Essays on Dynamic Macroeconomics

The Role of Demographics and Public Capital

Ward Romp

**Research School
Systems, Organisation
and Management**

Essays on Dynamic Macroeconomics

The Role of Demographics and Public Capital

Willem Eduard Romp

Publisher: Labyrint Publications
Postbus 334
2984 AX Ridderkerk
The Netherlands
Tel: +31 (0)180 463962

Printed by: Offsetdrukkerij Ridderprint B.V., Ridderkerk.

ISBN: 90-5335-112-4/978-90-5335-112-3

RIJKSUNIVERSITEIT GRONINGEN

Essays on Dynamic Macroeconomics
The Role of Demographics and Public Capital

Proefschrift

ter verkrijging van het doctoraat in de
Economische Wetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. F. Zwarts,
in het openbaar te verdedigen op
donderdag 15 februari 2007
om 16.15 uur

door

Willem Eduard Romp

geboren op 24 januari 1976
te Harderwijk

Promotores:	Prof. dr. B.J. Heijdra
	Prof. dr. J. de Haan
Beoordelingscommissie:	Prof. dr. A.C. Meijdam
	Prof. dr. E. Sterken
	Prof. dr. J.E. Sturm

Acknowledgements

This thesis is the product of five years of labour, both by me and others. First of all, I could not have done it without my supervisors, Ben Heijdra and Jakob de Haan. I enjoyed working with Ben as he always pushed me to the limit. The many discussions were sometimes explosive, but did result in new insights and the first three chapters of this thesis. I wrote chapter 5 together with Jakob and he made sure I did not lose sight of the ultimate objective: to finish this thesis.

The whole project would not have been possible without the initiators Jan Jacobs, Peter Groote, and Jan-Egbert Sturm. They went through all bureaucratic procedures to collect the required funds. During the project Jan's door was always open for conversations, work-related and otherwise. I thank Jan Oosterhaven for showing me how intriguing science can be. He hired me as a research assistant and motivated me to choose for a scientific career.

I am indebted to the reading committee, Elmer Sterken, Lex Meijdam en Jan-Egbert Sturm for their willingness to read the final manuscript and for their comments. I enjoyed the hospitality of Jan-Egbert during my visit to Konstanz in the autumn of 2004, where we laid the foundations for chapter 6. Robert Inklaar provided us with the data and finally helped me finish this, sometimes frustrating, chapter.

I apologise to all people who were ever forced to share their room with me and my noisy habits. A special word of thanks must go to Mark. The days in 505 would not have been the same without the small blue ball, the large coloured ball, and the bouncing ball. To Abdul and the neighbours in those days: sorry for that!

For dinner at the ACLO I enjoyed the company of Adriaan, Pim, Robert, Macro, Bart and many others. It always was a pleasant break and the coffee afterwards gave me enough energy to continue working late.

Finally I especially thank Marijke for her patience and support. From now on I hopefully have more time and I hereby promise to spend it with her.

Ward Romp

Contents

Chapter 1

Introduction

The trade-off between welfare of future and current generations is one of the key issues in macroeconomic research (Royal Swedish Academy of Sciences, 2006). Economists and policy makers have long realised that many decisions have a dynamic and intergenerational nature and that merely focusing on either the immediate effects or – at the other extreme – the long-run effects of different policy options does not reveal potentially important transitory effects. To understand the implications of supposedly welfare enhancing decisions made today, we must know what the effects are on economic performance, *and* who reaps the rewards.

Dynamic macroeconomics is by now a well-developed field. Just a few of the questions that have been studied are how to optimise economic growth (e.g. Phelps, 1961), what the determinants of economic growth are (e.g. Barro, 1991 and Sala-i-Martin, 1997) and how different generations interact (e.g. Diamond, 1965).

This dissertation covers two of the many different aspects of dynamic macroeconomics. Part I focuses on the role of demographics and demographic changes as one of the determinants of intergenerational transfers. Part II looks at the impact of public capital on economic growth. Both parts are self-contained and can for this reason be read separately.

1.1 Realistic demographics in overlapping generations models

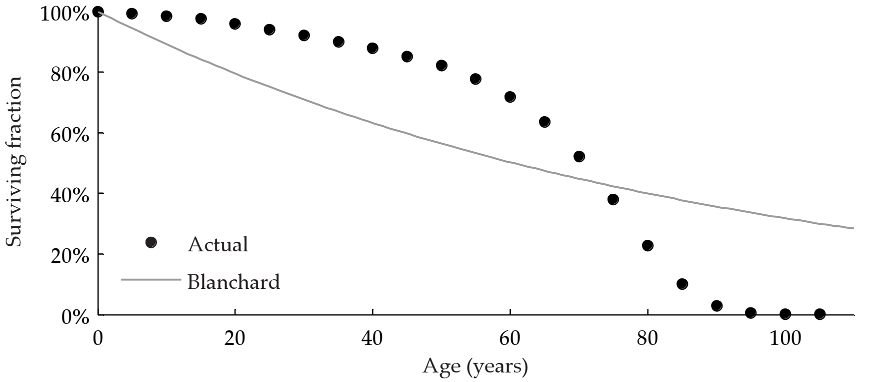
Death is the only certainty in life. Unfortunately, for planning purposes, most of us do not know when the Grim Reaper will make his one and final call. The microeconomic implications of this *lifetime uncertainty* were first studied in the context of a dynamic consumption-saving model by Yaari (1965). He showed that, faced with a positive mortality rate, individual agents will discount future felicity more heavily due to the uncertainty of survival. Furthermore, with lifetime uncertainty the consumer faces not only the usual solvency condition ('living within one's means') but also a constraint prohibiting negative net wealth at any time – the agent is simply not allowed by capital markets to die indebted. To solve this problem, Yaari assumes that an individual can purchase (annuity) or sell (life insurance) actuarial notes at an actuarially fair interest rate. A 'rational' individual will use such notes to fully insure itself against the adverse effect of lifetime uncertainty.

Yaari's insights were embedded in a general equilibrium growth model by Blanchard (1985). Because of its flexibility, the Blanchard-Yaari model and the closely connected Buiter (1988) and P. Weil (1989) model have achieved workhorse status in macroeconomic analysis during the last two decades. One typical area of application of these models is intergenerational welfare analysis of various policies (e.g. Bovenberg, 1993, 1994, and Bettendorf and Heijdra, 2001a). Other areas where the Blanchard-Yaari models have been used are demographic change and economic growth (e.g. de la Croix and Licandro, 1999; Kalemni-Ozcan et al., 2000), and social security and ageing (e.g. Bettendorf and Heijdra, 2006). Most intermediate and advanced macroeconomic textbooks nowadays contain a description of the Blanchard-Yaari overlapping generations models (e.g. Blanchard and Fischer, 1989 and Heijdra and van der Ploeg, 2002).

Although the Blanchard-Yaari framework is quite flexible, it has at least one major drawback. In order to allow for exact aggregation of individual decision rules, Blanchard simplified the Yaari model by assuming a constant death probability, i.e. the probability of dying does not depend on the age of the individual. A direct implication of this assumption is that individuals enjoy a perpetual youth. No matter what their age is, the expected remaining lifetime remains the same. This makes exact aggregation of individual choices possible since the propensity to consume out of total wealth is the same for everybody. This propensity to consume is age-independent because everybody faces the same mortality rate and expected remaining lifetime and uses the same mortality-risk adjusted discount factor.

