as illustrated by type $\hat{w}$, and so can be thought of as being rationed, since they would like to earn more gross income if this could be taxed at the rate t_1 , but are not prepared to do so if it is taxed at the rate t_2 .

In the figure, we move from the linear to the piecewise linear system by reducing the lump sum from L to a , reducing the marginal tax rate for the first bracket, and increasing it for the second. This has two kinds of incentive effect and three kinds of distributional effect. All those consumers in the lower tax bracket, except those initially at $\hat{y}$, will have increased their gross income and effort levels because their marginal tax rate has fallen. On the other hand, all the consumers in the upper bracket will have reduced income and effort because their marginal tax rate has risen.²¹ Which effect dominates overall depends on the proportions of the population in each of the two brackets, i.e. on the density function of consumer types, and on the relative sizes of the increases and decreases in income of each type. The more elastic the labour supply of the consumers in the lower bracket relative to those in the upper bracket, the more likely that the overall net efficiency effect will be positive.

In terms of distribution, note that those types shown in the figure as $\tilde{w}_0$ and $\tilde{w}_1$ are just indifferent between the two systems. Then, all those wage types in the interval $(\tilde{w}_0, \tilde{w}_1)$ are strictly better off and all those in the intervals $[w_0, \tilde{w}_0)$ and $(\tilde{w}_1, w_1]$ are strictly worse off, since they are on higher and lower budget constraints respectively. Thus there is redistribution from both ends towards the middle. The net effect will again depend on the proportions of households in these groups, and the planner’s evaluation of their respective gains and losses of utility.

Let us now assume that convex piecewise linear taxation is indeed the best system short of fully non-linear taxation, and consider the determination of the optimal values of the parameters of this system,²² and in particular t_1 , t_2 , and $\hat{y}$. Figure 6.6 shows how the tax threshold $\hat{y}$ is determined.

²¹ Note the usefulness of the assumption of vertically parallel indifference curves – quasilinear utility functions, zero income effects – as well as the single crossing condition, in obtaining these and subsequent results.

²² Again this is conducted at the intuitive level, and the reader is referred to chapter 7 for a more rigorous treatment.

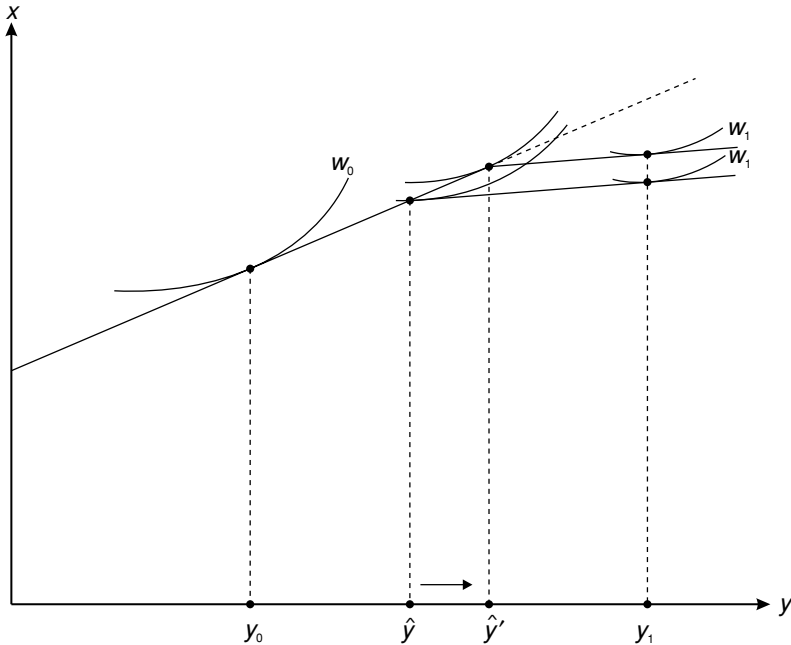


Figure 6.6 Determination of the optimal tax threshold

Suppose that the kink is initially at $\hat{y}$ in the figure, and consider the effects of a small increase to $\hat{y}'$, holding tax rates constant. This clearly leaves the choices of all consumers initially in equilibrium to the left of $\hat{y}$ unaffected, and so we can ignore these. From the planner's point of view, the beneficial effect is that all consumers at $\hat{y}$ who were previously rationed can now increase their labour supply and gross income, and this yields them a net gain in units of consumption of approximately $[(1 - t_1) - \psi_y(\hat{y}, w)](\hat{y}' - \hat{y}) > 0$. Their marginal increase in net income, $(1 - t_1)$, exceeds the marginal disutility of the extra effort, $\psi_y(\hat{y}, w)$. Moreover, their extra income $(\hat{y}' - \hat{y})$ brings an increase in tax revenue of $t_1(\hat{y}' - \hat{y})$. So these will be summed over all the consumers in this group, with their consumption gain being evaluated at the weight the planner attaches to this, their marginal social utility of income. The cost of this arises because all consumers previously to the right of $\hat{y}$ have a vertical shift upwards in their budget constraint, as shown, for example, at y_1 . In other words, the upper-bracket consumers obtain a lump sum consumption gain, and this represents for the planner a loss of tax revenue, while, since these are the better-off consumers, the social value of their utility gain does not compensate the planner for this. The optimal $\hat{y}$ then balances these two sets of effects at the margin.

Given the optimal choice of tax brackets, the problem of the optimal choice of the tax rates is quite straightforward.²³ Essentially, we have two linear tax problems, one

²³ Of course, in the mathematics, optimal threshold and tax rates are determined simultaneously. See chapter 7 for a formal treatment.

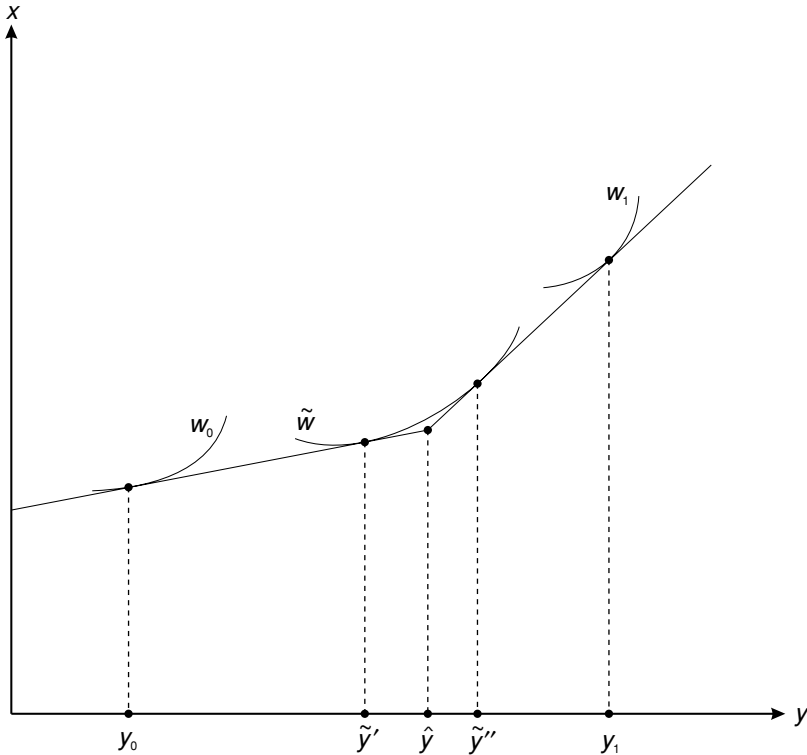


Figure 6.7 A non-convex tax system

for the consumers in the lower tax bracket, one for those in the upper, and indeed the conditions determining the optimal tax rates are very similar to those for the optimal rate in a simple linear tax system.²⁴ The optimal tax rate in each bracket trades off the efficiency losses from a rise in the tax rate aggregated over the consumers in the relevant bracket, against the income distributional effects of the tax within that bracket. Finally, the choice of the lump sum payment a determines the general level of tax rates, with higher a requiring higher tax rates and achieving more redistribution in general, but also creating higher efficiency losses.

All this, however, *assumes* that it is indeed optimal to have a convex tax system. It is possible that the optimal piecewise linear tax system is non-convex, and figure 6.7 shows what this involves.

We can describe the tax system formally as before, except that now we have $t_1 > t_2$. Each consumer then faces the budget constraint shown in the figure. Again consumers are continuously distributed at points of tangency around the budget constraint, with one important difference to the convex case: there is a unique wage type, say $\tilde{w}$, which is just indifferent between earning gross income $\tilde{y}'$ and being in the lower bracket, and

²⁴ For which see chapter 7 below.

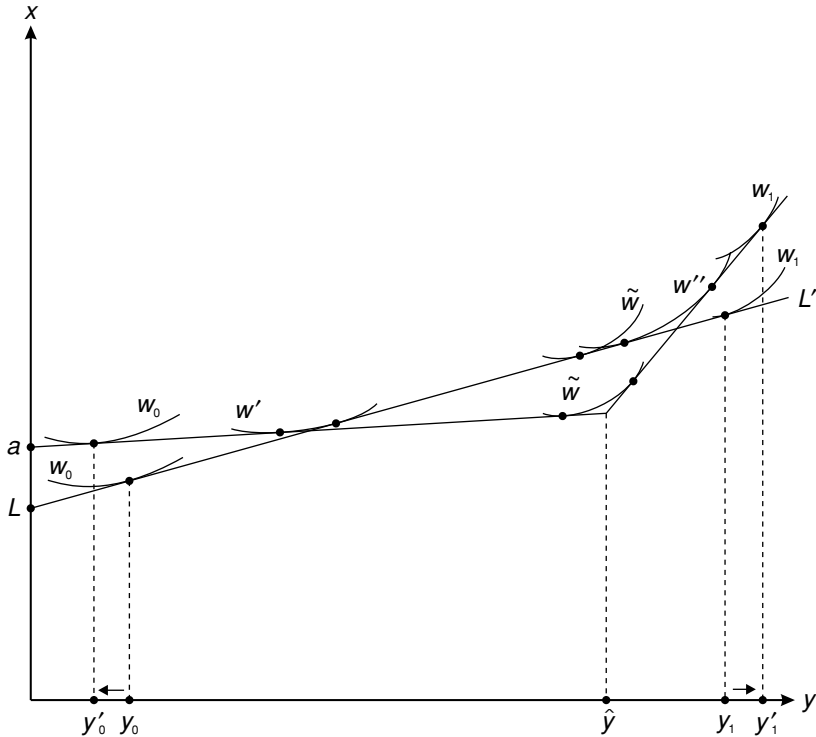


Figure 6.8 Effects of a non-convex piecewise linear tax

earning gross income $\tilde{y}''$ and being in the upper bracket. We cannot say which of these two incomes they will actually choose. However, if we were to increase t_1 slightly, any consumers choosing $\tilde{y}'$ would jump to $\tilde{y}''$, while if we were to increase t_2 slightly, any consumers choosing $\tilde{y}''$ would jump to $\tilde{y}'$. This introduces a discontinuity in the tax revenue function which has to be taken into account in the formal analysis. However, at the intuitive level of the discussion here we can still give an account of the reason we might want to move from an optimal linear tax to a non-convex piecewise linear tax, which will also suggest how we can compare the likely optimality of one or other of the two piecewise linear tax systems.

Thus in figure 6.8 suppose we initially have the linear income tax LL' as shown, and then replace it with the non-convex piecewise linear tax. Raising the tax rate on consumers in the initial bracket will have an adverse incentive effect, causing all consumers with wage rates in the interval $[w_0, \tilde{w}]$ to reduce their gross income and labour supply (assuming for convenience that all $\tilde{w}$ -types choose $\tilde{y}'$). On the other hand, reducing the tax rate in the upper bracket has the reverse effect on consumers in the interval $(\tilde{w}, w_1]$. In distributional terms, there is a redistribution of consumption and utility away from the middle, and toward the lower and higher incomes. Specifically, all those with wage rates in the intervals $[w_0, w']$ and $(w'', w_1]$ are strictly better off

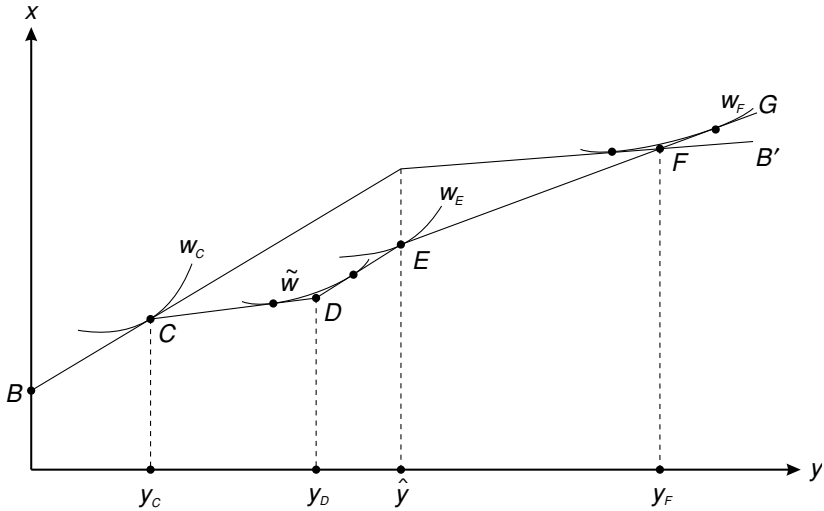


Figure 6.9 Phasing out a universal benefit

than before, those in the interval (w', w'') strictly worse off, while types w', w'' are indifferent. Thus the move to a non-convex piecewise linear system redistributes income away from the middle, and towards the higher and lower income groups.

Choice of the optimal parameters of the non-convex tax system is analogous to that in the convex case, with the important difference that there is a discontinuity involved in the choice of the tax threshold $\hat{y}$. Given choice of this threshold, we have two linear tax problems to determine the optimal tax rates t_1 and t_2 . Then, a small increase $d\hat{y}$ will cause a downward jump in income and effort and a loss of utility for that wage type which now becomes the marginal type, indifferent between the two brackets,²⁵ and so a discrete loss of tax revenue. Against this there is a gain in tax revenue $(t_1 - t_2)d\hat{y}$ from the types in the upper income (lower tax rate) bracket, since they are faced with a lower (parallel) budget constraint as a result of the increase in $\hat{y}$. In other words they now have to pay the higher tax rate on a slightly larger amount of income. The optimal threshold balances these two effects.²⁶

6.3.3.1 Universal benefits, means testing and convexity Consider a system in which there initially exists a universal untaxed benefit B paid to all consumers, together with a convex piecewise linear tax system with tax rates t_1 , t_2 , and $t_1 < t_2$. The implied budget constraint is BB' in figure 6.9. Now suppose it is argued that this universal benefit 'should be made available only to those who really need it', and should be withdrawn above a certain income threshold, y_C in the figure, at a rate of r cents per dollar

²⁵ On the assumption that when indifferent everyone chooses the lower bracket. Nothing essential changes if the opposite assumption is taken.

²⁶ For further discussion of this case and a computational approach to determining the optimum, see Slemrod *et al.* (1994).

of the benefit. The savings resulting from this will be used in a revenue neutral way to reduce the top rate of tax, t_2 , in the ‘interests of improving incentives’. This effectively amounts to changing a convex into a non-convex piecewise linear tax system, with a welfare gain only for the top subset of higher rate taxpayers. Thus, in the figure, assuming the amount of benefit and the tax brackets stay the same, the effect of phasing out the benefit with income is to raise the effective tax rate in the lower bracket to $t_1 + r$, up to the income level, y_D , at which the benefit is fully phased out. The marginal tax rate then falls to t_1 (the new budget constraint is parallel to the previous one over the range DE) and we have the non-convexity. Suppose the upper bracket marginal tax rate is then reduced from t_2 to t'_2 with $t_1 + r > t'_2 > t_1$. Then the budget constraint becomes that labelled $BCDEFG$ in the figure.

The efficiency effects are: gross income and labour supply remain unchanged for wage types in the interval $[w_0, w_C]$, fall for those in $(w_C, \tilde{w})$, remain unchanged in the interval $[\tilde{w}, w_E]$, and increase for the types in $(w_E, w_1]$. Everyone in the interval (w_C, w_F) is worse off; only those in $(w_F, w_1]$ are better off.

Interestingly, this is only a subset of the upper bracket taxpayers, the higher income subset. The lower income subset of higher rate taxpayers lose because the fall in their marginal rate is insufficient to compensate for the higher rate they paid in the phase-out range, or, equivalently, for the loss of the universal benefit. Thus the welfare gains of the policy are relatively narrowly concentrated, and to support the policy it must be argued that the efficiency gains of the lower tax rate in the upper bracket more than outweigh the losses over the interval $(w_C, \tilde{w})$, as well as the adverse distributional effects across the income distribution from y_C to y_F .

6.3.4 Taxing two-earner households

6.3.4.1 Linear taxation In discussing the equity and efficiency effects of taxation across households, it is useful to focus on the case of linear taxation, since the main points can be made most simply in this context. We begin with the analysis by Boskin and Sheshinski (1983), which is generally regarded as providing the ‘conventional wisdom’ on the subject.²⁷ They consider a population of two-person households, with a household’s type defined by a wage pair (w_f, w_m) , $w_i \in [w_0, w_1]$. Each individual divides total time between work and leisure, and we again assume the household’s utility function²⁸ takes the quasilinear form²⁹ $x - \psi_f(y_f, w_f) - \psi_m(y_m, w_m)$, where x is total consumption. Each household receives a lump sum a and pays a tax $T(y_f, y_m) = t_f y_f + t_m y_m$. The problem is to determine the optimal values of the tax parameters (a, t_f, t_m) , given a joint density function $g(w_f, w_m)$ strictly positive on

²⁷ See also Feldstein and Feenberg (1996) who confirm this conventional wisdom.

²⁸ This can be thought of as being derived from the maximisation of a household welfare function, a Samuelsonian HWF, as discussed in chapter 3. Note that this choice of utility function implies that we are ignoring within-household distributional effects in order to focus on across-household effects. This is further discussed in chapter 7.

²⁹ Boskin and Sheshinski take a more general form that does not rule out income effects.

$[w_0, w_1] \otimes [w_0, w_1]$, and a standard social welfare function giving the social planner's preferences over the utility levels of *individuals*. We can regard the main purpose of the analysis as exploring the validity of the following intuition.³⁰ If we accept the stylised fact that women have a higher compensated labour supply elasticity than men, then using standard Ramsey arguments, on efficiency grounds women should pay a lower tax rate than men, i.e. $t_f < t_m$.³¹ The optimal linear tax system is selective, in terms of the classification set out in the previous section,³² with women paying the lower tax rate.

Note that this is stronger than simply arguing for individual rather than joint income as the tax base, which is how the discussion is sometimes presented.³³ Indeed, the case for selective taxation in general is rather easy to establish, since we can think of joint taxation in this linear tax case as simply imposing the constraint on the optimisation problem that $t_f = t_m$, and such a constraint cannot increase and may reduce the value of the maximand. Thus the problem should be initially formulated, as in Boskin and Seshinski, as a selective taxation problem, and the solution will determine which tax rate should be lower, with, in the special case that the optimal tax rates turn out to be equal, joint taxation as a possible outcome.

However, Boskin and Sheshinski's formal analysis does not actually establish this conventional wisdom. In the next chapter, we show that the optimal marginal tax rates can be written as

$$t_i^* = \frac{Cov[s, y_i]}{\bar{y}_i} \quad i = f, m \quad (6.25)$$

where the numerator is the covariance across households between a household's marginal social utility of income, s , and the gross market income of the individual i within it. We can take this covariance to be negative. The denominator is the average value of the compensated derivative of gross income of males or females with respect to their respective tax rate, which is also negative. The stylised fact says that $|\bar{y}_{t_f}| > |\bar{y}_{t_m}|$, but of course this only implies a lower tax rate for women if their covariance term in the numerator is not sufficiently greater in absolute value as to outweigh this. In other words, although taxing women brings with it a higher deadweight loss than taxing men, it could be the case that the tax rate on women is a sufficiently better instrument for redistributing income across households, that overall the planner would still prefer a higher tax rate on women.

³⁰ Already expressed in earlier, non-technical discussions in, for example, Munnell (1980) and Rosen (1977).

³¹ Note that, unless it can be shown that the within-household income distribution significantly favours women in a way disapproved of by the social planner, taking account of within-household distributional effects will not change this result qualitatively.

³² Recall that in a linear tax system joint and independent taxation are equivalent.

³³ See, for example, Piggott and Whalley (1996). Rosen's question 'Is it time for separate filing?' should be understood as being posed in the context of a move to an individual progressive piecewise linear tax system, in which case women would be taxed in general at lower rates than men because their incomes are lower. Likewise, the argument that 'gender-based taxation does not exist' loses much of its force if we would really be implementing tax reform in the context of this kind of tax system. For further discussion of 'gender-based taxation' see Alesina *et al.* (2007).

While not explicitly acknowledging this problem, Boskin and Sheshinski address it by constructing a numerical example of an optimal tax problem and showing that its solution implies a lower tax rate for women. But a single example does not seem to us to be a solid basis for an entire conventional wisdom. Fortunately, as we show in the next chapter, a more general argument can be constructed. If there is positive assortative matching, i.e. the wage rates of men and women who form households together are strongly positively correlated, and if the ratio of male gross income to female gross income rises as we move through the wage distribution,³⁴ then the covariance term for men will be greater in absolute value than that for women. Since the empirical evidence strongly supports these conditions, we can conclude that in this linear taxation framework, and, as we show in the next chapter, in the piecewise linear case as well, given the stylised facts, selective taxation with lower tax rates for women is optimal.

6.3.4.2 *Household production and female labour supply heterogeneity*

Sandmo (1990) analyses optimal linear income and commodity taxation in the presence of household production in the context of single-person households.³⁵ Perhaps the most striking result is how little, at least in the formal expressions determining the optimal tax rate, seems to change. Interpreting the expression in (6.25) now as relating to a household consisting of a single individual, a similar expression characterises the optimal tax rate in Sandmo's model. The gross income derivative will now be determined by the elasticity of substitution between market work and household production, but this makes little difference to the formal results. The main qualitative difference arises in relation to the numerator, the equity term. Sandmo shows that underlying this covariance will now be the relationship between productivity in household production, the market wage rate and the marginal social utility of income. If household productivity and the market wage are positively correlated, then this will tend to increase the absolute value of the covariance between the marginal social utility of income and labour income, and so increase the marginal tax rate, while if the correlation is negative then this has the reverse effect.

The implications of household production increase sharply in significance, however, when we consider the problem of the taxation of couples in economies in which there is specialisation in domestic work by gender, but with a high degree of heterogeneity across households. The question now becomes that of the relation between the tax rates on men's and women's incomes, and how this is influenced by domestic productivity variation across households.

As we have argued at some length in chapter 5, specialisation in household production and female labour supply heterogeneity, which set in when children are first present in the household, have important consequences for the evaluation of tax systems. The

³⁴ This is clearly supported by the data, as indicated by the hours and earnings profiles in figures 1.5 and 1.6.

³⁵ An earlier paper by Boskin (1975) analysed a two-sector general equilibrium model consisting of a taxed market production sector and an untaxed household production sector and estimated the efficiency losses from the resulting distortion, which were significant. Apps and Jones (1986) also examined optimal taxation of couples with household production, in the context of a trade-tax model.

reasons are clear. Household production cannot be included in the tax base. The arrival of children creates a high demand for child care that can be met either by parental care or by working in the market and buying it in. Couples who choose to work in the market and buy in child care will pay more tax than those who specialise in the household production of child care and related services, at any given pair of wage rates (innate productivities). Under non-selective linear taxation, a household in which both partners work full-time in the market will pay up to twice as much tax as one in which only one partner works full-time in the market and the other full-time at home. Under progressive piecewise linear joint taxation, the two-earner couple may pay more than twice as much tax.³⁶

The assumption needed to support this is that single and two-earner couples with the same *household* income are equally well-off. This implies that home child care can be ignored, either because it is totally unproductive, or because the allocation of time to this activity is perfectly negatively correlated with the wage. Neither of these conditions is supported empirically.

More formally, it can be shown that higher taxation of two-earner couples at a given household income is appropriate only if there is a positive association between the second earner's labour supply and the household's utility possibilities. Whether this is the case depends on the relationship between domestic productivity and female labour supply, *inter alia*. We showed in chapter 3 that, theoretically, this could be either positive or negative. Higher domestic productivity reduces the time required to produce a given domestic output, thus freeing time for market work, but also reduces the implicit prices of the household goods and therefore increases demand and the resulting time requirement. Moreover, the higher the husband's wage the higher his demand for domestic output and, other things equal, the less time supplied by the wife to the market. If there is positive associative matching, this could imply that higher-wage wives work more at home.

Clearly, productivity in household production is only one variable that, after controlling for wage rates and demographics, could influence female labour supply. Other important variables are, as discussed in earlier chapters, the price of non-parental child care, the perceived rate of depreciation of work-related human capital and the expected future wage path as a function of current labour supply decisions. In a simple static framework, only the first of these can really be taken into account. We now discuss the effects of the price of non-parental child care in the context of piecewise linear taxation.

6.3.4.3 Piecewise linear taxation and the price of market child care Here we show in a simple model how a non-convex piecewise linear tax system and variation in the price of market child care as well as domestic productivity can generate large differences in female labour supply across households, which do not reflect correspondingly large differences in household utility. We assume that the male partner in the household

³⁶ See section 6.4, where this is illustrated with examples of actual tax systems.

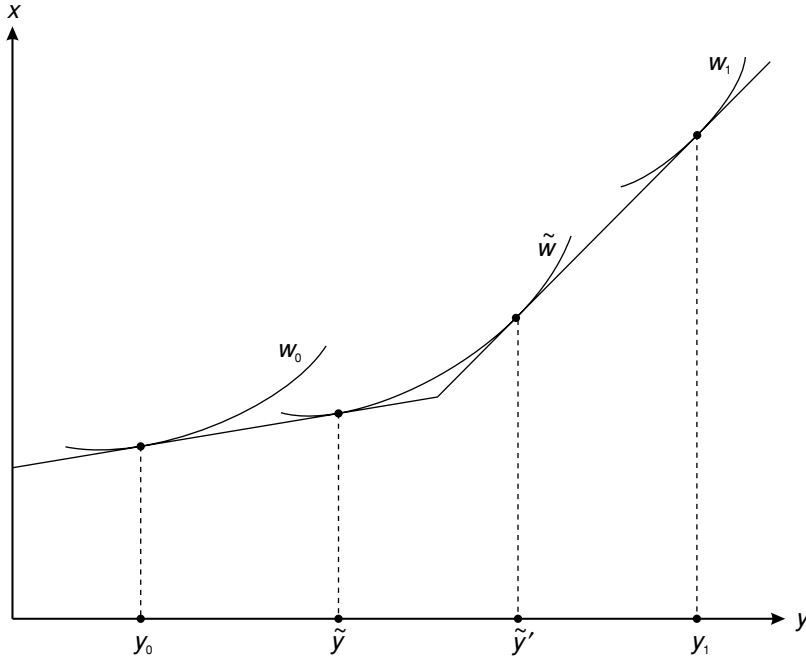


Figure 6.10 Small differences in type, large differences in income and consumption

has a fixed labour supply, and the female divides her time between market labour supply l and child care, c . The household's utility is assumed to depend only on child care and so we can regard its objective as maximising this.³⁷ Child care is produced according to the production function $c(h, m, k)$, assumed concave in h, m for given k , where h is parental time, m a bought-in market input and k an exogenous productivity parameter, distributed on the interval $[k_0, k_1]$. The budget and time constraints are respectively

$$pm = x \quad (6.26)$$

and

$$h + l = 1 \quad (6.27)$$

where p , the market price of the child care input, is distributed on $[p_0, p_1]$, and x is net income, gross income being denoted by $y = wl$, with w the given gross wage. Writing $m = x/p$ and $h = 1 - y/w$, we can write the production function as $\gamma(x, y; k, p)$. For fixed k, p , we can generate from $\gamma(\cdot)$ a family of indifference curves in the (y, x) -plane, and it is straightforward to show that they will have the slope and curvature of those shown in figure 6.10.

³⁷ Introducing consumption into the utility function would just complicate the analysis without adding anything of substance.

In the figure, it is assumed that the piecewise linear tax system has two brackets, with the upper bracket having the lower tax rate. At any point (y, x) we can show that reducing p with k constant reduces the slope of an indifference curve, and so the households are distributed continuously around the non-convex budget constraint with p falling and household utility increasing as we move from left to right. Thus a ranking of households by gross household income would correspond to the ranking by utility levels. Let $\tilde{p}$ denote the type of the household which is just indifferent between income levels $\tilde{y}$ and $\tilde{y}'$ in the figure. Then, note that if we take small intervals of p -values just below and just above $\tilde{p}$ respectively, these households will differ quite considerably in gross and net income levels although they are quite close to each other in terms of type. The jump discontinuity in labour supply and gross income created by the non-convex budget constraint produces large differences in behaviour in households which are in fact quite similar in type.

The effects of variations in the productivity parameter k are more complex, since they depend on whether a change in k at a given level of output leads to an increase in the h -intensity or in the m -intensity of child care production. In the former case, this leads to a decrease in l and y . Thus, in figure 6.10, indifference curves would correspond to higher values of k , and therefore higher utility, the further to the *left* they are. This is the case in which a ranking on gross income actually is negatively associated with a ranking on utility levels. On the other hand, if increasing k reduces the h -intensity of child care production, l and $\hat{y}$ would increase with k and the association between gross income and utility would be positive. Again, however, because of the non-convexity in the tax system, households close together in utility can be widely apart in terms of gross and net income.

In reality, both productivity in household production and the price of market child care will vary across households, and clearly the nature of their covariation will determine the relationships among household utility possibilities, female labour supply and household utility possibilities. For example, if, for whatever reason, there is a strong positive correlation between the price of bought-in child care and household productivity, at any given wage rate more productive wives are more likely to work at home, and this will tend to make the correlation between household income and utility possibilities negative. Since we know virtually nothing empirically about these relationships, the best we can do until we have such information is to analyse alternative logical possibilities.

6.4 Income tax systems in practice

In this section we apply the foregoing discussion to an evaluation of the income tax systems of Australia, Germany, the UK and the US. We focus on the tax treatment of couples with dependent children, where at least one partner is in work and earning an income that places the family outside the wider welfare system. Our aim is to identify the effective marginal and average tax rates faced by 'in-work' families in each country, and to assess whether the structure of rates in each case could be considered to be

consistent with the objective of achieving a more equal distribution of welfare, at a minimum cost to efficiency.

We examine first the formal tax system, as defined earlier, and then extend the analysis to include child-related benefit payments and tax credits. We base the calculations for each country as follows:

- for **Australia**, the personal income tax (including the low income tax offset), the Medicare levy, family tax benefit part A and family tax benefit part B;
- for **Germany**, the income tax when couples file jointly, child benefit (*Kindergeld*) or tax allowance (*Kinderfreibetrag*), depending on which provides the greater advantage, and the unification (*Solidaritätszuschlag*) and church taxes;
- for the **UK**, the income tax, universal child benefit, child tax credit and working tax credit;
- for the **US**, the federal income tax when couples file jointly and the earned income tax credit.

To identify as clearly as possible the *structure* of effective rates under these systems, we find it useful to compare marginal and average tax rates faced by the male partner as primary earner and by the female partner as secondary earner, in a household that changes *type* by changing female labour supply. We compare three types:

SE A single-earner household in which the male works full-time in the market and the female works full-time at home.

PT A two-earner household in which the male works full-time in the market and the female divides her time equally between market work and household production, and earns half the market income of the male.

FT A two-earner household in which both partners work full-time in the market and earn the same income.³⁸

We present results for a representative household with two dependent children under 13 (one under 5 in Australia and under 1 in the UK), zero non-labour incomes, and no intra-household gender wage gap. We compute the marginal tax rate (MTR) and average tax rate (ATR) faced by each partner as functions of *primary income*, and then as functions of *household income*, under the tax and family benefit/credit systems that applied in the 2007 or 2007/2008 tax year in each country. The results are presented diagrammatically as profiles of MTRs and ATRs with respect to primary and household income. We also include in part C of the appendix to this chapter a brief discussion of the impact of UK national insurance contributions and the US payroll tax on effective tax rates in those countries.

We first discuss the results for the two countries with individual formal income tax systems, Australia and the UK, beginning with the strictly progressive, piecewise linear system of the UK. We then compare the results with those we obtain for the two countries

³⁸ Typically, of course, a full-time female worker earns less than her husband, but this assumption is a useful simplification that helps to bring out more clearly the central features of the tax systems considered.

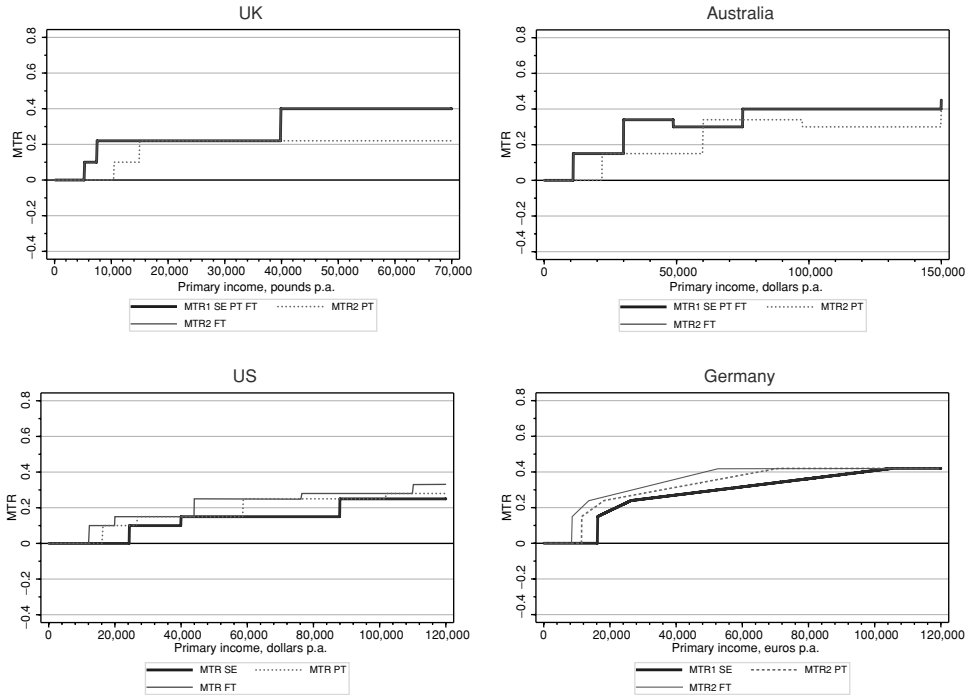


Figure 6.11 Formal income tax system: MTRs by primary income

with joint taxation, Germany and the US. We discuss the piecewise linear system of the US first, followed by the non-linear German system. The details of the formal income tax system of each country that are taken into account in these computations are set out in part A of the appendix to this chapter, and those of the family benefits and tax credits, in part B of the appendix.

6.4.1 Formal income tax systems: MTRs and ATRs by primary income

Figure 6.11 depicts the profiles of MTRs that apply to the incomes of partners, as the income of the primary earner rises, under the formal tax systems of the four countries. Under the independent systems of the UK and Australia, each partner's MTR depends only on own income. The figures for these countries therefore show separate MTR profiles for each partner. The MTR profile of the primary earner of each type is denoted by MTR1 SE, MTR1 PT and MTR1 FT, respectively. Similarly, the MTR profiles of the second earner are denoted by MTR2 SE, MTR2 PT and MTR2 FT. Under the US system of 'full' joint taxation, primary and second earners face the same MTR at any given level of income. We therefore label the profiles MTR SE, MTR PT and MTR FT, respectively. Strictly speaking, only the US federal income tax can be classified as a system of full joint taxation. The German income tax can be viewed as one of 'partial' joint taxation because primary and second earners face slightly different MTRs due to

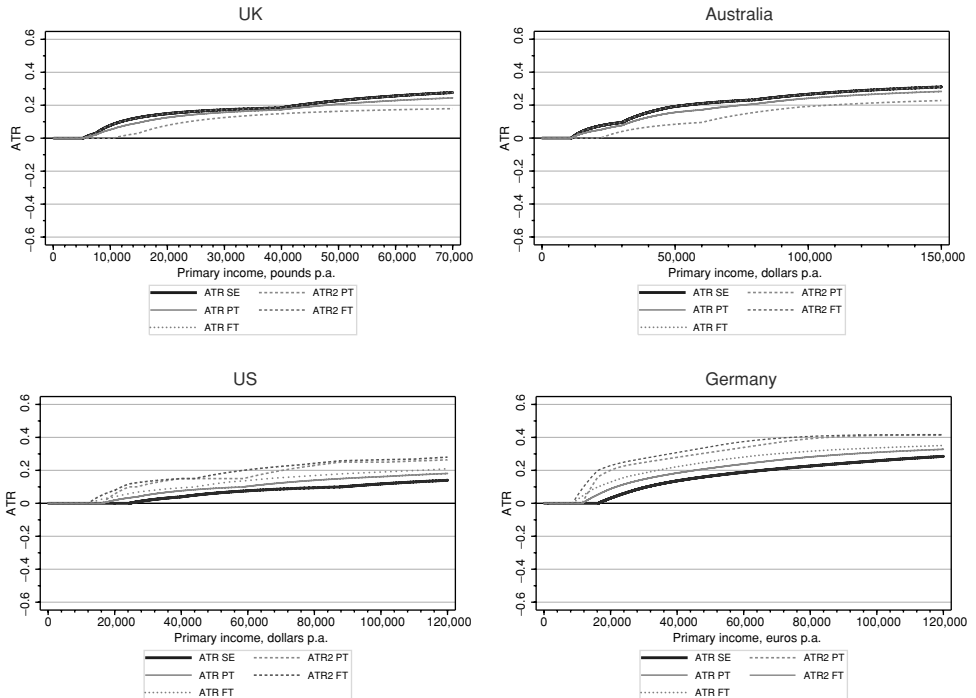


Figure 6.12 Formal income tax system: ATRs by primary income

the allowance of €920 per gainfully employed person for the cost of going out to work. Because the allowance is very small, the MTR profiles of the primary earners in the PT and FT households closely approximate those of the respective second earners, and for simplicity we omit them. Thus the figure for Germany presents only the MTR profile of the primary earner of the SE household, MTR1 SE, and those of the second earners in the PT and FT households, MTR2 PT and MTR2 FT.

Figure 6.12 presents the resulting ATR profiles. The graph for each country plots the ATR profile of each household type, ATR SE, ATR PT and ATR FT. The figures also include the ATR profiles of the second earner in the PT and FT households, ATR2 PT and ATR2 FT. Under the UK and Australian systems of individual taxation, these can be calculated independently of primary earnings. Under the joint tax systems of the US and Germany, the ATR on the second income depends on primary earnings, which we calculate in terms of the additional tax a household must pay due to its second earnings. This implies a model of the household in which only the second earner's labour supply is variable. We discuss this in more detail below.

Since individuals with the same income face the same MTR under independent taxation, the MTR profiles of the primary earners of each household type coincide under the UK and Australian formal income tax systems, as indicated by MTR1 SE PT FT in the figures for these countries. The second earner in the FT household also has the same MTR profile because she is assumed to earn the same income. The second earner in the PT household has a lower income and therefore faces a lower MTR at

any given level of primary income, until her income reaches the lower limit of the next tax band. If the second earner has a more responsive labour supply, her lower tax rate under a progressive individual income tax is consistent with the Ramsey rule, as shown by Boskin and Sheshinski (1983).

While the UK and Australian formal income tax systems are similar, in that they are both based on individual incomes, they differ in an important respect. Only the formal UK income tax system has a strictly progressive MTR scale across all income brackets. The Australian scale is not strictly progressive across all tax brackets due to the low income tax offset (LITO), which raises the rate on incomes from \$30,000 to \$48,750 to 34 cents in the dollar. The rate on income from \$48,750 to \$80,000 then falls to 30 cents in the dollar.³⁹

The MTR profiles shown for partners in the US illustrate the effect of applying a progressive piecewise linear tax schedule to joint incomes. Partners in the two-earner households face higher MTRs at each level of primary income until they reach the top bracket (apart from where they fall within the bands of income across which the same MTR applies), and the higher the income of the second earner, the higher the MTR at any given level of primary income. Thus their MTR profiles are to the left of that of the primary earner in the SE household. Under the non-linear German income tax system MTRs rise continuously within the second and third bands rather than in steps, as in the US piecewise linear system. As a result, the MTR profiles of partners in the two-earner households are consistently above that of the SE household across these bands.