Figure 1.1 shows the fit of the predicted surviving fraction of a of Blanchard's perpetual youth model. The stars denote the observed – and for the ages above 85 the predicted – surviving fraction of the cohort born in the Netherlands in 1920. The line shows the best possible fit of Blanchard's surviving fraction. The figure clearly shows the poor fit of Blanchard's perpetual youth model. Of the cohort under con-

Figure 1.1. Actual versus estimated survival fractions for Blanchard's perpetual youth model



Notes: Data for the cohort born in the Netherlands, 1920 (male and female). Observed survival rates for ages 0–85, projected survival rates otherwise. *Source:* Human Mortality Database (2006) and own calculations.

sideration, 70% actually reached the age of 60; according to the perpetual youth model, only 50% should have survived that long. After 77 years, the perpetual youth model grossly overestimates the surviving fraction. According to the demographic projections, only 0.03% of the 1920-cohort will reach the age of 100, whereas the perpetual youth model predicts 30%! Blanchard's modelling dilemma is clear: exact aggregation is 'bought' at the expense of a rather unrealistic description of the demographic process. Of course, in a closed-economy context the aggregation step is indispensable because equilibrium factor prices (wages and interest rates) are determined in the aggregate factor markets. However, in the context of a small open economy, factor prices are typically determined in world markets so that the aggregation step is not necessary.

The main objective in Part I of this dissertation is to extend the standard overlapping generations (OLG) model of a small open economy with a realistic demographic process, to compare the results of the extended OLG-model with the results of the standard perpetual youth model, and finally to use our extended OLG model to analyse the effects of ageing shocks on economic growth and retirement. In Chapter 2 we show that, provided we restrict attention to the case of a small open economy, it is quite feasible to construct and analytically analyse a Blanchard-Yaari type overlapping-generations model with a realistic description of the mortality process. One main effect of a realistic mortality process is that it results in a humped shaped savings profile, as in Modigliani's classic life cycle model.

As in Blanchard's OLG model, it is possible to analytically determine the steady state comparative effects of our extended model. Our comparative static analysis shows that the long-run effects are qualitatively equal to the effects in the standard Blanchard-Yaari models. However, there are major differences in the transition paths. We show that a realistic mortality process gives rise to drastically different impulse-response functions associated with various macroeconomic shocks. The transition period of our more realistic model is much shorter than in the standard Blanchard-Yaari models. These old models sometimes predict transition periods of one, maybe two centuries, in our models this time is reduced to 40 to 60 years. This is because in our models individuals age as time goes by. During the transition periods, the fluctuations are larger than in the standard models, sometimes even twice as large.

Chapters 3 and 4 use the framework developed in Chapter 2 to study two questions. In Chapter 3 we introduce a schooling decision. Following *inter alia* Bils and Klenow (2000), and Kalemni-Ozcan et al. (2000) agents engage in educational activities at the start of life and thus create human capital to be used later on in life for production purposes. Individuals have no preference for either school or work, so utility maximisation ensures that the schooling period is chosen in such a way that it maximises the present value of after tax income. The schooling externality is based on the 'standing on the shoulders of'–type as proposed by Azariadis and Drazen (1990). The effect of schooling on an individual's knowledge and earnings potential depends on the knowledge of the teacher. An individual's education decision affects future generations because the students of today are tomorrow's teachers. Depending on the strength of the intergenerational externality in the accumulation of knowledge, the model gives rise to exogenous growth or, in a knife-edge case, endogenous growth.

We then use this model to analyse the effects of demographics shocks, i.e. a longevity shock and a baby bust, on the main economic indicators like economic growth during transition and in the long run. Even though the model is more complex than the basic model in Chapter 2, we are still able to analytically characterise the long-run effects. One of the more surprising results is that there is a highly non-linear, even non-monotonic relation between longevity and economic growth; lower mortality may lead to lower per capita growth, even in the long run. Depending on the type of shock, a change in the mortality process or in the birth rate, economic growth changes gradually or with jumps. In both cases the transition process is non-monotonic and the transition periods extend over decades or even centuries. This may explain the results of Kelley and Schmidt (1995). Based on a panel consisting of 89 countries and covering 30 years they found that the link between population growth, birth rates and the mortality process changes over time and the transition profile depends heavily on the type of demographic shock.

In Chapter 4 we extend the basic framework of Chapter 2 with a retirement decision similar to the one proposed by Sheshinski (1978), and a pension system. We use this model at a microeconomic level to study the effects of ageing shocks and public pension system reforms on the retirement decision of an individual who faces lifetime uncertainty. We extend the literature in two directions. First, we show how to transform the individual optimisation problem into a problem with a convenient and intuitive graphical representation. As always, a two dimensional visualisation of the originally infinite dimensional optimisation problem greatly simplifies the comparative static analysis. Using this graphical apparatus we can easily explain why most people in the Western world retire at the youngest age that (early) retirement benefits become available and why the optimal retirement age seems to be insensitive to fiscal stimuli (Gruber and Wise, 1999; Duval, 2003). Second, we show the analytical comparative static effects that describe how rational individuals will react to demographic shocks and fiscal stimuli of the pension system.

We use our retirement model at the macroeconomic level to determine the required system reforms to keep the pension systems sustainable in the long run. In the current ageing societies of the Western world, the sustainability of the social security and retirement systems is questionable (Gruber and Wise, 1999, 2004, 2005; OECD, 1998, 2005). The problem is twofold. On the one hand, people live longer but do not tend to work longer. This increases the number of people that receive old age benefits. On the other hand, due to lower birth rates, the number of productive people decreases. The pressure on the retirement system calls for reforms, usually painful ones (see Lindbeck and Persson, 2003 for a literature overview).

Confronting our model with demographic shocks of the type and magnitude that have hit the Western world in the post-war period, we find some remarkable results. First, a baby bust immediately puts pressure on the pension system, whereas the effects of a mortality shock only show up after 60 years. Second, the negative welfare effects are much smaller if the policy reform includes raising the retirement age instead of only raising tax revenues.

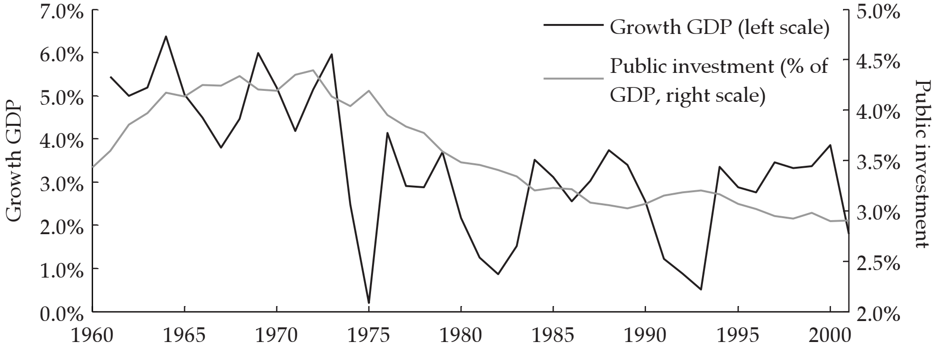
Two main conclusions can be drawn from Part I. First, we cannot ignore the transitional dynamics in OLG models. The transition periods are usually very long and the transition paths are non-monotonic. Focusing solely on the comparative statics implies that one neglects the interests of the people that live during the transition period. Second, it is quite feasible to extend the standard Blanchard-Yaari type OLG models for the small open economy with a realistic description of the mortality process and the demographic realism matters. The extended models predict very different impulse-response functions, on both the individual and the aggregate level and, maybe even more important, different welfare effects.

1.2 Public capital and economic growth: an empirical analysis

In Part II of this dissertation we ignore intergenerational issues and focus entirely on public capital, one of the determinants of economic growth. Public capital, and especially infrastructure, is central to the activities of households and firms. According to the World Bank (1994), public capital represents the 'wheels' – if not the engine – of economic activity. Input-output tables for example show that telecommunications, electricity, and water are used in the production process of nearly every sector, while transport is an input for every commodity.

As Gramlich (1994) notes, it is surprising how long economists have ignored the impact of infrastructure on aggregate productivity. Macroeconomists have long felt that the stock of public capital is an important factor input in the production of total output, but no one had empirically linked the movements of infrastructure and productivity until Aschauer (1989). In his seminal contribution Aschauer estimated the output elasticity of public capital and concluded that a 10 percent increase of the public capital stock resulted in a 4 percent increase in total output, an output elasticity of 40%! His paper hit a magic button and ever since a substantive research effort focused on estimating the contribution of public capital to the productivity of private factors and to economic growth. A reason why Aschauer's work received this much attention is that it provided an explanation for the productivity slowdown of the 1970s and 1980s in the US, as well in other OECD countries. As Figure 1.2 shows, average output growth in 22 OECD countries (left scale) dropped dramatically in the 1970s. The drop in output growth more or less coincides with a reduction of public investment (right scale). The policy message was simple: increase public capital spending to stimulate productivity. A message loved by policy makers (Gramlich, 1994).

Figure 1.2. Average economic growth and public investment in 22 OECD countries



Source: Kamps (2006) and own calculations.

However, several economists questioned Aschauer's estimates on the grounds that they were implausibly high (see, for instance, Gramlich, 1994). Furthermore, the early studies were fraught with methodological and econometric difficulties. Issues ranking high on the list of potential problems include reverse causation from productivity to public capital and a spurious correlation due to non-stationarity of the data. In their survey of the earlier literature, Sturm et al. (1998) show that the literature contained a relatively wide range of estimates, with a marginal product of public capital that is much higher than that of private capital (e.g., Aschauer, 1989), roughly equal to that of private capital (e.g., Munnell, 1990a), well below that of private capital (e.g., Eberts, 1986) and, in some cases, even negative (e.g., Hulten and Schwab, 1991). The wide range of estimates makes the results of these older studies almost useless from a policy perspective.

The contribution of the second part of this dissertation is twofold. Chapter 5 provides an up-to-date overview of the modern empirical literature on public capital and economic growth. Chapter 6 provides new estimates of the growth-enhancing effects of public capital and tests whether one of the main tools in the literature on growth enhancing effects of public capital, the production function approach, can actually provide robust estimates.

In our survey in Chapter 5 we focus on two questions. First, can a robust empirical relationship be found that an increase in public capital spurs economic growth? And second, to what extent do conclusions on the effect of increases of public capital change once it is taken into account that infrastructure construction diverts resources from other uses. We observe that most studies have tackled the methodological issues that hampered the early literature and that there seems to be more consensus in the more recent literature about the effects of public capital on economic performance. Most recent studies generally suggest that public capital may, under specific circumstances, raise real income per capita.

Another conclusion of our survey is that estimates of the impact of infrastructure investment on economic growth differ substantially, depending on the countries and time period covered, the level of aggregation and the econometric methods employed. In Chapter 6 we attempt to arrive at a set of robust estimates of the output elasticity of public capital using internationally comparable aggregate and industry data for a considerable number of developed countries and a substantial number of years.

As mentioned by Munnell (1992), many studies use growth rates to identify the effect of infrastructure on productivity, thereby destroying the long-run relationship. Infrastructure investment however mostly consists of projects with long durations and effectiveness. In Chapter 6 we use the Pooled Mean Group (PMG) estimator proposed by Pesaran et al. (1999) that avoids this problem by identifying the long-run relationship between variables in an error-correction framework. In cross-country and cross-industry estimates, efficiency gains are possible by restricting the parameter of interest to be equal across countries and/or industries. However, the PMG estimator only restricts the long-run parameter to be the same across countries or industries, while allowing for heterogeneity in the adjustment to this long run.

Chapter 6 contains estimates of the impact of infrastructure on economic growth based both on aggregate and industry data. But while data and econometric methodology are state-of-the-art and counter much, if not all, criticism raised in this literature, stable output elasticity estimates are elusive. Indeed, the estimated parameters of the output elasticities at the industry level vary wildly between equally plausible econometric specifications, ranging from -2 to 2. The aggregate estimates tend to be more stable, but even here, output elasticities range between 0.04 and 1.13. While it is hard to discount cross-country variation, the cross-specification variation we find suggests extreme sensitivity to conceptually innocuous specification choices. Overall, this suggests that production function estimates of the impact of infrastructure are not well-suited to be used for infrastructure policy recommendations.

Part I

Realistic Demographics in Overlapping Generations Models

Chapter 2

The Basic Model

> It is possible that death may be the consequence of two generally co-existing causes; the one, chance, without previous disposition to death or deterioration; the other, a deterioration or an increased inability to withstand destruction.
> *(Gompertz, 1825)*

2.1 Introduction

The opening quotation is a verbal introduction to a phenomenon that is now often called Gompertz' law of mortality. In his path-breaking paper, Benjamin Gompertz[1] (1825) identified two main causes of death, namely one due to pure chance and another depending on the person's age. He pointed out that if only the first cause were relevant, then 'the intensity of mortality' would be constant and the surviving fraction of a given cohort would decline in geometric progression. In contrast, if only the second cause would be relevant, and 'if mankind be continually gaining seeds of indisposition, or in other words, an increased liability to death' then the force of mortality would increase with age. Gompertz' law was subsequently generalized by Makeham (1860) who argued that the instantaneous

---

This chapter is based on joint work with Ben Heijdra, 'A Life-Cycle Overlapping-Generations Model of the Small Open Economy', forthcoming in the Oxford Economic Papers.

[1] As Hooker (1965) points out, Benjamin Gompertz can be seen as one of the founding fathers of modern demographic and actuarial theory. See also Preston et al. (2001, p. 192). Blanchard (1985, p. 225) and Faruqee (2003, p. 301) incorrectly refer to the non-existing 'Gomperty's law'.

mortality rate depends both on a constant term (first cause) and on a term that is exponential in the person's age (second cause).

The microeconomic implications for consumption behaviour of lifetime uncertainty—resulting from a positive death probability—were first studied in the seminal paper by Yaari (1965). He showed that, faced with a positive mortality rate, individual agents will discount future felicity more heavily due to the uncertainty of survival. Furthermore, with lifetime uncertainty the consumer faces not only the usual solvency condition but also a constraint prohibiting negative net wealth at any time—the agent is simply not allowed by capital markets to expire indebted. Yaari assumes that the household can purchase (annuity) or sell (life insurance) actuarial notes at an actuarially fair interest rate. In the absence of a bequest motive, the household will use such notes to fully insure against the adverse effect of lifetime uncertainty.

The Yaari insights were embedded in a general equilibrium growth model by Blanchard (1985). In order to allow for exact aggregation of individual decision rules, Blanchard simplified the Yaari model by assuming a constant death probability, i.e. only the first cause of death is introduced into the model and households enjoy a perpetual youth. Because of its flexibility, the Blanchard-Yaari model has achieved workhorse status in the last two decades.[2] As Blanchard himself points out, his modelling approach has the disadvantage that it cannot capture the life-cycle aspects of consumption and saving behaviour—the age-independent mortality rate ensures that the propensity to consume out of total wealth is the same for all households.[3]

Blanchard's modelling dilemma is clear: exact aggregation is 'bought' at the expense of a rather unrealistic description of the demographic process.[4] Of course, in a closed-economy context, the aggregation step is indispensable because equilibrium factor prices are determined in the aggregate factor markets. However, in the context of a small open economy, factor prices are typically determined in world markets so that the aggregation step is not necessary and life-cycle effects can be modelled. The main objective in this chapter is to elaborate on exactly this point.

[2] For our purposes, the most important extension is due to Buiter (1988) who allows for non-zero population growth by using the insights of P. Weil (1989). For a textbook treatment of the Blanchard-Yaari model, see Blanchard and Fischer (1989, ch. 3) or Heijdra and van der Ploeg (2002, ch. 16).

[3] Blanchard shows that a 'saving-for-retirement' effect can be mimicked by assuming that labour income declines with age. Faruqee and Laxton (2000) use this approach in a calibrated simulation model.