The two types of formal income tax systems that apply in the four countries, progressive individual and progressive joint, give rise to very different ATR profiles, as shown in figure 6.12. Under the individual-based systems of the UK and Australia, partners with the same income pay the same tax and therefore have the same ATR. Thus, the ATR profiles of the SE and FT households, ATR SE and ATR FT, and of the second earner in the FT household, ATR2 FT, coincide at any given level of primary income. That of the PT household, ATR PT, is lower because the second earner has a lower income and the rate scale is progressive.⁴⁰

The systems of joint filing in the US and Germany yield opposite results. Not only does the ATR profile of the FT household rise above that of the SE household, but the profile of the PT household switches from being below to above that of the SE household. This illustrates the shift in the tax burden from single to two-earner households under a reform that involves switching from an individual to a joint-income-based tax system.

It is important to keep in mind that even under an individual system, the FT household pays twice as much tax as the SE household at any given level of primary income, because the second earner in the former works in taxed market production, whereas in

³⁹ This is due to the withdrawal of the LITO at a rate of 4 cents in the dollar from a lower threshold of \$30,000. The LITO is, in fact, an entirely redundant policy instrument that serves only to reduce the transparency of the true rate scale.

⁴⁰ Overall, the UK system tends to be less progressive, and so the ATR profiles rise less steeply and are more widely spaced vertically. The Australian system is more progressive due primarily to the relatively low rate of 15 cents in the dollar up to \$30,000 (≈£14,000). In the UK system, the rate of 20 pence in the pound begins at the relatively low income of £7,256.

the latter she works in untaxed home production. Under the joint tax systems of the US and Germany, the FT household pays more than twice as much tax as the SE household, at any given level of primary income, because of marginal rate interdependence.

It is of interest to consider the implications of this in terms of the tax treatment of the second earner in the context of a model presented in section 4.2.2, in which only her labour supply is variable. In this kind of model, joint taxation can be accurately characterised as a system under which ‘the first dollar earned by the wife is taxed at the same rate as the last dollar earned by the husband’. Her incremental tax burden or effective ATR profile can then be calculated as the increment in the total amount of tax the household pays resulting from her going out to work, taking account of whatever marginal rate interdependence (in the sense defined earlier in this chapter) that exists, expressed as a percentage of her earnings.⁴¹ The reason for doing this is that it is this effective ATR on the wife’s earnings that will determine, among other factors, whether she will in fact go out to work.

The profiles ATR2 PT and ATR2 FT in the figures for the US and Germany represent the results of this calculation. They are well above the ATR profile of the SE household. In this sense, the US federal income tax and the German income tax systems can be seen to discriminate against the second earner, and in turn the two-earner household, with the degree of discrimination increasing as the incomes of the two partners converge, at any given level of primary income.

The degree of discrimination is greater under the German system, as indicated by the wider gaps between the ATR profiles of the three household types, because the schedule is more progressive than that of the US income tax. This illustrates an important trade-off that arises under joint taxation. A country that chooses a joint tax system imposes upon itself a positive relationship between discrimination against the second earner – and therefore against the two-earner family – and overall progressivity. The higher the degree of progressivity, the greater the discrimination. Moreover, higher taxes on the second earner may, in the short run, appear to provide a source of revenue for financing moves toward reductions in the top rates of income tax, but if the labour supply of the second earner is sufficiently responsive to a higher tax rate, the reform may reduce the tax base and therefore prove to be counter-productive. In Germany in particular this is very likely to be the case.

6.4.2 *Effective MTRs and ATRs by primary income*

When we introduce family benefits, exemptions and tax credits, MTR and ATR profiles for three of the countries, the UK, Australia and the US, change dramatically. However, there is little change for Germany.

⁴¹ Effective tax rates on the second earner calculated in this way can be found in the annual OECD reports on ‘Taxing Wages’. See, for example, OECD (2006), and Jaumotte (2003). There is also an extensive literature that evaluates the US EITC programme in terms of its effective rates on second earners. See, for example, Eissa and Hoynes (2004), (2006), and Eissa and Liebman (1996).

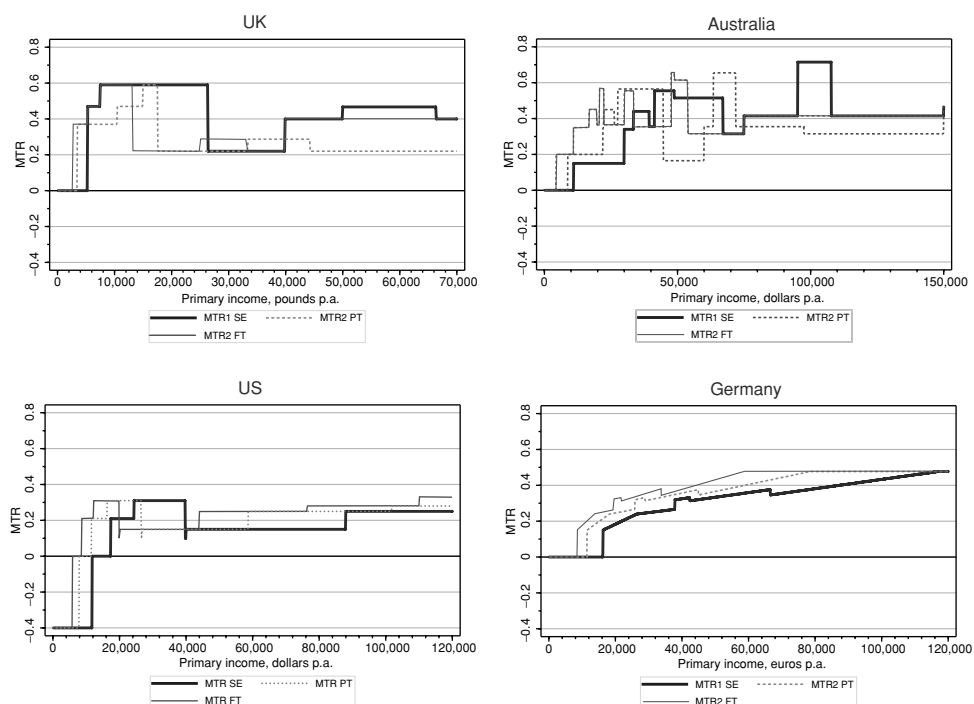


Figure 6.13 Effective MTRs by primary income

Family benefits and tax credits are essentially policy instruments that change the parameters of the effective income tax system, as set out earlier in this chapter. Some are universal and others are not. In subsection 6.3.3 we showed in some detail how replacing a universal benefit by an income-tested benefit can be equivalent to introducing a reform with a less progressive MTR schedule. If the benefit is sufficiently tightly income-tested, the reform can involve switching from a convex to a non-convex piecewise linear MTR schedule. In fact we obtain this result for the UK, Australia and the US over significant bands of income, when we calculate the effective MTRs of their formal income tax schedules in combination with child benefits and tax credits. We also show how in the first two of these countries income testing on household income causes a move away from an individual towards a joint income tax system. Both the UK and Australia now have, in effect, systems of *partial* joint taxation.

Figure 6.13 plots the effective MTR profiles that apply when tax credits and family benefits are combined with the formal income tax of the respective country, for the UK, Australia, the US and Germany. Since, under a system of partial joint taxation, the MTR profiles of primary and second earners are neither independent nor the same, for simplicity we show only the MTR profiles of the primary earner of the SE household, MTR1 SE, and of second earners in the PT and FT households, MTR2 PT and MTR2 FT (as in the graph for Germany in figure 6.11). Since, under the US system, partners

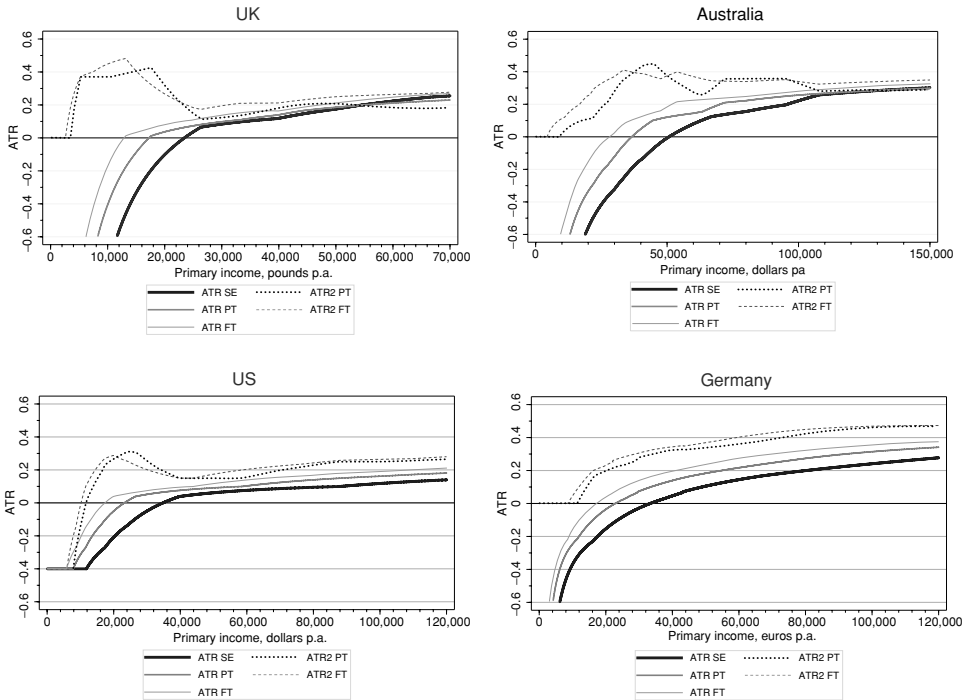


Figure 6.14 Effective ATRs by primary income

face the same MTR at any given level of income in each household type, the profiles are labelled MTR SE, MTR PT and MTR FT, respectively, for the three household types.

Figure 6.14 presents the resulting effective ATR profiles. As in figure 6.12, the graph for each country plots the ATR profile of each household type, ATR SE, ATR PT and ATR FT. The figures also include the ATR profiles of the second earners in the PT and FT households, ATR2 PT and ATR2 FT, calculated again in terms of the additional tax a household must pay due to its second earnings.

The effective MTR profiles in figure 6.13 bear no resemblance to those of the formal tax systems of the UK, Australia and the US depicted in figure 6.11. This is due to the withdrawal of tax credits, benefits and exemptions on the basis of joint income or, as in the case of Australia, on the income of the second earner.

The joint-income-tested child tax credit (CTC) and working tax credit (WTC) in the UK have two effects. First, they replace the progressive MTR schedule with one that has the highest marginal tax rates on low to average incomes – there is a shift towards a strongly inverted U-shaped schedule of MTRs. This is due to the withdrawal of the WTC and child element of CTC, totalling £7,825 for a family with two dependent children, at a rate of 37 pence in the pound from the lower threshold of £5,225. The payment is completely withdrawn at an upper income limit of £26,374. There is a much smaller inverted U-shaped rise in the MTR scale at a joint income of £50,000, the lower

income limit for the family element of CTC, which is withdrawn at 6.67 pence in the pound above this limit.

Second, as noted previously, because the credits are withdrawn on family income, they result in a shift toward joint taxation. The MTR profiles of partners of the PT and FT household shift to the left of that of the SE household, consistent with joint taxation. With the lower limit for the joint income test set at £5,225, second earners effectively start paying tax on their personal allowance under the income tax. The second earner in the FT household starts paying tax on the allowance when her income is equal to half the allowance, and in PT, when her income reaches two-thirds of the allowance. Thus the tax credits deny the second earner the full zero-rated threshold under the income tax. This is also a characteristic feature of a joint tax system. In contrast to the income targeted tax credits, the universal child benefit (CB) has no effect on MTRs.

When family tax benefit part A (FTB-A), family tax benefit part B (FTB-B) and the Medicare levy (ML) are added to the Australian personal income tax, effective MTR profiles exhibit initially an inverted U-shape, as in the UK, and there is a significant move towards joint taxation. The main difference between the two systems is that benefit withdrawal rates are lower under the Australian system, and therefore the phase-out range of income is wider. When first introduced in 2000, the 'maximum rate' of FTB-A was phased out at 30 cents in the dollar. Since then the rate has been lowered to 20 cents, and the threshold income at which the rate applies has been raised. These changes were made in response to concerns about high MTRs at low income levels, of the kind that we now see in the UK. The second rise in the MTR schedule is due to the phasing out of the 'base rate' of FTB-A at 30 cents in the dollar. Combining the US earned income tax credit (EITC) programme with the federal income tax gives a result that has strong similarities with the UK system up to the point where the credits are fully phased out, which is around average annual earnings.

Unlike the preceding three countries, Germany has not gone down the path of introducing family benefits highly targeted on joint income, but instead has retained the universal child benefit, which has no effect on MTRs until it is replaced by child tax deduction at the income level at which the latter has a greater tax advantage. At this point there is a small downward shift in the non-linear MTR profile. The small upward shift that appears in the profile before this point is due to phasing in the unification tax at the higher rate of 20 cents above the exemption limit.

Turning to the ATR profiles in figure 6.14, we can see that including family benefits and tax credits that are withdrawn on joint income makes the overall tax system of the respective country more progressive at low income levels, but also leads to a significant shift toward joint taxation. The effective ATR profiles depicted for the UK and Australia now exhibit differences that are characteristic of a system of joint taxation up to the income level at which the benefits and credits are fully withdrawn. In the UK, the higher MTRs on the incomes of partners in the PT and FT households, up to the upper joint income limit for the WTC and child element of the CTC, raise their ATR profiles to well above that of the SE household, ATR SE, in a way that is consistent with joint

taxation.⁴² The large gap between ATR profiles up to this income limit illustrates the way in which the overall tax burden can, in relative terms, be shifted to two-earner households by increasing the progressivity of a rate scale applying to joint income rather than individual incomes. When we consider these reforms in the context of a model in which the primary earner's labour supply is fixed, they can be seen to raise selectively the MTRs on the incomes of the second earner in the PT and FT households, MTR2 PT and MTR2 FT, and, in turn, their ATR profiles, ATR2 PT and ATR2 FT.

The ATR profiles under the Australian system tend to match those of a joint tax system across a wider range of income. For example, the profile of the FT household, ATR FT, is well above that of the SE household, ATR SE, up to an income level of around \$100,000. In contrast, the UK system has the characteristic features of joint taxation up to a combined income of £26,374, which is below average annual earnings. However, if there are further attempts to ease the very high rates that now apply at low income levels in the UK, this can be expected to lead to ATR profiles that more closely match those of the Australian system.

Because the US federal income tax is already one of joint filing, the effect of phasing out the EITC at high rates on low joint incomes is to reinforce joint taxation at more progressive rates at the bottom of the distribution, as indicated by the wider gap between the ATR profiles of the three household types at low income levels. However, because the rate scale of the federal income tax is not very progressive, the ATR profiles of the three household types have much in common with those of the UK and Australia.

The ATR profiles for Germany differ very little from those of its formal income tax because it has not introduced targeted family benefits. We therefore do not see much higher ATRs on the second income at low income levels, as in the other three countries. However, the vertical gaps between the ATR profiles of the three household types are greater than in the US system further along the income distribution because, as already noted, the German income tax is more strongly progressive than the US federal income tax.

6.4.3 *Effective MTRs and ATRs by household income*

Under a system of full joint taxation, MTR and ATR profiles with respect to household income will be identical for the three household types. To illustrate further the extent to which the UK and Australia have shifted towards systems of joint taxation, we compare effective MTRs and ATRs by household income. The results are presented in figures 6.15 and 6.16.

Figure 6.15 confirms that the UK system has shifted towards one of joint taxation, but only up to a combined income of £26,374. Thereafter, as figure 6.16 shows, there is a significant vertical gap between the ATR profiles of the household types, indicating that it has some way to go before it reaches a full joint tax system. Australia is much

⁴² Note that while the universal CB does not alter MTRs, it widens the gap between the ATR profiles of the three household types because all three receive the same universal benefit but do not contribute equally to financing it.

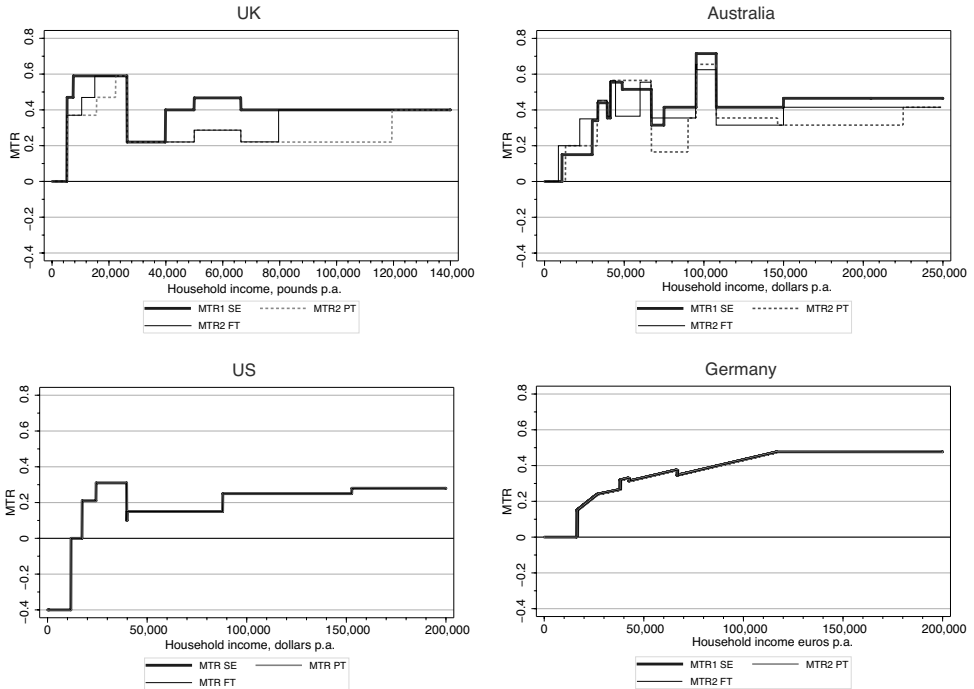


Figure 6.15 Effective MTRs by household income

closer, with ATR profiles for the three household types that are almost identical up to around \$100,000, which is almost twice average annual earnings. Thereafter there is a slightly larger gap.

Australia is the only country of all four that has selective taxation. This is due to the withdrawal of FTB-B on the income of the second earner only. It raises her MTR by an additional 20 cents in the dollar across an income band of \$4,381 to \$17,356. The effect of this is to shift the lower segment of her MTR profile to the left of that of SE, the single earner household, as indicated. This contributes significantly towards achieving a movement towards joint taxation below average earnings.

The country that most closely approximates full joint taxation is Germany. When we omit the allowance of €920 for work-related expenses, the MTR and ATR profiles of the three household types become identical, in a way that is consistent with full joint taxation. Thus, the German and the US income tax systems impose almost identical or identical tax burdens on families with the same joint income, irrespective of whether that income is earned by one or both parents working full-time in the market, and therefore irrespective of large wage differentials between families. Since a household's utility possibilities are basically determined by the wage rates of its members, this implies a significant degree of vertical inequity.

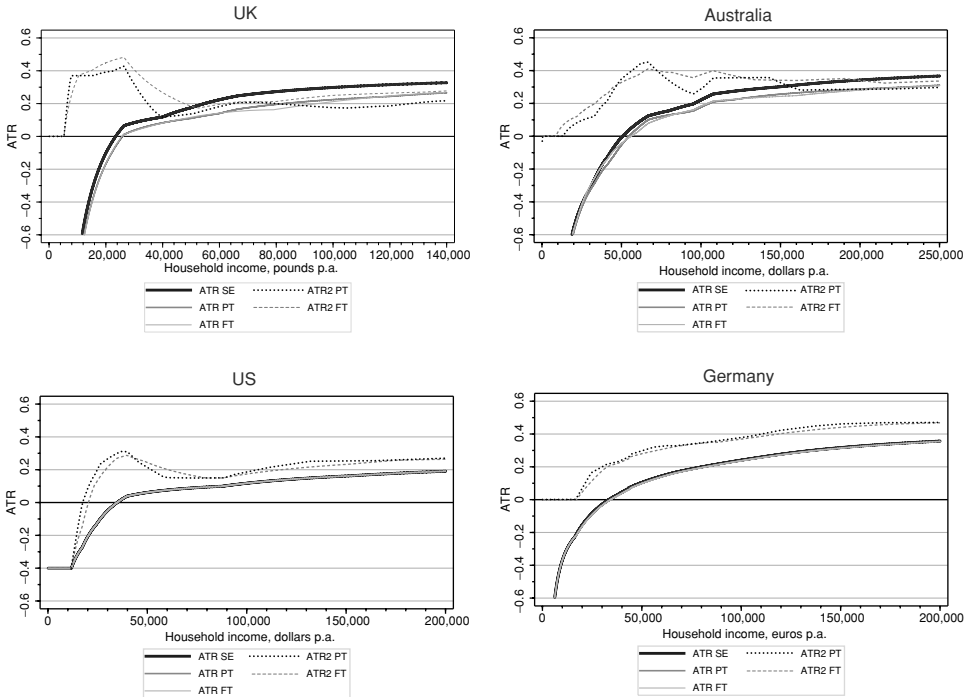


Figure 6.16 Effective ATRs by household income

6.4.4 *Effective ATRs on second earner's income*

We now compute ATRs on the incomes of second earners at selected levels of primary income, for each of the four countries, on the assumption that the male labour supply, and therefore primary income, is fixed at these levels. The results are presented in figure 6.17.

For the UK, the ATR profiles of the second earner are computed for four levels of primary income, £15,000, £20,000, £25,000 and £30,000. We can see immediately that the second earner faces very high ATRs in households with low primary incomes, and that her ATR can fall dramatically at higher levels of primary income. This is due to the withdrawal of the CTC and WTC at 37 pence in the pound on a joint income above £5,220, with both fully withdrawn at £26,374. When the primary earner's income reaches £30,000, the second earner's ATR profile is that of the formal income tax, as depicted in figure 6.11.

The second earner ATR profiles for Australia are computed for primary income levels of \$30,000, \$40,000, \$50,000 and \$60,000. While Australia initially introduced tightly income-tested benefits based on joint income as in the UK, it has since lowered the withdrawal rates and shifted the thresholds, as noted previously. These changes were a response to concerns about incentive effects on female labour supply. However, we can see from the figure that this may be no solution to the problem of labour supply incentives, because the effect is simply to widen the range of second-earner income

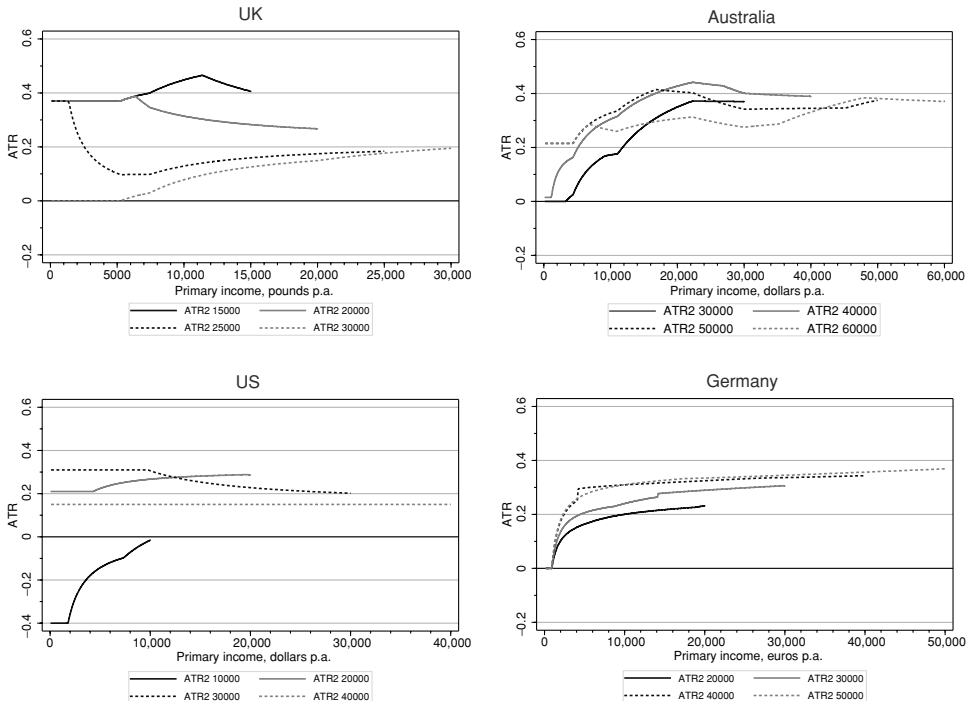


Figure 6.17 ATRs on second income

over which higher MTRs, and therefore higher ATRs, apply. It also has the effect of moving more strongly towards a system of joint filing under a schedule of effective marginal rates that exhibits an inverted U-shaped profile.

The US provides a more extreme example of income-tested credits. The phase-in rate of the EITC begins at 40 cents in the dollar where the MTR of the federal income tax is zero. It is, in effect, a wage subsidy for employers of low-wage labour. At a fixed primary income of \$10,000, the ATR is initially -40 per cent, and starts to rise at the point at which joint income reaches the upper threshold of the phase-in rate of the EITC, \$11,326, after which the ATR rises. In the case of a primary income of \$20,000, the second earner first faces an MTR and ATR of 21 cents in the dollar, the withdrawal rate of the EITC, until she reaches the lower income threshold of the 10 cents rate of the formal income tax. At this point her ATR begins to rise. At a primary income of \$30,000, the second earner initially faces an ATR of 31 cents on her first dollar because her income, when added to primary income, falls within the phase-out range of the EITC and also within the 10 per cent rate range of the formal income tax. Her rate then begins to drop at an income of \$9,783, because the EITC is fully withdrawn on a joint income of \$39,783. At \$40,000 of primary income, the second earner faces the 15 cents in the dollar rate that applies to the second taxable income band of the federal income tax for couples filing jointly. The US EITC can be seen essentially as a poverty

alleviation approach to income redistribution, with minimal support for the ‘middle’ and relatively low tax rates at the ‘top’.

Finally, for Germany we compute the second earner’s ATR profile for primary incomes of €20,000, €30,000, €40,000 and €50,000. From the results we can see that a formal joint income tax systematically taxes the second earner at higher rates at any given level of primary income, but nevertheless applies a progressive MTR scale to that income. The result is that lower-income second earners married to lower-income primary earners typically face lower ATR profiles.

These results illustrate the high ATRs that relatively low-income earners can face due to income testing of benefits and credits on the basis of combined income, in the UK, Australia and the US. In Germany, second earners face high effective tax rates above a minimal threshold across the entire distribution of primary income. Policies of these kinds, by selectively taxing two-earner households, and therefore, effectively, the second earner at higher rates, increase the net-of-tax gender wage gap. This may have consequences not only for female labour supply, but also for the intra-household shares of income and consumption, if these are a function of outside opportunities.⁴³

The changes also have implications for the across-household distribution of income. If the ‘new’ systems involve joint taxation under a rate schedule that exhibits an inverted U-shaped profile, the second earner’s ATR will reach high rates at low income levels, and then begin to decline. In this case, when these reforms are combined with reductions in the top tax rates of the formal income tax, which has been the case in the UK, Australia and the US, they not only increase the net-of-tax gender wage gap, they also have the effect of shifting a greater share of the tax burden to relatively low and average wage workers. This direction of reform has been pursued in labour markets with rising average wages but increasing earnings inequality, and is therefore difficult to rationalise in terms of conventional equity and efficiency criteria.

6.5 Conclusions

We began by setting out a classification of possible formal tax systems, according to the structure of marginal tax rates and the treatment of individual incomes. Although this leads in principle to a large number of logically possible systems, in practice almost all are actually piecewise linear and convex, and the main issue is whether joint or individual incomes should be the tax base, and, if the latter, whether the tax schedules applied to the two incomes should be the same (individual taxation) or different (selective taxation). Also important, however, is the degree of marginal rate progressivity, and the fact that effective tax systems often introduce significant non-convexities and elements of joint taxation into ostensibly individual tax systems.

⁴³ For further discussion see Apps (2006).

There are three main differences between joint and individual taxation. The first is that under the former, marginal tax rates on the incomes of two earners in the household are equalised; under the latter they reflect individual income differences. This is more important the greater the variation in marginal tax rates with income and the larger the differences in individual incomes. In general, women as second earners face higher tax disincentives to work under a joint than under an individual system, other things being equal. If we believe that the intra-household income distribution depends on outside opportunities, for example as indicated by net-of-tax wage rates,⁴⁴ then this suggests that women are worse off and men better off under a joint taxation system. If we accept the stylised fact that women's compensated labour supply elasticities are higher than men's, then this suggests that a move to individual taxation, given progressive tax rates, or to selective taxation, would reduce deadweight efficiency losses by reducing marginal tax rates on women's incomes.

The second main difference is that joint taxation has marginal rate interdependence, and individual taxation has marginal rate independence. Under joint taxation, given that the man would in any case work full-time, the calculation of the increment in household net income resulting from the woman taking a job must rationally take into account the effect of this on the man's tax rate and the resulting increase in his tax liability. This will be larger the narrower the tax brackets over which marginal tax rates vary, the more progressive the marginal rate structure, and the larger the wife's income. Expressing the increment in the household's total tax burden as a percentage of the wife's income shows that, in typical joint taxation systems, this is much larger than would be suggested by taking the ratio of her formal tax liability alone to her income. Thus, these interdependence effects are strong in joint tax systems while they are absent in systems based on individual income, adding further to the disincentive effects on female labour supply the former create.

Third, a joint tax system implies higher relative tax burdens on two-earner than on single-earner households, for a given aggregate tax revenue, than does a system based on individual incomes. This follows directly from basing tax liability on total household market income. The usual concept of equity in income distribution argues for relating tax liability to the utility possibilities or real standard of living of the household. To the extent that these are not strongly positively correlated with household income, and there are several reasons for arguing that this is the case, then individual or selective taxation is also likely to be superior on grounds of equity of the across-household welfare distribution.

It has long been recognised that, whatever may be the separation between the formal institutional channels through which income taxation and other forms of income redistribution are carried out, it is important to combine at least the most important of them to calculate the effective structure of income taxation. Our work on this shows that at least for three of the countries considered – Australia, the UK and the US – we

⁴⁴ As is suggested by the models reviewed in chapter 3.

observe the highest marginal tax rates at low to average income levels. Furthermore, in the first two of these countries, the withdrawal of benefits on the basis of joint incomes moves ostensibly individual tax systems much closer to a joint tax system. In Australia, there are elements of selective taxation, but with women facing the higher marginal tax rates. Driving these tendencies appears to be the desire to reduce tax rates at the top while providing incentives for low-skilled individuals to accept low-wage jobs rather than remaining unemployed and living on state transfers, and at the same time to keep budget deficits low. The lack of transparency of the results, in terms of the consequences for the equity and efficiency of the resulting effective tax systems, should probably be regarded as intentional.

Appendix: Part A

UK income tax – individual based

Under the UK income tax the schedule of rates and bands applies to individual incomes. The 2007–8 financial year schedule specifies three marginal rates: 0.10, 0.22 and 0.40, applying to taxable income bands of £1–£2,230, £2,231–£34,600 and £34,600+, respectively. There is also a personal allowance of £5,225 for each individual under 65, which is deducted from total income to give taxable income. In effect, there is a zero-rated income band from £1 to £5,225, with the three rates applying to bands of income of £5,226 to £7,255, £7,256 to £39,825, and over £39,825, respectively.

Australian personal income tax – individual based

The formal schedule of the Australian personal income tax is specified as a progressive schedule, beginning with a \$6,000 zero-rated threshold followed by a rate of 15 cents in the dollar on incomes up to \$30,000, 30 cents in the dollar on incomes from \$30,001 to \$75,000, 40 cents in the dollar on incomes from \$75,001 to \$150,000, and 45 cents in the dollar on incomes above \$150,000. However, there is a low income tax offset (LITO) of \$750 for individuals with incomes below \$11,000. The offset is withdrawn at a rate of 4 cents in the dollar on incomes above \$30,000. The LITO is an entirely redundant policy instrument that serves only to reduce the transparency of the true tax rate schedule, which is as follows:

Taxable income bands	MTR (+LITO)
\$0–\$11,000	0.00
\$11,001–\$30,000	0.15
\$30,001–\$48,750	0.34
\$48,751–\$75,000	0.30
\$75,001–\$150,000	0.40
\$150,000+	0.45

US federal income tax – joint filing

The US federal income tax is a joint tax system of the income-splitting form. Couples have the option⁴⁵ of computing their income tax liability on the basis of splitting their combined taxable income equally. Thus i 's taxable income, y_i , $i = m, f$, is $y_i = (y_m + y_f)/2$ and the household's tax burden, T , is computed by applying the following piecewise linear income tax schedule:

Taxable income bands	MTR
\$0–\$15,650	0.10
\$15,651–\$63,700	0.15
\$63,701–\$128,500	0.25
\$128,501–\$195,850	0.28
\$195,851–\$349,700	0.33
\$349,700+	0.35

For couples filing jointly there is a \$10,000 standard deduction, a \$3,200 personal exemption for every taxpayer and a dependency exemption of \$3,200 for each dependent child that can be subtracted from income to obtain taxable income to which the rates apply. For a two-parent family with two dependent children, the total amount allowed as a deduction is therefore \$22,800. The tax schedule on income has, in effect, a zero-rated threshold of this amount, with the rates in the table applying to bands that are in addition to this amount.

German income tax – joint filing

The formal German income tax is also a system of joint taxation of the income-splitting form. Couples can compute their income tax liability on the basis of splitting their combined income equally. The schedule of marginal rates has non-linear and piecewise linear segments. The tax burden on partner i 's income, T_i , is calculated by applying the following tax functions to taxable income bands as follows:

Taxable income bands	Tax function
€0–€7,664	$T_i = 0$
€7,665–€12,739	$T_i = (883.7x + 1,500)x$
€12,740–€52,151	$T_i = (228.74z + 2,397)z + 989$
€52,151+	$T_i = 0.42y_i - 7,914$

⁴⁵ Couples can also file separately. However, there is no tax advantage in doing so because the income bands for separate filing are exactly half those of joint filing. It is of interest to note that the rate schedule for singles is not the same as that for 'married filing separately'. The schedule for singles is far more generous above an income of \$32,550, and the difference between the two schedules gives some indication of the degree of tax discrimination against the two-earner couple.

where $x = (y_i - 7664)/10,000$, and $z = (y_i - 12,739)/10,000$. The quadratic functions in the second and third bands give MTRs that increase continuously and linearly with taxable income, and are obtained by differentiating the function with respect to income. There is also a work-related allowance, which we treat here as part of the formal income tax. Taxable income can be reduced by €920 per gainfully employed person, as an allowance for the cost of going out to work.

Appendix: Part B

The following is an outline of the child benefits and tax credits we include for each country, for the representative two-child family.

UK: universal child benefit (CB), child tax credit (CTC) and working tax credit (WTC)

The 2007–8 CB for a family with two dependent children is £30.20 per week. The CTC comprises a family element and a child element. The 2007–8 family element is £545, and is paid to any family with at least one dependent child under 16. A family with a child under 1 receives twice the amount. The child element is £1,845, and is paid in respect of each dependent child. Thus the total amount payable to a two-parent family with two children, one under 1 year, is £4,780. The WTC is based on earned family income and hours. It consists of a basic element of £1,730, a second adult element of £1,700, and an element payable for working 30 hours per week or more (jointly for couples) of £705. The total amount for a qualifying family is £4,135. Families with a joint income below £5,225 are eligible for full CTC and WTC. Above this lower income limit, WTC plus the child element of CTC is reduced by 37p for every £1 of family income. The lower income limit for the family element of CTC is £50,000, and is reduced by one pound for every 15 pounds (or 6.67p per £1).

Australia: Medicare levy (ML), family tax benefit part A (FTB-A) and family tax benefit part B (FTB-B)

The ML is calculated at 1.5 per cent of taxable income. There are exemption categories or reductions based on family income. For a family with two dependent children, the lower family income limit for exemption/reduction is \$33,841. A rate of 10 cents in the dollar applies above this limit until the exemption is fully withdrawn.⁴⁶

Family tax benefit part A (FTB-A) provides a cash transfer of \$4,460.30 per child under 13 years. There are two components: a ‘maximum rate’ and a ‘base rate’. The maximum rate refers to the total amount less the base rate of \$1,890.70. For a child under 13, the maximum rate is withdrawn on joint family income above \$41,318 at a

⁴⁶ There is also a surcharge for individuals and families on higher incomes who do not have private patient hospital cover, calculated at an additional 1 per cent of taxable income.

rate of 20 cents in the dollar. The base rate is withdrawn at a rate of 30 cents in the dollar on joint income above \$95,192. Family tax benefit part B (FTB-B) is a cash transfer of \$3,584.30 for a child under 5 and is withdrawn on second earner's income above \$4,380.00. The benefit is fully withdrawn at \$22,301.50.

US: earned income tax credit (EITC)

The US EITC is a refundable earnings tax credit. The 2007 maximum credit for a two-parent family with two dependent children is \$4,716. The credit is phased in at a rate of 40 cents in the dollar up to a threshold of \$11,790. The credit is then withdrawn at a rate of 21 cents in the dollar when income exceeds \$17,326, and is completely phased out at \$39,783. In 2007 there was also a child tax credit (CTC) for each qualifying child under the age of 17 years. This credit has not been taken into account in the construction of the tax profiles.

Germany: child benefit (CB) (Kindergeld) or child tax allowance (CTA) (Kinderfreibetrag), unification tax (UT) and church tax (CT)

The 2007 CB for each dependent child up to age 25, for families with up to three children, is €1,848.00. The family can receive either CB or, alternatively, a CTA, which provides a deduction from taxable income of €5,808.00 for each child. The tax office determines which is the greater and makes a corresponding adjustment to the tax liability at the end of the tax year. The unification tax is levied at 5.5 per cent on the income tax liability, when this liability exceeds an exemption limit of €1,944 for couples. Once the income tax liability exceeds this limit, the surcharge is phased in at a higher rate of 20 cents of the difference between the income tax liability and the exemption limit until it equals 5.5 per cent of the total liability. The church tax is 8.0 per cent and is computed in the same way. It is possible to opt out of paying the church tax, but most households do not do so.

Appendix: Part C

Social security and compulsory pension schemes

The countries we have discussed have additional social security taxes or, alternatively, compulsory contributions to pension schemes that are effectively taxes. These can add significantly to the effective MTRs and ATRs we have computed for the overall income tax systems. Social security taxes differ from purely redistributive income taxation in that they are usually, at least partly, contributory, and for this reason we view them as outside the scope of the tax problems we consider in this book. Nevertheless, it is worth briefly noting that contribution rates can be highly regressive, as in the UK and US, and can therefore have the effect of raising significantly the MTRs that apply at relatively low to average earnings.

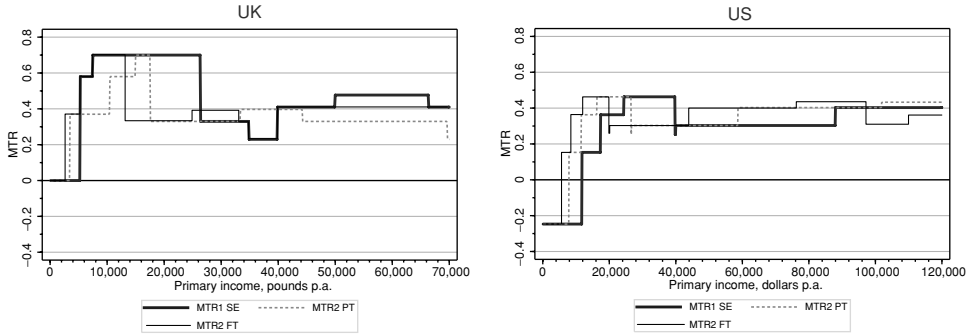


Figure 6.18 MTRs with social security taxes, by primary income

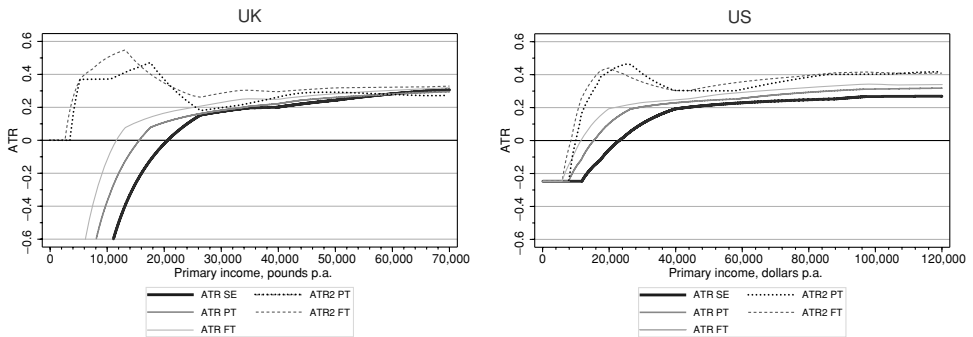


Figure 6.19 ATRs with social security taxes, by primary income

UK employees pay national insurance (NI) contributions of 11 per cent on earnings from £5,225 to £34,840, and only 1 per cent on earnings over £34,840.⁴⁷ The 2007 US employee rate of payroll tax for old age, survivors, and disability insurance was 6.2 per cent of earned income up to a threshold of \$97,500, with an additional employee rate of 1.46 per cent on all earned income.⁴⁸

Figure 6.18 graphs the new MTR profiles that apply in the UK and US with respect to primary income, when the employee components of NI contributions in the UK and of payroll tax in the US are included in the calculation of tax rates for these countries. Figure 6.19 shows the impact on ATR profiles.