[4] Blanchard suggests that a constant mortality rate may be more reasonable if the model is applied to 'dynastic families' rather than to individual agents (1985, p. 225, fn.1). Under this interpretation the mortality rate refers to the probability that the dynasty literally becomes extinct or that the chain of bequests between the generations is broken.

As we demonstrate below, provided we restrict attention to the case of a small open economy, it is quite feasible to construct and analytically analyse a Blanchard-Yaari type overlapping-generations model incorporating a realistic description of demography. In addition we show that such a model gives rise to drastically different impulse-response functions associated with various macroeconomic shocks—the demographic realism matters.

The remainder of this chapter is organized as follows. Section 2.2 sets out the model. Following Calvo and Obstfeld (1988) and Faruqee (2003), we assume that the mortality rate is age-dependent and solve for the optimal decision rules of the individual households.[5] We establish that the propensity to consume out of total wealth is an increasing function of the individual's age, provided the mortality rate is non-decreasing in age. Next, we postulate a constant birth rate and characterize both the population composition and the implied aggregate population growth rate associated with the demographic process. Still using the general demographic process we characterize the steady-state age-profiles for consumption, human wealth, and asset holdings.

In Section 2.3 we employ actual demographic data for the Netherlands to estimate the parameters of various demographic processes, among which the Blanchard and Gompertz-Makeham demographic models. Not surprisingly, the latter model provides by far the superior fit with the data. Interestingly, the estimated Gompertz-Makeham (G-M hereafter) model distinguishes two 'phases' of life, namely youth and old-age. During youth, Gompertz' first cause of death dominates and the mortality rate is virtually constant, but during old-age it rises exponentially with age. In our view, the G-M model is interesting for at least two reasons. First, it presents a continuous-time generalization of the Diamond (1965) model, allowing individuals to differ even within each 'phase' of life. Second, it gives rise to relatively simple analytical expressions for the propensity to consume and the steady-state age profiles for consumption, human wealth, and financial assets. In the remainder of the section we show that the G-M model also gives rise to a bell-shaped age profile for financial assets (Modigliani's life-cycle pattern).

In Section 2.4 we compute and visualize the effects on the key variables of three typical macroeconomic shocks affecting the small open economy, namely a

[5] The relationship between these papers and this chapter is as follows. Calvo and Obstfeld (1988) recognize age-dependent mortality but do not solve the decentralized model. Instead, they characterize the dynamically consistent social optimum in the presence of time- and age-dependent lumpsum taxes. Faruqee (2003) models age-specific mortality in a decentralized setting but is ultimately unsuccessful. Indeed, he confuses the cumulative density function with the mortality rate (by requiring the death rate to go to unity in the limit; see Faruqee (2003, p. 302)). Furthermore, he is unable to solve the transitional dynamics.

balanced-budget spending shock, a temporary tax cut (Ricardian equivalence experiment), and an interest rate shock. We compare and contrast the results obtained for the Blanchard and G-M models. In the second part of Section 2.4 we also present the welfare effects associated with the shocks and demonstrate that the G-M model may give rise to non-monotonic welfare effects on existing generations, something which is impossible in the Blanchard case. We conclude Section 2.4 by showing that the two models also give rise to significantly different impulse-response functions for the aggregate variables (especially for asset holdings)—the heterogeneity does not 'wash out' in the aggregate. Finally, in Section 2.5 we mention a number of possible model applications and extensions and we draw some conclusions.

2.2 The model

2.2.1 Households

Individual consumption

From the perspective of birth, the expected lifetime utility of an agent is given by:

$$\Lambda(v,v) \equiv \int_v^{\infty} [1-\Phi(\tau-v)]U[\bar{c}(v,\tau)]e^{-\theta\cdot(\tau-v)}d\tau, \tag{2.1}$$

where the first argument of $\Lambda(\cdot,\cdot)$ denotes the the birth date, the second denotes the moment of evaluation, $U[\cdot]$ is 'felicity' (or instantaneous utility), $\bar{c}(v,\tau)$ is consumption of a vintage-v agent at time τ ($\geq v$), and θ is the constant pure rate of time preference ($\theta > 0$). Intuitively, $1-\Phi(\tau-v)$ is the probability that an agent born at time v is still alive at time τ (at which time the agent's age is $\tau - v$). The instantaneous mortality rate (or death probability) of an agent of age s is given by the hazard rate of the stochastic distribution of the age of death $m(s) \equiv \phi(s)/[1-\Phi(s)]$, where $\phi(s)$ and $\Phi(s)$ denote, respectively, the density and distribution (or cumulative density) functions. These functions exhibit the usual properties, i.e. $\phi(s) \geq 0$ and $0 \leq \Phi(s) \leq 1$ for $s \geq 0$. Since, by definition, $\Phi'(s) = \phi(s)$ and $\Phi(0) = 0$, it follows that the first term on the right-hand side of (2.1) can be simplified to:

$$1-\Phi(\tau-v) = e^{-M(\tau-v)}, \tag{2.2}$$

where

$$M(\tau-v) \equiv \int_0^{\tau-v} m(s)ds \tag{2.3}$$

is the cumulative mortality factor. By using (2.2) in (2.1) we find that the utility function of a newborn agent can be written as:

$$\Lambda(v,v) \equiv \int_v^\infty U[\bar{c}(v,\tau)]e^{-\theta\cdot(\tau-v)-M(\tau-v)}d\tau. \tag{2.4}$$

As was pointed out by Yaari (1965), future felicity is discounted both because of pure time preference (as $\theta > 0$) and because of life-time uncertainty (as $M(\tau - v) > 0$).[6]

From the perspective of some later time period t ($> v$), the utility function of the agent born at time v takes the following form:

$$\Lambda(v,t) \equiv e^{M(t-v)} \int_t^\infty U[\bar{c}(v,\tau)]e^{-\theta\cdot(\tau-t)-M(\tau-v)}d\tau, \tag{2.5}$$

where the discounting factor due to life-time uncertainty ($M(\tau - v)$) depends on the age of the household at time τ.[7] The felicity function is iso-elastic:

$$U[\bar{c}(v,\tau)] = \begin{cases} \dfrac{\bar{c}(v,\tau)^{1-1/\sigma}-1}{1-1/\sigma} & \text{for } \sigma \neq 1 \\ \ln \bar{c}(v,\tau) & \text{for } \sigma = 1 \end{cases}, \tag{2.6}$$

where σ is the constant intertemporal substitution elasticity ($\sigma \geq 0$). As explained in more detail in Box 2.1, for the Blanchard model ($m(\cdot)$ constant) the choice of σ is far from innocuous in an open economy model with an exogenous interest rate. If the intertemporal substitution elasticity is too high, the model has no solution.

The household budget identity is given by:

$$\dot{\bar{a}}(v,\tau) = [r + m(\tau - v)]\bar{a}(v,\tau) + \bar{w}(\tau) - \bar{z}(\tau) - \bar{c}(v,\tau), \tag{2.7}$$

where $\bar{a}(v,\tau)$ is real financial wealth, r is the exogenously given (constant) world rate of interest, $\bar{w}(\tau)$ is the real wage rate, and $\bar{z}(\tau)$ is the lumpsum tax (the latter two variables are assumed to be independent of age and we assume that $\bar{w}(\tau) > \bar{z}(\tau)$). Labour supply is exogenous and each household supplies a single unit of labour. As usual, a dot above a variable denotes that variable's time rate of change,

[6] Yaari (1965, p. 143) attributes the latter insight to Fisher (1930, pp. 216–7).

[7] The appearance of the term $e^{M(t-v)}$ in front of the integral is a consequence of the fact that the distribution of expected remaining lifetime is not memoryless in general. Blanchard (1985) uses the memoryless exponential distribution for which $M(s) = \mu_0 s$ (where μ_0 is a constant) and thus $M(t-v) - M(\tau - v) = -M(\tau - t)$. Equation (2.5) can then be written in a more familiar format as $\Lambda(v,t) \equiv \int_t^\infty U[\bar{c}(v,\tau)]e^{-(\theta+\mu_0)(\tau-t)}d\tau$.

e.g. $\dot{\bar{a}}(v,\tau) \equiv d\bar{a}(v,\tau)/d\tau$. Following Yaari (1965) and Blanchard (1985), we postulate the existence of a perfectly competitive life insurance sector which offers actuarially fair annuity contracts to the households. Since household age is directly observable, the annuity rate of interest faced by a household of age $\tau - v$ is equal to the sum of the world interest rate and the instantaneous mortality rate of that household.[8]

Abstracting from physical capital, financial wealth can be held in the form of domestic government bonds ($\bar{d}(v,\tau)$) or foreign bonds ($\bar{f}(v,\tau)$).

$$\bar{a}(v,\tau) \equiv \bar{d}(v,\tau) + \bar{f}(v,\tau). \tag{2.8}$$

The two assets are perfect substitutes in the households' portfolios and thus attract the same rate of return.

In the planning period t, the household chooses paths for consumption and financial assets in order to maximize lifetime utility (2.5) subject to the flow budget identity (2.7) and a solvency condition, taking as given its initial level of financial assets $\bar{a}(v,t)$. The household optimum is fully characterized by:

$$\frac{\dot{\bar{c}}(v,\tau)}{\bar{c}(v,\tau)} = \sigma\cdot(r-\theta), \tag{2.9}$$

$$\Delta(u,r^*)\bar{c}(v,t) = \bar{a}(v,t) + \bar{h}(v,t), \tag{2.10}$$

$$\bar{h}(v,t) \equiv e^{ru+M(u)} \int_u^\infty [\bar{w}(s+v) - \bar{z}(s+v)] e^{-rs-M(s)} ds, \tag{2.11}$$

where $u \equiv t - v$ is the age of the household in the planning period, $r^* \equiv r - \sigma\cdot(r-\theta)$, and $\Delta(u,\lambda)$ is defined in general terms as:[9]

$$\Delta(u,\lambda) \equiv e^{\lambda u+M(u)} \int_u^\infty e^{-\lambda s-M(s)} ds, \qquad \text{(for } u \geq 0\text{)}. \tag{2.12}$$

Equation (2.9) is the 'consumption Euler equation', relating the optimal time profile of consumption to the difference between the interest rate and the pure rate of time preference. The instantaneous mortality rate does not feature in this expression because households fully insure against the adverse effects of lifetime uncertainty (Yaari, 1965). In order to avoid having to deal with a taxonomy of different cases, we restrict attention in the remainder of this chapter (and throughout Part I of this

[8] See Mitchell et al. (1999) for a discussion on the availability of 'actuarial fair' annuities.

[9] As we demonstrate below, $\Delta(u,\lambda)$ plays a very important role in the model. Proposition 2.1 covers its main properties.

dissertation) to the case of a nation populated by patient agents, i.e. $r > \theta$.[10] Equation (2.10) shows that consumption in the planning period is proportional to total wealth, consisting of financial wealth ($\bar{a}(v,t)$) and human wealth ($\bar{h}(v,t)$). The marginal (and average) propensity to consume out of total wealth equals $1/\Delta(u,r^*)$, where r^* can be seen as the 'effective' discount rate facing the consumer. Clearly, $\Delta(u,r^*)$ depends only on the household's age in the planning period and not on time itself, because r and $M(\cdot)$ are not time dependent. For future reference, Proposition 2.1 establishes the important properties of the $\Delta(u,\lambda)$ function. Finally, human wealth is defined in (2.11) and represents the market value of the unit time endowment, i.e. the present value of after-tax wage income, using the annuity rate of interest for discounting purposes. Unless after-tax wage income is time-invariant, human wealth depends on both time and on the household's age in the planning period.

Proposition 2.1. *Let $\Delta(u,\lambda)$ be defined as in* (2.12) *and assume that the mortality rate is non-decreasing, i.e. $m'(s) \geq 0$ for all $s \geq 0$. Then the following properties can be established for $\Delta(u,\lambda)$:*

(i) *decreasing in λ, $\partial\Delta(u,\lambda)/\partial\lambda < 0$;*

(ii) *non-increasing in household age, $\partial\Delta(u,\lambda)/\partial u \leq 0$;*

(iii) *upper bound, $\Delta(u,\lambda) \leq 1/[\lambda + m(u)]$ (if $\lambda + m(u) > 0$);*

(iv) *$\Delta(u,\lambda) > 0$;*

(v) *$\lim_{\lambda\to\infty} \Delta(u,\lambda) = 0$;*

(vi) *for $m'(s) > 0$ and $m''(s) \geq 0$, the inequalities in (ii)-(iii) are strict and $\lim_{u\to\infty} \Delta(u,\lambda) = 0$.*

Proof. By definition, $M(u) \equiv \int_0^u m(s)ds$ so that $M(0) = 0$, $M'(u) = m(u) > 0$, and $M''(u) = m'(u) \geq 0$. First consider $\lambda + m(u) > 0$. Since $M(s)$ is a convex function of s we have $M(s) \geq M(u) + m(u)\cdot(s-u)$ and thus:

$$\Delta(u,\lambda) \leq e^{\lambda u + M(u)} \int_u^\infty e^{-\lambda s - m(u)\cdot(s-u) - M(u)} ds = \frac{1}{\lambda + m(u)}. \tag{2.13}$$

This establishes part (iii). Part (i) follows by straightforward differentiation:

$$\frac{\partial\Delta(u,\lambda)}{\partial\lambda} = -e^{\lambda u + M(u)} \int_u^\infty (s-u) e^{-\lambda s - M(s)} ds < 0. \tag{2.14}$$

[10] The results for the other cases (with $r < \theta$ or $r = \theta$) are easily deduced from our mathematical expressions.

Similarly, part (ii) is obtained by differentiating $\Delta(u,\lambda)$ with respect to u:

$$\frac{\partial \Delta(u,\lambda)}{\partial u} = [\lambda + m(u)]\Delta(u,\lambda) - 1 \leq 0, \tag{2.15}$$

where the sign follows from (2.13). For the alternative case, with $\lambda + m(u) < 0$, (2.13) no longer holds but (2.14)–(2.15) do. For $m'(u) > 0$ the inequalities in (2.14)–(2.15) are strict. Parts (iv)–(vi) are obvious. ■

BOX 2.1

Iso-elastic felicity and intertemporal optimisation

The apparent simplicity of the iso-elastic felicity function as defined in Equation (2.6) can give quite unexpected results. Specifically, if the intertemporal elasticity of substitution is too high, optimal consumption is not defined. As an example, consider a simplified version of the model developed so far. If we look at the optimisation problem for a newborn, but impose Blanchard's perpetual youth model ($m(u) = m$) we have

$$\begin{aligned} &\max_{\bar{c}(u)} \int_0^\infty \frac{\bar{c}(s)^{1-1/\sigma} - 1}{1 - 1/\sigma} e^{-(\theta+m)s} ds \\ &\text{s.t.} \quad \dot{\bar{a}}(u) = (r+m)\bar{a} - \bar{c}(u), \\ &\qquad \bar{a}(0) = \bar{a}_0,\ \bar{a}(u) \geq 0,\ \bar{c}(u) \geq 0 \text{ for all } u. \end{aligned}$$

Where we assumed that an individual has no income, but that he receives the present value of his future after tax earnings at birth. This problem is mathematically equivalent to the infinite horizon optimal growth problem with a constant capital–output ratio, which is analysed in depth by, among others, Tinbergen (1956) (See Takayama, 1985, ch. 5, sec. D for a discussion).