In both countries the effect is to accentuate the inverted U-shaped profile of MTRs. As figure 6.18 indicates, relatively low income earners in the UK can face an effective MTR of 70 per cent, the sum of the 59 pence in the pound depicted in figure 6.13

⁴⁷ Employers also make NI contributions at the rate of 12.4 per cent on all earnings of their employees.

⁴⁸ Employers pay the same rates on the earnings of their employees. Thus the combined employee–employer rate is 12.4 per cent of earned income up to the threshold of \$97,500, and an additional rate of 2.92 per cent on earned income.

for the UK, and the additional NI contribution of 11 pence in the pound. Similarly, in the US, the employee rate of payroll tax raises effective MTR for relatively low and average wage earners from 38 cents in the dollar to around 50 cents in the dollar. Thus the overall systems of both countries become much less progressive.

Australia has a compulsory superannuation contribution rate of 9 per cent of earnings. These contributions represent, in effect, a flat rate tax of 9 per cent on earnings. Unlike most taxes, the revenue is used to finance private sector pensions and saving schemes.

7 Optimal linear and piecewise linear income taxation

7.1 Introduction

The point of departure of the theory of optimal income taxation is the proposition that, ideally, a tax should be levied on an individual's innate productivity endowment, which determines the utility level he can achieve on the labour market. Since this is, however, unobservable, a tax must instead be levied on money income. The underlying model of behaviour, whether in the theory of non-linear taxation first developed by Mirrlees (1971), or in the theory of linear taxation formulated by Sheshinski (1972), is that of a utility-maximizing individual who divides his time optimally between market labour supply and leisure, given his net wage. The gross wage measures his productivity. There is a given distribution of wage rates over the population, and the problem is to maximize some social welfare function defined on individual utilities.

In Mirrlees's non-linear tax analysis, the problem is seen as one in mechanism design. An optimally chosen menu of marginal tax rates and lump sum tax/subsidies is offered, and individuals select from this menu in a way that reveals their productivity type. As well as the government budget constraint, therefore, a key role is played by incentive compatibility or self-selection constraints.

In Sheshinski's linear tax analysis, the attempt to solve the mechanism design problem is abandoned. All individuals are pooled, and the problem is to find the optimal marginal tax rate and lump sum subsidy over the population as a whole, subject only to the government budget constraint. Although more restrictive than Mirrlees's formulation, therefore, and more tightly constrained in terms of the possibilities of redistribution, the linear tax system is more capable of being applied in practice. In each case, the theory provides an analysis of how concerns with the equity and efficiency effects of a tax system interact to determine the parameters of that system, and in particular its marginal rate structure and degree of progressivity.

However, when we come to consider debates on actual tax policy, it becomes clear that a central issue, that of how *couples* should be taxed, cannot be addressed directly by these models, not only for the obvious reason that they are based on the labour supply decisions of a single individual, but also because they incorporate the simple dichotomisation of time into market labour supply and leisure: the direct consumption

of own time. In actual fact, as we have stressed repeatedly in this book, individuals divide their time between market work and household production as well as 'pure leisure'. The formation of a household allows specialisation and exchange between its members, in direct analogy to a small economy. This is a crucial aspect of the discussion of the taxation of couples, because the empirical evidence¹ shows that households with the same wage rates, demographic characteristics and non-wage incomes make widely different choices in their allocation of the female partner's time between market and domestic work. This implies that the value of (untaxable) household production may vary substantially across households, quite possibly inversely with its market income, and the question then arises of how this affects the utility possibilities of the household and the design of a tax structure based on market income. In the earlier public finance literature these issues were recognised as being of central importance, for example by Munnell (1980) and Apps (1981), and discussion of them was taken further by Apps and Jones (1986), Apps and Rees (1988) and Sandmo (1990).

In Boskin and Sheshinski (1983), progress was made in the direction of generalising the optimal tax literature to take account of the two-person nature of households, by specifying a household utility function defined on a single consumption good and two types of leisure. This allows the following intuition to be explored. In an economy consisting of two-person households, it is possible to distinguish costlessly between the male primary earner and the female secondary earner. Taking also the stylised fact that female labour supply elasticities are significantly greater than those of males,² on standard Ramsey principles we would expect that optimally, at any given income, women should face lower tax rates than men.³ Boskin and Sheshinski claimed to have confirmed this intuition. However, their theoretical analysis left the conclusion open, as we show below.

A limitation of the analysis is that household production is ignored, and, as a result, household welfare depends only on market labour income. The point about household production is that it raises a question mark about this relationship.⁴ In constructing a model to analyse this issue further, at its core must be some hypothesis that explains how household production, female labour supply, household utility possibilities and market income are related to each other. Empirically very little is known about this. Here, we view variations in female labour supply at a given wage rate as resulting from varying productivities in household production.⁵ The purpose of this chapter is to show

¹ See, for example, chapter 1.

² Although it should be noted that Heckman (1993) questions whether this stylised fact really is supported in its entirety by a careful analysis and interpretation of the evidence.

³ This intuition is clearly stated by both Rosen (1977) and Munnell (1980).

⁴ There are also, as we have discussed in earlier chapters, other factors determining female labour supply which also raise the issue of the relationship between household utility possibilities and market income, and for which 'household production' could stand as a general representation. For example, what if the cost of non-parental child care is positively correlated with domestic productivity, so that at any given wage rate wives with high domestic productivity are more likely to work at home? We simply need to know more about the determinants of female labour supply heterogeneity.

⁵ As we discussed in chapter 5, at least two other factors could be expected to be important: as just mentioned, the availability, cost and quality of non-parental child care; and the perceived depreciation in market human capital

that this productivity variation has important implications for the horizontal equity of a tax system with money incomes as the tax base, and for the structure of an optimal tax system. In chapter 9 we pursue the same point in the context of tax reform.

The importance of household production for optimal tax analysis was emphasised by Sandmo (1990), who derived results for optimal indirect and direct taxation in a model of single-person households which produce a good that is a perfect substitute for a market good. In particular he shows how the nature of the correlation between productivities in domestic and market work (the gross market wage rate) will affect the marginal tax rate in a linear income tax system. This chapter adopts a somewhat different model, in which domestic goods are exchanged only within the household, but the main departure is to have the household consist of two individuals, so that the issues of the taxation of couples can be directly discussed. We extend Sandmo's results to deal with the issue of horizontal inequity inherent in tax systems based on observed money income in the presence of significant variations in market labour supply of at least one of the household members.

7.2 Taxation and the within-household income distribution

The models of the household surveyed in chapters 3 and 4 raise two closely related issues for the theory of income taxation. First, how should we formulate the model of the household which specifies the relationships among tax rates, labour supplies, income and utilities on which the optimal tax analysis is based? Second, how might taxation be used as an instrument to influence the within-household distribution of utilities? We consider these in turn.

Every optimal tax model begins with a social welfare function (SWF) defined on the utilities of *individuals* in the economy. This reflects the view that only individuals can have utilities, and economic policy is concerned with individual welfare, defined in terms of the achieved values of individual utility functions. However, if individuals form households, the effects on their utilities of policy instruments such as taxes will be mediated through the household decision process, and so the connection between policy instruments and individual welfares must be established by models that specify the results of this process.⁶

We can illustrate by taking the simplest kind of optimal tax problem, that of lump sum income redistribution.⁷ Suppose there are two households, indexed $j = L, H$, each consisting of two individuals, indexed $i = f, m$ with μ_{ij} the exogenously given income of individual i in household j , and $\mu_{iH} > \mu_{iL}$. The utility functions are $u(x_{ij})$, where x is consumption. Note that all utility functions are identical, a standard assumption in

resulting from withdrawal from the labour market. For some approaches to analysis of these, see Apps and Rees (2004) and Attanasio *et al.* (2003).

⁶ Gugl (2005) and Alesina *et al.* (2007) use a bargaining model approach to analyse the effects of taxation on the intra-household resource allocation.

⁷ The discussion generalises easily to the various forms of income taxation considered later in this chapter and the next.

optimal tax theory. The SWF is $S(u(x_{fL}), u(x_{mL}), u(x_{fH}), u(x_{mH}))$, and is assumed to be concave and increasing in its arguments. If the ‘social planner’ can determine each individual’s consumption directly, then she can solve the problem

$$\max_{x_{ij}} S(u(x_{fL}), u(x_{mL}), u(x_{fH}), u(x_{mH})) \quad (7.1)$$

$$\text{s.t.} \quad \sum_{i=f,m} \sum_{j=L,H} x_{ij} \leq \sum_{i=f,m} \sum_{j=L,H} \mu_{ij} \equiv \mu \quad (7.2)$$

implying the first best allocations x_{ij}^* satisfying equality of the marginal social utilities of income $s_{ij}^* \equiv S_{ij}u'(x_{ij}^*)$ across individuals.

More realistically, however, suppose that she cannot control the individuals’ actual consumptions, but only their incomes μ_{ij} . In a world in which households consist only of single individuals there would be no problem with this, but in the economy we are dealing with, we have to specify how individual incomes determine individual consumptions and utilities through the household decision process.

Drawing on the extension of Samuelson’s approach discussed in chapter 3, we can do this by representing the household as choosing its internal distribution of consumption by solving the problem

$$\max_{x_{ij}} H_j(u(x_{fj}), u(x_{mj}), \mu_{fj}, \mu_{mj}) \quad (7.3)$$

$$\text{s.t.} \quad \sum_{i=f,m} x_{ij} \leq \sum_{i=f,m} \mu_{ij} \equiv \mu_j \quad j = L, H \quad (7.4)$$

where H_j is its generalised household welfare function (GHWF). The solution to this problem yields each household’s *sharing rule*, which can be written as $x_{ij}(\mu_{fj}, \mu_{mj})$. The planner must then formulate the second-best problem

$$\max_{\mu_{ij}} S(u(x_{fL}), u(x_{mL}), u(x_{fH}), u(x_{mH})) \quad (7.5)$$

$$\text{s.t.} \quad \sum_{i=f,m} \sum_{j=L,H} x_{ij} \leq \sum_{i=f,m} \sum_{j=L,H} \mu_{ij} \quad (7.6)$$

$$x_{ij} = x_{ij}(\mu_{fj}, \mu_{mj}) \quad i = f, m \quad j = L, H. \quad (7.7)$$

The additional constraints (7.7) reflect the fact that the planner is constrained to act through the household’s allocation process, whatever this may be. Solving this problem by substituting the last four constraints into the corresponding utility functions, the optimal solution $(\hat{x}_{ij}, \hat{\mu}_{ij})$ will be characterised by the first-order conditions

$$\sum_{i=f,m} S_{ij}u'(\hat{x}_{ij}) \frac{\partial x_{ij}}{\partial \mu_{kj}} = \lambda \quad k = f, m \quad j = L, H. \quad (7.8)$$

Note that the partial derivatives of the sharing rules necessarily enter these conditions, though knowing these alone will not be sufficient to solve for the optimal incomes $\hat{\mu}_{ij}$. We also have to know the sharing rules themselves, since these determine the values of the marginal social utilities of income. Note also that this places formidable additional informational requirements on the implementation of optimal tax theory. As

well as having to know individual utility functions (or equivalently demand functions), incomes and the social welfare function, we also have to know the household sharing rules.

Is this problem really a second-best problem, or can the planner achieve the first-best? The planner can achieve the first-best if and only if the equations

$$x_{ij}(\mu_{fj}^*, \mu_{mj}^*) - x_{ij}^* = 0 \quad i = f, m \quad j = L, H \quad (7.9)$$

have a solution vector $[\mu_{ij}^*]$ which is non-negative and satisfies the constraint $\sum_{i=f,m} \sum_{j=L,H} \mu_{ij}^* = \mu$. In that case she can find the optimum directly from (7.9), rather than solving the first-order conditions (7.8). Intuitively, the first-best consumptions x_{ij}^* have to be capable of being generated via the household sharing rules by some choice of feasible lump sum redistribution.

Given the wide variety of possible SWFs and GHWFs this is by no means guaranteed. For example, suppose that the planner is a Rawlsian and so wishes to equalise utilities both within and across households, but each household has a weighted utilitarian GHWF

$$H_j = \alpha_j(\mu_{fj}, \mu_{mj})u(x_{fj}) + [1 - \alpha_j(\mu_{fj}, \mu_{mj})]u(x_{mj}) \quad (7.10)$$

such that $\alpha_j < \frac{1}{2}$ for all feasible μ_{fj}, μ_{mj} . Then every household equilibrium gives $x_{mj} > x_{fj}$ in this model.

Following Apps and Rees (1988), we say that there is *dissonance* when the planner cannot achieve the first best, i.e. when the constraint of having to act via the household's allocation process creates a real second-best, and *non-dissonance* when in fact she can achieve the first-best. In either case conditions (7.8) characterise her optimum.

One very simple but useful case of non-dissonance⁸ arises as follows. Assume that the SWF takes the weakly separable form

$$S = S[S_L(u(x_{fL}), u(x_{mL})), S_H(u(x_{fH}), u(x_{mH}))] \quad (7.11)$$

so that the marginal rate of substitution between utilities of the members of one household is independent of the utility levels of the members of any other household. In addition, the household welfare functions take the form

$$H_j = S_j(u(x_{fj}), u(x_{mj})) \quad j = L, H. \quad (7.12)$$

Thus, not only does the household's preferences over the utility pairs of its members *not* depend on the distribution of (non-wage) income, but also these preferences are identical to those of the planner. Then, given any aggregate household income μ_j , the household will distribute it exactly as the planner would wish, since the planner's and the household's sharing rules must be identical. In that case, the planner can use a two-stage budgeting procedure: first decide on the allocation of aggregate income between households, and then leave each household to allocate this income optimally among its members.

⁸ Discussed in Apps and Rees (1988).

Now recall from chapter 3 that Samuelson's theorem implies that, for a HWF of the form in (7.12), there exists a household utility function defined on aggregate consumption, say $u_j(x_j)$. In other words, we can take the household allocation process as given and determine its outcome *as if* its decisions were taken by a single (fictitious) individual. In this case, therefore, we can write the planner's problem as

$$\max_{x_j} S(u_L(x_L), u_H(x_H)) \quad (7.13)$$

$$\text{s.t.} \quad \sum_{j=L,H} x_j \leq \mu. \quad (7.14)$$

The disadvantage of doing this is that we have completely abstracted from the within-household income distribution and can say nothing about it. The advantage is that, when we in fact are interested only in the *across-household* welfare distribution, we have a simple way of avoiding the complications and suppressing the terms, such as sharing-rule derivatives, that arise when within-household distribution is taken into account. At the very least, this discussion clarifies what we are in effect assuming when we do this.

We now turn to the second question. In fact the preceding discussion has shown how it could be answered. The problem in (7.5) shows how a planner can use lump sum taxes to determine the intra-household welfare distribution, and we have also given necessary and sufficient conditions under which the first-best is available.

However, it is not very realistic to assume that it is feasible for a planner to design lump sum taxes or subsidies specifically for each individual in each type of household. One problem is that of asymmetric information, and this is taken up in chapter 8, which analyses non-linear tax systems. More generally, the parameters of linear and piecewise linear tax systems are not household-specific, and are therefore clumsy instruments with which to try to influence the within-household welfare distribution. Other instruments exist for this purpose, for example:

- *In-kind transfers*, such as food stamps or housing grants, which seek to ensure an allocation of consumption within the household that is fairer or socially more desirable than a cash transfer is expected to be.⁹
- *Targeting cash transfers or tax allowances to specific household members* when this is expected to increase the probability of achieving the aims of the transfer programme. The famous example here is that of the switch in the UK child benefit system from a per child tax allowance on the primary earner's income to a flat rate per child cash payment. It was argued by the proponents of this policy that tax allowances were received by the husband, whereas cash payments would be collected by the wife, and the policy change would make it more likely that the transfer would actually benefit the children. This became a much discussed example in the household economics literature,¹⁰ since of course if anonymity (in the sense defined in chapter 3) holds,

⁹ Attempts to explain why governments use in-kind transfers rather than apparently more efficient cash transfers that are based on single-person household models miss this rather obvious feature of many real-world in-kind transfer programmes.

¹⁰ This policy change was first discussed in the context of household models in Apps and Rees (1988).

the policy would have no effect. The empirical evidence produced by Lundberg *et al.* (1997) therefore, that the policy change led to a small but significant increase in household expenditure on women's and children's clothing, was important both in supporting models of the household in which anonymity does not hold (for example Lundberg and Pollak's (1993) 'separate spheres bargaining' model) and in showing that the policy at least to some extent achieved its aim.

- *The legal and institutional environment*, for example the laws on domestic violence, divorce, property ownership both within marriage and following divorce, and the treatment of spouses for social insurance purposes.

We shall therefore adopt the following approach in the optimal tax analysis of this and the next chapters. We will assume complete non-dissonance, in the special form just defined, throughout this analysis. From the technical point of view, this allows us to use a household utility function to focus on the effects of taxation on the across-household welfare distribution, without having to consider terms in the optimal tax functions relating to the within-household income distribution, which would only clutter them up needlessly. More substantively, it reflects our view that we have too little robust information about within-household income distribution to be able to incorporate parameters reflecting this into a general tax system,¹¹ which is in any case not a good *specific* instrument with which to try to meet these concerns. As it happens, however, we do believe that the types of tax system which best meet the goals of equity and efficiency across households, namely piecewise linear tax systems with marginal rate progressivity and with individual incomes as the tax base, would also have the best general effects on the gender distribution of utilities within the household, essentially because they do not discriminate so heavily against female labour supply as joint taxation systems, and therefore provide better outside options.

7.3 Optimal linear taxation

7.3.1 The Boskin–Sheshinski model

This model, based on the optimal linear income tax analysis of Sheshinski (1972), could be viewed as making the smallest possible extension to the model of the individual worker/consumer just necessary to analyse taxation of two-person households. Its main contribution is to make precise the intuition that selective taxation could be optimal because the elasticity of female labour supply is higher than that of male labour supply.

A household has the utility function $u(y, l_f, l_m)$, where x is a market consumption good, and $l_i \geq 0$, $i = f, m$, is the labour supply of household member¹² i . The

¹¹ See, for example, our discussion in chapter 4 of the empirical work on household sharing rules in connection with the collective model.

¹² Although it could just as well be thought of as referring to a single individual with two sorts of labour supply or leisure.

household faces the budget constraint

$$x = a + \sum_{i=f,m} (1 - t_i) y_i$$

where a is the lump sum transfer in a linear tax system and t_i is the marginal tax rate on i 's gross income $y_i \equiv w_i l_i$, with w_i the exogenously given gross market wage. Thus a household is characterised by a pair of wage rates (w_f, w_m) ; otherwise households are identical. Since this is a linear tax problem, with all households facing the same tax rate, we do not have to assume that a household's wage pair is observable. There is a given population joint density function $f(w_f, w_m)$, everywhere positive on $\Omega = [w_f^0, w_f^1] \otimes [w_m^0, w_m^1] \subset \mathbf{R}_+^2$, which tells us how households are distributed according to the innate productivities in market work of their members, as measured by their market wage rates.

To focus attention on what we regard as the most important aspects of the results, we assume that the household utility function¹³ takes the quasilinear form

$$u = x - u_f(l_f) - u_m(l_m) \quad u'_i > 0 \quad u''_i < 0 \quad i = f, m$$

which, however, we find more convenient to write in terms of gross incomes

$$u = x - \psi_f(y_f) - \psi_m(y_m); \quad \psi_i(y_i) \equiv u_i\left(\frac{y_i}{w_i}\right) \quad i = f, m.$$

Solving the household's utility maximisation problem yields demands $x(a, t_f, t_m)$, $y_i(t_i)$ and the indirect utility function $v(a, t_f, t_m)$ such that

$$\frac{\partial v}{\partial a} = 1; \quad \frac{\partial v}{\partial t_i} = -y_i.$$

Note that

$$y'_i(t_i) = w_i \frac{dl_i}{dt_i}$$

is a compensated derivative, because of the absence of income effects. For the same reason, it is straightforward to confirm that labour supplies and gross incomes are strictly increasing in the wage rate and decreasing in the tax rate. Thus household utility is strictly increasing in household income, which is therefore a good indicator of household welfare. Note that the choice of utility function sets the effects of one partner's wage on the labour supply of the other to zero. This makes it much easier to derive the main insights of the analysis without doing too much injustice to the facts.¹⁴

To find the optimal tax system we introduce the social welfare function $W(\cdot)$, which is strictly increasing, strictly concave and differentiable in the utility of every household,

¹³ Clearly the model can say nothing about the within-household welfare distribution, as discussed in the previous section.

¹⁴ Empirical evidence seems to suggest no significant effects of a wife's wage on her husband's labour supply and only very weak negative effects of a husband's wage on his wife's labour supply.

and the planner's problem is then

$$\max_{a, t_f, t_m} \iint_{\Omega} W[v(a, t_f, t_m)] f(w_f, w_m) dw_f dw_m$$

subject to the tax revenue constraint

$$\iint_{\Omega} [t_f y_f + t_m y_m] f(w_f, w_m) dw_f dw_m - a - G \geq 0$$

where $G \geq 0$ is a per household revenue requirement. The first-order condition with respect to the lump sum a can be written

$$\iint_{\Omega} \frac{W'}{\lambda} f(w_f, w_m) dw_f dw_m = 1$$

where $\lambda > 0$ is the marginal social cost of tax revenue and W'/λ the marginal social utility of income to a household with characteristic (w_f, w_m) . Thus the optimal a equates the average marginal social utility of income to the marginal cost of the lump sum. We denote a household's marginal social utility of income W'/λ by s , and its mean by $\bar{s}$. Thus the condition sets $\bar{s} = 1$. Because of the assumptions on $W(\cdot)$, households with relatively low wage pairs will have values of s above the average, and those with relatively high wage pairs below.

The first-order conditions on the marginal tax rates, using the above condition, can be written as

$$t_i^* = \frac{Cov[s, y_i]}{\bar{y}_i'} \quad i = f, m$$

where

$$Cov[s, y_i] = \iint_{\Omega} \left(\frac{W'}{\lambda} - 1 \right) y_i f(w_f, w_m) dw_f dw_m$$

is the covariance of the marginal social utility of household income and the gross household income of individual i , and

$$\bar{y}_i' = \iint_{\Omega} y_i'(t_i^*) f(w_f, w_m) dw_f dw_m$$

is the average compensated derivative of gross income with respect to the tax rate, and is negative.

Now the argument that $t_f^* < t_m^*$ is based on the empirical evidence suggesting that $|\bar{y}_f'| > |\bar{y}_m'|$, but this clearly considers only part of the optimal tax formula, and is in general neither necessary nor sufficient for the result. In other words, though taxing women at a given rate creates a higher average deadweight loss than taxing men at the same rate, the policymaker's willingness to trade off efficiency for equity might imply

that the tax rate on women could optimally be higher than that on men, if the covariance between the marginal social utility of household income and women's gross income is in absolute value sufficiently higher than that of men, so that the corresponding redistributive effects make that worthwhile.

This indeterminacy is also present in Boskin and Sheshinski's paper, though the greater generality and complexity of their model perhaps makes it less obvious. In order to be able to say something more definite, they take a model based on specific social welfare and household utility functions, a joint wage distribution and 'empirically plausible' parameter values, and solve numerically for the marginal tax rates. The result is that the male marginal tax rate is higher than the female.

One example, however, seems to us to constitute a very inadequate basis for an entire conventional wisdom. It is certainly true that equality of the marginal tax rates appears as a highly special case, requiring equality of the ratios of equity and efficiency terms in each case, and so joint taxation is almost certain to be suboptimal, but the results of this model so far do not make a conclusive case for taxing women at a lower rate than men. The optimal tax analysis suggests a departure from income-splitting, but it does not tell us much about the appropriate direction of this departure. In fact, the analysis is unnecessary to give the result that males and females should be taxed at different rates, since joint taxation amounts to imposing on the optimal tax problem the constraint that the marginal tax rates be equal, and such a constraint cannot increase the value of the objective function at the optimum.

To explore the robustness of the conclusion further, write

$$\text{Cov}[s, y_i] = \rho_i \sigma_i \sigma_s \quad i = f, m$$

with ρ_i the correlation coefficient between s and y_i , σ_i the standard deviation of y_i , and σ_s the standard deviation of s . Then we have

Result 7.1

$$t_f^* < t_m^* \Leftrightarrow \frac{\rho_f \sigma_f}{\rho_m \sigma_m} < \frac{\bar{y}'_f}{\bar{y}'_m}.$$

It seems to be an open question empirically, whether this condition is satisfied. All we can really conclude from Boskin and Sheshinski's example is that the assumed functional forms and parameter values lead to satisfaction of this condition.

However, we can take the discussion further and put the conventional wisdom on a firmer foundation if we assume:

- *assortative matching*: across households, the female wage is a monotonic increasing function of the male wage;
- *diverging incomes*: as the male wage increases, the couple's gross income difference $y_m - y_f$ increases monotonically.

Then we have

Result 7.2 Assortative matching and diverging incomes are sufficient (given $|\bar{y}'_f| > |\bar{y}'_m|$) to ensure $t_f^* < t_m^*$.

Proof: We can write

$$\begin{aligned} Cov[s, y_m] - Cov[s, y_f] &= Cov[s, y_m - y_f] \\ &= \iint_{\Omega} (s - 1)[y_m - y_f] f(w_f, w_m) dw_f dw_m. \end{aligned} \quad (7.15)$$

□

Given assortative matching, we know that s is falling monotonically with w_m , while given diverging incomes we know that $y_m - y_f$ is increasing monotonically with w_m , thus $Cov[s, x_m - x_f] < 0$, and so

$$|Cov[s, y_m]| > |Cov[s, y_f]| \quad (7.16)$$

and the male covariance is higher in absolute value. Hence the male tax rate is both more effective as a redistributive instrument and less costly in terms of deadweight loss, and so it will be optimally higher than the female.¹⁵

An important limitation of the Boskin–Sheshinski model, as our discussion in the introduction suggests, is that it omits household production. Why should this matter? After all, it could be argued, all that is really important are the labour supply (gross income) derivatives and the covariance of gross income with the marginal social utility of household income. Whether substitution at the margin is between market work and leisure, or market work and household production is, on this argument, just a matter of detail that does not really have substantive implications.

What makes this argument untenable is the large variation across households in female labour supply¹⁶ and the implication that market income may well not correctly reflect utility possibilities. In the Boskin–Sheshinski model, the household's utility possibilities necessarily increase with household market income, which is therefore an appropriate welfare measure for purposes of income taxation. A central consequence of taking account of household production, in a way that also explains the empirical evidence on female labour supply, is that household income may be a poor, and possibly negative, indicator of household welfare, which in turn should have important policy implications. In the next section we set up a simple household model incorporating household production, and use it in an extension of the optimal linear taxation model.

7.3.2 Optimal linear taxation and household production

We now consider the analysis of optimal linear taxation based on the household model of section 3.2, which incorporates household production. In that model, we assumed that only f engages in household production, and the output of the domestic good

¹⁵ For further discussion see Apps and Rees (2008). ¹⁶ Extensively discussed in chapter 1.

is simply proportional to her time input, where $k > 0$ is the factor of proportionality. It follows that the implicit price or marginal opportunity cost of the household good is $p = (1 - t_f)w_f/k$, so that the household is better off the higher is k . We showed in chapter 3 that female labour supply could either be increasing or decreasing as a function of k . Where it is increasing, household utility and female wage income will be positively correlated, and so this will strengthen the negative covariance between the marginal social utility of income and female income, thus tending to increase the optimal tax rate t_f^* . Here, we explore more formally the implications of the converse case, in which, other things being equal, female labour supply decreases as k increases.

A household is now characterised by a triple (w_f, w_m, k) , and we let $g(w_f, w_m, k)$ denote the joint density function of these three variates. Denoting the domain of this density function by $\Gamma = [w_f^0, w_f^1] \otimes [w_m^0, w_m^1] \otimes [k_0, k_1] \subset \mathbf{R}_+^3$, we can formulate the social welfare function and government budget constraint for the optimal tax problem respectively as

$$S = \int \int \int_{\Gamma} S[v(a, t_f, t_m)]g(w_f, w_m, k)dw_m dw_f dk \quad (7.17)$$

and

$$\int \int \int_{\Gamma} [t_f y_f + t_m y_m]g(w_f, w_m, k)dw_m dw_f dk - a - G \geq 0 \quad (7.18)$$

where $G \geq 0$ is again exogenous net revenue per household.

Formally, the results look very similar to those derived for the Boskin–Sheshinski model

Result 7.3 The optimal lump sum a satisfies

$$\int \int \int_{\Gamma} \frac{S'}{\lambda} g(w_f, w_m, k)dw_m dw_f dk = 1$$

Result 7.4 The marginal tax rates satisfy

$$t_i^* = \frac{Cov[s, y_i]}{\bar{y}_i'} \quad i = f, m$$

where

$$Cov[s, y_i] = \int \int \int_{\Gamma} \left(\frac{S'}{\lambda} - 1 \right) y_i g(w_f, w_m, k)dw_m dw_f dk \quad i = f, m$$

and $\bar{y}_i'$ is again the average compensated derivative of gross income with respect to the tax rate t_i . However, the inclusion of household production, together with the assumption that female labour supply falls with k , does make an important difference to the interpretation of the results. This can be brought out most clearly if we again make the assumption of *perfect assortative matching*, and write $w_f = \alpha w_m$, $\alpha \in (0, 1)$.

Consider now any given subset of households with the same wage rates $(\alpha \hat{w}_m, \hat{w}_m)$. Across this ‘wage subset’, as k varies from k_0 to k_1 , male income y_m and tax paid $t_m^* y_m$

both remain constant, while female income y_f falls as k increases. Then household income and the amount of tax paid on the female income both fall. On the other hand, the marginal social utility of income also falls, since households with higher k have higher utility. Thus, within each wage subset at a given lump sum and pair of marginal tax rates there is effectively a regressive income redistribution from worse-off to better-off households, essentially because household market income is *positively* associated with the marginal social utility of income for households with a *given wage*. Other things being equal, this implies that the female tax rate will be a less satisfactory instrument of redistribution than the male tax rate. Indeed if, across the entire set of households, this effect were sufficiently strong relative to the effect of an increasing wage, the female tax rate could be negative. The empirical evidence on the high variation in female labour supply after controlling for both male and female wage rates suggests that this effect is too important to be ignored.

To put this a little more formally, on the assumption of perfect assortative matching we can write

$$\begin{aligned} Cov[s, y_i] &= \int_{w_m^0}^{w_m^1} \int_{k_0}^{k_1} (s - 1)y_i g(w_m, k) dk dw_m \\ &= \int_{w_m^0}^{w_m^1} \int_{k_0}^{k_1} s y_i g(w_m, k) dk dw_m - \bar{y}_i. \end{aligned} \quad (7.19)$$

Then the counterpart to condition (7.16) in the Boskin–Sheshinski model is

$$\bar{y}_m - \bar{y}_f > \int_{w_m^0}^{w_m^1} \int_{k_0}^{k_1} s[y_m - y_f] g(w_m, k) dk dw_m. \quad (7.20)$$

Again this is more likely to be satisfied if households with relatively high s -values are associated with smaller differences in gross incomes and conversely. The effect of household production under the assumption that female labour supply falls with k is in fact to guarantee that this is the case for each given wage value $\hat{w}_m$, for as k increases with $\hat{w}_m$ given, s falls and $[y_m - y_f]$ rises.

7.4 Piecewise linear taxation

The first theoretical contribution on this subject was made by Sheshinski (1989), who formulated and solved the problem of the optimal two-bracket piecewise linear tax system for an individual worker/consumer. Unfortunately, he claims to have proved that, under standard assumptions, marginal rate progressivity, here referred to as the convex case, always holds: at the social optimum the tax rate on the higher income bracket must always exceed that on the lower. However, Slemrod *et al.* (1996) show that there is a mistake in Sheshinski's proof. They then carry out simulations which, again on standard assumptions, in all cases produce the converse result – the upper-bracket marginal tax rate is optimally always lower.

This is, however, also somewhat problematic. In general non-parameterised models there is no reason to rule out the convex case, and it may be that the specific functional forms and parameter values chosen by Slemrod *et al.* for their simulations do not in fact capture reality particularly well. They may imply compensated labour supply elasticities for higher income earners that are too high, and an insufficiently inequality-averse social welfare function. One could also point to the fact that in reality virtually all *formal* tax systems are in fact convex – marginal tax rates increase with gross income.¹⁷ In the following analysis we focus on the derivation of the optimal convex piecewise linear tax system in the case of joint taxation. We then discuss the equity and efficiency gains that can arise when we move from this system to a convex piecewise linear tax system based on individual incomes. We provide a formal analysis of the effects of changing the parameters of this kind of tax system, viewed as a problem in tax reform, in chapter 9. There we also show again the relevance of household production, as a proxy for any kind of significant determinant of female labour supply heterogeneity, for tax policy.

7.4.1 Solutions to the household choice problem

We again take the quasilinear household utility functions introduced in section 7.3, but make a slight change in notation. Given the underlying utility function

$$u = x - u_f(l_f) - u_m(l_m) \quad u'_i > 0, u''_i < 0 \quad i = f, m \quad (7.21)$$

we redefine it as:

$$u = x - u_f\left(\frac{y_f}{w_f}\right) - u_m\left(\frac{y_m}{w_m}\right) \quad (7.22)$$

$$\equiv x - \psi^f(y_f, w_f) - \psi^m(y_m, w_m) \quad (7.23)$$

with $\psi_y^i > 0$, $\psi_{yy}^i < 0$, $\psi_{yw}^i < 0$, $i = f, m$. The slope of i 's indifference curve in the (y, x) -space is $\psi_y^i(y_i, w_i) > 0$ and it decreases with w_i , i 's type.

We assume a continuous joint density function $f(w_f, w_m)$, strictly positive for all $(w_f, w_m) \in [w_f^0, w_f^1] \times [w_m^0, w_m^1]$. An important aspect of this leisure–labour supply model is that the wage rate and gross income are perfect indicators of the achieved utility level of the household: the higher are w_i and therefore y_i , other things equal, the higher is the household's utility. We now look at the household's labour supply problem under a convex tax system.

Under a joint tax system the parameters are $(a, t_1, t_2, \hat{y})$, with a the lump sum payment to all households, t_1 and t_2 the marginal tax rates in the first and second brackets respectively, and $\hat{y}$ the level of total household income determining the upper

¹⁷ On the other hand, as we saw in the previous chapter, non-convexities over specific income ranges are created by additional rules concerning the withdrawal of cash benefits and tax subsidies, suggesting that at least over some range of the income distribution policymakers may want non-convexities.

limit of the first bracket. The household faces the budget constraint

$$x \leq a + (1 - t_1)y \quad y \leq \hat{y} \quad (7.24)$$

$$x \leq a + (t_2 - t_1)\hat{y} + (1 - t_2)y \quad y > \hat{y} \quad (7.25)$$

where $y = \sum_{i=f,m} y_i$. We assume the convex case, so that $t_1 < t_2$.

There are three solution possibilities:

(i) Optimal household income $y^* < \hat{y}$. In that case we have

$$\psi_y^i(y_i^*, w_i) = 1 - t_1 \Rightarrow y_i^* = y_i^*(t_1, w_i) \quad i = f, m \quad (7.26)$$

$$x^* = a + (1 - t_1) \sum y_i^*(t_1, w_i) \quad (7.27)$$

$$u = a + (1 - t_1)y^* - \sum \psi^i(y_i^*(t_1, w_i), w_i) \equiv v(a, t_1, w_f, w_m) \quad (7.28)$$

and the derivatives of the indirect utility function are

$$\frac{\partial v}{\partial a} = 1; \quad \frac{\partial v}{\partial t_1} = -y^*. \quad (7.29)$$

(ii) Optimal income $y^* > \hat{y}$. In that case we have

$$\psi_y^i(y_i^*, w_i) = 1 - t_2 \Rightarrow y_i^* = y_i^*(t_2, w_i) \quad i = f, m \quad (7.30)$$

$$x^* = a + (t_2 - t_1)\hat{y} + (1 - t_2) \sum y_i^*(t_2, w_i) \quad (7.31)$$

$$u = a + (t_2 - t_1)\hat{y} + (1 - t_2) \sum y_i^*(t_2, w_i) - \sum \psi^i(y_i^*(t_2, w_i), w_i) \quad (7.32)$$

$$\equiv v(a, t_1, t_2, \hat{y}, w_f, w_m) \quad (7.33)$$

and the derivatives of the indirect utility function are

$$\frac{\partial v}{\partial a} = 1; \quad \frac{\partial v}{\partial t_1} = -\hat{y}; \quad \frac{\partial v}{\partial t_2} = -(y^* - \hat{y}); \quad \frac{\partial v}{\partial \hat{y}} = (t_2 - t_1). \quad (7.34)$$

(iii) Optimal income $y^* = \hat{y}$. In that case we have

$$\psi_y^i(\hat{y}, w) = 1 - t_1 - \rho^* \quad i = f, m \quad (7.35)$$

$$x^* = a + (1 - t_1)\hat{y} \quad (7.36)$$

$$u = a + (1 - t_1)\hat{y} - \sum \psi^i(\hat{y}, w_i) \equiv v(a, t_1, \hat{y}, w_f, w_m) \quad (7.37)$$

and the derivatives of the indirect utility function are

$$\frac{\partial v}{\partial a} = 1; \quad \frac{\partial v}{\partial t_1} = -\hat{y}; \quad \frac{\partial v}{\partial \hat{y}} = (1 - t_1) - \psi_y^i(\hat{y}, w_i) = \rho^*. \quad (7.38)$$

Here, $\rho^* \geq 0$ is a shadow price expressing the fact that the household is implicitly rationed at $\hat{y}$. It would like to earn more income if it could do so at the marginal tax rate $(1 - t_1)$, but increasing income would cause its tax rate to rise to t_2 .

Note that in all three cases, for every household type (w'_f, w'_m) we have $\psi_y^f(y_f^*, w'_f) = \psi_y^m(y_m^*, w'_m)$. This simply says that a necessary condition for the efficient production of

gross income is that marginal utility costs of doing so must be equalised, since the net marginal return to each gross income under joint taxation is always the same. If these were unequal it would always be possible to increase utility while keeping consumption constant by increasing gross income with the lower marginal cost and reducing that with the higher.