Following Takayama (1985) we obtain the Euler equation (2.9), which must hold for any *feasible* Euler path of consumption. Solving the linear differential equation (2.9) gives the equivalent condition

$$\bar{c}(u) = \bar{c}_0 e^{\sigma \cdot (r-\theta)} \tag{2.16}$$

where $\bar{c}_0$ is initial consumption. Substitution into the budget identity gives for the path of assets

$$\bar{a}(u) = e^{(r+m)u}\left[\bar{a}_0 - \bar{c}_0 \int_0^u e^{-[r-\sigma\cdot(r-\theta)+m]s} ds\right] \tag{2.17}$$

The integrand is positive, so non-negativity of $\bar{a}(u)$ implies $\bar{c}_0 = \bar{a}_0/\Delta(0, r^*)$, with $\Delta(\cdot)$ defined in (2.12) and $r^* = r - \sigma\cdot(r-\theta)$, provided that the integral converges as $u \to \infty$.

This is exactly the problem. The integral diverges for $\sigma\cdot(r-\theta) \geq r + m$, so $\bar{c}_0 > 0$ is not eligible since assets will eventually become negative. This leaves just one option, $\bar{c}_0 = 0$. However, this solution implies that $\bar{c}(u) = 0$ for all u, which is the worst possible solution. The only conclusion is that the solution path for the infinite horizon problem does not exist for $\sigma\cdot(r-\theta) \geq r+m$. There are two solutions to this problem: (1) impose an (usually) arbitrary upper limit to consumption, or (2) use an upward sloping mortality function.

Upper bound on consumption Suppose that for some reason, individual consumption has an upper bound $\bar{c}_M$ (exogenous or from another part of the model). To prevent trivial solutions, assume that $\bar{c}_M$ is too high to be able to afford consuming $\bar{c}_M$ the entire life, but low enough that if the consumption profile is declining, that for young ages the upper bound is binding.

If $r > \theta$, consumption increases at rate $\sigma\cdot(r-\theta)$ as long as the upper bound is not binding. If consumption hits the upper bound, it stays at this level until the individual dies. Note that these results hold for all $r > \theta \geq 0$, there is no restriction on the parameters as in the model without an upper bound. In the linear utility model, $\sigma \to \infty$, we have a bang-bang solution. Consumption is zero until total wealth is just sufficient to finance maximum consumption indefinitely.

If $r < \theta$, consumption will start at its maximum until a certain age. After this age consumption will decrease at rate $|\sigma\cdot(r-\theta)|$. Again we have a bang-bang solution if felicity is linear in consumption. Consumption is first maximal until the consumer cannot afford to consume any more. Consumption jumps to zero and the consumer spends the rest of his life paying off his debt.

Increasing mortality Another possibility to overcome the rather strange implication that consumption is postponed forever is to assume that the mortality rate $m(u)$ increases with age and has no upper bound (this excludes the perpetual youth model). Note that these assumptions imply that the mortality rate becomes infinite, possibly for a finite age as in Boucekkine et al. (2002). The optimisation problem for a newborn becomes

$$\begin{aligned} &\max_{\bar{c}(u)} \int_0^\infty \frac{\bar{c}(s)^{1-1/\sigma}-1}{1-1/\sigma} e^{-\theta s - M(s)} ds \\ &\text{s.t.} \quad \dot{\bar{a}}(u) = [r + m(u)]\bar{a} - \bar{c}(u), \\ &\qquad\quad \bar{a}(0) = \bar{a}_0,\ \bar{a}(u) \geq 0,\ \bar{c}(u) \geq 0 \text{ for all } u. \end{aligned}$$

which gives the same Euler equation (2.16) and (2.17) becomes

$$\bar{a}(u) = e^{ru+M(u)} \left[\bar{a}_0 - \bar{c}_0 \int_0^u e^{-[r-\sigma\cdot(r-\theta)]s - M(s)} ds \right] \tag{2.18}$$

The integral in Equation (2.18) always converges as $u \to \infty$ (as long as $\sigma \ll \infty$), so we do not have the restriction on the parameters as in the Blanchard model. To show this, note that by assumption there always exists an age s^* above which the cumulative mortality rate $M(s)$ is higher than $-r^*u$, so we can split the integral in two parts, $0 \leq s \leq s^*$ and $s^* \leq s$. Boundedness of the integrand assures that the integral over $[0, s^*]$ exists and the second part exists because for $s > s^*$ the $e^{-M(s)}$ part suppresses the exponential term e^{-r^*s}, so the integral over $[s^*, \infty)$ converges.

Demography

In order to allow for non-zero population growth, we employ the analytical framework developed by Buiter (1988) which distinguishes the instantaneous mortality rate $m(s)$ and the birth rate b (> 0) and thus allows for net population growth or decline.[11] The population size at time t is denoted by $L(t)$ and the size of a newborn generation is assumed to be proportional to the current population:

$$L(v, v) = bL(v). \tag{2.19}$$

[11] The birth rate b is the crude birth rate, i.e. the number of newborns per capita. A more realistic assumption would be that only women between (approximately) 20 and 40 can give birth, but this makes the model intractable.

The size of cohort v at some later time τ is:

$$L(v,\tau) = L(v,v)[1-\Phi(\tau-v)] = bL(v)e^{-M(\tau-v)}, \tag{2.20}$$

where we have used (2.2) and (2.19). The aggregate mortality rate, $\bar{m}$, is defined by

$$\bar{m}L(t) = \int_{-\infty}^{t} m(t-v)L(v,t)dv, \tag{2.21}$$

and it is assumed that the system is in a 'demographic steady state' so that $\bar{m}$ is constant. Despite the fact that the expected remaining lifetime of each individual is stochastic, there is no aggregate uncertainty in the economy. In the absence of international migration, the growth rate of the aggregate population, n, is equal to the difference between the birth rate and the aggregate mortality rate, i.e. $n \equiv b - \bar{m}$. It follows that $L(v) = A_0 e^{nv}$, $L(t) = A_0 e^{nt}$ and thus $L(v) = L(t)e^{-n\cdot(t-v)}$. Using this result in (2.20) we obtain the generational population weights:

$$l(v,t) \equiv \frac{L(v,t)}{L(t)} = be^{-n\cdot(t-v)-M(t-v)}, \qquad \text{for } t \geq v. \tag{2.22}$$

The key thing to note about (2.22) is that the population proportion of generation v at time t only depends on the age of that generation and not on time itself.

The growth rate of the population in the demographic steady state is computed by combining (2.21) and (2.22) and simplifying:

$$\frac{1}{b} = \Delta(0,n). \tag{2.23}$$

For a given birth rate b, eq. (2.23) implicitly defines the coherent solution for n and thus for the aggregate mortality rate, $\bar{m} \equiv b - n$.[12]

Per capita household sector

Per capita variables are calculated as the integral of the generation-specific values weighted by the corresponding generation weights. For example, per capita consumption, $c(t)$, is defined as:

$$c(t) \equiv \int_{-\infty}^{t} l(v,t)\bar{c}(v,t)dv, \tag{2.24}$$

[12] For a constant mortality rate m, we have $1/\Delta(0,n) = n + m$ so that (2.23) implies $n = b - m$. Blanchard (1985) sets $b = m$ so that $n = 0$ (constant population).

where $l(v,t)$ and $\bar{c}(v,t)$ are defined in, respectively, (2.22) and (2.10) above. Exact aggregation of (2.10) is impossible because both $\Delta(u,r^*)$ and the wealth components, $\bar{a}(v,t)$ and $\bar{h}(v,t)$, depend on the generations index v. The 'Euler equation' for per capita consumption can nevertheless be obtained by differentiating (2.24) with respect to time and noting (2.9) and (2.22):

$$\dot{c}(t) = b\bar{c}(t,t) + \sigma\cdot(r-\theta)c(t) - \int_{-\infty}^{t} [n + m(t-v)]l(v,t)\bar{c}(v,t)dv. \tag{2.25}$$

Per capita consumption growth is boosted by the arrival of new generations who start to consume out of human wealth (first term on the right-hand side) and by individual consumption growth (second term). The third term on the right-hand side of (2.25) corrects for population growth and (age-dependent) mortality.[13]

Per capita financial wealth is defined as $a(t) \equiv \int_{-\infty}^{t} l(v,t)\bar{a}(v,t)dv$. By differentiating this expression with respect to t we obtain:

$$\dot{a}(t) = (r-n)a(t) + w(t) - z(t) - c(t), \tag{2.26}$$

where the wage rate $w(t) = \bar{w}(t)$, taxes $z(t) = \bar{z}(t)$, and we have used eq. (2.7) and noted the fact that newborns are born without financial assets ($\bar{a}(v,v) = 0$). The interest rate net of population growth is assumed to be positive, i.e. $r > n$. As in the standard Blanchard model, annuity payments drop out of the expression for per capita asset accumulation because they constitute transfers (via the life insurance companies) from those who die to agents who stay alive.