To summarise these results: the households can be partitioned into three groups according to their wage type and the total joint income that arises from their optimal choice of labour supply. Thus:

$$\Omega_0 = \{(w_f, w_m); \sum y_i^*(t_1, w_i) \equiv y(t_1, w_f, w_m) < \hat{y}\} \quad (7.39)$$

$$\Omega_1 = \{(w_f, w_m); \sum y_i^*(t_1, w_i) = \hat{y}\} \quad (7.40)$$

$$\Omega_2 = \{(w_f, w_m); \sum y_i^*(t_2, w_i) \equiv y(t_2, w_f, w_m) > \hat{y}\}. \quad (7.41)$$

Given the continuity of $f(w_f, w_m)$, consumers are continuously distributed around the budget constraint, with both maximised utility v and gross incomes y_i^* continuous functions of the w_i . Household utility is a strictly increasing function of each wage rate, while household income is also strictly increasing for all except the interval Ω_1 , where it is constant.¹⁸

7.4.2 The optimal convex joint tax system

The planner chooses the parameters of the tax system to maximise a social welfare function defined as

$$\int \int_{\Omega_0} S[v(a, t_1, w_f, w_m)] f(w_f, w_m) dw_f dw_m \quad (7.42)$$

$$+ \int \int_{\Omega_1} S[v(a, t_1, \hat{y}, w_f, w_m)] f(w_f, w_m) dw_f dw_m \quad (7.43)$$

$$+ \int \int_{\Omega_2} S[v(a, t_1, t_2, \hat{y}, w_f, w_m)] f(w_f, w_m) dw_f dw_m \quad (7.44)$$

where $S(\cdot)$ is a strictly concave and increasing social welfare function. The government budget constraint is

$$\int \int_{\Omega_0} t_1 y(t_1, w_f, w_m) f(w_f, w_m) dw_f dw_m \quad (7.45)$$

$$+ \int \int_{\Omega_1} t_1 \hat{y} f(w_f, w_m) dw_f dw_m \quad (7.46)$$

$$+ \int \int_{\Omega_2} [t_2 y(t_2, w_f, w_m) + (t_1 - t_2) \hat{y}] f(w_f, w_m) dw_f dw_m - a \quad (7.47)$$

$$\geq G \quad (7.48)$$

¹⁸ See the appendix to this chapter for a proof of this.

where $G \geq 0$ is a per capita revenue requirement. From the first-order conditions characterising a maximum of social welfare subject to the government budget constraint¹⁹ we derive the following:

Result 7.5

$$\bar{s} \equiv \int \int_{\Omega_0 \cup \Omega_1 \cup \Omega_2} \frac{S'}{\lambda} f(w_f, w_m) dw_f dw_m = 1 \quad (7.49)$$

where $\bar{s}$ is the average marginal social utility of income over the entire population and λ is the shadow price of tax revenue. Since the same lump sum a is paid to each consumer, this is essentially the same condition as for linear taxation. However, since it implies

$$\int \int_{\Omega_0 \cup \Omega_1} \left(\frac{S'}{\lambda} - 1 \right) f(w_f, w_m) dw_f dw_m = - \int \int_{\Omega_2} \left(\frac{S'}{\lambda} - 1 \right) f(w_f, w_m) dw_f dw_m \quad (7.50)$$

and S'/λ falls with each wage, the left-hand side must be positive and so the value of the integral on the right-hand side negative. That is, the consumers in Ω_2 , the higher tax bracket, on average have marginal social utilities of income below the population average, and the converse is true for consumers in the lower tax bracket. This is of course what we would expect.

The conditions characterising the optimal marginal tax rates yield²⁰

Result 7.6

$$t_1^* = \frac{\int \int_{\Omega_0} \left(\frac{S'}{\lambda} - 1 \right) [y^* - \hat{y}^*] f(w_f, w_m) dw_f dw_m}{\int \int_{\Omega_0} y_{t_1}(t_1^*, w) f(w_f, w_m) dw_f dw_m} \quad (7.51)$$

$$t_2^* = \frac{\int \int_{\Omega_2} \left(\frac{S'}{\lambda} - 1 \right) [y^* - \hat{y}^*] f(w_f, w_m) dw_f dw_m}{\int \int_{\Omega_2} y_{t_2}(t_2^*, w) f(w_f, w_m) dw_f dw_m}. \quad (7.52)$$

The denominator, the average (compensated) derivative of gross household income with respect to the marginal tax rate, which is negative, can be interpreted as the efficiency effect of the tax. The numerator is the equity effect. Since $y^* < \hat{y}^*$ for the subset Ω_0 , while marginal social utilities of income are above the average, the numerator will also be negative. Likewise $y^* > \hat{y}^*$ for the subset Ω_2 , while marginal social utilities will be below average, and so the numerator here is also negative.

Note the strong formal similarity with the results for optimal linear taxation. The welfare gain of piecewise linear over linear taxation arises out of the fact that the marginal tax rates t_1^* , t_2^* reflect more closely the covariation of income with the marginal social utility of income, and the average compensated gross income derivatives, a measure of deadweight loss, within the respective subgroups.

¹⁹ In deriving these conditions, it must be taken into account that the limits of integration are functions of the tax parameters. Because of the continuity of optimal gross income in w_f , w_m , these effects all cancel and the first-order conditions reduce to those shown here.

²⁰ For convenience we change the partial derivative notation, writing y_{t_1} for $\partial \sum y_i / \partial t_1$ and so on.

The wholly new element is the determination of the optimal income threshold at which the tax brackets change, $\hat{y}^*$. The condition for optimal choice²¹ of $\hat{y}$ is:

Result 7.7

$$\begin{aligned} & \int \int_{\Omega_1} \left\{ \frac{S'}{\lambda} v_{\hat{y}} + t_1^* \right\} f(w_f, w_m) dw_f dw_m \\ &= -(t_2^* - t_1^*) \int \int_{\Omega_2} \left(\frac{S'}{\lambda} - 1 \right) f(w_f, w_m) dw_f dw_m. \end{aligned} \quad (7.53)$$

The left-hand side gives the marginal social benefit of a relaxation of the constraint on the consumer types in Ω_1 who are effectively constrained by $\hat{y}$. First, for $w_f, w_m \in \Omega_1$ the marginal utility with respect to a relaxation of the gross income constraint is $v_{\hat{y}} > 0$, as shown earlier. This is weighted by the marginal social utility of income to these household types. Moreover, since they increase their gross income, this increases tax revenue at the rate t_1^* . The right-hand side is positive and gives the marginal social cost of increasing $\hat{y}$. Since $t_2^* > t_1^*$, this reduces the tax burden on the higher income group. This can be thought of as equivalent to giving a lump sum payment to higher rate taxpayers proportionate to the difference in marginal tax rates, and this is weighted by the sum of net marginal social utilities of income to consumers in this group, which is negative, as we just showed.

7.4.2.1 Taxation of individual incomes Just as in the case of optimal linear taxation, we can regard joint taxation as equivalent to imposing on the optimal tax problem the constraint that marginal tax rates be equalised across members of any given household, and since such a constraint cannot increase the maximised value of the objective function, selective taxation is necessarily weakly superior to joint taxation. The general problem then is to find two optimal piecewise linear tax schedules for males and females respectively. Given the similarity of the expressions determining the optimal tax rates in each bracket with those in the optimal linear tax system just pointed out, it is easy to see how Boskin and Sheshinski's proposition would be extended to the argument that the tax rates in the piecewise linear tax system for women should be lower than those for men: higher average compensated labour supply derivatives, and a lower covariance (in absolute value) between the marginal social utility of income and gross income, in both tax brackets, would give this result. The optimal bracket thresholds for each system would be determined along exactly the same lines as in the above analysis. It would also be reasonable to allow the intercept term or lump sum transfer to differ between the male and female tax schedules, since this must also be weakly welfare-improving.

Of course, moving between the two systems would not satisfy the Pareto criterion for a policy change: some households would be better off, some worse off. It is well known that in a joint taxation system households with relatively higher primary incomes

²¹ See chapter 6 for a diagrammatic discussion of this condition.

and lower secondary incomes benefit from income-splitting, and this represents, given revenue neutrality, a shift in tax burdens from such households to households with higher second-earner incomes.²² In terms of the political economy of such changes in tax systems, it seems also to be the case that in some countries, for example Germany, wider political and religious views on the role and status of women and the social importance of maternal v. market child care play an important role in the discussion. It may be more feasible to obtain welfare improvements by changing the parameters of a tax system in which men and women are subject to the same tax system in principle, but where the differences in male and female earnings allow de facto gender-based taxation, than to try to construct parallel tax systems differentiated by gender. Overall, therefore, it seems to us that the most fruitful context in which to carry out a formal analysis of these questions is that of tax reform, and so we take up this topic again in chapter 9.

Appendix

Here we show that optimal household gross incomes and therefore utilities are strictly increasing functions of the wage rates, so that higher wage households will have lower marginal social utilities of income. We can solve the household's optimisation problem in two steps:

$$1. \quad \min_{y_i} \Psi = \sum_{i=f,m} \psi^i(y_i, w_i) \quad \text{s.t.} \quad \sum_{i=f,m} y_i \leq y \quad (7.54)$$

for any gross household income y . This yields the value function

$$\Psi(y, w_f, w_m) = \sum_{i=f,m} \psi^i(y_i^*, w_i)$$

and straightforward comparative statics shows that

$$2. \quad \Psi_y, \Psi_{yy} > 0, \Psi_{yw_i} < 0 \quad (7.55)$$

$$\max_{x,y} x - \Psi(y, w_f, w_m) \quad (7.56)$$

$$\text{s.t. } x \leq a + (1 - t_1)y \quad y \leq \hat{y} \quad (7.57)$$

$$x \leq a + (t_2 - t_1)\hat{y} + (1 - t_2)y \quad y > \hat{y}. \quad (7.58)$$

We have the three solution possibilities as shown earlier, and standard comparative statics on these shows that optimal gross income y^* is strictly increasing in the wage rates in those cases where $y^* \neq \hat{y}$.

²² See, for example, the analysis of empirical tax systems in chapter 6.

8 Optimal non-linear taxation

8.1 Introduction

The first part of this chapter gives a self-contained exposition of the theory of non-linear taxation, as a prelude to the extension of the theory to the case of two-person households. Our aim is to present as simply and clearly as possible the main ideas and results of the theory in a reasonably rigorous way. In doing this, we set up an analytical framework based on specific utility and social welfare functions. We provide references to the literature for those who would like to see a more general treatment.

8.2 The two-type case

In a seminal paper, the importance of which extends well beyond public economics,¹ James Mirrlees analysed the implications for optimal taxation of an informational constraint on the planner's ability to use lump sum taxation to redistribute income from high- to low-productivity individuals. One reason often given for the impossibility of lump sum taxation relates to the difficulty of finding a tax base which is truly non-distortionary. Taxing wage income distorts labour supply decisions; taxing wealth distorts saving and consumption decisions; taxing goods distorts the pattern of expenditures.² On the other hand, if we really know or can observe a consumer's innate productivity type,³ there is nothing to rule out the simple instruction: pay the tax collector T_i^* units of income. We do not need to relate the tax to anything except the productivity type of the individual. But this, according to Mirrlees, is the problem. If the planner cannot observe the productivity type of the individual, then lump sum taxation of this kind becomes infeasible. The tax has to be applied to something observable, most likely wage income, and so the trade-off between equity and efficiency comes

¹ Mirrlees (1971) formulated and solved a model of the adverse selection problem, which has since become of central importance in many areas of economics, for example contract theory, the theory of monopoly regulation, and the theory of non-linear pricing.

² In old English houses one often observes bricked-up windows, a consequence of a window tax.

³ Note the word *innate*. A problem here would be that productivity is the result of the decision to invest in one's human capital, so that these decisions could also be distorted by taxes on productivity type. This raises a question about Mirrlees's approach that we will discuss later in this chapter.

into play. We then have to find the second-best optimal tax policy. We now put this discussion more formally.

Define $y_i = w_i l_i$ as type i 's *gross wage income*, and, since $l_i = y_i / w_i$, we can, as in chapter 7 earlier, write the utility function as

$$u_i = u(x_i) - v\left(\frac{y_i}{w_i}\right) \equiv u(x_i) - \psi_i(y_i) \quad i = 1, 2 \quad (8.1)$$

with

$$\psi'_i(y_i) = \frac{v'\left(\frac{y_i}{w_i}\right)}{w_i} > 0 \quad (8.2)$$

$$\psi''_i(y_i) = \frac{v''\left(\frac{y_i}{w_i}\right)}{w_i^2} > 0. \quad (8.3)$$

In the absence of taxation, the individual's choice problem becomes:

$$\max_{x_i, y_i} u_i \quad \text{s.t. } x_i \leq y_i \quad (8.4)$$

with the first-order condition

$$\left[\frac{dx_i}{dy_i} \right]_{u_i^*} = \frac{\psi'_i(y_i^*)}{u'(x_i^*)} = \frac{v'(l_i^*)}{w_i u'(x_i^*)} = 1. \quad (8.5)$$

Figure 8.1 illustrates this. Both consumer types have the same budget constraint in the (y_i, x_i) -space, with a slope of 1, but they now have indifference curves which differ in a way that is very important for the subsequent analysis. Consider how the marginal rate of substitution *at a given point* (y_i^0, x_i^0) varies with w_i

$$\frac{\partial}{\partial w_i} \left[\frac{\psi'_i(y_i^0)}{u'(x_i^0)} \right] = \frac{\partial}{\partial w_i} \left[\frac{v'\left(\frac{y_i^0}{w_i}\right)}{w_i u'(x_i^0)} \right] \quad (8.6)$$

$$= \frac{\left(-\frac{y_i^0}{w_i^2}\right) w_i u' v'' - v' u'}{(w_i u')^2} < 0. \quad (8.7)$$

So at any given point in the (y_i, x_i) -plane, the indifference curve for the higher-productivity type, type 2, is flatter than that for the lower type, type 1. This is the *single-crossing condition*, and plays an important role in the subsequent analysis. Note that it implies that in figure 8.1, the tangency point for the type 2 consumer must lie on the common budget line to the right of that for type 1, implying that she has both higher gross income and consumption.

Note also that the utility level achieved by consumer 2 at the point (y_1^*, x_1^*) is higher than that of consumer 1. Thus, at this point

$$u_1^* = u(x_1^*) - v\left(\frac{y_1^*}{w_1}\right) \quad (8.8)$$

$$u_2 = u(x_1^*) - v\left(\frac{y_1^*}{w_2}\right) \quad (8.9)$$

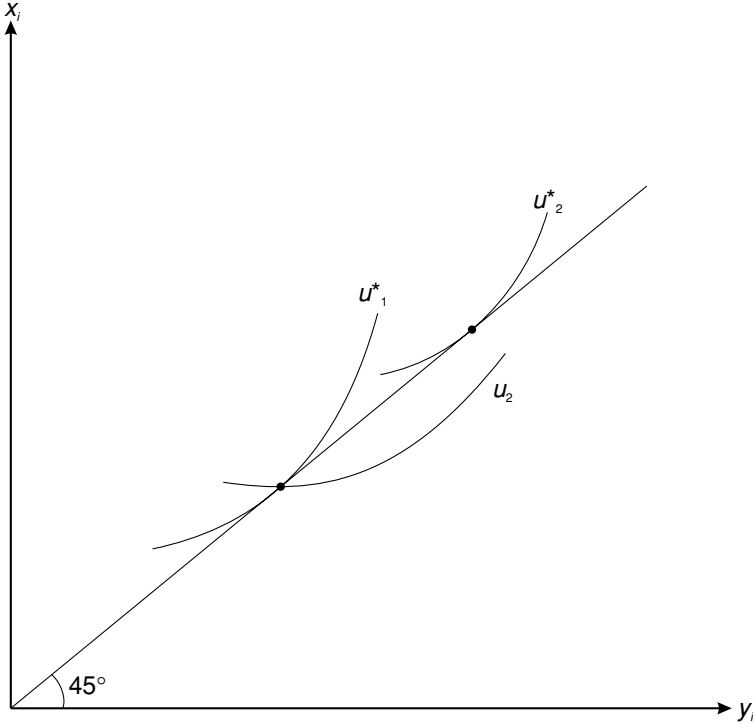


Figure 8.1 Single crossing

and so $w_2 > w_1 \Rightarrow u_2 > u_1^*$. From the figure it is then clear that $u_2^* > u_2 > u_1^*$. Thus, high productivity consumers are better off at the market equilibrium.

The intuition for the difference in slopes of the indifference curves is worth spelling out. Thus consider point (y_1^*, x_1^*) in figure 8.1. Suppose each type is at this point and is required to earn a slightly larger gross income $y_1^* + dy$, and we ask, how much extra consumption dx_i do we have to give each of them to leave them just as well off? Since $w_2 > w_1$, type 2 has to make a smaller increase in labour supply to generate the extra gross income, so suffers a lower loss of utility, and so needs a smaller increase in consumption to compensate her for this, than type 1.

We now derive briefly the optimal lump sum taxes for this model, because this will be useful in interpreting the second-best results. The consumer's budget constraint is now $x_i = y_i - T_i$, implying $T_i = y_i - x_i$, and so the utilitarian planner's problem is

$$\max_{x_i, y_i} W = \sum \phi_i [u(x_i) - \psi_i(y_i)] \quad (8.10)$$

subject to the government budget constraint

$$\sum \phi_i (y_i - x_i) \geq G \quad (8.11)$$

where $G \geq 0$ is the per capita revenue requirement. If $G = 0$, we call this the ‘pure redistribution’ case.

The first-order conditions are

$$\phi_i[u'(x_i^*) - \lambda^*] = 0 \quad i = 1, 2 \quad (8.12)$$

$$\phi_i[-\psi'_i(y_i^*) + \lambda^*] = 0 \quad i = 1, 2 \quad (8.13)$$

$$\sum \phi_i(y_i^* - x_i^*) - G = 0. \quad (8.14)$$

We note that non-satiation implies that the government budget constraint must bind. We can then quickly establish in this context the main results:

$$u'(x_1^*) = \lambda^* = u'(x_2^*) \Leftrightarrow x_1^* = x_2^* \quad (8.15)$$

$$\left[\frac{dx_i}{dy_i} \right]_{u_i^*} = \frac{\psi'_i(y_i^*)}{u'(x_i^*)} = 1 \quad i = 1, 2 \quad (8.16)$$

$$y_1^* < y_2^*. \quad (8.17)$$

Figure 8.2 illustrates the solution for the special case of pure redistribution and where $\phi_1 = \phi_2 = \frac{1}{2}$.

A key feature of the solution to the lump sum tax problem is that at the planner’s optimum, high-productivity types are worse off than low-productivity types. It is essential, therefore, that the planner can observe an individual’s type. Suppose that this is not the case. A consumer knows her own type, but this is her private information, unavailable to the planner. This information asymmetry then creates an adverse selection problem. All individuals would claim to be low-productivity types, and if the planner took this at face value and applied the tax accordingly, the budget constraint would be violated.

The solution to this adverse selection problem is (now) well known. We introduce a *self-selection* or *incentive compatibility (IC) constraint*,⁴ which requires that the equilibrium allocation be such that the high-productivity type has no incentive to lie. Thus we formulate the planner’s problem as:

$$\max_{x_i, y_i} W = \sum \phi_i [u(x_i) - \psi_i(y_i)] \quad (8.18)$$

$$\text{s.t.} \quad \sum \phi_i (y_i - x_i) - G \geq 0 \quad (8.19)$$

$$u(x_2) - \psi_2(y_2) \geq u(x_1) - \psi_2(y_1). \quad (8.20)$$

The IC constraint (8.20) ensures that a type 2 consumer does not gain from choosing the type 1 allocation rather than the allocation designed for her.⁵

⁴ In accordance with the Revelation Principle, which says that by inducing consumers to reveal their type, the planner achieves an allocation that is at least as good as could be achieved by any other tax design.

⁵ It can be shown that the low productivity consumer would never prefer the high productivity type’s allocation in equilibrium, due to the single crossing condition, and so we only need to impose one incentive compatibility constraint.

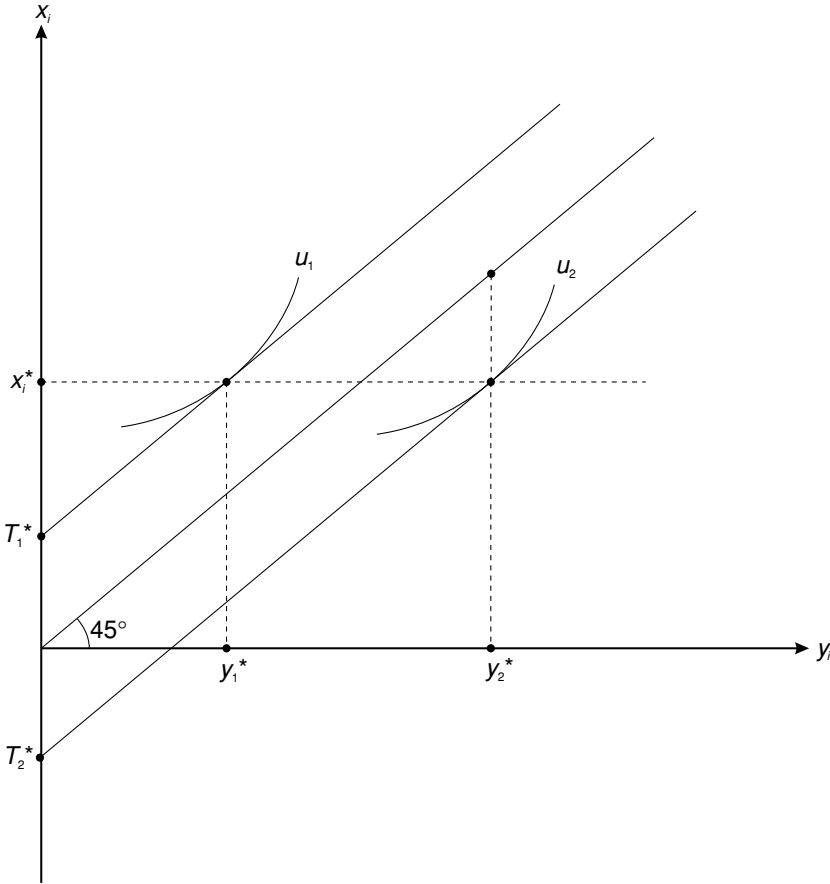


Figure 8.2 Lump sum redistribution

The first-order conditions (assuming an interior solution)⁶ are

$$\phi_1[u'(\hat{x}_1) - \hat{\lambda}] - \hat{\mu}u'(\hat{x}_1) = 0 \quad (8.21)$$

$$-\phi_1[\psi'_1(\hat{y}_1) - \hat{\lambda}] + \hat{\mu}\psi'_2(\hat{y}_1) = 0 \quad (8.22)$$

$$\phi_2[u'(\hat{x}_2) - \hat{\lambda}] + \hat{\mu}u'(\hat{x}_2) = 0 \quad (8.23)$$

$$-\phi_2[\psi'_2(\hat{y}_2) - \hat{\lambda}] - \hat{\mu}\psi'_2(\hat{y}_2) = 0 \quad (8.24)$$

together with the two constraints satisfied as equalities (if $\hat{\mu} = 0$ we must still have an adverse selection problem, so $\hat{\mu} > 0$. In that case non-satiation implies $\hat{\lambda} > 0$).

The main results of the model are:

⁶ As we show below, this is not an innocuous assumption, and in fact there is considerable interest in relaxing it.

- (i) ‘No distortion at the top’: The last two conditions imply

$$\left[\frac{dx_2}{dy_2} \right]_{\hat{u}_2} = \frac{\psi'_2(\hat{y}_2)}{u'(\hat{x}_2)} = 1 \quad (8.25)$$

This is precisely the condition on type 2’s allocation that results from the lump sum tax problem, hence the name of this result. It caused quite a stir, because it implies that the marginal rate of tax on the gross income of the high-productivity type is zero.⁷ This does not of course mean that she pays no tax. As we show below, she pays a lump sum tax $\hat{T}_2 = \hat{y}_2 - \hat{x}_2 > 0$. But it does conflict with conventional notions of the progressivity of the tax system, since, as discussed in chapter 6, virtually all formal tax schedules in practice have marginal tax rates increasing with taxable income, and so in interpreting this model it will be important to understand why this result holds.

- (ii) ‘Distortion at the bottom’: We can express conditions (8.21) and (8.22) in the following way. Define

$$\delta \equiv \psi'_1(\hat{y}_1) - \psi'_2(\hat{y}_1). \quad (8.26)$$

We can use (8.2) to show that $\delta > 0$. Thus we have

$$\delta \equiv \frac{v'(\frac{\hat{y}_1}{w_1})}{w_1} - \frac{v'(\frac{\hat{y}_1}{w_2})}{w_2}. \quad (8.27)$$

Then since $w_1 < w_2$, $\hat{y}_1/w_1 > \hat{y}_1/w_2$, and the convexity of $v(\cdot)$ implies $v'(\frac{\hat{y}_1}{w_1}) > v'(\frac{\hat{y}_1}{w_2})$, while dividing these by w_1 and w_2 respectively strengthens the inequality.

Next, we can write (8.21) as

$$(\phi_1 - \hat{\mu})\psi'_1(\hat{y}_1) = \phi_1\hat{\lambda} - \hat{\mu}\delta \quad (8.28)$$

and (8.22) as

$$(\phi_1 - \hat{\mu})u'(\hat{x}_1) = \phi_1\hat{\lambda}. \quad (8.29)$$

Dividing (8.29) into (8.28) gives

$$\left[\frac{dx_1}{dy_1} \right]_{\hat{u}_1} = \frac{\psi'_1(\hat{y}_1)}{u'(\hat{x}_1)} = 1 - \frac{\hat{\mu}\delta}{\phi_1\hat{\lambda}} < 1. \quad (8.30)$$

Given the strict convexity of the indifference curves, this implies that consumption and gross income (therefore labour supply) of the low-productivity type are reduced relative to the levels that correspond to the first-best condition. Thus we have $\hat{y}_1 < y_1^*$, $\hat{x}_1 < x_1^*$. Figure 8.3 illustrates this solution.

We again assume pure redistribution and equal proportions of types. High-productivity consumers are indifferent between their allocation $(\hat{y}_2, \hat{x}_2)$ and that of the low-productivity consumers $(\hat{y}_1, \hat{x}_1)$, while, because of the single-crossing

⁷ Mirrlees’s paper appeared at a time when the top marginal tax rate in the UK was about 90 per cent.

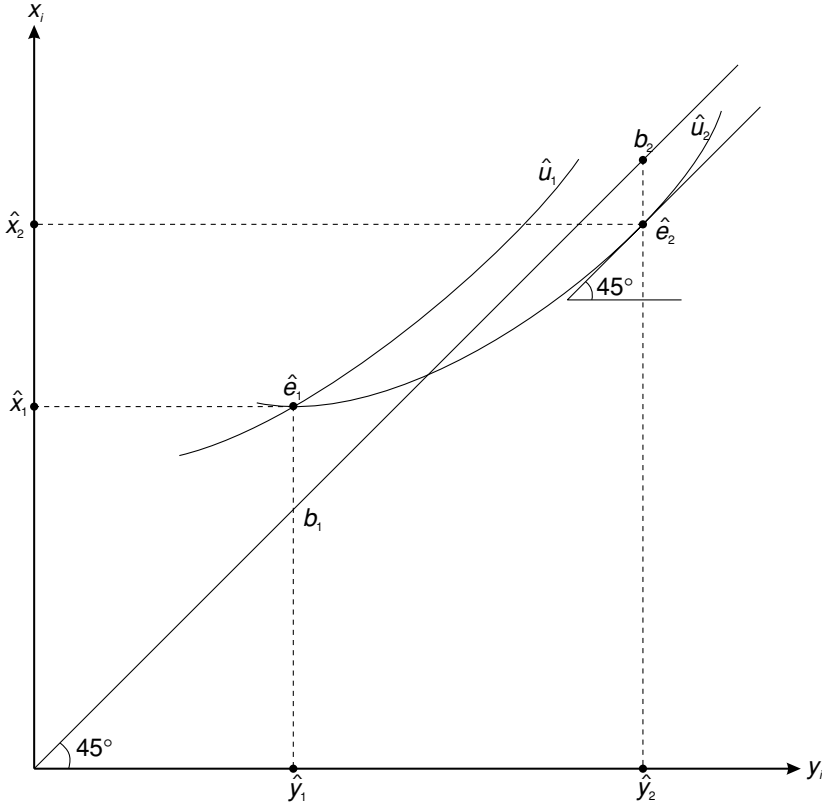


Figure 8.3 Optimal non-linear taxation

condition, low-productivity consumers strictly prefer their allocation.⁸ High types consume less than their gross income by the amount shown by $\hat{e}_2 b_2$; low types consume more than their gross incomes by the corresponding amount $\hat{e}_1 b_1$. It is important to note that at the allocation $(\hat{y}_1, \hat{x}_1)$ the slope of the low type's indifference curve is less than 1, as we have just proved.

In figure 8.4, we show how this second-best optimal allocation can be implemented by a tax system. As already suggested, the high-productivity consumers each pay a lump sum tax $\hat{T}_2 = \hat{y}_2 - \hat{x}_2$, which gives them the budget constraint $x_2 = y_2 - \hat{T}_2$ and 'guides' them to their allocation $(\hat{y}_2, \hat{x}_2)$ as shown in the figure. To induce low-productivity consumers to choose the second-best optimal allocation $(\hat{y}_1, \hat{x}_1)$, they have to be offered the budget constraint $x_1 = \hat{a} + (1 - \hat{t})y_1$, as shown in the figure. In that case they choose $(\hat{y}_1, \hat{x}_1)$ and satisfy the condition

$$\left[\frac{dx_1}{dy_1} \right]_{\hat{a}_1} = \frac{\psi'_1(\hat{y}_1)}{u'(\hat{x}_1)} = 1 - \hat{t}. \quad (8.31)$$

⁸ This is why we did not need an incentive compatibility constraint for this type.

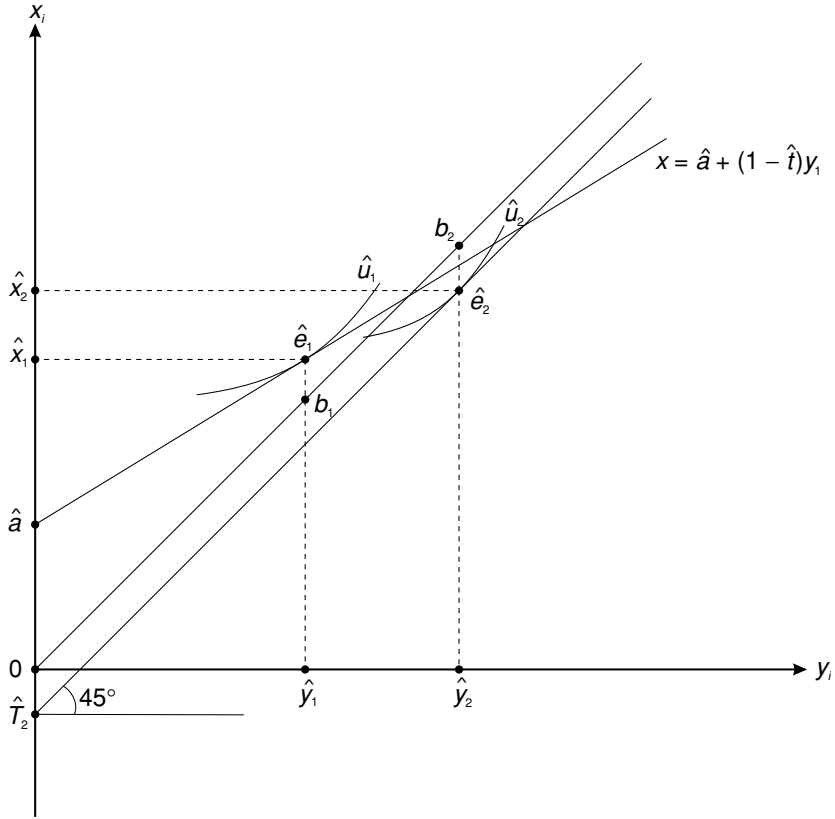


Figure 8.4 Implementation by taxes

We see then from (8.30) that the optimal marginal tax rate is

$$\hat{t} = \frac{\hat{\mu} \delta}{\phi_1 \hat{\lambda}}. \quad (8.32)$$

The tax must be chosen to create just the right amount of distortion in the choices of the low-productivity consumers. In order to ensure that they have the right amount of consumption they receive the lump sum payment

$$\hat{a} = (\hat{x}_1 - \hat{y}_1) + \hat{t} \hat{y}_1 \quad (8.33)$$

which more than repays them their tax bill $\hat{t} \hat{y}_1$. The positive marginal tax rate is necessary to induce the right choice of gross income (labour supply) even though they end up getting a net transfer, $(\hat{x}_1 - \hat{y}_1)$. Note also that when this tax function is offered, it must be specified to apply only to $y \leq \hat{y}_1$, since without such a quantity limitation the high-productivity type would choose it.

Given the importance of the result that the low-productivity type's equilibrium is distorted in this way, we should try to gain an intuitive understanding of why this

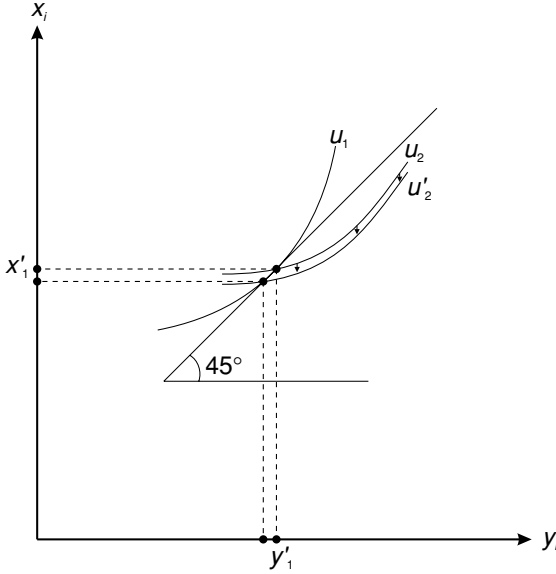


Figure 8.5 Distortion of the lower type is always optimal

is so. In fact it is based on a classic second-best argument. It is possible to achieve incentive compatibility while *not* distorting the type 1 equilibrium, as figure 8.5 shows. Here, we pass the type 2's indifference curve through a point at which the slope of the type 1's indifference curve is 1. The allocations (y'_1, x'_1) , (y'_2, x'_2) are feasible and incentive-compatible, but not optimal.

To see this, note first that at this allocation the utility level of the high-productivity consumers must be higher than that of the low-productivity consumers. That is

$$u(x'_2) - \psi_2(y'_2) \geq u(x'_1) - \psi_2(y'_1) > u(x'_1) - \psi_1(y'_1). \quad (8.34)$$

To show the last inequality, note that since $w_2 > w_1$ and $v' > 0$,

$$\psi_2(y'_1) = v \left(\frac{y'_1}{w_2} \right) < v \left(\frac{y'_1}{w_1} \right) = \psi_1(y'_1) \quad (8.35)$$

We know, therefore, from the analysis of the lump sum taxation case, that the planner would like to redistribute utility from the high- to the low-productivity consumers.

Next, referring to figure 8.5, we see that by moving a very small distance leftwards along the type 1 indifference curve, making very small changes $dx_1, dy_1 < 0$ in the type 1 allocation, such that

$$dx_1 = \frac{\psi'(y'_1)}{u'(x'_1)} dy_1 = dy_1 \quad (8.36)$$

we leave unchanged the first component of the planner's objective function

$$W = \phi_1[u(x_1) - \psi_1(y_1)] + \phi_2[u(x_2) - \psi_2(y_2)] \quad (8.37)$$

since

$$\phi_1[u'(x'_1)dx_1 - \psi'_1(y'_1)dy_1] = 0. \quad (8.38)$$

We also leave unchanged the first component of the government budget constraint

$$\phi_1[y_1 - x_1] + \phi_2[y_2 - x_2] = G. \quad (8.39)$$

However, doing this relaxes the incentive compatibility constraint, allowing us to shift downwards the indifference curve on which we place the type 2 consumers, and allowing an increase in tax revenue from them, since we can increase y_2 , or decrease x_2 , or both. With this net increase in tax revenue we can increase the consumption and/or reduce the gross income (labour supply) of the type 1 consumers, thus raising their utility levels. It follows that we can always increase the extent of redistribution from type 2 to type 1 consumers by making such a change. Thus the allocations (y'_1, x'_1) , (y'_2, x'_2) cannot be optimal. Because of the distortion, in the form of a limitation on the ability to redistribute, created by the IC constraint, it is second-best optimal to create a distortion in the allocation to the low-productivity consumers.

The second-best optimum can then be regarded as being found by trading off these distortions. Refer again to the figure. Suppose we continue to move leftwards along the indifference curve u_1 , pushing down the u_2 -indifference curve as we do so. As the slope of u_1 falls, so the reductions in y_1 become increasingly larger than the reductions in x_1 , implying increasing net losses in tax revenue from type 1 consumers. These eat into the tax revenue gains from type 2 consumers and imply a diminishing increment of tax revenue available for increasing redistribution towards type 1 consumers. Ultimately a point will be reached at which the marginal gain in redistribution is zero.

The figure is not well suited to show the determination of the final equilibrium (in the course of redistribution type 1 consumers will be placed on higher indifference curves than u_1 in the figure), but the expression for the marginal tax rate in (8.32) gives a nice summary of the determinants of the final position. First, it shows that the trade-off between distortions depends on the relative proportions of the two types. The larger is ϕ_1 , the greater the weight on the loss of net revenue from type 1, and the smaller is that on the gain in revenue from type 2, and so the smaller will be the overall final distortion, as expressed by the value of the tax rate.

Next, note that $\hat{\mu}$ is the shadow price attached to the IC constraint, and so it gives the value to the planner of a small loosening of this constraint. The greater this is, the greater will be the distortion. $\hat{\lambda}$, on the other hand, is the shadow price of the government budget constraint. When the distortion on type 1 consumers is increased, tax revenue, as we just saw, is reduced, and so the higher the marginal cost of tax revenue the smaller will be the final distortion.

To explain the role of δ , note that by dividing it by $u'(\hat{x}_1)$ we obtain

$$\frac{\psi'_1(\hat{y}_1)}{u'(\hat{x}_1)} - \frac{\psi'_2(\hat{y}_1)}{u'(\hat{x}_1)} \quad (8.40)$$

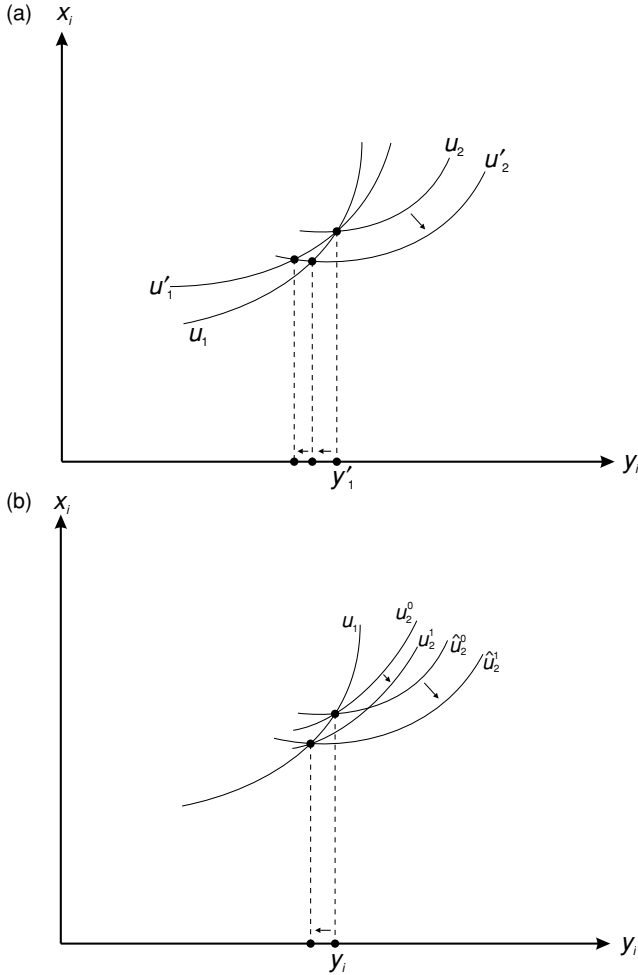


Figure 8.6 Relative slopes and the degree of distortion

which is the difference in slopes of the type 1 and 2 indifference curves at the point $(\hat{y}_1, \hat{x}_1)$. The larger this difference, the larger the distortion. The intuition for this can be seen in figure 8.6. In (a) we see that for a given downward shift in the type 2 indifference curve, the steeper the slope of the type 1 curve the smaller the loss of tax revenue, i.e. the less costly the redistribution. In (b) we see that for a given move along the type 1 indifference curve, the flatter the type 2 curve, the greater the reduction in x_2 , i.e. the greater the increase in taxation, that can be made. Thus, overall, the greater the difference in these two slopes, the more it pays to distort the type 1 equilibrium.

Now look again at condition (8.30), and notice a possible mathematical difficulty. The left-hand side of this condition must always be positive, but, as we make ϕ_1 arbitrarily small, what is there to rule out the possibility that the right-hand side becomes negative?

In that case the condition cannot hold. This purely mathematical observation has an interesting economic implication. It may in fact be optimal, if ϕ_1 is sufficiently small, to set a marginal tax rate high enough to exclude type 1 workers from the labour force, while paying them an unemployment consumption benefit. First we provide the mathematics for this case, then we can look at the economic intuition.

We explicitly introduce the constraint that $y_1 \geq 0$ into the above problem (in principle we should do that for all four variables but this is the only one for which it might bind). We then replace (8.21) and (8.22) by the appropriate Kuhn–Tucker conditions

$$-\phi_1[\psi'_1(\hat{y}_1) - \hat{\lambda}] + \hat{\mu}\psi'_2(\hat{y}_1) \leq 0 \quad \hat{y}_1 \geq 0 \quad (8.41)$$

$$\{-\phi_1[\psi'_1(\hat{y}_1) - \hat{\lambda}] + \hat{\mu}\psi'_2(\hat{y}_1)\}\hat{y}_1 = 0. \quad (8.42)$$

This says that if at the optimum $\hat{y}_1 > 0$, then the equality in the condition must hold as before, but the optimum could be at $\hat{y}_1 = 0$, in which case $-\phi_1[\psi'_1(\hat{y}_1) - \hat{\lambda}] + \hat{\mu}\psi'_2(\hat{y}_1) \leq 0$. This is the corner solution case. Rearranging the conditions as we did before gives in this case

$$\frac{\psi'_1(\hat{y}_1)}{u'(\hat{x}_1)} \geq 1 - \frac{\hat{\mu}\delta}{\phi_1\hat{\lambda}} \quad (8.43)$$

which is now mathematically correct if the left-hand side must be positive and the right-hand side could be negative.

The economic intuition for this case is given by recalling the intuition underlying the distortion in type 1's equilibrium away from the first-best condition. The idea is that we reduce x_1 and y_1 away from the point at which type 1's marginal rate of substitution equals 1, because this allows us to tax type 2 more heavily, and thus redistribute more, while maintaining incentive compatibility. This involves creating a distortion for type 1 consumers, driving a wedge between their marginal rate of substitution and the wage rate, and so we are sacrificing efficiency. How far we go depends on the trade-off between the marginal efficiency loss and the marginal distributional gain. This is the form the equity efficiency trade-off takes in this model. But now suppose there are very few type 1 and many type 2 consumers. Then creating a marginal distortion adds a lot to redistribution and costs only a little in aggregate efficiency loss. The terms of the trade-off are very favourable: so favourable indeed that we would like actually to have negative y_1 . Since this is not possible, the best we can do is to set y_1 at zero. Here, the marginal benefit from creating the distortion still exceeds its marginal cost (in all but a trivial special case), but this is as far as we can go. Note that the IC constraint still binds – type 2 is indifferent between not working and consuming $\hat{x}_1$, on the one hand, and getting her own allocation $(\hat{y}_2, \hat{x}_2)$ on the other. Figure 8.7 illustrates the nature of the equilibrium in this case.

We turn now to the intuition of the 'no distortion at the top' result. As already mentioned, since at the type 2 allocation the marginal rate of substitution equals one, there is a zero *marginal* tax rate. The reason for this is that as long as incentive compatibility is achieved, keeping type 2 consumers indifferent between their allocation and that of

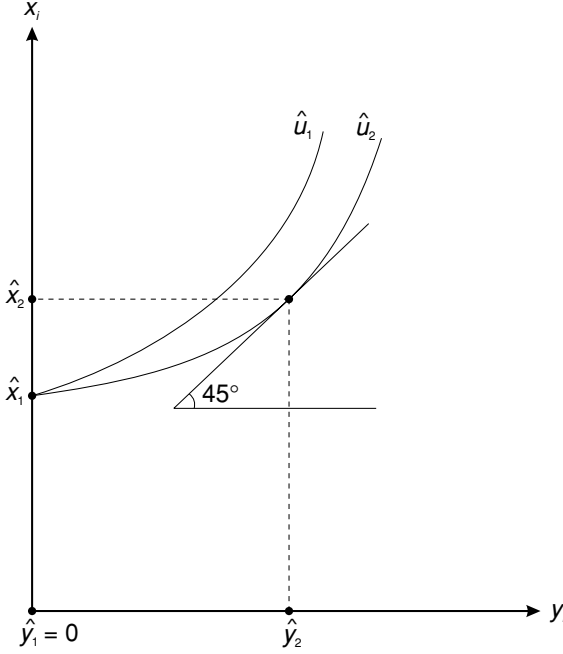


Figure 8.7 The corner solution

type 1, there is no reason to create a distortion. There are no higher productivity types whose allocation can be changed to increase the degree of redistribution.

As an alternative explanation, suppose, as in figure 8.3, we have found the optimal type 1 allocation $(\hat{y}_1, \hat{x}_1)$. Then as long as type 2 receives an allocation on the corresponding indifference curve $\hat{u}_2$, incentive compatibility is satisfied, and the planner is free to choose any point on this indifference curve. Since she wants to redistribute from type 2 to type 1, we could regard her as solving the problem of maximising the tax revenue from type 2 consumers while remaining on this indifference curve, i.e. solving the problem

$$\max_{x_2, y_2} (y_2 - x_2) \quad \text{s.t.} \quad u(x_2) - \psi_2(y_2) = \hat{u}_2 \quad (8.44)$$

It is then easy to see that the solution requires the allocation $(\hat{y}_2, \hat{x}_2)$, where $[dx_2/dy_2]_{\hat{u}_2} = \psi_2'(\hat{y}_2)/u'(\hat{x}_2) = 1$.

Finally, we can pose the question: to what extent does the existence of asymmetric information limit the possibilities of redistribution? The answer is: quite severely. Under lump sum taxation the planner wanted to achieve a position on the utility possibility frontier UPF such as U in figure 8.8, at which type 1 consumers have higher utility than type 2. But the need to respect the IC constraint means that at the planner's second-best optimum, $u(\hat{x}_2) - \psi_2(\hat{y}_2) > u(\hat{x}_1) - \psi_1(\hat{y}_1)$. (To show this just replace (y_1', x_1') , (y_2', x_2') by $(\hat{y}_1, \hat{x}_1)$, $(\hat{y}_2, \hat{x}_2)$ in (8.34)). Thus the planner is at a point such as S, below the UPF and above the equality line, in figure 8.8. The distortion to the type 1 allocation

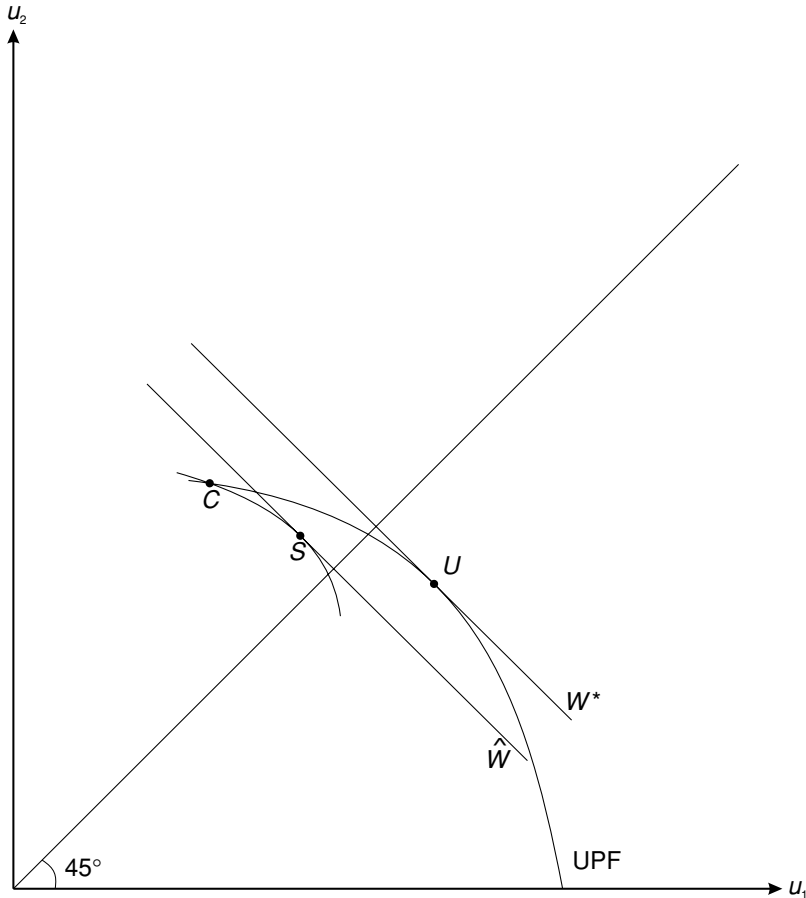


Figure 8.8 Limits to redistribution

creates an efficiency loss which pushes the economy inside the UPC, and the planner has to trade off the distributional gain against the efficiency loss. Thus the maximised value of the SWF is $\hat{W} < W^*$.

This relatively simple two-type version of Mirrlees's analysis is sufficient to bring out most, but not quite all, of the main results. It does not allow us to analyse the way in which the marginal tax rate, the degree of distortion of the allocation for types below the highest productivity type, varies with gross income, i.e. it does not allow us to study the structure of the optimal tax function. For this we need a model with a continuum of types, and this is presented in the next section.

8.3 A continuum of types

In this section we use the methods of optimal control theory to solve the problem of optimal non-linear taxation with a continuum of types. Thus we assume individuals have

values of w in the finite interval $[w_0, w_1]$ with differentiable distribution function $F(w)$ and continuous density $f(w) = F'(w)$ everywhere positive on $[w_0, w_1]$. We assume $w_0 > 0$.

Using $v(l)$ we now define

$$v(l) = v\left(\frac{y}{w}\right) \equiv \psi(y, w) \quad (8.45)$$

with

$$\psi_w = \frac{\partial v\left(\frac{y}{w}\right)}{\partial w} = -\frac{v'y}{w^2} < 0 \quad (8.46)$$

and

$$\psi_{yw} = -\frac{\left[\frac{v'y}{w} + v'\right]}{w^2} < 0 \quad (8.47)$$

so that increasing w reduces both the total and marginal utility cost of generating gross income.

The consumer now faces a tax function $T(y)$, which we assume is differentiable, and must take into account the effect of varying her gross income (labour supply) on her tax liability and therefore net income. Thus her choice problem is

$$\max_{x, y} u(x) - \psi(y, w) \quad \text{s.t. } x \leq y - T(y) \quad (8.48)$$

with first-order condition

$$\left[\frac{dx_i}{dy_i}\right]_{u_i^*} = \frac{\psi'_i(y_i^*)}{u'(x_i^*)} = 1 - t(y_i^*) \quad (8.49)$$

where $t \equiv T'(y)$ is her marginal tax rate. An important part of the analysis is to characterise how this marginal tax rate varies optimally with income.

We again assume the planner is utilitarian. As a benchmark we assume first that she can observe a consumer's type, and so solves the problem of optimal lump sum taxation, which can be formulated as

$$\max_{x(w), y(w)} \int_{w_0}^{w_1} \{u[x(w)] - \psi[y(w), w]\} f(w) dw \quad (8.50)$$

$$\text{s.t. } \int_{w_0}^{w_1} [y(w) - x(w)] f(w) dw \geq G. \quad (8.51)$$

In optimal control theory it is usual to express the integral constraint by introducing a state variable $z(w)$ and writing the constraints

$$z'(w) = [y(w) - x(w)] f(w) \quad (8.52)$$

$$z(w_0) = 0 \quad (8.53)$$

$$z(w_1) \geq G \quad (8.54)$$

which are clearly equivalent to (8.51). Thus x and y are the control variables and z is the state variable. The Hamiltonian function for the problem is

$$H = \{u[x(w)] - \psi[y(w), w]\}f(w) + \lambda(w)[y(w) - x(w)]f(w) \quad (8.55)$$

with λ the costate variable. Note that the Hamiltonian is concave in x and y , and so the following conditions are both necessary and sufficient for optimal functions $x^*(w)$, $y^*(w)$, $z^*(w)$:

$$\frac{\partial H}{\partial x} = \{u'[x^*(w)] - \lambda\}f(w) = 0 \quad (8.56)$$

$$\frac{\partial H}{\partial y} = \{-\psi_y[y^*(w), w] + \lambda\}f(w) = 0 \quad (8.57)$$

$$\lambda' = -\frac{\partial H}{\partial z} \quad (8.58)$$

$$z^*(w_0) = 0 \quad (8.59)$$

$$z^*(w_1) \geq G \quad \lambda(w_1) \geq 0 \quad \lambda(w_1)[z^*(w_1) - G] = 0. \quad (8.60)$$

Since $\partial H/\partial z = 0$, we have that λ is constant, and the first condition implies $\lambda > 0$. Thus the budget constraint is binding as an equality. For any pair of values w , $\hat{w} \in [w_0, w_1]$, the first condition implies $x^*(w) = x^*(\hat{w})$, while the second implies

$$\psi_y[y^*(w), w] = \psi_y[y^*(\hat{w}), \hat{w}]. \quad (8.61)$$

Given $w < \hat{w}$, and recalling the definition of $\psi(\cdot)$, we have

$$\frac{v'[l(w)]}{w} = \frac{v'[l(\hat{w})]}{\hat{w}} \Leftrightarrow l(\hat{w}) > l(w) \quad (8.62)$$

by the convexity of $v(\cdot)$. It then follows that utility is decreasing in w , while gross income y increases with w .

These results are essentially familiar from the two-type case discussed in this chapter. Figure 8.9 illustrates the optimal functions in the case of pure redistribution, $G = 0$. Redistribution takes place from higher- to lower-productivity consumers.

In the case of asymmetric information, the incentive compatibility (IC) constraint now requires that, for every pair of values w , $\hat{w} \in [w_0, w_1]$, the functions $x(\cdot)$, $y(\cdot)$ must satisfy

$$u[x(w)] - \psi[y(w), w] \geq u[x(\hat{w})] - \psi[y(\hat{w}), w] \quad (8.63)$$

to ensure that type w chooses the allocation designed for it. The problem is that this is not a tractable formulation of the planner's optimisation problem – there is an uncountable infinity of such constraints – and the trick is to find a way of expressing the constraint so that it becomes tractable. This can be done as follows. For *given* w , and given that she may choose any pair of functions $x(\hat{w})$, $y(\hat{w})$ on offer, we can write the type- w consumer's utility as a function of $\hat{w}$,

$$u(\hat{w}) = u[x(\hat{w})] - \psi[y(\hat{w}), w] \quad (8.64)$$

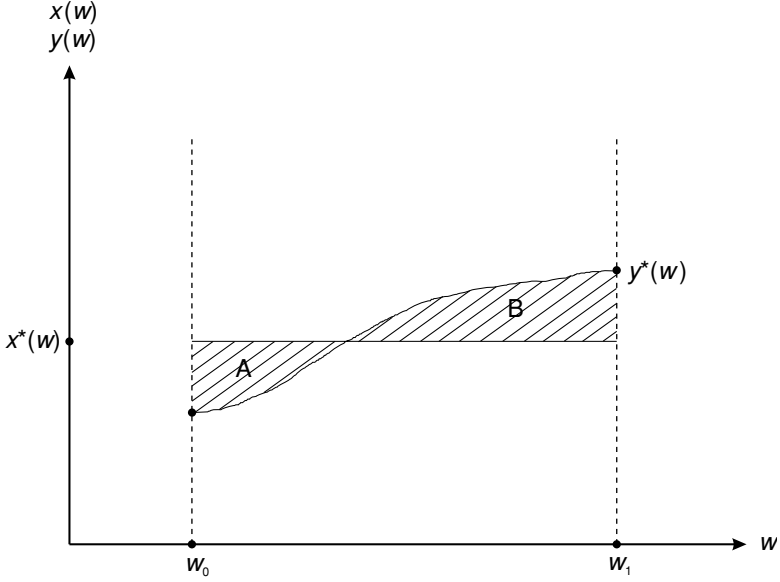


Figure 8.9 Optimal lump sum redistribution

with the derivative

$$u'(\hat{w}) = u'[x(\hat{w})]x'(\hat{w}) - \psi_y[y(\hat{w}), w]y'(\hat{w}). \quad (8.65)$$

Then incentive compatibility will be satisfied if $x(\cdot)$, $y(\cdot)$ are chosen in such a way that $u(\hat{w})$ is maximized at $\hat{w} = w$, since that is exactly equivalent to (8.63). The first-order condition for this maximisation problem is

$$u'[x(w)]x'(w) - \psi_y[y(w), w]y'(w) = 0 \quad (8.66)$$

and the second-order condition is

$$u''(\hat{w}) \leq 0 \quad \text{at } \hat{w} = w. \quad (8.67)$$

If $x(\cdot)$, $y(\cdot)$ are chosen in such a way that these conditions hold for all w then incentive compatibility is satisfied.

Now the total derivative of utility with respect to w is

$$u'(w) = u'[x(w)]x'(w) - \psi_y[y(w), w]y'(w) - \psi_w[y(w), w]. \quad (8.68)$$

Thus from (8.66) incentive compatibility implies

$$u'(w) = -\psi_w[y(w), w]. \quad (8.69)$$

Because in the second-best optimal control problem we are going to take utility $u(w)$ as a state variable, we take (8.69) as the *first-order* IC constraint. Note that, since $\psi_w < 0$, it has the immediate implication that incentive compatibility requires utility

to be increasing with type, so again we see that the utilitarian planner is going to be very constrained in the extent of redistribution she can undertake.

It can be shown⁹ that the second-order condition (8.67) is equivalent to the condition $y'(w) \geq 0$. However, for technical reasons – the avoidance of singularities – we write this constraint in the equivalent form

$$g[y'(w)] \geq 0 \quad (8.70)$$

where $g(\cdot)$ has the properties $g(0) = 0$, $g'(\cdot) > 0$, $g''(\cdot) \neq 0$, but is otherwise arbitrary. Finally, we can simplify the problem by eliminating the variable x . From

$$u = u(x) - \psi(y, w) \quad (8.71)$$

we obtain

$$x = u^{-1}[u + \psi(y, w)] \equiv \varphi(u, y, w) \quad (8.72)$$

with, using the implicit function theorem,

$$\varphi_u = \frac{1}{u'(x)}; \quad \varphi_y = \frac{\psi_y(y, w)}{u'(x)}; \quad \varphi_w = \frac{\psi_w(y, w)}{u'(x)}. \quad (8.73)$$

We are now in a position to formulate and solve the optimal control problem. The state variables are u , y and z , and the control variable is $c(w) = y'(w)$. Thus the problem is

$$\max_{c(w)} \int_{w_0}^{w_1} u(w) f(w) dw \quad (8.74)$$

subject to

$$z'(w) = [y(w) - \varphi(u, y, w)] f(w) \quad (8.75)$$

$$y'(w) = c(w) \quad (8.76)$$

$$u'(w) = -\psi_w[y(w), w] \quad (8.77)$$

$$g[c(w)] \geq 0 \quad (8.78)$$

and the endpoint conditions

$$z(w_0) = z(w_1) = 0 \quad (8.79)$$

$$y(w_0), \quad y(w_1) \quad \text{free} \quad (8.80)$$

$$u(w_0), \quad u(w_1) \quad \text{free}. \quad (8.81)$$

The first constraint, with its endpoint condition, is once again equivalent to the government budget constraint, where for convenience we set $G = 0$, and (8.77), (8.78) are the first- and second-order IC constraints respectively.

We introduce the costate variables λ , ν and μ and form the Hamiltonian

$$H = uf + \lambda(y - \varphi)f + \nu c + \mu(-\psi_w) \quad (8.82)$$

⁹ See Mirrlees (1971).

and then the Lagrangean

$$L = H + \gamma g \quad (8.83)$$

where, for convenience, we suppress the arguments of the functions.

The first-order conditions are then

$$\frac{\partial L}{\partial c} = v + \gamma g' = 0 \quad (8.84)$$

$$\lambda' = -\frac{\partial L}{\partial z} = 0 \quad (8.85)$$

$$v' = -\frac{\partial L}{\partial y} = -[\lambda f(1 - \varphi_y) - \mu \psi_{wy}] \quad (8.86)$$

$$\mu' = -\frac{\partial L}{\partial u} = -[f(1 - \lambda \varphi_u)] \quad (8.87)$$

with the condition on the Lagrange multiplier γ

$$g[c(w)] \geq 0 \quad \gamma \geq 0 \quad g[c(w)]\gamma = 0 \quad (8.88)$$

and the transversality conditions

$$v(w_0) = v(w_1) = 0 \quad (8.89)$$

$$\mu(w_0) = \mu(w_1) = 0. \quad (8.90)$$

Note first that if the constraint on $g(\cdot)$ (the second-order constraint) is non-binding for all w , then $\gamma = 0$ for all w , and so, from the first condition, $v = 0$ for all w also. But then we must have $v' = 0$. This is the so-called ‘first-order condition case’ investigated by Mirrlees. We shall obtain the results for this case, and then examine heuristically what happens if the second-order constraint is binding somewhere.

In this case, with $v' = 0$, the third condition can be written (given $\lambda \neq 0$) as

$$\varphi_y = \frac{\psi_y(y, w)}{u'(x)} = 1 - \frac{\mu \psi_{wy}}{\lambda f}. \quad (8.91)$$

The left-hand side is the marginal rate of substitution between y and x , the slope of an indifference curve. This condition is the counterpart in continuous type space of the condition (8.30) we derived for the two-type case. If we set the marginal tax rate $t(w)$ as

$$t(w) = \frac{\mu \psi_{wy}}{\lambda f} \quad (8.92)$$

we see that (8.91) then gives the first-order condition for the consumer’s utility maximisation problem. That is, the optimal tax function can be implemented by setting the marginal tax rate as in (8.92). Since $\psi_{wy} < 0$, to show that the tax rate is non-negative we have to show that the costate variable $\mu \leq 0$, while $\lambda > 0$.

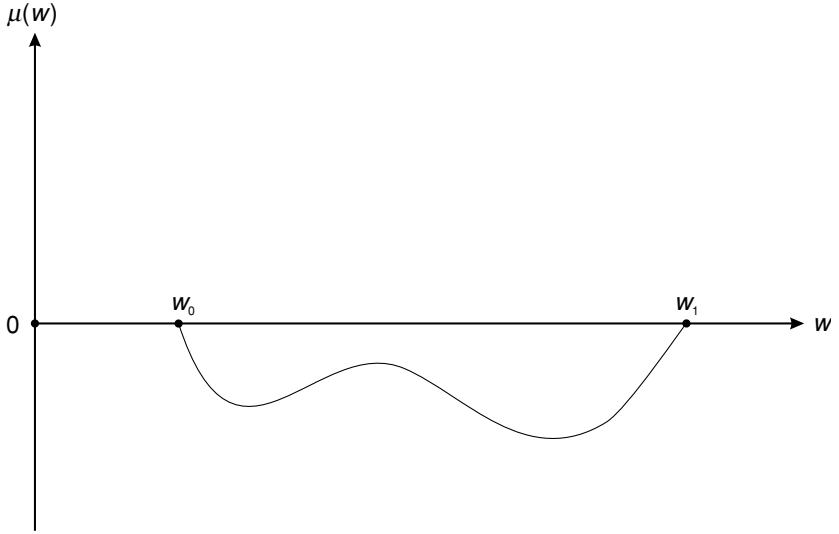


Figure 8.10 $\mu(w) < 0$, $w \in (w_0, w_1)$

To do this, note first that from the first-order condition, and the fact that the second-order condition is non-binding, we have

$$x' = \frac{\psi_y(y, w)}{u'(x)} y' > 0. \quad (8.93)$$

Since $u(x)$ is concave, this implies that u' is falling as w increases. Now write condition (8.87) as

$$\mu' = \left(\frac{\lambda}{u'} - 1 \right) f. \quad (8.94)$$

Since $\mu(w_0) = \mu(w_1)$ integrating gives

$$\int_{w_0}^{w_1} \mu' dw = \int_{w_0}^{w_1} \left(\frac{\lambda}{u'} - 1 \right) f dw = 0. \quad (8.95)$$

Since $u' > 0$ this establishes that $\lambda > 0$.

Now consider the bracketed term in (8.94), which determines the sign of μ' , since $f > 0$. With λ constant and u' falling, this term is monotonically increasing. The value of μ must go from 0 at w_0 to 0 at w_1 . Therefore the bracketed term cannot be non-negative at w_0 , for then this could never happen. So it starts off negative, but then must become positive to bring μ back up to zero. But this implies that $\mu < 0$ on the open interval (w_0, w_1) . Figure 8.10 illustrates this.

Having established that the marginal tax rate will be positive for $w \in (w_0, w_1)$, it is also easy to show that it cannot exceed 1, since

$$x' = (1 - t)y' \quad (8.96)$$

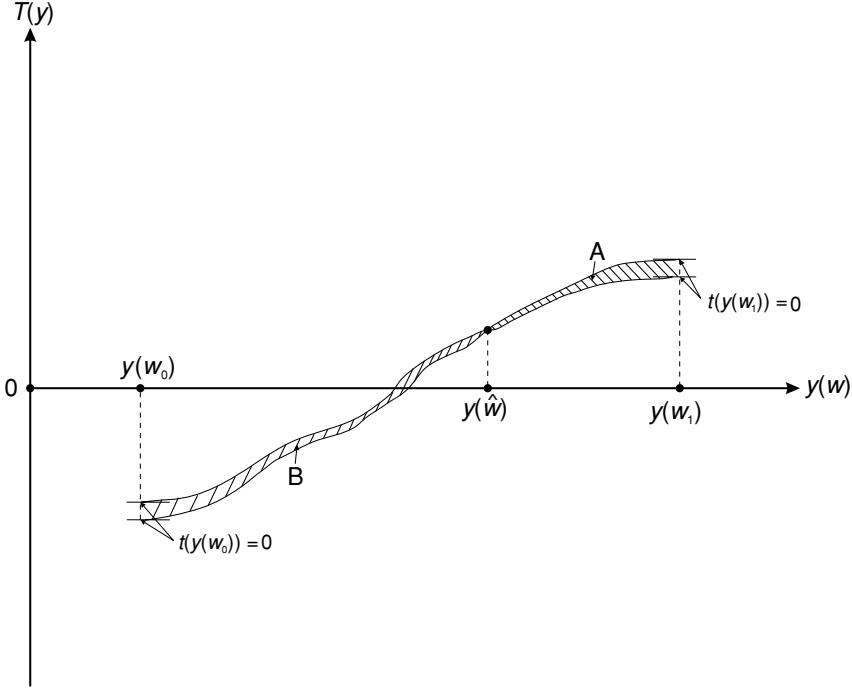


Figure 8.11 No distortion at the top and bottom

and so $x', y' > 0 \Rightarrow t < 1$.

Now, from the transversality condition $\mu(w_1) = 0$ we have

$$\frac{\psi_y(y(w_1), w_1)}{u'(x(w_1))} = 1. \quad (8.97)$$

This is the ‘no distortion at the top’ condition, also already familiar from the two-type case. Moreover, we have from the transversality condition (8.90) $\mu(w_0) = 0$, and so

$$\frac{\psi_y(y(w_0), w_0)}{u'(x(w_0))} = 1. \quad (8.98)$$

This is something we would certainly not have expected from the two-type case. It says that there is also no distortion at the bottom: $t(w_0) = 0$. The intuition for this is symmetric to that for the ‘no distortion at the top’ result, as we now show.

In Figure 8.11, $T(y)$ is a tax function with zero slope at both endpoints, representing the above transversality conditions. Assume it also satisfies the government budget constraint. The redistribution is again from higher to lower productivities and gross incomes. Suppose we make an increase in the marginal tax rate at an interior point such as $y(\hat{w})$. The effect of this is to increase tax revenue, by the area marked A, from all consumers with incomes greater than $y(\hat{w})$. This then implies that we can reduce the tax schedule at lower incomes while still meeting the constraint, say by the area given as B. The optimal marginal tax rate at each gross income point can be thought of as trading

off the distortion created above that point by the positive marginal tax rate against the benefit of redistribution to poorer consumers below that point.

Now consider the effect of increasing the slope of the curve above zero at $y(w_1)$. This creates a distortion for these consumers, but no increased revenue for redistribution because there are no consumers with incomes above that point. On the other hand, raising the marginal tax rate at $y(w_0)$ increases the distortion and increases tax revenue at every higher income, but this brings no corresponding benefit because there are no consumers with incomes below $y(w_0)$ who can be made better off. Thus the optimal marginal tax rate at this point is also zero. This simple point does not emerge in the discrete two-type case.

Note that this ‘no distortion at the bottom’ argument in a certain sense does not hold when there is an interval of types $[w_0, w_m)$ for which the consumers are optimally unemployed, so that $y(w) = 0$ for $w \in [w_0, w_m)$, with w_m the lowest type at which y becomes positive (this is the counterpart of the corner solution in the two-type case). In this case a small increase in the marginal tax rate at w_m generates additional revenue which can be redistributed to the unemployed, and so there is a positive marginal tax rate at the lowest gross income. This kind of possibility is referred to as ‘bunching’, and is discussed more fully below.

One important implication of the result that marginal tax rates are zero at the top and bottom of the productivity distribution is that the marginal tax rate cannot be monotonically increasing. This conflicts with the idea of a tax system being progressive, in the sense of having a strictly increasing marginal tax rate. Unfortunately, at this level of generality, it is hard to say anything more specific about the shape of the optimal tax function than this, together with the comforting conclusion that the marginal rate lies between zero and one. As (8.92) shows, the tax function will depend on the form of the utility function, the shape of the productivity distribution, and, something which in our analysis is concealed by the assumption of a utilitarian planner, the form of the social welfare function. For this reason a number of authors have made specific assumptions about these functions in order to calculate numerical examples of optimal tax functions. Though these are useful in increasing our understanding of the model, they do not provide robust recommendations for practical tax policy.

One important numerical example is that constructed by Ebert (1992), who shows that it cannot simply be taken for granted that the second-order condition is non-binding at all points along the optimal tax function. Taking a quasilinear utility function and a permissible, though not particularly plausible, productivity density function, Ebert derived an optimal relationship between consumption and gross income such as that depicted in figure 8.12. The key point about it is the fact that as w increases, the curve in the (y, x) -plane describes a loop backwards between y_a and y_b . Between points a and b in the figure $y' < 0$, thus violating the second-order condition. The problem then is that although the points on the curve are all local optima, since they satisfy the first-order condition, they are not all global optima, and so the tax function underlying this curve cannot be implemented.

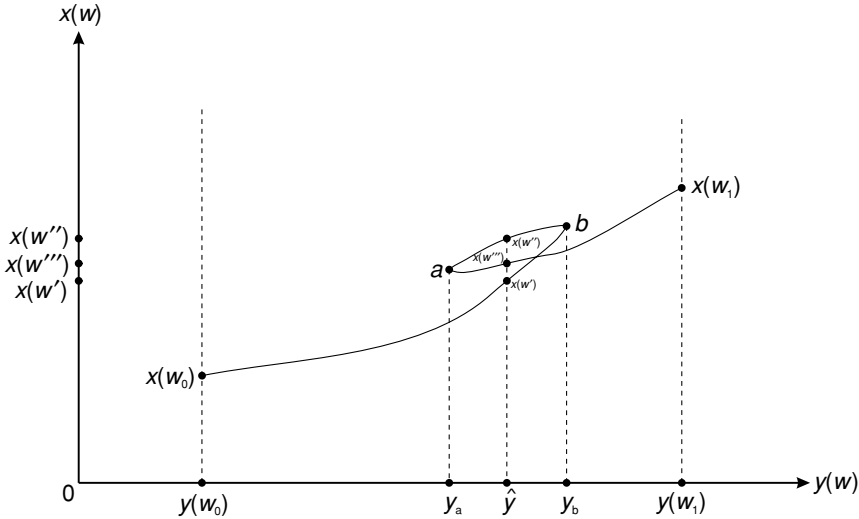


Figure 8.12 Second-order condition binding

Thus consider the gross income $\hat{y}$. Corresponding to this are optimal consumptions $x(w')$, $x(w'')$ and $x(w''')$ for the three productivity types, with $w' < w'' < w'''$. But obviously, confronted with this schedule, each type will choose $x(w'')$. That is true everywhere between a and b : only the points on the upper portion of the loop would be chosen. The ‘solution’ ignores the fact that the second-order constraint must be binding over this range.

Ebert then shows that when the second-order condition is taken into account in solving for the optimal path, we obtain the phenomenon known as ‘bunching’, which is illustrated in figure 8.13. The optimal function relating x and y has a kink at some value of gross income, corresponding to the fact that, as shown in part (b) of the figure, there is an interval of w -values such that $y' = x' = 0$. These consumers are being treated equally by the tax system, even though they have differing productivities. In fact, the higher the value of w in this interval, the less labour has to be supplied, and so utility is still increasing over the interval.

It is also possible to have bunching over an initial interval of the productivity distribution, say $[w_0, w_m)$, with the consumers in this range unemployed:¹⁰ they choose zero labour supply and receive consumption x_0 , and so are bunched at $(0, x_0)$. This is the corner solution case. We can show that at w_m , the lowest productivity at which gross income (labour supply) becomes positive, the marginal tax rate is also strictly positive, so that in this sense the ‘no distortion at the bottom’ result does not apply. Thus suppose to the contrary the tax rate were zero, so that the consumer at w_m would be in equilibrium with a marginal rate of substitution, $\varphi_y = 1$. The single crossing condition then implies that as we move leftwards through the interval $[w_0, w_m)$, the slopes of the

¹⁰ As first shown by Seade (1977).

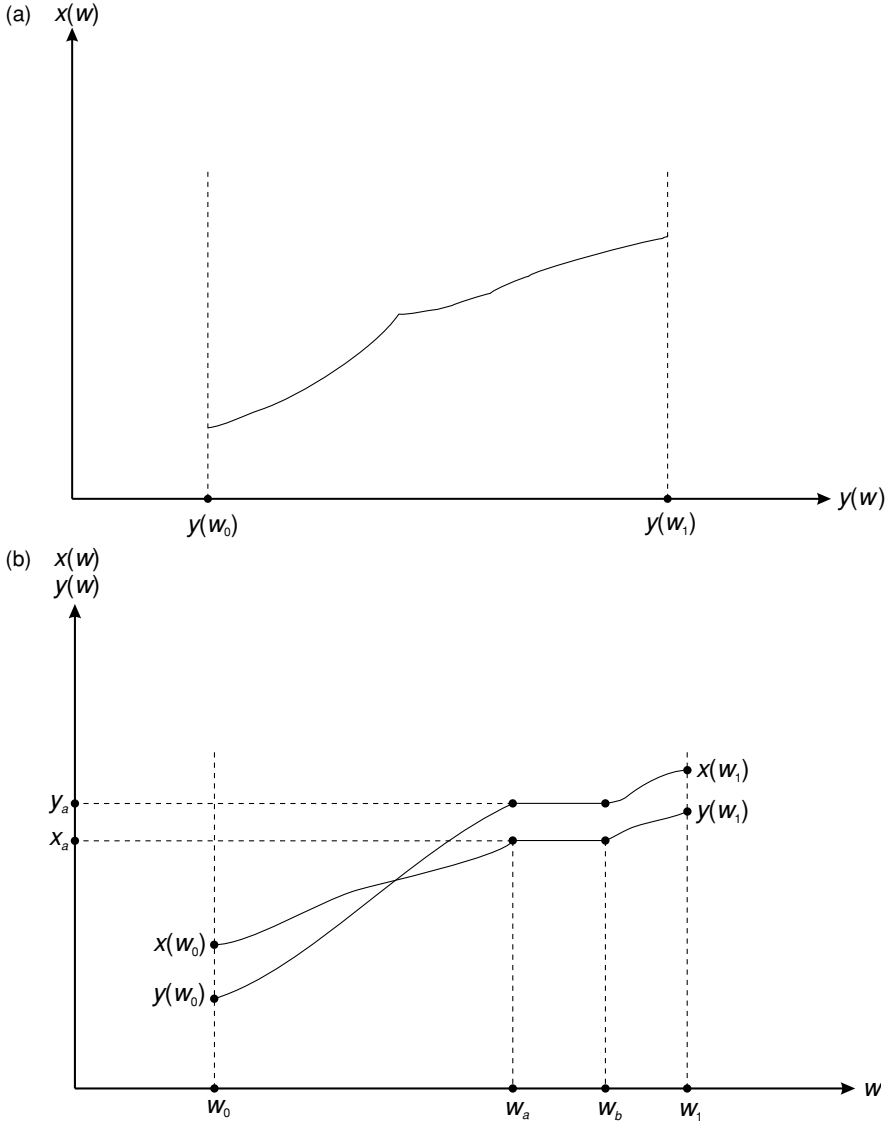


Figure 8.13 Bunching

corresponding indifference curves at the corner solution are increasing (refer to figure 8.14), implying that at w_0 , the slope $\varphi_y(w_0) > 1$. However, we now show that in fact $\varphi_y(w_0) = 1$, giving a contradiction. Thus the slope at w_m must be less than 1, implying a positive marginal tax rate.

Let $\phi[c(w)] \equiv \gamma g'$ in (8.84), so that $v' = -\phi'c'(w)$. Inserting this into (8.86) and rearranging gives

$$\lambda f(1 - \varphi_y) = \mu \psi_{wy} + \phi'c'(w). \quad (8.99)$$

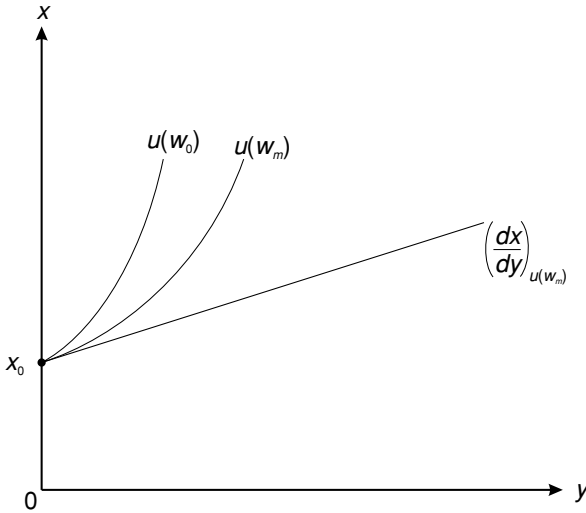


Figure 8.14 Bunching at $y = 0$

At $w = w_0$ we have from the transversality conditions that $\mu = 0$. Moreover, over the bunching interval $[w_0, w_m)$ we have that $y(w)$ is constant at zero, implying $c(w_0) = c'(w_0) = 0$. Thus at w_0 the right-hand side of (8.99) is zero, which gives the result.

Finally, a simple argument establishes that bunching cannot occur at the top end of the distribution, w_1 . We know that at this point the optimal tax rate is zero. Suppose that over some interval $[w_l, w_1]$ there is bunching. Then as we move leftwards from w_1 in this interval, since w is falling the slopes of the corresponding indifference curves at the bunched (y, x) point must be increasing (single crossing). But this implies that the marginal tax rate is falling from zero, i.e. is negative, which we ruled out earlier as not possible. Thus no such bunching interval exists.

From the point of view of the concerns of this book, a limitation of the Mirrlees tax analysis is that it is based on the standard single-consumer model. Unfortunately, extension of the approach to a two-person household is far from being a trivial matter, as we see in the next section.

A second issue concerns the ‘innateness’ of productivity and its identification with the market wage in the model. As a modelling strategy this is extremely useful, since it captures in a simple and tractable way the fundamental idea underlying this approach to the trade-off between efficiency and equity in income taxation. There is some innate characteristic that determines the achieved utility differences between otherwise identical individuals, and ideally a planner would like to redistribute utilities by taxing this characteristic. Since she cannot observe it, she does the next best thing, designing the tax system in such a way as to maximise the extent of redistribution subject to the IC constraint. The Revelation Principle tells us that there is nothing better she can do.

However, in reality the productivity differences that drive wage differences¹¹ are not innate, but the result of investment in human capital, which is an endogenous decision variable. Thus using a mechanism to induce revelation of one's 'wage type' and taxing accordingly is not after all taxing an innate characteristic, and can be expected to distort choices of investment in human capital formation.

A further issue arises out of the implicit assumption in the model, that the planner can commit herself not to use the information about the taxpayer's type that is revealed at the optimal solution. Thus suppose, as is in reality the case, that taxes are levied at regular intervals of time. If in the first period the planner uses the above mechanism, she then knows each consumer's type. She would then want to switch to lump sum taxation in the second period. If the high productivity type expects her to do this, she will not behave as the model predicts. It is quite possible to extend the model to analyse this case, but the resulting tax systems are complex and typically involve a great deal of pooling – equal tax rates for consumers of different types, at the optimum.¹² The assumption that the planner can commit not to use any information revealed by an individual's choice of tax rate is therefore an important one.

8.4 Two-person households

We now make the simplest possible extension to the two-type model to take account of two-person households.¹³ Suppose households consist of two individuals who may be of either productivity type, so that there are four possible household types.¹⁴ Call the first individual f and the second m . A household's type is then described by the pair (w_i, w_j) , $i, j = H, L$, with w_i the wage of f and w_j that of m . A household of type (w_i, w_j) has the utility function

$$u_{ij} = u(x_{ij}) - v\left(\frac{y_i}{w_i}\right) - v\left(\frac{y_j}{w_j}\right) \equiv u(x_{ij}) - \psi_i(y_i) - \psi_j(y_j) \quad i, j = H, L. \quad (8.100)$$

All individuals in all households have identical preferences, and the properties of the functions v and ψ are just as in the earlier discussion of the two-type model. This assumption on the form of the household utility function introduces a form of symmetry that is very powerful in simplifying the analysis, as we see below.