Finally, per capita human wealth is defined as $h(t) \equiv \int_{-\infty}^{t} l(v,t)\bar{h}(v,t)dv$ so that $\dot{h}(t)$ can be written as:

$$\dot{h}(t) = (r-n)h(t) + b\bar{h}(t,t) - w(t) + z(t). \tag{2.27}$$

In the standard Buiter model per capita human wealth is the same for all generations and accumulates at the constant annuity rate of interest ($r+m$). In contrast, in the present model the effects of the net interest rate ($r-n$) and the birth rate (b) are separate, with the former applying to per capita human wealth and the latter applying to the human wealth of newborn generations.

[13] If the mortality rate were constant, as in Blanchard (1985) and Buiter (1988), then $n \equiv b - m$ and Equation (2.25) would simplify to $\dot{c}(t) = \sigma\cdot(r-\theta)c(t) - b[c(t) - c(t,t)]$.

2.2.2 Firms, government, and foreign sector

Following Buiter (1988) we keep the production side of the model in this chapter as simple as possible by abstracting from physical capital altogether.[14] Competitive firms face the technology $Y(t) = k_0 L(t)$ where k_0 is an exogenous productivity index and $L(t)$ is the aggregate supply of labour. The real wage rate is then given by $w(t) = k_0$.

The government budget identity is given by:

$$\dot{d}(t) = (r-n)d(t) + g(t) - z(t), \tag{2.28}$$

where $d(t) \equiv \int_{-\infty}^{t} l(v,t)\bar{d}(v,t)dv$ is the per capita stock of domestic bonds, and $g(t)$ is per capita government goods consumption. The government solvency condition is $\lim_{\tau \to \infty} d(\tau)e^{[r-n][t-\tau]} = 0$, so that the intertemporal budget constraint of the government can be written as:

$$d(t) = \int_{t}^{\infty} [z(\tau) - g(\tau)] e^{-(r-n)(\tau - t)} d\tau. \tag{2.29}$$

To the extent that there is outstanding debt (positive left-hand side), it must be exactly matched by the present value of current and future primary surpluses (positive right-hand side), using the net interest rate $(r-n)$ for discounting.

Finally, the evolution of the per capita stock of net foreign assets is explained by the current account:

$$\dot{f}(t) = (r-n)f(t) + w(t) - c(t) - g(t), \tag{2.30}$$

where we have used that $y(t) \equiv Y(t)/L(t) = w(t)$ and where $f(t)$ is defined as usual as $f(t) \equiv \int_{-\infty}^{t} l(v,t)f(v,t)dv$ denotes the per capita stock of net foreign bonds in the hands of domestic households.

2.2.3 Steady-state equilibrium

It is relatively straightforward to characterize the steady state of the model. The steady-state values for all variables are designated by means of a hat overstrike, e.g. $\hat{c}$ is steady-state per capita consumption. Where no confusion can arise, the

[14] In the context of a small open economy with firms facing convex investment adjustment costs, our approach does not entail much loss of generality because the investment and savings systems decouple in that case. See Matsuyama (1987); Bovenberg (1993, 1994); Heijdra and Meijdam (2002) and Heijdra and van der Ploeg (2002, pp. 571-581).

time index is also suppressed. Since technology is held constant, the steady-state wage rate is time-invariant, i.e. $w(t) = \hat{w} = k_0$. If the government variables are also held constant, so that $z(t) = \hat{z}$, $g(t) = \hat{g}$, and $d(t) = \hat{d} \equiv (\hat{z} - \hat{g})/(r - n)$, then the economy settles into a unique saddle-point stable steady-state equilibrium in which $c(t) = \hat{c}$, $h(t) = \hat{h}$, $a(t) = \hat{a}$, and $f(t) = \hat{f}$.[15]

In the steady-state equilibrium, all variables applying to individuals can be re-written solely in terms of their age, $u \equiv t - v$ (as is also the case outside the steady state for $\Delta(u, r^*)$—see eq. (2.12) above). After some straightforward substitutions we find:

$$\hat{\hat{h}}(u) \equiv \hat{\hat{h}}(v,t) = (\hat{w} - \hat{z})\Delta(u,r), \tag{2.31}$$

$$\hat{\hat{c}}(u) \equiv \hat{\hat{c}}(v,t) = \frac{\hat{\hat{h}}(0)}{\Delta(0,r^*)} e^{\sigma\cdot(r-\theta)u}, \tag{2.32}$$

$$\hat{\hat{a}}(u) \equiv \hat{\hat{a}}(v,t) = \Psi(u,r,r^*)\hat{\hat{h}}(0), \tag{2.33}$$

where $r^* \equiv r - \sigma\cdot(r - \theta)$, $\Delta(u, \lambda)$ is defined in eq. (2.12), and $\Psi(u,r,r^*)$ is given by:

$$\Psi(u,r,r^*) \equiv e^{ru+M(u)} \left[\frac{\int_u^\infty e^{-r^* s - M(s)} ds}{\Delta(0,r^*)} - \frac{\int_u^\infty e^{-rs-M(s)} ds}{\Delta(0,r)} \right]. \tag{2.34}$$

Deferring the economic intuition behind (2.31)–(2.33) to Section 2.3.2, we simply note that human wealth is positive (since $\hat{w} > \hat{z}$) and proportional to $\Delta(u,r)$, the properties of which are covered in Proposition 2.1. Human wealth at birth is an important determinant for the age profiles for both consumption and financial assets. In the absence of initial financial wealth (e.g. received bequests), $\hat{\hat{h}}(0)$ is the key 'initial condition' facing agents. Consumption rises monotonically with age but the age profile of financial assets depends critically on the demographic process, i.e. on the $\Psi(u,r,r^*)$ function. The main properties of this function are stated in Lemma 2.1. If rate of time preference θ equals the interest rate r, than individuals do not save or borrow (Lemma 2.1(i)). With a constant mortality rate, financial wealth rises monotonically with age (Lemma 2.1(iv)). When the mortality rate increases with age, however, the assets are positive and increasing early on in life, but return to zero at higher ages provided the condition in Lemma 2.1(iii) is satisfied. For the general case, the asset profile may display multiple peaks though there is only a

[15] Saddle-point stability follows trivially from the fact that all agents in the economy satisfy their respective solvency conditions. Consumption and human wealth are forward-looking variables (able to feature discrete jumps) whilst total financial assets and net foreign assets are predetermined (non-jumping) variables.

single peak for the empirically most relevant G-M model studied in Section 2.3.2 below.