¹¹ Wage differences may also reflect other factors such as labour market discrimination, market power and pure luck.

¹² For approaches to this problem see Apps and Rees (2006).

¹³ For the literature on this topic see Schroyen (2003), Apps and Rees (2006), Brett (2007), and Kleven *et al.* (2007). To keep the analysis relatively simple, we abstract here from issues of both within-household distribution and household production. Schroyen's model has the household members producing a household public good, but since productivity in this does not vary across households, this does not directly influence the structure of optimal taxes.

¹⁴ For some literature relevant to this case, see in particular Armstrong and Rochet (1999), and Rochet and Stole (2001).

Let ϕ denote the proportion of f s who are high-wage, and ϕ_H the proportion of their partners who are also high-wage, while ϕ_L denotes the proportion of partners of low-wage f s who are high-wage. Then we assume $1 > \phi_H, \phi_L > 0$. The probability distribution of types is then

$$\begin{aligned}(w_H, w_H) &: \phi\phi_H \equiv \phi_{HH} \\ (w_H, w_L) &: \phi(1 - \phi_H) \\ (w_L, w_H) &: (1 - \phi)\phi_L \\ (w_L, w_L) &: (1 - \phi)(1 - \phi_L) \equiv \phi_{LL}.\end{aligned}$$

Since ‘mixed’ households are essentially identical, because of the assumption of identical preferences, we group them together and let $\phi_M = \phi(1 - \phi_H) + (1 - \phi)\phi_L$. In effect, then, we reduce the number of possible household types to three, which considerably reduces the number and complexity of the incentive compatibility constraints that we will have to consider.¹⁵ Note that the allocation received by an individual of type $j = H, L$ may in general depend on the type of household to which he or she belongs, and so we attach a subscript to the y -variable to indicate this. Thus y_{Mj} is the gross income of a j -productivity individual in a mixed household, y_{jj} that of each individual in a matched household.

We again assume that the planner is utilitarian, and so the social welfare function W is

$$\begin{aligned}\phi_{HH}[u(x_{HH}) - 2\psi_H(y_{HH})] &+ \phi_M[u(x_M) - \psi_H(y_{MH}) - \psi_L(y_{ML})] \\ &+ \phi_{LL}[u(x_{LL}) - 2\psi_L(y_{LL})]\end{aligned}\quad (8.101)$$

while the government budget constraint is

$$\phi_{HH}(2y_H - x_{HH}) + \phi_M[y_{MH} + y_{ML} - x_M] + \phi_{LL}(2y_{LL} - x_{LL}) \geq G. \quad (8.102)$$

Under symmetric information, where the planner can observe everyone’s type, it is straightforward to show that the optimal lump sum taxes, found by maximising W subject to (8.102), imply a simple extension of the results for single-person households. Everyone receives the same consumption

$$x_{HH} = x_M = x_{LL} \quad (8.103)$$

and high-productivity individuals supply more labour, regardless of the type of their partner

$$y_{HH} = y_{MH} > y_{LL} = y_{ML}. \quad (8.104)$$

It follows that

$$u(x_{LL}) - 2\psi_L(y_{LL}) > u(x_M) - \psi_H(y_{MH}) - \psi_L(y_{ML}) > u(x_{HH}) - 2\psi_H(y_{HH}) \quad (8.105)$$

¹⁵ As Schroyen (2003) points out, this assumes that the tax authority is able to observe who forms a household with whom, something excluded in his model. Here we adopt this assumption as being reasonably realistic.

and under asymmetric information high-productivity individuals would want to mimic low-productivity individuals and all households would represent themselves as of type LL .

Under symmetric information, the ‘unit of taxation’ could be regarded either as the household, in which case taxes would be expressed as

$$2y_{HH} - x_{HH}, \quad y_{MH} + y_{ML} - x_M, \quad 2y_{LL} - x_{LL} \quad (8.106)$$

or equivalently as the individual, with taxes

$$y_{HH} - \frac{1}{2}x_{HH}, \quad y_{MH} - \frac{1}{2}x_M, \quad y_{ML} - \frac{1}{2}x_M, \quad y_{LL} - \frac{1}{2}x_{LL}. \quad (8.107)$$

Taxation is determined by individual productivity, but observability of household type means it can also be levied on household income.

Under asymmetric information, we have to introduce incentive compatibility constraints. A potential difficulty here is the multiplicity of logically possible constraints. Figure 8.15 shows the directions of possible incentive compatibility constraints, with an arrow indicating a possibly binding constraint. In (a) we show the ‘downward constraints’, in (b) the ‘upward constraints’. An important consequence of our assumptions is that there are no transverse constraints, imposing incentive compatibility between HL - and LH -type households, since they are in this model identical. This is a substantial simplification, though as the analysis of Brett (2007) makes clear,¹⁶ it also means that we rule out possible cases in which tax rates are negative at the optimum.

We can show, using the same type of argument as was applied in the single-person household case, that the upward constraints cannot bind at the optimum, since a standard single crossing condition holds between at least one member of the households HH , HL , LH , and LL in the usual way. Thus we solve the problem in the presence only of the three downward constraints:

$$u(x_{HH}) - 2\psi_H(y_{HH}) \geq u(x_M) - \psi_H(y_{MH}) - \psi_H(y_{ML}) \quad (8.108)$$

$$u(x_M) - \psi_H(y_{MH}) - \psi_L(y_{ML}) \geq u(x_{LL}) - \psi_H(y_{LL}) - \psi_L(y_{LL}) \quad (8.109)$$

$$u(x_{HH}) - 2\psi_H(y_{HH}) \geq u(x_{LL}) - 2\psi_H(y_{LL}). \quad (8.110)$$

In analysing these constraints it is very useful to have the

Lemma *Given that the $\psi(\cdot)$ functions are continuous, strictly convex and increasing, we have:*

$$\psi_L(y) - \psi_H(y) > (=)\psi_L(y') - \psi_H(y') \Leftrightarrow y > (=)y'. \quad (8.111)$$

Proof: It suffices just to prove the inequality. We can rewrite the proposition as

$$\int_{y'}^y [\psi'_L(y) - \psi'_H(y)]dy > 0 \Leftrightarrow y > y' \quad (8.112)$$

and this follows immediately from the fact, established in section 8.2, that $\psi'_L(y) - \psi'_H(y) > 0$, all y . $\square$

¹⁶ The present model could be regarded as nested in his model, since he does not make assumptions on household preferences which make mixed households identical. Consequently he has to consider IC constraints between the two types of mixed households, which considerably complicates the analysis.

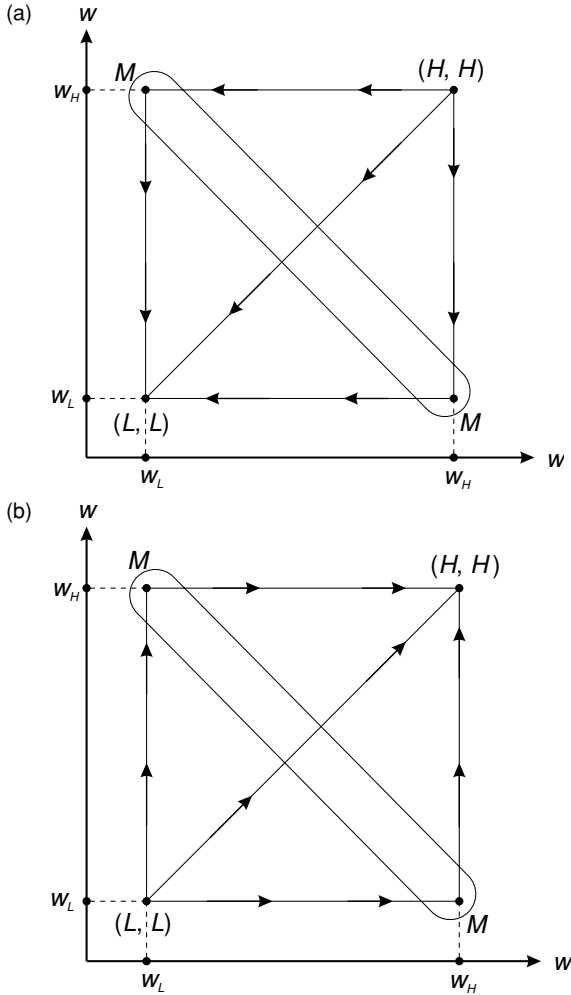


Figure 8.15 Directions of possible binding constraints

Using this lemma, the intuition for which is shown in figure 8.16, it is then straightforward to show the following:

- (i) the first two constraints binding and the third non-binding implies $y_{LL} < y_{ML}$;
- (ii) the first and third constraints binding and the second non-binding implies $y_{ML} < y_{LL}$;
- (iii) the second and third constraints binding and the first non-binding also implies $y_{ML} < y_{LL}$;
- (iv) all three constraints binding implies $y_{ML} = y_{LL}$.

These results are derived simply by setting the appropriate equalities or inequalities in the constraints in each case, rearranging and cancelling terms and then applying the above lemma. The intuition is that in each case there will be distortion of the optimum

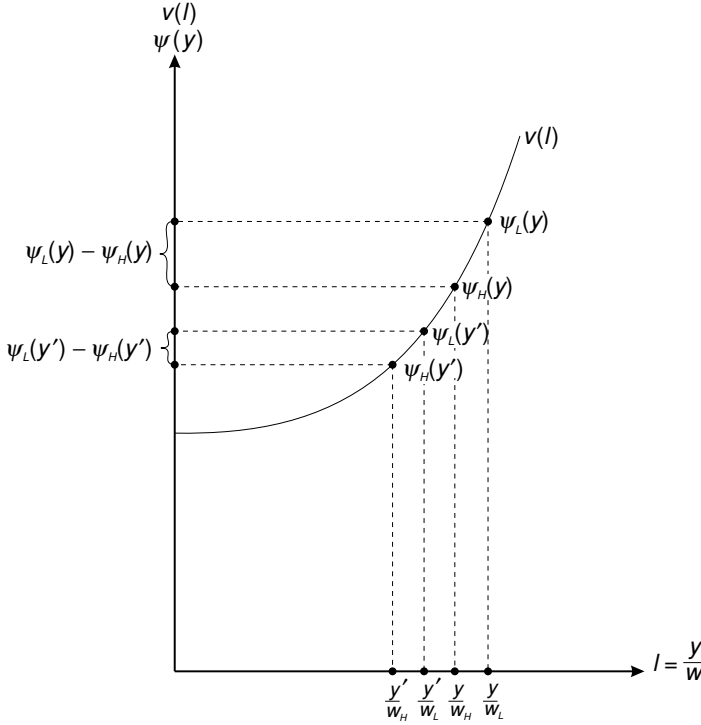


Figure 8.16 Illustration of the lemma

allocation, in the sense of section 8.2, only in the case of the low-productivity individuals in each type of household, and the relative extent of the distortions in the M - and LL -type households will depend on the parameters of the model, in particular the proportions ϕ_{ij} and the derivatives u' and ψ'_i . In general, each of these four cases is possible.

We formulate the second-best optimal taxation problem as that of maximising W subject to the budget constraint and the IC constraints (8.108)–(8.110).

The first-order conditions for this problem, with λ and μ_{HM} , μ_{ML} and μ_{HL} the multipliers attached to the budget and three incentive constraints respectively, are:

$$(\phi_{HH} + \mu_{HM}^* + \mu_{HL}^*)u'(x_{HH}^*) - \lambda^*\phi_{HH} = 0 \quad (8.113)$$

$$-(\phi_{HH} + \mu_{HM}^* + \mu_{HL}^*)2\psi'_H(y_{HH}^*) + \lambda^*2\phi_{HH} = 0 \quad (8.114)$$

$$(\phi_M + \mu_{ML}^* - \mu_{HM}^*)u'(x_M^*) - \lambda^*\phi_M = 0 \quad (8.115)$$

$$-(\phi_M + \mu_{ML}^* - \mu_{HM}^*)\psi'_H(y_{MH}^*) + \lambda^*\phi_M = 0 \quad (8.116)$$

$$-(\phi_M + \mu_{ML}^*)\psi'_L(y_{ML}^*) + \lambda^*\phi_M + \mu_{HM}^*\psi'_H(y_{ML}^*) = 0 \quad (8.117)$$

$$(\phi_{LL} - \mu_{HL}^* - \mu_{ML}^*)u'(x_{LL}^*) - \lambda^*\phi_{LL} = 0 \quad (8.118)$$

$$-(2\phi_{LL} - \mu_{ML}^*)\psi'_L(y_{LL}^*) + \lambda^*2\phi_{LL} + (2\mu_{HL}^* + \mu_{ML}^*)\psi'_H(y_{LL}^*) = 0 \quad (8.119)$$

together with the constraints.

From the first four conditions we immediately have that, in all cases, there is ‘no distortion at the top’:

$$\frac{\psi'_H(y_{HH}^*)}{u'(x_{HH}^*)} = \frac{\psi'_H(y_{MH}^*)}{u'(x_M^*)} = 1. \quad (8.120)$$

In no household type is the high-productivity individual subject to a non-zero marginal tax rate. The reason for this is that at any point (x^0, y^0) , the marginal rates of substitution $\psi'_H(y^0)/u'(x^0)$ are the same for both HH - and M -household types. Thus there is never any gain in terms of extra redistribution from distorting the equilibrium of the high-productivity type in the mixed household (refer back to the discussion of the intuition of the results in section 8.2). On the other hand, this non-distortion relates to the form of the equilibrium condition, not to the values at the optimum, which are different from those in the lump sum tax case, as we show below.

Moreover, it is straightforward to show from the constraints (using the fact that $\psi_H(y) < \psi_L(y)$ for all y) that in all cases

$$u(x_{LL}^*) - 2\psi_L(y_{LL}^*) < u(x_M^*) - \psi_H(y_{MH}^*) - \psi_L(y_{ML}^*) < u(x_{HH}^*) - 2\psi_H(y_{HH}^*) \quad (8.121)$$

precisely reversing the planner’s desired utility ordering.

A further important general result follows from noting that if at any point (y^0, x^0) ,

$$\frac{\psi'_H(y^0)}{u'(x^0)} = 1 \quad (8.122)$$

then

$$m(y^0, x^0) \equiv \psi'_H(y^0) - u'(x^0) = 0 \quad (8.123)$$

and from the implicit function theorem

$$\left[\frac{dx}{dy} \right]_{m(y^0, x^0)=0} = -\frac{m_y}{m_x} = \frac{\psi''_H(y^0)}{u''(x^0)} < 0. \quad (8.124)$$

Thus the set of points satisfying (8.122) forms a negatively sloped locus in the (y, x) -plane, and this in turn implies that we have

$$y_{MH}^* > y_{HH}^*, \quad x_{HH}^* > x_{MH}^* \quad (8.125)$$

since it is easy to show that we always have $y_{MH}^* > y_{ML}^*$.

The most interesting aspect of this result is that the gross income (labour supply) of the high productivity type in the mixed household is always above that in the HH -household. This follows because the IC constraint is relaxed as a result, enabling more redistribution away from the HH -household, by reducing the mixed household’s consumption level. As we shall see, given the no distortion at the top condition and the relationship in (8.125), this must imply a further reduction in the attractiveness of mimicking the mixed household by increasing its required gross income level y_{MH}^* .

To obtain further results we now need to work through the specific cases. Space precludes a treatment of them all, but it is sufficient to illustrate with just one of them. Thus we take the case in which $\mu_{HM}^* > 0$, $\mu_{ML}^* > 0$, $\mu_{HL}^* = 0$, $y_{ML}^* > y_{LL}^*$, so that the diagonal constraint is non-binding. From the constraints we have immediately that $x_M^* > x_{LL}^*$ since otherwise, with $y_{MH}^* > y_{ML}^* > y_{LL}^*$, the constraint (8.108) could not possibly be satisfied. Thus the orderings of consumptions and gross incomes¹⁷ are

$$x_{HH}^* > x_M^* > x_{LL}^* \quad (8.126)$$

$$y_{HL}^* > y_{HH}^* > y_{ML}^* > y_{LL}^*. \quad (8.127)$$

In this case, the LL -households are clearly the ‘most distorted’, in comparison with lump sum taxation. From the first-order conditions with the appropriate non-zero multipliers we have that L -types in M households have an allocation characterised by

$$\frac{\psi'_L(y_{ML}^*)}{u'(x_M^*)} = 1 - t_{ML}^* \quad (8.128)$$

where

$$t_{ML}^* = \frac{\mu_{ML}^* \delta_{ML}}{\phi_M \lambda^*} \quad (8.129)$$

and

$$\delta_{ML} = \psi'_L(y_{ML}^*) - \psi'_H(y_{ML}^*) > 0. \quad (8.130)$$

Thus there is a downward distortion in their labour supply as compared to the first best, brought about by a positive marginal tax rate.

For both members of LL households we have

$$\frac{\psi'_L(y_{LL}^*)}{u'(x_{LL}^*)} = 1 - t_{LL}^* \quad (8.131)$$

where

$$t_{LL}^* = \frac{\mu_L^* \delta_{LL}}{2\phi_{LL} \lambda^*} \quad (8.132)$$

and

$$\delta_{LL} = \psi'_L(y_{LL}^*) - \psi'_H(y_{LL}^*) > 0. \quad (8.133)$$

Thus again there is a downward distortion of labour supply as compared to the first best. From the relations between the consumptions and gross incomes and the strict concavity of $u(\cdot)$ and convexity of $\psi_L(\cdot)$ we have that

$$\frac{\psi'_L(y_{LL}^*)}{u'(x_{LL}^*)} < \frac{\psi'_L(y_{ML}^*)}{u'(x_M^*)} \quad (8.134)$$

¹⁷ In all other cases, the ordering of consumptions is the same, but the ordering of gross incomes may have to be changed, as we derived earlier.

implying

$$t_{LL}^* > t_{ML}^*. \quad (8.135)$$

This confirms the idea that *LL*-households are the ‘more distorted’ away from the allocation they would receive under lump sum taxation. Note that the possibility of this inequality, and therefore of this particular case, will be influenced by the relative values of the probabilities ϕ_M and ϕ_{LL} . Thus we would not expect it to hold if, other things being equal, the probability of mixed households was small relative to that of households in which both members are of low productivity. As we discussed in section 8.2, the lower the probability of a type, the lower the marginal cost of distorting its labour supply, and so the greater the distortion. Indeed, in this case, for ϕ_{LL} sufficiently low we could again have a corner solution in which low-productivity households would be paid not to work.

A final general point concerns the issue of the tax unit.¹⁸ In this analysis, the *tax base* is clearly the individual income, in the sense that tax rates vary with the individual’s type and are applied to the individual’s gross income. Nonetheless, it is also important to note that the tax rate applied to the income of an individual of a given type also depends on the type of household she is in. High-productivity types in mixed households do worse than those in *HH*-households, because they obtain lower consumption (abstracting from issues of within household distribution) and have to work harder – implicitly they pay a higher lump sum tax. Low-productivity types in mixed households, on the other hand, pay a lower marginal tax rate than their counterparts in *LL*-households.¹⁹ Taxes vary both with the individual and with household type. In this sense the household is the tax unit.

¹⁸ This point is also clear in the analyses of Schroyen and Brett.

¹⁹ In an analysis of the marriage market this type of tax system could well influence who forms households with whom.

9 Tax reform

9.1 Introduction

Any given tax system at any one time is unlikely to be optimal in the sense defined in the previous three chapters. It is also extremely unlikely that it would be politically and administratively feasible to replace an existing tax system with an optimal one, even if we knew with any degree of quantitative precision what it looked like. From the point of view of practical tax policy, therefore, the question of tax reform, the implementation of small, piecemeal changes to an existing tax system, is the most relevant one. In this chapter, we adopt the approach of tax reform to analyse the effects of tax policy in an economy consisting of two-person households with household production. We take account of the effects of tax reform on the within-household distribution of utilities, exploiting in doing so the idea of the household as a small economy, but our main concern, as in the previous chapters on optimal taxation, is with the effects on the across-household distribution and their implications for the tax structure.

The public economics literature on the theory of tax reform looks forbiddingly technical.¹ We believe, however, that the main general points of this literature can be made simply and non-technically. The formal analysis is only required when we begin to look at specific tax reform policies, as we do in the remainder of this chapter.

To begin with, we should take account of two traditional warnings on the possibility of misleading intuition when we are in the second-best world of piecemeal policy reform. First, a policy which replaces a Pareto-inefficient with a Pareto-efficient allocation may not satisfy the criterion of a Pareto improvement and may actually reduce social welfare, in the absence of the possibility of lump sum redistribution.

Figure 9.1 makes this point very simply. A move from inefficient allocation A to efficient allocation B makes consumer 1 worse off while making 2 better off, and lies below the indifference curves passing through A of all planners with inequality aversion

¹ See, for example, Guesnerie (1995) and Myles (1995) ch. 6. The reader for whom our discussion here is too intuitive should consult these.

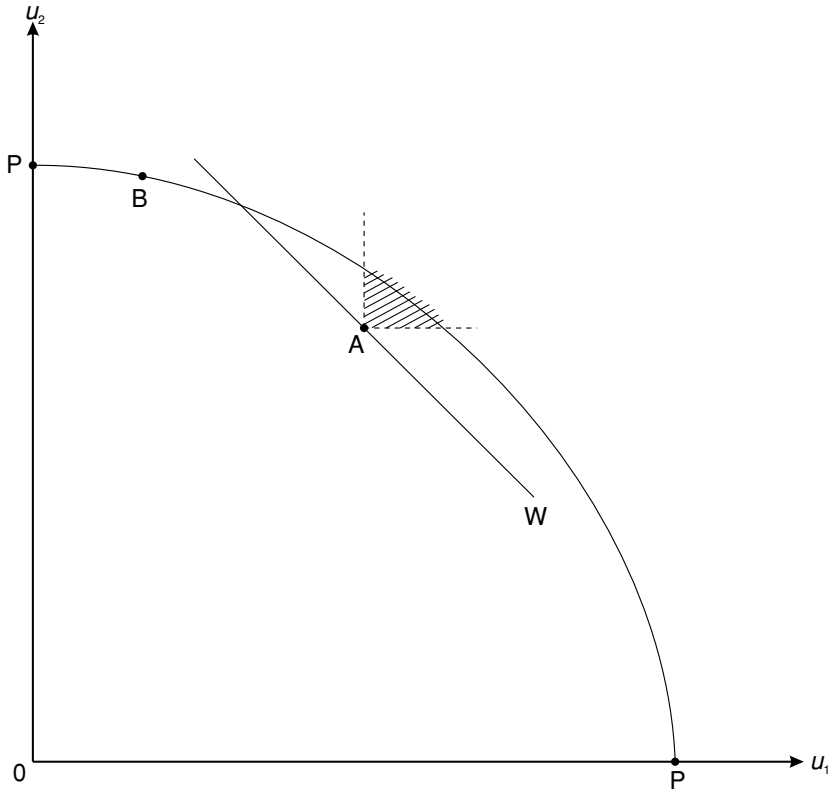


Figure 9.1 A move to Pareto efficiency to reduce social welfare

at least as large as that of the utilitarian planner. Concentrating only on ‘pure efficiency’ in making policy recommendations, and ignoring issues of income distribution, runs the risk of falling foul of this pitfall. The figure also makes clear that the requirement that a policy meet the Pareto criterion for improvement, moving the allocation from A to somewhere within the shaded area, is in general more stringent than that the policy increase social welfare. The willingness to trade off the utility gains and losses of different consumers widens the set of policies that can be compared.

Second, it is wrong to believe that reducing or eliminating a distortion from Pareto efficiency in some part of a second-best economy is the same thing as making a welfare improvement, since second-best optima typically involve trading off distortions throughout the economy. Although this has been well understood since the beginning of the second-best literature,² it continues to be ignored, even in recent contributions.

We take it that the underlying model of the tax problem is well formulated, in the sense that the objective function in the problem is continuous and the set of

² See, for example, Meade (1955) and Lipsey and Lancaster (1956).

feasible values of the policy instruments is non-empty, closed and bounded. Then a globally optimal tax policy, or vector of values of these instruments, certainly exists.³ Then if, as we assume, we are at a point in this set that is not a global optimum, it must always be possible to find directions of change that increase the value of the maximand.

An important caveat, however, concerns the distinction between local and global optima and the specification of what is meant by a tax reform. Typically, we consider only *localised tax reforms*, i.e. differentials of the policy instruments. Moreover, although it is reasonable to assume that the underlying maximisation problem is well formulated in the sense just defined, it is a characteristic of optimal tax problems that stronger requirements such as quasiconcavity of the objective function in the relevant tax instruments or convexity of the feasible set of these instruments do not hold.⁴ It follows that there may well be local optima that are not global optima, and if the economy is at one of these, and only local deviations are possible, then there will be no feasible welfare-improving tax reforms. Throughout the analysis, therefore, we will assume that the economy is not even at a local optimum with respect to the tax instruments with which we are concerned.

9.2 Tax reform policies

Tax reform⁵ is high on the policy agenda in many countries, for example in Australia, Canada, Germany, the US and the UK. A frequent theme is the proposal to bring down the highest marginal tax rates as well as to simplify the rate structure. Interestingly, as we discussed extensively in chapter 6, these countries have differing approaches to the taxation of couples. The US and Germany have joint taxation, under which primary and secondary earners in a household face the same marginal tax rates.⁶ In the remaining countries there is individual taxation, with marginal rates varying with individual income, so that primary and secondary earners in the same household may face different marginal rates. A question that arises, therefore, is whether the contemplated reforms could be expected to have the same effects on incentives and equity under the different systems, and a further question is that of how the reforms will affect the comparison of the systems in terms of their equity and efficiency. It has been strongly argued, for example, as we saw in chapter 7, that efficiency considerations support not just individual taxation, but individual taxation with a lower marginal rate schedule for women. Reducing higher marginal rates in an individual tax system would appear to

³ This follows from Weierstrass's theorem.

⁴ In this sense, the planner is at a real disadvantage in not being able to choose quantities directly, but having to work only through tax rates, since in the quantity space these further restrictions can reasonably be applied.

⁵ This section and the next draw heavily on Apps and Rees (1999a) and (1999b).

⁶ Though as Feldstein and Feenberg (1996) point out, working wives are in effect faced with higher marginal tax rates in the US system because they pay social security taxes far in excess of the actuarial value of the benefits they receive. The same is true of Germany.

move in the opposite direction to this, since it would bring the marginal rates faced by primary and secondary earners closer together.

As we have repeatedly stressed throughout this book, a comprehensive treatment of the issues concerning the taxation of couples involves the concepts of *production and trade within the household*. In the standard analysis a single consumer consumes an untaxed good, 'leisure', the time not allocated to market work, and the trades she makes with the rest of the economy, essentially exchange of labour time for market goods, are taxed. Income taxation can be viewed as a tax on trade between the market and household sectors, with all trades between individuals mediated through the market place. If, however, individuals form multi-person households, they may also allocate time for production for trade with other household members, and these trades are untaxed. The implicit wage incomes derived from production and trade with other household members cannot be taxed, and income taxation is then a *selective* tax on trade.

This has important implications for the analysis of reforms to the taxation of couples. First, as we showed in chapter 1, in many households partners specialise either in market or in household production. A tax reform may therefore induce an opposite welfare change for each partner by changing the intra-household terms of trade of household for market goods. Since we are concerned with the welfare of *individuals*, these effects need to be captured in a formal model. Second, the evidence presented in chapter 1 also shows that the degree of specialisation in domestic production of, typically, the female partner varies widely across households, even after controlling for demographic characteristics and wage rates. Households tend to divide into *traditional* households, where the female partner specialises almost entirely in domestic work, and *non-traditional* households, in which female market labour supply is significant, approaching that of the male partner in the upper half of the distribution. Since traditional households have higher levels of untaxed production and trade, a tax reform can have substantially different impacts on the different household types. This is reflected in the policy debate on the relative merits of joint versus individual taxation, extensively discussed in the previous two chapters. The difference between these systems only really matters when there is significant variation in female labour supply across households.

As we have seen, in optimal tax theory a central element is the idea that individuals differ in their innate productivities in market work, where these are measured by their market wages. It seems then only consistent to allow their productivities in domestic production also to differ. Moreover, this has the advantage of allowing us to *explain* the varying degrees of specialisation in domestic production as between traditional and non-traditional households in an analytically tractable way, without having to depart from the assumption of identical utilities, also a standard element of optimal tax models. Indeed, variation in domestic productivities is the *raison d'être* for introducing household production into the model. As we suggested in chapter 3, if productivities were identical across households then it would suffice for most purposes to regard household members as simply consuming 'leisure time'.

Here we begin by using one of the models presented in chapter 3 to analyse the effects of tax reform on individual welfares, taking account of the multi-person nature of households and the existence of two household types. There are two tax rates, one paid by primary earners, the other by secondary earners, and these rates may or may not be equal.⁷ Thus the initial situation could be one of either *selective taxation*, *progressive individual taxation* (both with unequal marginal tax rates) or *joint taxation* (equal marginal tax rates).

We consider two types of tax reform. First we take the case of a revenue-neutral reform that increases the rate paid by secondary earners and reduces that paid by primary earners. This could be thought of as a move toward a flat (or flatter) rate tax in a progressive individual system, such as that prevailing in Australia and the UK, or to a change in the rates in the selective individual tax system favoured by the economics literature. We analyse this tax reform first in an economy in which men and women may face different wage rates, but where there is no wage dispersion across households. This is to allow us to focus on the effects of introducing across-household variation in productivities in household work. We then extend the analysis of this tax reform to the case where wage rates vary across households. Finally, we go on to consider briefly the case of a revenue-neutral move away from a system of joint taxation toward selective individual taxation with a lower rate for secondary earners, in the model with across-household wage dispersion. In each case we characterise welfare-improving reforms. The approach can easily be extended to any other type of tax reform. To clarify as sharply as possible the implications of productivity variation we assume only two household types. This is a significant abstraction but allows us to focus on the central issues. The main purpose is to show how the welfare evaluation of the tax reforms in each case depends crucially on the way in which labour supplies vary with domestic productivities as well as with market wage rates.

9.2.1 Tax reform with no wage dispersion

First we review very briefly the household model we shall use (already covered in depth in chapter 3). Each member produces a domestic good with a constant returns technology. Because it is the variation in the *female's* allocation of time to household production across households that is important, we allow female domestic productivities to vary across households but assume those of men are all the same. This variation could be due to either or both differences in human capital among women and differences in household physical capital, which, since it is assumed exogenous, does not appear explicitly. Since preferences of all individuals are assumed identical, and women all receive the same market wage, as do all men, variations across households in female labour supply are due to this productivity variation.

⁷ As we pointed out in chapter 8, tax rates are not normally differentiated by gender. However, under a tax system where the marginal rates increase with income, women may well be relatively more concentrated in the lower tax bracket and men in the higher. It is then a useful simplification to take the extreme case here.

The gross market wage represents an individual's productivity in market production. There is a single market good, x , produced by both women and men. All men are equally productive, as are all women, but male and female productivities can differ. Since labour is the only input and there are constant returns to scale we can write the aggregate production function simply as

$$x = \sum_h n_h (w_m l_m^h + w_f l_f^h) \quad (9.1)$$

where w_i is the productivity parameter, l_i^h is the labour input while n_h is the number of households of type h . We assume $l_i^h > 0, \forall h, i$, though of course household types may differ in the precise amounts of labour they supply. The households differ in respect of f 's productivity in domestic production, as we shall shortly specify. We let x be the numeraire with price 1. The wage rates in this competitive economy are w_f and w_m . As usual in optimal tax models, wage rates are exogenous.

In a household, each partner produces a good which is partly consumed by her or himself and partly traded for the other goods. Since it is assumed that men in all households have the same productivity, their production functions can be written simply as

$$y_m^h = t_m^h, \quad h = 1, 2 \quad (9.2)$$

while those of women are

$$y_f^h = k_h t_f^h, \quad h = 1, 2 \quad (9.3)$$

where t_j^h is the time spent in producing the household good y_j^h by individual $j = f, m$ in household of type $h = 1, 2$. We assume $k_1 > k_2$, so type 1 (2) households will be referred to as the high (low) productivity households. The individuals have the strictly quasiconcave and increasing utility functions

$$u(x_i^h, y_{if}^h, y_{im}^h) \quad h = 1, 2 \quad i = f, m \quad (9.4)$$

where

$$\sum_{i=f,m} y_{ij}^h = y_j^h \quad j = f, m \quad (9.5)$$

and y_{ij}^h is the amount of the domestic good produced by individual j that is consumed by i .

The household's budget constraint is given by

$$\sum_{i=f,m} x_i^h = \alpha + \sum_{i=f,m} \beta_i w_i l_i^h \quad h = 1, 2 \quad (9.6)$$

where α is the lump sum payment made to each household and $1 - \beta_i$ is the marginal tax rate paid by $i = f, m$ under the given tax system.

Given the individual time constraints

$$l_i^h + t_i^h = 1 \quad i = f, m \quad (9.7)$$

where total available time is normalised to 1, if we define the implicit prices of the domestic goods as

$$p_f^h = \frac{\beta_f w_f}{k_h} \quad h = 1, 2 \quad (9.8)$$

$$p_m = \beta_m w_m \quad (9.9)$$

we can replace the separate constraints (9.2), (9.3), (9.6) and (9.7) with the single *full-income budget constraint*

$$\sum_{i=f,m} (x_i^h + p_f^h y_{if}^h + p_m y_{im}^h) = \alpha + \sum_i \beta_i w_i \equiv X \quad h = 1, 2. \quad (9.10)$$

Note also that only the price of the domestic good produced by f varies across household types, with of course more productive households having the lower price. These prices have a straightforward explanation: they are the marginal opportunity costs of the domestic outputs. All households have the same full income X , but high productivity households are clearly better off because they face a lower implicit price of the female domestic good, and so can enjoy higher total consumption at any given full income.

We regard the household as generating its consumption allocation by first choosing income shares s_i^h , such that

$$\begin{aligned} s_f^h &= \alpha_f^h + \beta_f w_f \\ s_m^h &= \alpha_m^h + \beta_m w_m. \\ \sum_{i=f,m} s_i^h &= X \quad h = 1, 2 \end{aligned} \quad (9.11)$$

where $\alpha_i^h \geq 0$ is a lump sum transfer to i , with $\sum_{i=f,m} \alpha_i^h = \alpha$. Thus each partner has a full income given by the value of the time endowment at his or her net market wage, plus a transfer. The sum of these transfers must equal the lump sum transfer from the tax system, α , but the intra-household transfer itself could be negative for one household member. This simply reflects the distributional choices of the household members. Then each solves the problem

$$\max u(x_i^h, y_{if}^h, y_{im}^h) \quad \text{s.t.} \quad x_i^h + p_f^h y_{if}^h + p_m y_{im}^h = s_i^h \quad i = f, m \quad (9.12)$$

to yield demand functions $x_i^h(p_f^h, p_m, s_i^h)$, $y_{ij}^h(p_f^h, p_m, s_i^h)$, together with the marginal utilities of individual income $\lambda_i^h(p_f^h, p_m, s_i^h)$, the shadow prices of the individual budget constraints.

The problem in (9.12) is a standard consumer problem and we can in the usual way define the indirect utility functions

$$v_i^h = v(p_f^h, p_m, s_i^h) \quad h = 1, 2 \quad i = f, m. \quad (9.13)$$

Then applying Roy's identity gives the partial derivatives

$$\frac{\partial v_f^h}{\partial p_f^h} = -\lambda_f^h (y_{ff}^h - y_f^h), \quad \frac{\partial v_m^h}{\partial p_f^h} = -\lambda_m^h y_{mf}^h \quad (9.14)$$

$$\frac{\partial v_f^h}{\partial p_m^h} = -\lambda_f^h y_{fm}^h, \quad \frac{\partial v_m^h}{\partial p_m^h} = -\lambda_m^h (y_{mm}^h - y_m^h) \quad (9.15)$$

$$\frac{\partial v_i^h}{\partial s_i^h} = \lambda_i^h \quad h = 1, 2 \quad i = f, m. \quad (9.16)$$

The derivatives in (9.14) and (9.15) reflect the significance of trade within the household. In general we expect an individual to consume less than the output of the good he or she produces, the resulting surplus being the amount traded with the other partner, at terms of trade given by the prices of the domestic goods. What is important here is that where partners specialise differentially in domestic production, changes in the terms of trade have greatly differing effects on the welfares of the individual household members. An increase (decrease) in the price of a good of which one is a net seller has a positive (negative) effect on one's utility that is proportionate to the amount sold.

From the solution to (9.12) we can also derive the market labour supplies as

$$l_f^h(p_f^h, p_m, s_f^h) = 1 - \frac{1}{k_h} \sum_{i=f,m} y_{if}^h(p_f^h, p_m, s_i^h) \quad (9.17)$$

$$l_m^h(p_f^h, p_m, s_m^h) = 1 - \sum_{i=f,m} y_{im}^h(p_f^h, p_m, s_i^h). \quad (9.18)$$

Thus an individual's labour supply is the difference between total time available and the amount of time required to meet total household demand for his or her domestic output.

As we saw in chapter 3, the effect of increasing domestic productivity on f 's market labour supply could be positive or negative. The intuitive reason for this ambiguity is that a change in the productivity has both a demand side and a supply side effect. The former results from the demand-increasing effect of a reduction in the implicit price of the domestic good following an increase in productivity, and is determined by the demand elasticity for this good. The latter results from the fact that with a higher productivity a given amount of domestic output can be supplied in less time. Whether females in high-productivity households supply more or less time to the market therefore depends, inter alia, on the relative strengths of these effects. This is important because *whether the households with higher female market labour supplies are also those with higher productivity and consequently higher utility possibilities is significant for the evaluation of the welfare effects of a given tax reform.*

Finally, we can note a perverse result of increasing the tax rate on women. Such an increase causes a larger shift in time allocation from the market to the household sector, the more productive the woman is in the market sector relative to the household. A tax increase causes a larger shift from more productive to less productive sectors

the greater the disparity in productivities. To see this, recall the expression for female labour supply in (9.17) and differentiate it with respect to the marginal tax rate⁸ to obtain

$$\frac{\partial l_f^h}{\partial \beta_f} = -\frac{\partial y_f^h}{\partial p_f^h} \frac{w_f}{k_h^2}. \quad (9.19)$$

This derivative is clearly larger the smaller is k_h , for two reasons. First, for a given demand derivative $\partial y_f^h / \partial p_f^h$, the lower the productivity the greater is the induced change in the domestic time allocation and so the larger the change in market labour supply. Second, the higher the ratio of productivities in the two sectors, w_f / k_h , the greater the effect of the tax change on the price of the domestic good and so the larger the change in market labour supply.

We now apply this model to an analysis of the welfare effects of tax reforms.

We characterise any tax reform as a *revenue-neutral* pair of infinitesimal changes in the marginal tax rates. We hold the lump sum α constant throughout. The government budget constraint is

$$\sum_{h=1,2} \sum_{i=f,m} n_h (1 - \beta_i) w_i l_i^h - n\alpha - R = 0 \quad (9.20)$$

where $R \geq 0$ is the aggregate revenue requirement, n_h is the number of households of type h and n is the total number of households. Revenue-neutrality requires

$$\sum_{h=1,2} \sum_{i=f,m} n_h \left[\sum_{j=f,m} (1 - \beta_j) w_j \frac{\partial l_j^h}{\partial \beta_i} - w_i l_i^h \right] d\beta_i = 0 \quad (9.21)$$

implying

$$d\beta_f = \mu d\beta_m \quad (9.22)$$

with

$$\mu \equiv -\frac{\sum_h n_h \left[\sum_{j=f,m} (1 - \beta_j) w_j \frac{\partial l_j^h}{\partial \beta_m} - w_m l_m^h \right]}{\sum_h n_h \left[\sum_{j=f,m} (1 - \beta_j) w_j \frac{\partial l_j^h}{\partial \beta_f} - w_f l_f^h \right]}. \quad (9.23)$$

It seems reasonable to assume that in both cases the incremental tax revenues resulting from the changes in labour supplies, the first terms in the numerator and denominator, do not exceed the respective tax bases, the second terms in the numerator and denominator, and so we will in everything that follows consider only the case where $\mu < 0$. A revenue-neutral tax reform that reduces one tax rate must increase the other.⁹

⁸ For simplicity we ignore the possible effect on f 's income share, which is unlikely to have much effect on the argument.

⁹ If both tax rates can be reduced in a revenue-neutral way costless welfare improvement is possible. We assume that, although not optimal, the tax system is sufficiently rational that no such opportunities are available.

Consider now the changes in individual utilities induced by a tax reform. To derive these we take the total differentials of the indirect utilities, using (9.14)–(9.16), and recalling that $p_f^h = \beta_f w_f / k_h$, and $p_m = \beta_m w_m$:

$$dv_f^h = \lambda_f^h \left[- (y_{ff}^h - y_f^h) \frac{w_f}{k_h} d\beta_f - y_{fm}^h w_m d\beta_m + w_f l_f^h d\beta_f + d\alpha_f^h \right] \quad (9.24)$$

$$dv_m^h = \lambda_m^h \left[- y_{mf}^h \frac{w_f}{k_h} d\beta_f - (y_{mm}^h - y_m^h) w_m d\beta_m + w_m l_m^h d\beta_m + d\alpha_m^h \right] \quad (9.25)$$

where $d\alpha_j^h = \sum_{i=f,m} (\partial \alpha_j^h / \partial \beta_i) d\beta_i$, $j = f, m$. We consider here the case of a revenue-neutral tax reform for which $d\beta_f < 0 < d\beta_m$, with female tax rates raised and male tax rates lowered. This corresponds, for example, to a move towards a flatter rate structure in a progressive individual tax system where women have lower wage incomes, or to a move toward joint taxation in a selective individual tax system in which women have lower rates. We interpret the expression in brackets in (9.24) and (9.25) term by term:

Intra-household terms of trade effects The first two terms in the brackets in (9.24) are negative, those in (9.25) are positive. The price of the household good women sell (men buy) has fallen, that of the good women buy (men sell) has risen, and so all women are worse off, all men are better off, in the absence of within-household redistribution. These effects will be greater:

- the larger is μ in absolute value, for a given $d\beta_m$;
- the larger is the volume of trade within the household;
- the lower is household productivity k_h , since this determines the size of the fall in the price of f 's domestic good. This implies that women's utility falls more in the low-productivity households, other things being equal. Since these are also the households with lower utilities, this could be thought of as a regressive effect of the tax reform.

Wage-income effects The term $w_f l_f^h d\beta_f$ will be negative and $w_m l_m^h d\beta_m$ will be positive, so this will be a further source of increase in male and decrease in female utilities. That is, before intra-household transfers, female full income falls and male full income rises. These effects will be greater the larger is μ in absolute value, given $d\beta_m$.

Household distributional effects The changes in net wage rates will cause changes in the intra-household income distribution. Without an explicit form for the sharing rule, or a specific household welfare function, we cannot characterise these even qualitatively. However, we can explore the conditions under which intra-household transfers could or could not at least potentially compensate women for their utility losses and still leave their partners with a net welfare gain.

We obtain a very intuitive result on the conditions under which both partners in a household can potentially be better off after the tax reform.

Proposition 1. *The male partner in the household can fully compensate the female for her utility loss and still be better off if and only if*

$$\sum_{i=f,m} w_i l_i^h d\beta_i > 0 \quad (9.26)$$

as a result of the tax reform.

Proof: A woman can be just compensated for her utility loss iff

$$d\alpha_f = (y_{ff}^h - y_f^h) \frac{w_f}{k_h} d\beta_f + y_{fm}^h w_m d\beta_m - w_f l_f^h d\beta_f. \quad (9.27)$$

Since $\sum_{i=f,m} d\alpha_i = 0$, substituting for $d\alpha_m = -d\alpha_f$ in (9.25) gives

$$dv_m^h = \lambda_m^h [w_m l_m^h d\beta_m + w_f l_f^h d\beta_f] \quad (9.28)$$

which gives the result. $\square$

The intuition underlying this result is straightforward, especially if we recall the analogy between the household and a small economy trading with the rest of the world at fixed prices, here the net of tax wage rates. A tax reform changes the household/economy's terms of trade, leading to changes in production, consumption and welfares within it, but it can only be better off in the aggregate if the change is equivalent to an increase in its net income at 'world prices'.

This result can be expressed in elasticity terms, also taking account of the revenue-neutrality requirement, as follows. Condition (9.26) holds iff

$$(w_m l_m^h + \mu w_f l_f^h) d\beta_m > 0 \quad (9.29)$$

implying

$$\frac{w_m l_m^h}{w_f l_f^h} > -\mu = \frac{\sum_h n_h w_m l_m^h \left[1 - \frac{\sum_{j=f,m} (1 - \beta_j) w_j}{w_m l_m^h} \frac{\partial l_j^h}{\partial \beta_m} \right]}{\sum_h n_h w_f l_f^h \left[1 - \frac{\sum_{j=f,m} (1 - \beta_j) w_j}{w_f l_f^h} \frac{\partial l_j^h}{\partial \beta_f} \right]}. \quad (9.30)$$

We can write the terms in square brackets in this condition as

$$\delta_i^h = 1 - \frac{(1 - \beta_j)}{\beta_i} \frac{w_j l_j^h}{w_i l_i^h} \sigma_{ij}^h - \frac{(1 - \beta_i)}{\beta_i} \sigma_i^h \quad h = 1, 2 \quad i, j = f, m \quad i \neq j \quad (9.31)$$

where σ_i^h is i 's (uncompensated) labour supply elasticity with respect to the net wage, and σ_{ij}^h is the (uncompensated) elasticity of j 's labour supply with respect to i 's net wage. Noting further that the wage rates do not depend on h , we can finally rewrite (9.29) as

$$\frac{l_m^h}{l_f^h} > \frac{\sum_h n_h l_m^h \delta_m^h}{\sum_h n_h l_f^h \delta_f^h}. \quad (9.32)$$

This way of expressing the condition allows derivation of some interesting special cases of the general result.

Proposition 2. *If the δ_i^h are all equal, then in the household whose ratio of male to female labour supply is above the ratio of average male to average female labour supply – the traditional household – the female can at least be compensated, while in the non-traditional household she cannot be.*

Proof: Just divide the δ s out of (9.32), define $\eta = n_1/n$ and let $h = 1$ denote the traditional household. Then we have to prove

$$\frac{l_m^1}{l_f^1} > \frac{l_m^2}{l_f^2} \Leftrightarrow \frac{l_m^1}{l_f^1} > \frac{\eta l_m^1 + (1 - \eta) l_m^2}{\eta l_f^1 + (1 - \eta) l_f^2} > \frac{l_m^2}{l_f^2}$$

which is straightforward to show by routine algebra. $\square$

The intuitive reason for this result is simply that in the traditional household the gain in after-tax male wage income more than compensates for the loss in after-tax female income because the female partner is supplying relatively little time to the market. By the same token, in households with a ratio of male to female labour supply below the ratio of the averages, the non-traditional households, the male cannot compensate the female without making himself worse off.

The empirical evidence on labour supply elasticities¹⁰ suggests, however, that $\delta_m^h > \delta_f^h$, implying that it is possible that in *no* household could the woman be compensated, if the ratio of male to female labour supplies in traditional households is sufficiently close to the ratio of the averages. This evidence suggests the following stylised facts:

- $\sigma_f^h > \sigma_m^h$ for each household type: women tend to have higher labour supply elasticities than men, especially in traditional households.
- $\sigma_{fm}^h \approx 0$. Variation in the after-tax wage rate of women has a negligible effect on male labour supply.
- $\sigma_{mf}^h < 0$. Variation in the male after-tax wage has a small negative effect on female labour supply.

To explore the implications of these stylised facts, suppose $\delta_m^h = \delta_m > \delta_f = \delta_f$. This says that labour supply responses to wage rate changes are essentially the same within gender, but differ between genders in the way suggested by the empirical evidence. Also, assume for simplicity, but quite realistically, that $l_m^1 = l_m^2$ so that there is no inter-household variation in male labour supplies. Then we have:

Proposition 3. *If $\delta_m^h = \delta_m > \delta_f^h = \delta_f > 0$, $h = 1, 2$ and $l_m^1 = l_m^2$, in no household can women be compensated for the tax change iff*

$$\frac{\delta_m}{\delta_f} > \frac{\eta l_f^1 + (1 - \eta) l_f^2}{l_f^1}.$$

¹⁰ For a survey see Heckman (1993).

Proof: We have to show that (9.32) cannot be satisfied for $h = 1$ given this condition. But this follows simply by dividing out l_m in (9.32), dividing through by n , recalling the definition of η and rearranging. $\square$

The result says that if the difference between male and female labour supply elasticities is sufficiently large relative to the across-household variation in female labour supply, then in no household is the gain in net wage income to men large enough to allow them to compensate women for the loss in net wage income. Intuitively, the more elastic is female labour supply relative to that of males, the greater the net loss of tax revenue induced by the changes in labour supplies, and so the greater must be the increase in the female tax rate relative to the decrease in the male tax rate. This makes it harder for the male within a household to compensate the female.

From the empirical standpoint the most relevant case to take seems to be that in which traditional households are better off and non-traditional households worse off as a result of the tax reform. The question then becomes one of evaluating this from the point of view of social welfare. The model suggests that the key factor in this evaluation is the nature of the relationship between the productivity parameter k_h and female labour supply, as analysed in chapter 3. If female labour supply falls with increasing productivity, corresponding to the case of high elasticity of demand for f 's domestic good, then traditional households will be those with the higher utilities of household members. It follows that the type of tax reform just considered is unambiguously regressive in its impact, since it is increasing the utilities of those who are already better off.

On the other hand, if female labour supply increases with domestic productivity, because the demand effect is weaker than the 'time-releasing' effect of greater efficiency, then non-traditional households are those with higher utilities, and the tax reform considered improves the welfare of the less well-off individuals, subject always to the proviso that the women in the traditional households are appropriately compensated for the fall in their full incomes by transfers from their partners.

This can be put more generally and precisely by introducing the social welfare function

$$W = W(v_f^1, \dots, v_m^1, \dots, v_f^2, \dots, v_m^2, \dots) \quad (9.33)$$

and taking its total differential

$$dW = \sum_h n_h \sum_i W_i^h dv_i^h \quad (9.34)$$

with the dv_i^h as given in (9.25) and (9.26).

Note that the social welfare function is defined, as it should be, on the utilities of all the *individuals* in the economy, and we assume the individual welfare weights $W_i^h > 0$. The expression for the change in social welfare involves both *within*-household and *across*-household distributional effects. There is no reason in general why the implicit welfare weights applied within a household would correspond to those that would be desired by

a ‘social planner’ – in chapter 7 earlier we called this the issue of *dissonance*. However, here we focus on the across-household distributional effects and so we assume that the distributional preferences of the household are precisely those the planner would wish. In other words given the tax reform contemplated here, men compensate their wives for the fall in after-tax incomes precisely to the extent that the planner would wish them to.

The marginal social utilities of income $W_i^h \lambda_i^h$ in that case become the same for individuals in a given household and vary only across household types, and we now denote these by ω_h . Then we have that the condition for social welfare to increase as a result of the tax reform becomes simply

$$dW = \sum_h n_h \omega_h (w_m l_m^h + \mu w_f l_f^h) d\beta_m > 0 \quad (9.35)$$

Since type 1 households have higher productivity and so are better off we expect that $\omega_2 > \omega_1$. If female labour supplies in traditional households are sufficiently low, then the term in brackets is positive for traditional and negative for non-traditional households. The condition is more likely to be satisfied if the term for traditional households has the higher welfare weight. This requires that domestic productivity and female labour supply are positively associated across households. That is, women who specialise in household production are less productive in that activity than their counterparts in non-traditional households. In that case household wage income is positively associated with the utilities of household members. On the other hand, in the case where domestic productivity and female labour supply are inversely related, higher wage income does not indicate higher utilities of household members.

Note that the strength of the relationship between domestic productivity and female labour supply is also important. Suppose, for example, that there is only a very small difference in the k_h values, but this difference induces a large increase in female labour supply. Then the welfare weights ω_h will be very close in value since the utility difference between the household types is small, and the condition in (9.35) is unlikely to be satisfied, since the negative value for non-traditional households will outweigh the positive value for traditional households. In that case the reverse tax reform – widening the gap between marginal rates – would be welfare-improving.

Finally let us suppose that the distribution of income *across households* is regarded as optimal, so that $\omega_h = 1$, $h = 1, 2$, that is, we consider only the pure efficiency effects of the tax reform. Then we have

Proposition 4. *If male and female labour supply elasticities are such that $\delta_m^h = \delta_m > \delta_f = \delta_f^h (> 0)$, then a revenue neutral tax reform with $d\beta_f < 0 < d\beta_m$ reduces social welfare.*

Proof: Given $\omega_h = 1$, $h = 1, 2$, condition (9.35) implies that welfare is reduced if

$$dW = (\mu w_f l_f + w_m l_m) d\beta_m < 0$$

where $l_i = \sum_h n_h l_i^h$, $i = f, m$. Thus for $d\beta_m > 0$ we will have $dW < 0$ iff

$$-\mu > \frac{w_m l_m}{w_f l_f} \quad (9.36)$$

that is if

$$-\mu = \frac{\sum_h n_h w_m l_m^h \delta_m}{\sum_h n_h w_f l_f^h \delta_f} > \frac{w_m l_m}{w_f l_f} \quad (9.37)$$

which is equivalent to

$$\frac{\delta_m}{\delta_f} > 1. \quad (9.38)$$

□

The condition in this proposition is of course a simple Ramsey-type condition on labour supply elasticities, and the result is fully to be expected when we are only concerned with allocative efficiency.

9.2.2 Tax reform with wage dispersion

We now extend the model of the previous section by allowing wage dispersion across households. We assume there are two wage rates, w_1 and w_2 , with $w_1 > w_2$, and these again represent productivities in market production. The rest of the household model is as before. Both men and women can be of either market productivity type. There are then in principle eight possible household types. So much diversity is unnecessary for our purposes and so we shall make the assumption of *perfect assortative matching*. A male and female in the same household have the same wage rate. There are then four possible household types each defined by a pair (w_i, k_j) , $i, j = 1, 2$. Clearly the household with (w_1, k_1) will have the highest utility possibilities (recall that $k_1 > k_2$) and that with (w_2, k_2) the lowest. We assume that an increase in the market wage has at least as large an effect on household utility possibilities as an increase in domestic productivity and so the household with (w_1, k_2) is no worse off and may be better off than that with (w_2, k_1) .

To explore the implications of the type of tax reform that was considered in the previous section we first number the household types as follows. Let:

$h = 1$ for (w_1, k_1)

$h = 2$ for (w_1, k_2)

$h = 3$ for (w_2, k_1)

$h = 4$ for (w_2, k_2)

and we again take a revenue neutral tax reform for which $d\beta_f < 0 < d\beta_m$. The assumption that women pay one tax rate and men another now requires either that we have a selective individual tax system, or a progressive individual tax system in which both

high- and low-wage men fall in one tax bracket and high- and low-wage women in another. Though restrictive, this is a useful simplification for present purposes.

The first three propositions of the preceding section can be easily extended with just a change in notation, so we proceed immediately to introduce a social welfare function defined on the utilities of the individuals in all four household types. We then obtain as the condition for a welfare improvement the counterpart of (9.35),

$$dW = \sum_{h=1}^4 n_h \omega_h w_h (l_m^h + \mu l_f^h) d\beta_m > 0 \quad (9.39)$$

where ω_h is again the welfare weight for household h , $w_h = w_1$ for $h = 1, 2$ and $w_h = w_2$ for $h = 2, 3$. In the pure efficiency case, with $\omega_h = 1$, it is straightforward to prove the counterpart of proposition 4: if within-gender elasticity differences are small enough to be ignored, so that $\delta_m^h = \delta_m$, $\delta_f^h = \delta_f$, $h = 1, \dots, 4$, while the stylised facts described in the previous section apply and $\delta_m > \delta_f > 0$, then a revenue-neutral tax reform raising tax rates on women and reducing those on men is welfare-reducing.

Maintaining this assumption on elasticities but returning to the case in which the welfare weights ω_h vary across households, define

$$\Gamma_h \equiv w_h (l_m^h + \mu l_f^h). \quad (9.40)$$

Γ_h may be positive or negative, since $\mu < 0$, and will be lower for non-traditional households with higher l_f^h and higher for traditional households with lower l_f^h , other things (i.e. w_h, l_m^h) being equal. Now note that we can write the welfare criterion in (9.39) as

$$\frac{1}{n} \sum_{h=1}^4 n_h \omega_h \Gamma_h = E[\omega_h \Gamma_h] > 0 \quad (9.41)$$

where $E[\cdot]$ is the expectations operator. It follows that we can write this condition as

$$Cov(\omega_h, \Gamma_h) + E[\omega_h]E[\Gamma_h] > 0. \quad (9.42)$$

We then have

Proposition 5. *If male and female labour supply elasticities are such that $\delta_m^h = \delta_m > \delta_f = \delta_f^h > 0$, then a necessary but not sufficient condition for a revenue-neutral tax reform with $d\beta_f < 0 < d\beta_m$ to increase social welfare is that $Cov(\omega_h, \Gamma_h) > 0$.*

Proof: $E[\omega_h] > 0$, and so we just have to show that $E[\Gamma_h] < 0$ on the given conditions. That is, we have to show

$$\sum_h \eta_h w_h (l_m^h + \mu l_f^h) < 0 \quad (9.43)$$

where $\eta_h \equiv n_h/n$. So, we have to show that

$$-\mu = \frac{\sum_h \eta_h w_h l_m^h \delta_m}{\sum_h \eta_h w_h l_f^h \delta_f} > \frac{\sum_h \eta_h w_h l_m^h}{\sum_h \eta_h w_h l_f^h} \quad (9.44)$$

which is equivalent to

$$\frac{\delta_m}{\delta_f} > 1 \quad (9.45)$$

as required. $\square$

For the covariance between ω_h and Γ_h to be positive, we essentially require that households with lower utility possibilities and therefore higher ω_h -values have sufficiently lower female labour supplies to allow Γ_h to rise despite the effect of the falling wage rate (and therefore possibly falling male labour supply) in tending to reduce Γ_h . Female market labour supply must fall sufficiently as the wage rate and/or domestic productivity fall. Thus traditional households must be the low-wage low-productivity households. If the converse is the case, then the postulated tax reform cannot be welfare-improving. That is, the condition in (9.39) could not be satisfied if women who specialise in household production, i.e. women in traditional households, are more productive in that activity than the women in non-traditional households.

To emphasise that this covariance condition is necessary but not sufficient let us take a case in which it is satisfied but the tax reform is not welfare-improving. Here we wish to explore the following intuition. Suppose that at each wage rate there is only a relatively small difference in domestic productivities, though this is enough to cause a significant increase in female labour supplies in the higher-productivity households. It follows that non-traditional households are only a little better off than traditional households at the same wage rate, but pay much more in tax. In that case the kind of tax reform considered here should be welfare-decreasing. A small decrease in the tax rate on men will require a large increase in the tax rate on women, giving a net decrease in overall labour supply. Traditional households are a little better off, non-traditional households a lot worse off, so, even though the welfare weights on the traditional households are higher, overall there could be a welfare decrease.

We show for a special case of this model, which can, however, easily be generalised, that this intuition is correct provided that the excess of female over male labour supply elasticities is sufficiently large relative to the differences in welfare weights between traditional and non-traditional households. Thus assume

- the traditional households, 2 and 4 (lower-productivity households), have zero female labour supplies;¹¹
- the non-traditional households 1 and 3 (higher-productivity households) have equal female labour supplies $l_f > 0$;

¹¹ Their labour supply elasticities are defined to be zero – a change in net-of-tax wage rates does not increase their labour supply.

- all male labour supplies are equal, at $l_m > 0$;
- the δ_m^h are all equal, as are the δ_f^h (defined only for non-traditional households) and their ratio $\delta_m/\delta_f = \delta > 1$;
- the welfare weights satisfy:

$$\omega_2 = \omega_1 + \epsilon_2$$

$$\omega_4 = \omega_3 + \epsilon_4$$

for $\epsilon_2, \epsilon_4 > 0$.

Thus traditional households have both high and low wage rates but lower domestic productivities. In general the results will depend on the relative number of households of each type, and so to clarify these effects we consider two cases.

Case A: Within a wage group there is the same number of households of each productivity type, so that $n_1 = n_2, n_3 = n_4$, but the number of high-wage households may differ from that of low-wage households, so that $n_3 = \nu n_1, \nu > 0$.

Case B: Within a productivity group there is the same number of households with each wage rate, so that $n_1 = n_3, n_2 = n_4$, but the number of households of each productivity type may differ, so that $n_2 = \nu n_1, \nu > 0$.

Then we have

Proposition 6. *Given these assumptions, the tax reform considered here is welfare-decreasing if and only if in case A:*

$$\delta > 1 + \frac{\epsilon_2 w_1 + \epsilon_4 \nu w_2}{\omega_1 w_1 + \omega_3 \nu w_2} \quad (9.46)$$

and in case B:

$$\delta > 1 + \frac{\nu}{(1 + \nu)} \frac{(\epsilon_2 w_1 + \epsilon_4 w_2)}{(\omega_1 w_1 + \omega_3 w_2)}. \quad (9.47)$$

Proof: In case A on the given assumptions we have

$$\mu = -2\delta \frac{l_m}{l_f}. \quad (9.48)$$

Then inserting this in (9.39) (with inequality reversed) and rearranging, recalling that $d\beta_m > 0$, gives the result.

In case B, we now have

$$\mu = -(1 + \nu)\delta \frac{l_m}{l_f} \quad (9.49)$$

and so inserting this in (9.39) and rearranging gives the corresponding result. $\square$

This result says that if female labour supply elasticities are sufficiently greater than male, and the welfare weights on traditional and non-traditional households at a given wage rate are sufficiently close together, then reducing the tax rate on men and increasing

that on women will be welfare-decreasing. It makes traditional households a little better off but non-traditional households so much worse off that this offsets the somewhat higher welfare weights given to the former and so is welfare-decreasing overall. These results do depend to some extent on the relative numbers of households of each type. If in (9.47) ν , the ratio of low-wage to high-wage households, increases, the inequality becomes less likely to be satisfied if and only if $\varepsilon_4/\omega_3 > \varepsilon_2/\omega_1$, but there is no particular reason for this to be the case. If in (9.48) ν , now the ratio of low- to high-productivity households, increases, then the condition becomes less likely to be satisfied, since the ratio of traditional to non-traditional households is increasing as a result.

9.3 Tax reform, joint taxation and the tax unit

In Chapter 7, we examined at some length the proposition that the optimal income taxation of couples implies that joint taxation is likely to be inferior to selective taxation, under which men and women are taxed at different marginal rates, with that on women's income being lower. Piggott and Whalley (1996) challenged what they call this 'conventional wisdom', and argued instead for joint taxation.

They begin by pointing out, quite correctly, that the literature on which this wisdom is based, in particular Boskin and Sheshinski (1983), ignores the existence of household production. Taking account of household production leads to the proposition that, given a perfectly competitive economy, with no taxation of market wage income the marginal rates of technical substitution between primary and secondary labour in market and in household production respectively would be equated to each other, since they are both equal to the ratio of the wage rates, while the marginal rate of substitution between the market and household good would be equal to the marginal rate of transformation between them, thus achieving Pareto efficiency.

The introduction of income taxation leads first to a distortion of the relation between the marginal rate of substitution between household and market goods and their marginal rate of transformation. This distortion is the focus of concern in the literature,¹² and, as we saw in chapter 7, with differing elasticities of male and female labour supplies, standard Ramsey rule considerations would argue for taxing primary and secondary workers at different marginal rates. However, income taxation also introduces a distortion between the marginal rates of technical substitution in household and market production, since the former will be equal to the ratio of net-of-tax wages of the individuals, while the latter is given by the ratio of gross wages – recall the aggregate production function in (9.1). This distortion would be eliminated if the marginal tax rates were equalised, since then the ratios of net and gross wage rates would be equal.

Thus, Piggott and Whalley argue, moving from individual to joint taxation, which equalises the marginal tax rates, could on balance increase welfare, because the welfare gain from the elimination of the distortion in production could outweigh the loss

¹² In which 'household production' is to be read as 'leisure'.

from increasing the consumption distortion. In their paper this general argument is substantiated by using a computable general equilibrium (CGE) model to show that for specific functional forms and particular regions of the parameter space, a movement from individual to joint taxation does indeed generate net welfare gains.

Now, we have just demonstrated in the context of tax reform that the conventional wisdom is essentially correct, even in the presence of domestic production. It could be argued, however, that the model we have used to do this, involving separate goods produced by each household member with a simple proportional production function, cannot capture the distortion in the marginal technical rate of substitution in household production, since this requires us to model the inputs of the household members as substitutes in producing a single household good. We now show, by taking the alternative model of household production,¹³ in which the household production function contains both time inputs as arguments, that the previous conclusions continue to hold.

9.3.1 *Substitute time inputs*

The economy now consists of n identical households,¹⁴ each of which produces a domestic good y according to a standard constant returns to scale production function

$$y = h(t_f, t_m). \quad (9.50)$$

The domestic good is consumed by partner i in the amount y_i . The utility functions, time and budget constraints are all as before. For simplicity we assume both household members have a positive market labour supply.

In determining their domestic labour supplies the individuals minimise total cost. The marginal cost of the domestic good to the household is $c(\beta_f w_f, \beta_m w_m)$, which, because of the constant returns assumption, is independent of output. Its derivatives c_i are the amounts of the respective time inputs required per unit of output. The implicit price of the domestic good at the household equilibrium is set equal to this marginal cost

$$p = c(\beta_f w_f, \beta_m w_m) \quad (9.51)$$

implying

$$\frac{\partial p}{\partial \beta_i w_i} = c_i. \quad (9.52)$$

Given income shares s_i , each individual then solves the problem

$$\max u^i(x_i, y_i) \quad \text{s.t.} \quad x_i + p y_i = s_i \quad i = f, m \quad (9.53)$$

¹³ Model 2 in chapter 3.

¹⁴ Since the proposition concerns pure efficiency and not distribution, we can move to a 'representative household' model for purposes of the present discussion, something which in general we would discourage, because distributional effects should not be ignored in tax analysis.

to yield demand functions $x_i(p, s_i)$, $y(p, s_i) = \sum_{i=f,m} y_i(p, s_i)$, the individual marginal utilities of income λ_i and the indirect utility functions $v_i(p, s_i)$ with partial derivatives (obtained by straightforward application of Roy's identity)

$$\frac{\partial v_i}{\partial p} = -\lambda_i y_i \quad (9.54)$$

$$\frac{\partial v_i}{\partial s_i} = \lambda_i \quad i = 1, 2. \quad (9.55)$$

Note that the domestic time input requirement is

$$t_i = c_i(\beta_f w_f, \beta_m w_m) y(p, s_i) \quad i = f, m \quad (9.56)$$

and market labour supply is $l_i = 1 - t_i$.

We again characterise the tax reform as a *revenue-neutral* pair of infinitesimal changes in the marginal tax rates. We hold the lump sum α constant throughout. The government budget constraint is

$$n \sum_{i=f,m} (1 - \beta_i) w_i l_i - n\alpha - R = 0 \quad (9.57)$$

where $R \geq 0$ is the revenue requirement. Revenue neutrality requires

$$\sum_{i=f,m} \left[\sum_{j=f,m} (1 - \beta_j) w_j \frac{\partial l_j}{\partial \beta_i} - w_i l_i \right] d\beta_i = 0 \quad (9.58)$$

implying

$$d\beta_m = \mu d\beta_f \quad (9.59)$$

with

$$\mu \equiv - \frac{\left[\sum_{j=f,m} (1 - \beta_j) w_j \frac{\partial l_j}{\partial \beta_f} - w_f l_f \right]}{\left[\sum_{j=f,m} (1 - \beta_j) w_j \frac{\partial l_j}{\partial \beta_m} - w_m l_m \right]} \quad (9.60)$$

just as in the previous section. Again we assume $\mu < 0$.

Consider now the welfare effects of a tax reform. There is a social welfare function $W(v_f, v_m)$ and we can write its total differential as

$$dW = \sum_i W_i dv_i = \sum_i W_i \lambda_i [-y_i(c_f w_f d\beta_f + c_m w_m d\beta_m) + ds_i] \quad (9.61)$$

with $W_i > 0$ the welfare weight on the i th individual.

Since the Piggott-Whalley paper was concerned solely with efficiency we can simply assume that the distributional preferences of the planner and of households and the resulting distribution of income among individuals and households is such that $W_i \lambda_i = 1$, $i = f, m$. Now summing through (9.61) gives exactly the same result as we derived in the previous model:

The tax reform is welfare improving, given $W_i \lambda_i = 1$, $i = f, m$, iff

$$dW = \sum_i w_i l_i d\beta_i > 0 \quad (9.62)$$

and we obtain this simply by summing through (9.61), recalling (9.54), (9.55) and that

$$\sum_i ds_i = \sum_i w_i d\beta_i.$$

The intuition underlying this result is just as before. The tax reform just changes the household's terms of trade. It will cause reallocations of production and consumption within the household, but we are ignoring the distributional effects of these. The household *overall* can only be better off if the value of its income at 'world prices' increases, and this is precisely condition (9.62). The remainder of the discussion then follows just as before.

Thus we see that the previous result is very robust and does not depend on the specific model of household production used. The Piggott–Whalley reasoning fails to take account of the basic message of the theory of the second-best. In general, a second-best optimum will involve a trade-off between the two distortions, rather than the complete elimination of one of them. Thus joint taxation is unlikely to be a second-best optimum.

It could of course be argued that the achievement of optimal taxes is in any case an impossibility and we should instead be concerned with looking simply for welfare improvements to existing policies – this is the motivation of the tax reform approach adopted in this chapter. However, the fact remains that if joint taxation is not a (local) second-best optimum, it will always be possible to find (local) welfare-improving departures from such a position, which must therefore imply individual taxation. The Piggott–Whalley computations show that a move (not necessarily local) from an existing tax structure with unequal marginal tax rates for primary and secondary workers, to joint taxation, which equalises these tax rates, *could* increase welfare – all we have to do is find a sufficiently bad initial allocation with unequal tax rates, since the proposition does not say that all allocations with unequal tax rates must be better than every allocation with equal tax rates. Nevertheless, once we have equalised tax rates, it must then be possible to achieve a further welfare improvement by departing in some direction from the joint taxation structure reached, as long as it is not a local second-best optimum.

9.4 Conclusions

In this chapter we have considered, in quite a general setting, some questions of tax reform from the point of view of its effects on the welfare of males and females in households with significantly differing female labour supplies. Once again we see how important the relationship between domestic productivity and female labour supply is in determining the distributional effects of tax changes. The analysis also tried to show, however, that even where the correlation between household productivity and female

labour supply is positive, the efficiency effects of relatively heavy taxes on working women, or, more generally, lower-paid secondary earners,¹⁵ can be sufficiently high that social welfare is reduced by such a policy. This repeats, in a different setting and from a different perspective, the lesson from the optimal tax analysis: the taxation of secondary earners is an insufficiently effective instrument for redistributing income across households to justify using it in this way, and incurring the efficiency losses that result.

¹⁵ Of the kinds identified in chapter 6.

Bibliography

- Abreu, D. (1986). 'Extremal equilibria of oligopolistic supergames'. *Journal of Economic Theory*, 39, 191–235.
- Abreu, D., D. Pearce and E. Stacchetti (1986). 'Optimal cartel equilibria with incomplete monitoring'. *Journal of Economic Theory*, 39, 251–69.
- Aguiar, M. and E. Hurst (2005). 'Consumption vs expenditure'. *Journal of Political Economy*, 113(5), 919–48.
- Alesina, A., A. Ichino and L. Karabarbounis (2007). 'Gender based taxation and the division of family chores'. Harvard University discussion paper.
- Altug, S. and R. A. Miller (1998). 'The effect of work experience on female wages and labor supply'. *Review of Economic Studies*, 56, 45–85.
- Ando, A. and F. Modigliani (1963). 'The "life cycle" hypothesis of saving: aggregate implications and tests'. *American Economic Review*, 53, 55–84.
- Angrist, J. D. and A. B. Krueger (1992). 'The effect of age at school entry on educational attainment: an application of instrumental variables with moments from two samples'. *Journal of the American Statistical Association*, 87, 328–36.
- Apps, P. F. (1981). *A Theory of Inequality and Taxation*. Cambridge: Cambridge University Press.
- (1982). 'Institutional inequality and tax incidence'. *Journal of Public Economics*, 18, 217–42.
- (2006). 'Female labour supply, taxation, and the new discrimination'. Presidential address, XX Annual Conference of the European Society for Population Economics, Verona, 22–24 June.
- Apps, P. F. and G. S. Jones (1986). 'Selective taxation of couples'. *Journal of Economics*, Supp. 5, 63–74.
- Apps, P. F. and R. Rees (1988). 'Taxation and the household'. *Journal of Public Economics*, 35, 355–69.
- (1996). 'Labour supply, household production and intra-family welfare distribution'. *Journal of Public Economics*, 60, 199–219.
- (1997a). 'The optimal taxation of couples'. CES working paper 145, University of Munich.
- (1997b). 'Collective labor supply and household production'. *Journal of Political Economy*, 105, 178–90.
- (1999a). 'Joint vs individual taxation in models of household production'. *Journal of Political Economy*, 107, 178–90.
- (1999b). 'On the taxation of trade within and between households'. *Journal of Public Economics*, 75, 241–63.
- (2001). 'Household saving and full consumption over the life cycle'. IZA discussion paper 280, Bonn.
- (2002). 'Household production, full consumption and the costs of children'. *Labour Economics*, 8, 621–48.

- (2003). 'Life cycle time allocation and saving in an imperfect capital market'. NBER Summer Institute Session on Aggregate Implications of Microeconomic Consumption Behavior, Boston, 21–25 July.
- (2004). 'Fertility, taxation and family policy'. *Scandinavian Journal of Economics*, 106(4), 745–65.
- (2005). 'Gender, time use and public policy over the life cycle'. *Oxford Review of Economic Policy*, 21(3), 439–61.
- (2006). 'Two extensions to the theory of optimal income taxation', mimeo.
- (2007a). 'Cooperative household models'. IZA discussion paper 3127.
- (2007b). 'The taxation of couples'. IZA discussion paper 2910.
- Apps, P. F. and E. Savage (1989). 'Labour supply, welfare rankings and the measurement of inequality'. *Journal of Public Economics*, 39, 335–64.
- Arellano, M. and C. Meghir (1992). 'Female labour supply and on-the-job search: an empirical model using complementary data sets'. *Review of Economic Studies*, 59, 537–59.
- Armstrong, M. and J.-C. Rochet (1999). 'Multi-dimensional screening: a user's guide'. *European Economic Review*, 43, 959–79.
- Arrufat, J. L. and A. Zabalza (1986). 'Female labor supply with taxation, random preferences, and optimization errors'. *Econometrica*, 54, 47–63.
- Ashworth, J. S. and D. T. Ulph (1981). 'Household models', in C. V. Brown (ed.), *Taxation and Labour Supply*. London: Allen and Unwin, 117–33.
- Atkinson, A. B. and J. E. Stiglitz (1988). *Lectures on Public Economics*. Maidenhead: McGraw-Hill.
- Attanasio, O. P. and J. Banks (1998). 'Trends in household saving: a tale of two countries'. Working paper series W98/15, London: Institute of Fiscal Studies.
- Attanasio, O. P., J. Banks, C. Meghir, and G. Weber (1999). 'Humps and bumps in lifetime consumption'. *Journal of Business and Economic Statistics*, 17, 22–35.
- Attanasio, O. P. and M. Browning (1999). 'Consumption over the life cycle and over the business cycle'. *American Economic Review*, 85, 1118–37.
- Attanasio, O. P., H. Low and V. Sanchez-Marcos (2003). 'Explaining changes in female labour supply in a life-cycle model'. Mimeo.
- Australian Bureau of Statistics (2001). *Government Benefits, Taxes and Household Income*, Cat. No. 6537.0. Canberra: ABS.
- Balestrino, A. (2001). 'On the optimal fiscal treatment of family size'. *FinanzArchiv*, 58(2), 140–57.
- Balestrino, A., A. Cigno and A. Pettini (2002). 'Endogenous fertility and the design of family taxation'. *International Tax and Public Finance*, 9, 175–93.
- Banks, J. and R. Blundell (1994). 'Household saving behavior in the United Kingdom', in J. Poterba (ed.), *International Comparisons of Household Saving*. Chicago: Chicago University Press, 169–206.
- Banks, J., R. Blundell and S. Tanner (1998). 'Is there a retirement saving puzzle?' *American Economic Review*, 88 (4), 769–88.
- Basu, K. (2006). 'Gender and say: a model of household behavior with endogenously-determined balance of power'. *Economic Journal*, 116, 558–80.
- Baxter, M. and U. J. Jermann (1999). 'Household production and the excess sensitivity of consumption to current income'. *American Economic Review*, 89 (4), 902–20.
- Becker, G. S. (1965). 'A theory of the allocation of time'. *Economic Journal*, 75, 493–517.
- (1973). 'A theory of marriage (Part I)'. *Journal of Political Economy*, 81, 813–46.
- (1974). 'A theory of marriage (Part II)'. *Journal of Political Economy*, 82, 511–26.
- (1981). *A Treatise on the Family*, 1st edn. Cambridge, MA: Harvard University Press.
- (1991). *A Treatise on the Family*, 2nd edn. Cambridge, MA: Harvard University Press.
- Becker, G. S. and R. T. Michael (1973). 'On the new theory of consumer behavior'. *Swedish Journal of Economics*, 75, 578–96.

- Benhabib, J., R. Rogerson and R. Wright (1991). 'Homework in macroeconomics: household production and aggregate fluctuations'. *Journal of Political Economy*, 99, 1166–87.
- Benoit, J.-P. and V. Krishna (1985). 'Finitely repeated games'. *Econometrica*, 53, 905–22.
- Bergson, A. (1938). 'A reformulation of certain aspects of welfare economics'. *Quarterly Journal of Economics*, 52, 30–4.
- Bergstrom, T. (1997). 'A survey of theories of the family'. in M. Rosenzweig and O. Stark (eds.), *Handbook of Population and Family Economics*. Amsterdam: North-Holland, 21–79.
- Bergstrom, T., L. Blume and H. Varian (1987). 'On the private provision of public goods'. *Journal of Public Economics*, 29, 25–49.
- Bernal, R. and M. P. Keane (2006). 'Child care choices and children's cognitive achievement: the case of single mothers'. Mimeo.
- Bernheim, B., J. Skinner and S. Weinberg (2001). 'What accounts for the variation in retirement wealth among US households'. *American Economic Review*, 91 (Sept), 832–57.
- Bester, H. and R. Strausz (2001). 'Contracting with imperfect commitment and the Revelation Principle. The single agent case'. *Econometrica*, 69, 1077–98.
- Binmore, K., A. Rubinstein and A. Wolinsky (1986). 'The Nash bargaining solution in economic modelling'. *Rand Journal of Economics*, 17, 176–88.
- Blomquist, N. S. (1983). 'The effects of income taxation on the labour supply of married men in Sweden'. *Journal of Public Economics*, 22, 169–97.
- Blundell, R., M. Browning and C. Meghir (1994). 'Consumer demand and the life-cycle allocation of household expenditures'. *Review of Economic Studies*, 61, 57–80.
- Blundell, R., P.-A. Chiappori, T. Magnac and C. Meghir (2007). 'Collective labour supply: heterogeneity and non-participation'. *Review of Economic Studies*, 74, 417–45.
- Blundell R. and T. MaCurdy (1999). 'Labor supply: a review of alternative approaches', in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. III. Amsterdam: North-Holland, 1560–1693.
- Blundell, R., C. Meghir, E. Symons and I. Walker (1988). 'Labour supply specification and the evaluation of tax reforms'. *Journal of Public Economics*, 36, 23–52.
- Blundell, R. and I. Walker (1982). 'Modelling the joint determination of household labour supplies and commodity demands'. *Economic Journal*, 92, 351–64.
- Boadway, R. and N. Bruce (1984). *Welfare Economics*. Oxford: Basil Blackwell.
- Boskin, M. J. (1975). 'Efficiency aspects of the differential tax treatment of market and household economic activity'. *Journal of Public Economics*, 4, 1–25.
- Boskin, M. J. and E. Sheshinski (1983). 'Optimal tax treatment of the family: married couples'. *Journal of Public Economics*, 20, 281–97.
- Brett, C. (2007). 'Optimal nonlinear taxes for families'. *International Tax and Public Finance*, 14, 225–61.
- Brett, C. and J. Weymark (2005). 'Optimal nonlinear taxation of income and savings in a two-class economy'. Mimeo.
- Browning, M. (1992). 'Children and household economic behaviour'. *Journal of Economic Literature*, 30, 1434–75.
- (2000). 'The saving behaviour of a two-person household'. *Scandinavian Journal of Economics*, 102 (2), 235–51.
- Browning, M., F. Bourguignon, P.-A. Chiappori and V. Lechene (1994). 'Income and outcomes: a structural model of intra-household allocation'. *Journal of Political Economy*, 102(6), 1067–96.
- Browning, M. and P.-A. Chiappori (1998). 'Efficient intra-household allocation: a characterisation and tests'. *Econometrica*, 66(6), 1241–78.
- Browning, M. and T. Crossley (2001). 'The life cycle model of consumption and saving'. *Journal of Economic Perspectives*, 15 (3), 46–60.

- Browning, M. and M. Ejrnaes (2002). 'Consumption and children'. CAM discussion paper 2002–06, University of Copenhagen.
- Browning, M. and A. Lusardi (1996). 'Household saving: micro theories and micro facts'. *Journal of Economic Literature*, 34, 1797–855.
- Browning, M. and C. Meghir (1991). 'The effects of male and female labor supply on commodity demands'. *Econometrica*, 59, 925–51.
- Carroll, C. D. (1992). 'The buffer stock theory of saving: some macroeconomic evidence'. *Brookings Papers on Economic Activity*, 61–156.
- (1994). 'How does future income affect current consumption?'. *Quarterly Journal of Economics*, 111, 111–48.
- (1997). 'Buffer stock saving and the life cycle/permanent income hypothesis'. *The Quarterly Journal of Economics*, 112, 1–55.
- (2001). 'A theory of the consumption function, with and without liquidity constraints'. *Journal of Economic Perspectives*, 15 (3), 23–46.
- Carroll, C. D. and M. Kimball (1996). 'On the concavity of the consumption function'. *Econometrica*, 64 (4), 981–92.
- Chen, Z. and F. Woolley (2001). 'A Cournot-Nash model of family decision making'. *Economic Journal*, 111 (474), 722–48.
- Chiappori, P.-A. (1988). 'Rational household labor supply'. *Econometrica*, 56 (1), 63–90.
- (1992). 'Collective labor supply and welfare'. *Journal of Political Economy*, 100, 437–67.
- (1997). 'Introducing household production in collective models of labor supply'. *Journal of Political Economy*, 105, 191–209.
- Cigno, A. (1986). 'Fertility and the tax-benefit system: a reconsideration of the theory of family taxation'. *Economic Journal*, 96, 1035–51.
- (1993). 'Intergenerational transfers without altruism'. *European Journal of Political Economy*, 9, 505–15.
- (1996). 'Cost of children, parental decisions and family policy'. *Labour*, 10(3), 461–74.
- (2001). 'Comparative advantage, observability, and the optimal tax treatment of families with children'. *International Tax and Public Finance*, 8, 451–66.
- Cigno, A., A. Luporini and A. Pettini (2000). 'Transfers to families with children as a principal-agent problem'. *Journal of Public Economics*, 87, 1165–72.
- Cigno, A. and A. Pettini (2001). 'Taxing family size and subsidizing child-specific commodities'. *Journal of Public Economics*, 84, 75–90.
- Colombino, U. (2000). 'The cost of children when children are a choice'. *Labour*, 14, 79–96.
- Cremer, H., A. Dellis and P. Pestieau (2003). 'Family size and optimal income taxation'. *Journal of Population Economics*, 6, 37–54.
- Crossley, T. and W. Yu (2004). 'Exploring the returns to scale in food preparation'. McMaster Economics discussion paper, 2004–6. McMaster University, Hamilton, Ont.
- Deaton, A. (1991). 'Saving and liquidity constraints'. *Econometrica* 59, 1221–48.
- (1992). *Understanding Consumption*. Oxford: Clarendon Press.
- Deaton, A. and J. Muellbauer (1980). *Economics and Consumer Behaviour*. Cambridge: Cambridge University Press.
- Deaton, A. and C. Paxson (1998). 'Economies of scale, household size and the demand for food'. *Journal of Political Economy*, 106, 879–930.
- Del Boca, D. and M. Locatelli (2008). *Motherhood and Participation*. Cambridge: Cambridge University Press.
- Del Boca, D. and A. Lusardi (2002). 'Credit market constraints and labor market decisions'. IZA discussion paper 598, Bonn.
- Del Boca, D. and D. Vuri (2007). 'The mismatch between labor supply and child care'. *Journal of Population Economics*, 20(4), 805–32.

- Donni, O. (2003). 'Collective household labour supply: nonparticipation and income taxation'. *Journal of Public Economics*, 87 (5–6), 1179–98.
- (2007). 'Collective female labour supply: theory and application'. *Economic Journal*, 117(516), 94–119.
- Ebert, U. (1992). 'A reexamination of the optimal nonlinear income tax'. *Journal of Public Economics*, 49, 47–73.
- Eckstein, Z. and K. I. Wolpin (1989). 'Dynamic labor force participation of married women and endogenous work experience'. *Review of Economic Studies*, 56, 375–90.
- Eissa, N. and H. Hoynes (2004). 'Taxes and the labor market participation of married couples: the earned income tax credit'. *Journal of Public Economics*, 88, 1931–58.
- (2006). 'Behavioral responses to taxes: lessons for the EITC and labor supply', in J. M. Poterba (ed.), *Tax Policy and the Economy*, 2nd edn. Cambridge, MA: MIT Press, 163–92.
- Eissa, N. and J. B. Leibman (1996). 'Labor supply responses to the earned income tax credit'. *Quarterly Journal of Economics*, 112(2), 603–37.
- Farrell, J. and E. Maskin (1989). 'Renegotiation in repeated games'. *Games and Economic Behavior*, 1, 327–60.
- Feldstein, M. and D. R. Feenberg (1996). 'The taxation of two-earner families', in M. Feldstein and J. M. Poterba (eds.), *Empirical Foundations of Household Taxation*. Chicago: University of Chicago Press, 39–76.
- Fenge, R. and Ochel, W. (2001). 'Die Vereinbarkeit von Familie und Beruf: Schlüssel für eine Kinderreiche Gesellschaft'. *ifo Schnelldienst*, 12, 17–29.
- Fisher, I. (1907). *The Rate of Interest*. New York: Macmillan.
- (1930). *The Theory of Interest*. New York: Macmillan.
- Fortin, B. and G. Lacroix (1997). 'A test of the unitary and collective models of household labour supply'. *Economic Journal*, 107, 933–55.
- Friedman, J. (1977). *Oligopoly and the Theory of Games*. Amsterdam: North-Holland.
- Friedman, M. (1957). *A Theory of the Consumption Function*. Princeton, NJ: Princeton University Press for NBER.
- Goode, A. and N. Watson (eds.) (2007). *HILDA User Manual – Release 5*. Melbourne Institute of Applied Economic and Social Research, University of Melbourne.
- Gorman, W. M. (1956). 'A possible procedure for analysing quality differentials in the egg market'. Ames: Iowa State College, mimeo. (Reissued as discussion paper B4, London School of Economics Econometrics Programme, London 1976).
- (1959). 'Separable utility and aggregation'. *Econometrica*, 27, 469–81.
- (1961). 'On a class of preference fields'. *Metroeconomica*, 13, 53–6.
- Gourinchas, P.-O. and J. A. Parker (2002). 'Consumption over the life cycle'. *Econometrica*, 70, 47–89.
- Graham, J. W. and C. A. Green (1984). 'Estimating the parameters of a household production function with joint products'. *Review of Economics and Statistics*, 66, 277–83.
- Green, E. and Porter, R. (1984). 'Noncooperative collusion under imperfect price information'. *Econometrica*, 52, 87–100.
- Greenwood, J. and Z. Hercowitz (1991). 'The allocation of capital and time over the business cycle'. *Journal of Political Economy*, 99, 1188–214.
- Gronau, R. (1973). 'The intrafamily allocation of time: the value of housewives' time'. *American Economic Review*, 63(4), 634–51.
- (1977). 'Leisure, home production and work – the theory of the allocation of time revisited'. *Journal of Political Economy*, 85(6), 1099–1123.
- (1980). 'Home production – a forgotten industry'. *Review of Economics and Statistics*, 62(3), 26–42.

- (1986). 'Home production – a survey', in O. Ashenfelter and R. Layard (eds.), *Handbook of Labour Economics*, vol. I. Amsterdam: North-Holland, 273–304.
- (1991). 'The intrafamily allocation of goods – how to separate the adult from the child'. *Journal of Labor Economics*, 9, 207–35.
- Gronau, R. and D. Hamermesh (2006). 'Time vs. goods: the value of measuring household production technologies'. *Review of Income and Wealth*, 52 (1), 1–16.
- Grossbard, S. (1976). 'An economic analysis of polygamy: the case of Maidiguru'. *Current Anthropology*, 17, 701–7.
- Grossbard-Schechtman, S. (1984). 'A theory of the allocation of time in markets for labor and marriage'. *Economic Journal*, 94, 863–82.
- (2003). 'A consumer theory with competitive markets for work in marriage'. *Journal of Socio-Economics*, 31, 609–45.
- Guesnerie, R. (1995). *A Contribution to the Pure Theory of Taxation*. Cambridge: Cambridge University Press.
- Gugl, E. (2005). 'Intrafamily distribution and taxation'. Proceedings of the 97th Annual Conference on Taxation, Minneapolis, November 11–13, 2004, National Tax Association, Washington DC, 344–9.
- Gustafsson, B. and U. Kjulin (1994). 'Time use in child care and housework and the total cost of children'. *Journal of Population Economics*, 7 (3), 287–306.
- Hart, O. (1995). *Firms, Contracts and Financial Structure*. Oxford: Oxford University Press.
- Hausman, J. A. (1981). 'The effect of taxes on labour supply', in H. Aaron and J. Pechman (eds.), *How Taxes Affect Economic Behavior*. Washington, DC: Brookings Institution, 27–72.
- (1985). 'Taxes and labour supply'. in A. J. Auerbach and M. Feldstein (eds.), *Handbook of Public Economics*, vol. I. Amsterdam: North-Holland, 213–63.
- Hausman, J. A. and P. Ruud (1984). 'Family labor supply with taxes'. *American Economic Review*, 74, 242–8.
- Heckman, J. (1974). 'Life cycle consumption and labor supply'. *American Economic Review*, 64, 188–94.
- (1980). 'Sample selection bias as a specification error with an application to the estimation of labor supply functions', in J. P. Smith (ed.), *Female Labor Supply, Theory and Estimation*. Princeton, NJ: Princeton University Press, ch. 5, 206–47.
- (1993). 'What has been learned about labour supply in the last twenty years?' *American Economic Review, Papers and Proceedings*, 83, 116–21.
- Hirshleifer, J. (1958). 'On the theory of optimal investment decision'. *Journal of Political Economy*, 66 (4), 329–52.
- Hurst, E. (2008). 'The retirement of a consumption puzzle'. National Bureau of Economic Research, working paper W13789.
- Imai, S. and M. P. Keane (2004). 'Intertemporal labor supply and human capital accumulation'. *International Economic Review*, 45 (2), 601–41.
- Jaumotte, F. (2003). 'Female labour force participation: past trends and main determinants in OECD countries'. *OECD Economic Studies*, 37, 51–108.
- Jehle, G. A. and P. J. Reny (2001). *Advanced Microeconomic Theory*, 2nd edn. Boston: Addison-Wesley.
- Kalenkoski, C., D. Ribar and L. S. Stratton (2005). 'Parental child care in single parent, cohabiting and married couple families: time diary evidence from the United Kingdom'. *American Economic Review Papers and Proceedings*, 95 (2), 194–8.
- Killingsworth, M. (1983). *Labor Supply*. New York: Cambridge University Press.
- Killingsworth, M. and J. Heckman (1986). 'Female labor supply: a survey', in O. Ashenfelter and R. Layard (eds.), *Handbook of Labor Economics*, vol. I. Amsterdam: North-Holland, 103–204.

- Kimball, M. S. (1990). 'Precautionary savings in the small and in the large'. *Econometrica*, 58, 53–73.
- Kleven, H. J., C. T. Kreiner and E. Saez (2007). 'The optimal income taxation of couples as a multi-dimensional screening problem'. CESifo, working paper 2092.
- Konrad, K. A. and K. E. Lommerud (1995). 'Family policy with non-cooperative families'. *Scandinavian Journal of Economics*, 97, 581–601.
- (2000). 'The bargaining family revisited'. *Canadian Journal of Economics*, 33 (2), 471–87.
- Kooreman, P. and A. Kapteyn (1986). 'Estimation of rationed and unrationed household labour supply functions using flexible functional forms'. *Economic Journal*, 96, 398–412.
- (1987). 'A disaggregated analysis of the allocation of time within the household'. *Journal of Political Economy*, 95, 223–49.
- (1990). 'On the empirical implementation of some game-theoretic models of household labor supply'. *Journal of Human Resources*, 25, 584–98.
- Kreps, D. M., P. Milgrom, J. Roberts and R. Wilson (1982). 'Rational cooperation in the finitely repeated prisoners' dilemma'. *Journal of Economic Theory*, 27, 245–52.
- Lancaster, K. J. (1966). 'A new approach to consumer theory'. *Journal of Political Economy*, 74, 132–57.
- Leuthold, J. H. (1968). 'An empirical study of formula income transfers and the work decision of the poor'. *Journal of Human Resources*, 3, 312–23.
- Li, G. and V. Vernon (2003). 'Testing the Barten model of economies of scale in household consumption: toward resolving a paradox by Deaton and Paxson'. *Journal of Political Economy*, 111 (6), 1361–77.
- Lipsey, R. and K. Lancaster (1956). 'The general theory of the second best'. *Review of Economic Studies*, 24, 11–32.
- Lundberg, S. (1988). 'Labor supply of husbands and wives: a simultaneous equation approach'. *Review of Economics and Statistics*, 70, 224–35.
- Lundberg, S. and R. A. Pollak (1993). 'Separate spheres bargaining and the marriage market'. *Journal of Political Economy*, 101, 988–1010.
- (1994). 'Noncooperative bargaining models of marriage'. *American Economic Review*, 84 (2), 132–7.
- (2003). 'Efficiency in marriage'. *Review of Economics of the Household*, 1(3), 153–67.
- Lundberg, S., R. A. Pollak, and T. J. Wales (1997). 'Do husbands and wives pool their resources? Evidence from the UK child benefit'. *Journal of Human Resources*, 32 (3), 463–80.
- Lundberg, S., R. Startz and S. Stillman (2003). 'The retirement-consumption puzzle: a marital bargaining approach'. *Journal of Public Economics*, 87, 1199–218.
- McElroy, M. B. (1990). 'The empirical content of Nash bargained household behaviour'. *Journal of Human Resources*, 25 (4), 559–83.
- McElroy, M. B. and M. Horney (1981). 'Nash-bargained decisions: towards a generalization of the theory of demand'. *International Economic Review*, 22, 333–49.
- MaCurdy, T., D. Green and H. Paarsch (1990). 'Assessing empirical approaches for analyzing taxes and labor supply'. *Journal of Human Resources*, 25, 415–90.
- Manser, M. and M. Brown (1980). 'Marriage and household decision making: a bargaining analysis'. *International Economic Review*, 21 (1), 31–4.
- Mas-Colell, A., M. D. Whinston and J. R. Green (1995). *Microeconomic Theory*. Oxford: Oxford University Press.
- Meade, J. (1955). *The Theory of Customs Unions*, vol. I. Amsterdam: North-Holland.
- Mincer, J. (1963). 'Market prices, opportunity costs and income effects', in C. Christ (ed.), *Measurement in Economics*. Stanford, CA: Stanford University Press, 67–82.
- Mirrlees, J. A. (1971). 'An exploration in the theory of optimum income taxation'. *Review of Economic Studies*, 38, 175–208.

- Modigliani, F. and R. Brumberg (1954). 'Utility analysis and the consumption function: an interpretation of cross-section data', in K. H. Kurihara (ed.), *Post-Keynesian Economics*. New Brunswick, NJ: Rutgers University Press, 388–436.
- Mroz, T. A. (1987). 'The sensitivity of an empirical model of married women's hours of work to economic and statistical assumptions'. *Econometrica*, 55, 765–99.
- Munnell, A. (1980). 'The couple versus the individual under the Federal Personal Income Tax', in H. Aaron and M. Boskin (eds.), *The Economics of Taxation*. Washington, DC: Brookings Institution, 247–80.
- Muth, R. F. (1966). 'Household production and consumer demand'. *Econometrica*, 34, 699–708.
- Myles, G. (1995). *Public Economics*. Cambridge: Cambridge University Press.
- Nelson, J. A. (1988). 'Household economies of scale in consumption: theory and evidence'. *Econometrica*, 56 (6), 1301–14.
- (1993). 'Household equivalence scales: theory versus policy?' *Journal of Labor Economics*, 11, 471–93.
- Nerlove, M. (1974). 'Toward a new theory of population and economic growth', in T. W. Schultz (ed.), *Economics of the Family*. Chicago: University of Chicago Press, 527–45.
- Olivetti, C. (2006). 'Changes in women's hours of market work: the role of returns to experience'. *Review of Economic Dynamics*, 9, 557–87.
- Ott, N. (1992). *Intrafamily Bargaining and Household Decisions*. Berlin: Springer-Verlag.
- Pencavel, J. (1987). 'Labor supply of men: a survey'. in O. Ashenfelter and R. Layard, (eds.), *Handbook of Labor Economics*, vol. I. Amsterdam: North-Holland, 3–102.
- Piggott, J. and J. Whalley (1996). 'The tax unit and household production'. *Journal of Political Economy*, 104, 398–418.
- Pollak, R. A. (1977). 'Price dependent preferences'. *American Economic Review*, 67, 64–75.
- (1985). 'A transaction cost approach to families and households'. *Journal of Economic Literature*, 23 (2), 581–608.
- (2003). 'Gary Becker's contribution to family and household economics'. *Review of Economics and the Household*, Jan/April.
- Pollak, R. A. and M. L. Wachter (1975). 'The relevance of the household production function and its implications for the allocation of time'. *Journal of Political Economy*, 68, 349–59.
- Pollak, R. A. and T. J. Wales (1979). 'Welfare comparisons and equivalence scales'. *American Economic Review*, 69, 216–21.
- Pudney, S. (1989). *Modelling Individual Choice: The Econometrics of Corners, Kinks and Holes*. Oxford: Basil Blackwell.
- Ransom, M. R. (1987). 'An empirical model of discrete and continuous choice in family labor supply'. *Review of Economics and Statistics*, 59, 465–72.
- Rees, R. (1985). 'Cheating in a duopoly supergame'. *Journal of Industrial Economics*, 23, 387–400.
- Rios-Rull, J.-V. (1993). 'Working in the market, working at home and the acquisition of skills: a general equilibrium approach'. *American Economic Review*, 83, 892–907.
- Rochet, J.-C. and L. A. Stole (2001). 'The economics of multidimensional screening'. Mimeo.
- Rosen, H. (1977). 'Is it time to abandon joint filing?'. *National Tax Journal*, 30, 423–8.
- Rotemberg, J. and G. Saloner (1986). 'A supergame-theoretic model of business cycles and price wars during booms'. *American Economic Review*, 76, 390–407.
- Rupert, P., R. Rogerson and R. Wright (1995). 'Estimating substitution elasticities in household production models'. *Economic Theory*, 6, 179–93.
- (2000). 'Homework in labor economics, household production and intertemporal substitution'. *Journal of Monetary Economy*, 46, 557–79.
- Salanié, B. (2002). *The Economics of Taxation*. Cambridge, MA: MIT Press.
- Samuelson, P. A. (1947). *Foundations of Economic Analysis*. Cambridge, MA: Harvard University Press.

- (1956). 'Social indifference curves'. *Quarterly Journal of Economics*, 70, 1–22.
- (1958). 'An exact consumption-loan model with or without the social contrivance of money'. *Journal of Political Economy*, 66, 467–82.
- Sandmo, A. (1990). 'Tax distortions and household production'. *Oxford Economic Papers*, 42, 78–90.
- Schroyen, F. (2003). 'Redistributive taxation and the household: the case of individual filings'. *Journal of Public Economics*, 87, 2527–47.
- Schultz, T. P. (1990). 'Testing the neoclassical model of family labour supply and fertility'. *Journal of Human Resources*, 25, 599–634.
- Seade, J. (1977). 'On the shape of optimal tax schedules'. *Journal of Public Economics*, 7, 203–36.
- Shaw, K. (1989). 'Life-cycle labor supply with human capital accumulation'. *International Economic Review*, 30, 431–56.
- (1994). 'The persistence of female labor supply: empirical evidence and implications'. *Journal of Human Resources*, 29, 348–78.
- Sheshinski, E. (1972). 'The optimal linear income tax'. *Review of Economic Studies*, 39, 297–302.
- (1989). 'Note on the shape of the optimum income tax schedule'. *Journal of Public Economics*, 40, 201–15.
- Singh, I., L. Squire and J. Strauss (eds.) (1986). *Agricultural Household Models: Extensions, Applications, and Policy*. Baltimore, MD: Johns Hopkins University Press (for the World Bank).
- Slemrod, J., S. Yitzhaki, J. Mayshar and M. Lundholm (1994). 'The optimal two-bracket linear income tax'. *Journal of Public Economics*, 53, 269–90.
- Soest, A. van (1993). 'Structural models of family labor supply: a discrete choice approach'. *Journal of Human Resources*, 30, 63–88.
- Thomas, D. (1990). 'Intra-household resource allocation: an inferential approach'. *Journal of Human Resources*, 25, 635–64.
- Thurow, L. (1969). 'The optimum lifetime distribution of consumption expenditure'. *American Economic Review*, 59, 324–30.
- Udry, C. (1996). 'Gender, agricultural production, and the theory of the household'. *Journal of Political Economy*, 104, 1010–46.
- Ulph, D. (1988). 'A general non-cooperative Nash model of household consumption behaviour'. Working paper 88–205, Department of Economics, University of Bristol.
- Vermeulen, F. (2002). 'Collective household models: principles and main results'. *Journal of Economic Surveys*, 16, 533–64.
- Warr, P. (1983). 'The private provision of a public good is independent of the distribution of income'. *Economics Letters*, 13, 207–11.
- Williamson, O. E. (1989). 'Transactions cost economics', in R. Schmalensee and R. D. Willig (eds.), *The Handbook of Industrial Organisation*, vol. I. Amsterdam: North-Holland, 135–82.
- Woolley, F. (1988). 'A non-cooperative model of family decision making'. TIDI discussion paper 125, London School of Economics.
- Zeldes, S. P. (1989). 'Optimal consumption with stochastic income: deviations from certainty equivalence'. *Quarterly Journal of Economics*, 104, 275–98.

Index

- age
 - age structure of populations 3
 - family life cycle defined on age of head of household 137–9
- aggregation
 - Samuelson's theorem 41, 42–3, 44
- Aguiar, M. 132
- Alesina, A. 38
- altruism 97
- anonymity 80
 - cooperative models 45–6
- Apps, P. F. 38, 39, 60, 75, 80, 94, 203, 206
- Arrufat, J. L. 93
- Ashworth, J. S. 38, 76, 77
- Attanasio, O. P. 128
- Australia 4
 - child benefits 198–9
 - labour supply 5, 6
 - social security system 201
 - taxation 3, 158, 162, 163, 181–2, 196
 - effective ATRs on second earner's income 192–4
 - effective MTRs and ATRs by household income 190–1
 - effective MTRs and ATRs by primary income 186–90
 - MTRs and ATRs by primary income 183–6
 - reforms 256
- Banks, J. R. 132
- bargaining
 - Nash bargaining model 38–9, 43, 61, 67–9, 77
 - with non-cooperative threat points 69–72
 - non-cooperative games 76
- basic needs 97
- Basu, K. 38, 77, 103
- Baxter, M. 116
- Becker, Gary S. 20, 23, 31, 36, 37, 38, 42, 67, 80, 132
- Becker model 23–6
 - Pollak/Wachter critique 26–30
- benefits *see* social security systems
- Bergstrom, T. 70
- Binmore, K. 76
- Blundell, R. 92, 93, 107, 128, 137, 138
- borrowing
 - cost of 117
 - family life cycle and 145
- Boskin, M. J. 162, 163, 176, 177, 185, 203, 208, 212, 214, 219, 272
- Brown, M. 38, 67, 68
- Browning, M. 38, 73, 97, 101, 102, 104, 106, 128
- budget constraints 48
- buffer stock saving 120, 125–8
- Burkina Faso 82
- business cycles 116
- Canada
 - taxation 3
 - reforms 256
- capital market
 - imperfect 112, 116–20, 133–5, 150–1, 152
 - perfect 133, 135, 150, 152, 153–4
- caring preferences 77
- Carroll, C. D. 126
- Chen, Z. 38, 71
- Chiappori, P.-A. 38, 72–3, 74, 75, 95, 101, 104, 106, 107
- children 10
 - child benefits 198–9
 - child support 4
 - cooperative models 58–61
 - costs of 96–101
 - family life cycle and 139–40, 141–2, 144–5
 - female employment and 7, 8–9, 11–12, 18
 - taxation and price of child care 179–81
- choice
 - household choice problem 215–17
- collective model 72–5
 - empirical work on 101–8
- collective rationality 73
- consumption 97–8
 - costs of children and 99, 100
 - family life cycle models 129–35, 142–3, 144–5
 - intertemporal profiles 151–4
 - full 116
 - in households 2
 - standard consumer model 20–3

- contracts
 - incompleteness 97
- convex joint tax system 217–20
- cooperative models 38, 39
 - across-household heterogeneity and utility possibilities 51–4
 - anonymity 45–6
 - children and child care 58–61
 - collective model 72–5
 - empirical work on 101–8
 - effects of wage changes on labour supply 54–6
 - generalising household welfare function 61–7, 68, 72
 - household production 47–58
 - Nash bargaining 67–9
 - with non-cooperative threat points 69–72
 - pooling 39, 45–6
 - Samuelson's theorem 39–44
 - symmetry 44–5
- corner solutions 91, 107
- Cournot–Nash equilibrium 70
- Deaton, A. 95, 116, 125
- decentralisation 41
- demand
 - standard consumer model 20–3
- distribution of income
 - taxation and 204–8, 263
- division of labour 2, 18
 - gender roles 71
- Donni, O. 107
- Ebert, U. 242
- efficiency
 - Pareto *see* Pareto efficiency
- Ejmaes, M. 128
- empirical models 88, 108
 - collective model 101–8
 - family life cycle models 149–51
 - household utility function model
 - with fixed labour supplies 92–4
 - theoretical underpinning 88–92
 - time-use data 94
 - comparison of 'exchange' and 'transfer' models 94–6
 - estimating costs of children 96–101
- employment *see* labour market
- endogenous income 112
- endogenous labour supply 113–16
- equilibrium models 38, 80–1
 - Nash equilibrium 70, 76, 77, 78
 - 'exchange' models 94–6
- expenditures
 - family life cycle 136–7
- extra-household environmental factors (EEPs) 82, 102, 103
- fairness 66–7
- family life cycle 3, 112–13, 128–9
 - average household 140–6
 - consumption 129–35, 142–3, 144–5
 - intertemporal profiles 151–4
 - empirical specification of model 149–51
 - evidence 135–54
 - household expenditure survey data 136–7
 - life cycle defined on age of head of household 137–9
 - life cycle defined on children 139–40
 - time-use in 129–35, 136–7
 - within-phase heterogeneity 146–8
- fertility
 - falls in 3
- Fisher, I. 116
- food consumption 132–3
- formation of households 67
- Fortin, B. 106, 107
- full consumption 116
- game theory 77
- gender roles 71
- general equilibrium theory 2
- generalised household welfare function (GHWF) 61–7, 68, 72, 79, 81–3, 106
- Germany 4
 - child benefits 199
 - labour supply 5
 - female employment 5, 7, 10–12, 13
 - taxation 3, 158, 160, 161, 163, 181–3, 197–8
 - effective ATRs on second earner's income 192–4
 - effective MTRs and ATRs by household income 190–1
 - effective MTRs and ATRs by primary income 186–90
 - MTRs and ATRs by primary income 183–6
 - reforms 256
- Gorman aggregation 42–3
- Gourinchas, P.-O. 128
- Gronau, R. 30, 31
- Grossbard, S. 39, 80, 81
- Heckman, J. 91
- heterogeneity 92
 - across-household heterogeneity and utility possibilities 51–4
 - female labour supply 178–9
 - within-phase heterogeneity 146–8
- Hirshleifer, J. 116
- Horney, M. 38, 44, 67, 68
- households 36–9
 - cooperative models 38, 39
 - across-household heterogeneity and utility possibilities 51–4
 - anonymity 45–6
 - children and child care 58–61
 - collective model 72–5, 101–8
 - effects of wage changes on labour supply 54–6
 - generalising household welfare function 61–7, 68, 72
 - household production 47–58
 - Nash bargaining 67–72
 - pooling 39, 45–6

- households (*cont.*)
 - Samuelson's theorem 39–44
 - symmetry 44–5
 - empirical models 88, 108
 - collective model 101–8
 - family life cycle models 149–51
 - household utility function model 88–94
 - time-use data 94–101
 - equilibrium models 38, 80–1
 - formation 67
 - household choice problem 215–17
 - household utility function (HUF) 42
 - with fixed labour supplies 92–4
 - theoretical underpinning of model 88–92
 - household welfare function (HWF) 39–44, 42, 43–4, 48, 67, 77
 - generalising 61–7, 68, 72, 79, 81–3, 106
 - income 14–18
 - life cycle models 109–13, 155–6
 - capital market imperfections 112, 116–20, 133–5, 150–1, 152
 - endogenous labour supply 113–16
 - family life cycle 3, 112–13, 128–54
 - perfect capital markets and 133, 135, 150, 152, 153–4
 - uncertainty 112, 120–8
 - non-cooperative models 38, 75–80
 - production in 2, 4–14, 34–5, 37, 116, 257
 - cooperative models 47–58
 - family life cycle models 132
 - optimal linear taxation and 212–14
 - taxing two-income households and 178–9
 - value of 32–4
 - standard consumer model 1, 20–3
 - taxation of two-income households 176–8, 202–3, 257
 - optimal non-linear taxation 246–53
 - within-household income distribution and 204–8, 263
 - traditional models 3
 - utility function models 2
 - Walrasian exchange model 39
- housing
- family life cycle and 145–6, 151–2
- Hurst, E. 132
- implicit prices 30–2, 132
- incentives
- taxation and 166
- income and wages 3, 116
- effects of wage changes on labour supply 54–6
 - endogenous income 112
 - family life cycles 148
 - household income 14–18
 - taxation and within-household income distribution 204–8, 263
 - own-wage effects 65–6
 - standard consumer model 20–1
 - taxation reforms with no wage dispersion 258–68
 - taxation reforms with wage dispersion 268–72
 - women 10, 55
- income tax 74, 181–3, 257, 272
- effective ATRs on second earner's income 192–4
 - effective MTRs and ATRs by household income 190–1
 - effective MTRs and ATRs by primary income 186–90
 - individual taxation 161–2
 - joint taxation 160–1
 - MTRs and ATRs by primary income 183–6
 - individual-based models 1, 20–3, 37
- information
- asymmetric 207
 - completeness of 79
- intertemporal choice 97
- intertemporal profiles of consumption 151–4
- Jermann, U. J. 116
- joint production 24
- joint taxation 160–1, 195, 256
 - optimal convex joint tax system 217–20
 - reforms 272–5
- Jones, G. S. 203
- Kapteyn, A. 76, 77
- Keynes, J. M. 122
- Killingsworth, M. 92
- Kimball, M. S. 122, 126
- Konrad, K. A. 38, 71
- Kooreman, P. 76, 77
- labour market 3
 - employment of women 3, 5–14, 18, 56–8, 138–9, 147–8
 - taxing two-income households and 178–9
- labour supply 4–14, 37
 - effects of wage changes 54–6
 - empirical collective model 106–8
 - endogenous 113–16
 - generalised household welfare function (GHWF) and 63–7
 - heterogeneity 178–9
 - household utility function (HUF) model and 92–4
 - life cycle models 113–16
 - standard consumer model 20–3
- Lacroix, G. 106, 107
- land 82
- leisure time 23, 37, 257
- Leuthold, J. H. 38, 76
- life cycle models 3, 10, 109–13, 155–6
 - capital market imperfections 112, 116–20, 133–5, 150–1, 152
 - endogenous labour supply 113–16
 - family life cycle 3, 112–13, 128–9
 - average household 140–6
 - consumption 129–35, 142–3, 144–5, 151–4
 - empirical specification of model 149–51
 - evidence 135–54
 - household expenditure survey data 136–7
 - life cycle defined on age of head of household 137–9
 - life cycle defined on children 139–40
 - time-use in 129–35, 136–7
 - within-phase heterogeneity 146–8

- perfect capital markets and 133, 135, 150, 152, 153–4
- uncertainty 112, 120–2, 123–4
 - buffer stock saving 120, 125–8
 - liquidity constraints 116, 124–5
 - precautionary saving 120, 122–4
- linear taxation systems 158–9, 202
 - optimal linear taxation
 - Boskin–Sheshinski model 208–12
 - household production and 212–14
 - piecewise linear taxation 214–20
 - piecewise systems 159
 - optimal piecewise linear taxation 214–20
 - redistribution and 169–76, 179–81
 - redistribution and 164–6
 - piecewise linear taxation 169–76, 179–81
 - two-income households 176–8
- liquidity constraints 116, 124–5
- Lommerud, K. E. 38, 71
- Lundberg, S. 38, 69, 70, 77, 79, 81, 208
- McElroy, M. B. 38, 44, 67, 68, 82
- male chauvinist model 92
- Manser, M. 38, 67, 68
- market substitutes 30–2
- marriage 67, 80, 82
- Mirrlees, James A. 157, 159, 166, 202, 221, 245
- moral hazard 4
- Muellbauer, J. 95
- Munnell, A. 203
- Nash bargaining model 38–9, 43, 61, 67–9, 75, 77
 - with non-cooperative threat points 69–72
- Nash equilibrium 70, 76, 77, 78
- Nelson, J. A. 97, 98
- non-cooperative models 38, 75–80
- non-linear taxation
 - optimal non-linear taxation 221
 - continuum of types 234–46
 - two-person households 246–53
 - two-type case 221–34
- non-linear taxation systems 159–60, 202
 - redistribution and 166–8
- optimal tax theory 62, 166–8, 202–4, 257, 275
 - linear taxation
 - Boskin–Sheshinski model 208–12
 - household production and 212–14
 - piecewise linear taxation 214–20
 - non-linear taxation 221
 - continuum of types 234–46
 - two-person households 246–53
 - two-type case 221–34
- Ott, N. 38
- own-wage effects 65–6
- Pareto efficiency 38, 72, 73, 77, 81–3, 254–5
 - commitment case 84
 - non-commitment case 84–5
- Parker, J. A. 128
- pension systems 3, 199–201
 - family life cycle and 151–2
- piecewise linear taxation systems 159
 - optimal piecewise linear taxation 214–15
 - optimal convex joint tax system 217–20
 - solutions to household choice problem 215–17
 - redistribution and 169–76
 - taxing two-income households 179–81
- Piggott, J. 272, 274
- Pollak, R. A. 20, 26, 29, 30, 32, 38, 63, 69, 70, 77, 79, 81, 98, 208
- pooling
 - cooperative models 39, 45–6
- poverty 4
- precautionary saving 120, 122–4
- prices
 - Becker model 24–5
 - implicit 30–2, 132
- production in households 2, 4–14, 34–5, 37, 116, 257
 - Becker model 23–6
 - cooperative models 47–58
 - family life cycle models 132
 - optimal linear taxation and 212–14
 - Pollak/Wachter critique 26–30
 - taxing two-income households and 178–9
 - value of 32–4
- public economics 41, 254
- public goods 70
- rationality
 - collective 73
 - non-cooperative models 77
- redistributive taxation systems 163–4
 - linear taxation 164–6
 - two-income households 176–8
 - non-linear taxation 166–8
 - taxing two-income households
 - household production and female labour supply
 - heterogeneity 178–9
 - linear taxation 176–8
- Rees, R. 38, 60, 75, 94, 203, 206
- renegotiation process 78–9
- salaries *see* income and wages
- Samuelson, Paul A. 36, 37, 38, 39, 40, 41, 42, 43, 46, 48, 61, 62, 63, 65, 66, 67, 68, 73, 205, 207
- Samuelson's theorem 39–44, 72
- Sandmo, A. 178, 203, 204
- saving
 - buffer stock 120, 125–8
 - family life cycle and 145, 148
 - precautionary 120, 122–4
- scale economies
 - Pollak/Wachter critique of Becker model 26–30
- selective taxation 162–3, 257
- Sheshinski, E. 157, 158, 162, 163, 176, 177, 185, 202, 203, 208, 212, 214, 219, 272
- Slemrod, J. S. 214
- social indifference curves 36
- social security systems 3, 175–6, 199–201, 207–8
 - child benefits 198–9
 - family life cycle models and 143–4
- social welfare function (SWF) 39–40, 62, 73, 204–6

- standard consumer model 20–3
- symmetry
 - compensated labour supply derivatives 64–5
 - cooperative models 44–5
- taxation 3, 13–14, 148, 157–8, 194–8
 - income *see* income tax
 - individual 161–2
 - joint 160–1, 195, 256
 - optimal convex joint tax system 217–20
 - reforms 272–5
 - linear systems 158–9, 202
 - optimal linear taxation 208–14
 - piecewise systems 159, 169–76
 - redistribution and 164–6, 169–76
 - two-income households 176–8
 - non-linear systems 159–60, 202
 - redistribution and 166–8
 - optimal tax theory 62, 166–8, 202–4, 257, 275
 - optimal linear taxation 208–14
 - piecewise linear taxation systems 159
 - optimal piecewise linear taxation 214–20
 - redistribution and 169–76, 179–81
 - redistributive systems 163–4, 176–81
 - linear taxation 164–6
 - piecewise linear taxation 169–76
 - taxing two-income households 176–81
 - reforms 254–8, 275–6
 - joint taxation and tax unit 272–5
 - policies 256–72
 - substitute time inputs 273–5
 - with no wage dispersion 258–68
 - with wage dispersion 268–72
 - selective 162–3, 257
 - two-income households 176–8, 202–3, 257
 - optimal non-linear taxation 246–53
 - within-household income distribution and 204–8, 263
- threat points 68, 69–72
- time-use 2, 10–11
 - Becker model 23–6
 - Pollak/Wachter critique 26–30
 - empirical household models 94
 - comparison of ‘exchange’ and ‘transfer’ models 94–6
 - estimating costs of children 96–101
 - family life cycle models 129–35, 136–7
 - standard consumer model 20–3
- traditional household models 3
- ‘transfer’ models 94–6
- Udry, Christopher 82
- Ulph, D. T. 38, 69, 76, 77
- uncertainty 97
 - life cycle models 112, 120–2
 - buffer stock saving 120, 125–8
 - liquidity constraints 116, 124–5
 - precautionary saving 120, 122–4
- United Kingdom 4
 - age of head of household 137
 - child benefits 198
 - households in 3
- labour supply 5
 - female employment 5, 6, 13, 138–9
- social security system 199, 200
- taxation 3, 158, 162, 181–3, 196
 - effective ATRs on second earner’s income 192–4
 - effective MTRs and ATRs by household income 190–1
 - effective MTRs and ATRs by primary income 186–90
 - MTRs and ATRs by primary income 183–6
 - reforms 256
- wages 17
- United States of America 4
 - child benefits 199
 - food consumption 133
 - households in 3
 - labour supply 5, 6
 - social security system 199, 200
 - taxation 3, 158, 160–1, 181–3, 197
 - effective ATRs on second earner’s income 192–4
 - effective MTRs and ATRs by household income 190–1
 - effective MTRs and ATRs by primary income 186–90
 - MTRs and ATRs by primary income 183–6
 - reforms 256
 - wages 17
- utility
 - across-household heterogeneity and utility possibilities 51–4
 - household utility function (HUF) 42
 - theoretical underpinning of model 88–92
 - with fixed labour supplies 92–4
 - possibilities 52
 - utility function models 2
- voluntary-contribution public goods 70
- Wachter, M. L. 20, 26, 29, 30, 32
- wages *see* income and wages
- Wales, T. J. 98
- Walrasian exchange model 39
- Warr, P. 70
- Weber, G. 128
- welfare 41
 - household welfare function (HWF) 39–44, 42, 43–4, 48, 67, 77
 - generalising 61–7, 68, 72, 79, 81–3, 106
 - social welfare function (SWF) 39–40, 62, 73, 204–6
- welfare economics 2
- welfare systems *see* social security systems
- Whalley, J. 272, 274
- within-phase heterogeneity 146–8
- women
 - employment of 3, 5–14, 18, 56–8, 138–9, 147–8
 - taxing two-income households 178–9
 - gender roles 71
 - wages 10, 55
- Woolley, F. 38, 69, 71
- Zabalza, A. 93