Lemma 2.1. *Let $\Psi(u,r,r^*)$ be defined as in* (2.34) *and note that* $r^* = r \Leftrightarrow \sigma\cdot(r-\theta) = 0$. *The following properties can be established for $\Psi(u,r,r^*)$:*

(i) $\Psi(u,r,r) = 0$ *for all* $u \geq 0$*;*

(ii) for $r > r^*$, $\Psi(u,r,r^*) \geq 0$ *with the equality sign only holding for* $u = 0$*;*

(iii) if $r > r^*$ *then* $\lim_{u\to\infty} \Psi(u,r,r^*) = 0$ *if and only if* $\lim_{u\to\infty} e^{(r-r^*)u}/[r+m(u)] = 0$*;*

(iv) if $m(u) = m_0$ *(Blanchard) then* $\Psi(u,r,r^*) \equiv e^{\sigma\cdot(r-r^*)u} - 1$ *is a strictly increasing function in* u*.*

Proof. We denote the term in square brackets on the right-hand side of (2.34) by $\Omega(u,r,r^*)$ and note that, $\Omega(0,r,r^*) = \lim_{u\to\infty} \Omega(u,r,r^*) = 0$. Taking the derivative with respect to u we find:

$$\Omega'(u,r,r^*) = e^{-M(u)} \left[\frac{e^{-ru}}{\Delta(0,r)} - \frac{e^{-r^*u}}{\Delta(0,r^*)} \right], \tag{2.35}$$

which clearly has a single root (at $\bar{u} \equiv 1/(r-r^*) \ln\left(\Delta(0,r)/\Delta(0,r^*)\right) > 0$) and satisfies $\Omega'(0,r,r^*) > 0$ (for $r > r^*$). This in combination with continuity of $\Omega(u,r,r^*)$ shows that $\Omega(u,r,r^*) > 0$ for $u > 0$ (and $r > r^*$). Since $\Psi(u,r,r^*) \equiv e^{ru+M(u)}\Omega(u,r,r^*)$, this proves part (ii).

To show part (iii), rewrite $\lim_{u\to\infty} \Psi(u,r,r^*)$ and use l'Hopital's rule

$$\begin{aligned}
\lim_{u\to\infty} \Psi(u,r,r^*) &= \lim_{u\to\infty} \left\{ \frac{1}{\Delta(0,r^*)} \frac{\int_u^\infty e^{-r^*s-M(s)}ds}{e^{-ru-M(u)}} - \frac{1}{\Delta(0,r)} \frac{\int_u^\infty e^{-rs-M(s)}ds}{e^{-ru-M(u)}} \right\} \\
\ldots &= \lim_{u\to\infty} \left\{ \frac{1}{\Delta(0,r^*)} \frac{e^{-r^*u-M(u)}}{[r+m(u)]e^{-ru-M(u)}} - \frac{1}{\Delta(0,r)} \frac{e^{-ru-M(u)}}{[r+m(u)]e^{-ru-M(u)}} \right\} \\
\ldots &= \lim_{u\to\infty} \left\{ \frac{1}{\Delta(0,r^*)} \frac{e^{(r-r^*)u}}{r+m(u)} - \frac{1}{\Delta(0,r)} \frac{1}{r+m(u)} \right\},
\end{aligned}$$

from which (iii) follows immediately. Part (iv) is obvious. ■

Simple expressions for the steady-state per capita variables can also be found:

$$\hat{c} = \hat{\bar{c}}(0) \frac{\Delta(0,n^*)}{\Delta(0,n)}, \tag{2.36}$$

$$\hat{h} = \frac{\hat{w} - \hat{z}}{r - n} \left[1 - \frac{\Delta(0, r)}{\Delta(0, n)} \right] \tag{2.37}$$

$$\hat{a} \equiv \hat{d} + \hat{f} = \frac{\hat{w} - \hat{z}}{r - n} \left[\frac{\Delta(0, r)}{\Delta(0, r^*)} \frac{\Delta(0, n^*)}{\Delta(0, n)} - 1 \right], \tag{2.38}$$

where $n^* \equiv n - \sigma \cdot (r - \theta)$ and the term in square brackets on the right-hand side of (2.38) is positive. Not surprisingly, per capita consumption exceeds consumption by newborns (because $n > n^*$ so that $\Delta(0, n^*) > \Delta(0, n)$), and both per capita human and financial wealth are positive.

Armed with these expressions it is straightforward to derive the long-run effects of various shocks impacting the economy.[16] A *balanced-budget increase in government consumption* ($d\hat{z} = d\hat{g} > 0$) leads to a decrease in steady-state human wealth and consumption for all cohorts:

$$\frac{d\hat{\bar{h}}(u)}{d\hat{z}} = -\Delta(u, r) < 0, \tag{2.39}$$

$$\frac{d\hat{\bar{c}}(u)}{d\hat{z}} = -\frac{\Delta(0, r)}{\Delta(0, r^*)} e^{\sigma \cdot (r - \theta) u} < 0. \tag{2.40}$$

Obviously, per capita steady-state consumption and human wealth also fall (see eqs (2.36) and (2.37)). It follows from (2.38) that per capita steady-state financial assets decline:

$$\frac{d\hat{a}}{d\hat{z}} = \frac{1}{r - n} \left[1 + \frac{d\hat{c}}{d\hat{z}} \right] < 0. \tag{2.41}$$

This implies that consumption is crowded out more than one for one. Finally, since government debt is unchanged (by design) it follows from the first equality in (2.38) that $d\hat{f}/d\hat{z} = d\hat{a}/d\hat{z}$. The balanced-budget increase in government consumption thus leads to a long-run reduction in financial assets and a reduction in net imports, just as in the standard open-economy Blanchard model with $r > \theta$ (1985, p. 230-1).

A *long-run tax-financed increase in public debt* ($[r - n] d\hat{d} = d\hat{z} > 0$) leads to a decrease in generation-specific and per capita steady-state consumption and human wealth (see (2.39)–(2.40)). It follows from (2.38) that:

$$(r - n) \frac{d\hat{f}}{d\hat{z}} \equiv -[r - n] \frac{d\hat{d}}{d\hat{z}} + \frac{d\hat{c}}{d\hat{z}} + 1 = \frac{d\hat{c}}{d\hat{z}} < -1. \tag{2.42}$$

As in the standard Blanchard model (with $r > \theta$), government debt more than displaces foreign assets in the households' portfolios (1985, p. 242).

[16] The impact and transitional effects of these shocks are studied in Section 2.4.

An *increase in the world interest rate* leads to higher discounting of after-tax wages and a reduction in both individual and aggregate human wealth:

$$\frac{d\hat{\bar{h}}(u)}{dr} = [\hat{w} - \hat{z}]\frac{\partial\Delta(u,r)}{\partial r} < 0, \tag{2.43}$$

$$\frac{d\hat{h}}{dr} = \int_0^\infty l(u)\frac{d\hat{\bar{h}}(u)}{dr}du < 0, \tag{2.44}$$

where we have used Proposition 2.1(i) to establish the sign in (2.43). The interest elasticity of individual consumption is given by:

$$\frac{r}{\hat{\bar{c}}(u)}\frac{d\hat{\bar{c}}(u)}{dr} = \frac{r}{\hat{\bar{h}}(0)}\frac{d\hat{\bar{h}}(0)}{dr} + r\sigma u - (1-\sigma)\frac{r}{\Delta(0,r^*)}\frac{\partial\Delta(0,r^*)}{\partial r^*}. \tag{2.45}$$

The effect on individual consumption is ambiguous in general because it results from the interplay of three effects, namely the (initial) human-wealth effect (HWE), the consumption-growth effect (CGE), and the (initial) consumption-propensity effect (CPE). The HWE is represented by the first term on the right-hand side of (2.45) and is negative as after-tax income is discounted more heavily. The CGE effect (the second term on the right-hand side) is positive and increasing in the household's age. An increase in the interest rate causes agents to adopt a steeper age-profile for consumption. Finally, the third term on the right-hand side represents the CPE, i.e. the effect of the interest rate change on a newborn's propensity to consume, $1/\Delta(0,r^*)$. In the empirically plausible case, with $\sigma < 1$, the CPE is positive, thus partially offsetting the negative HWE. For the case with a logarithmic felicity function, which we focus on from Section 2.3.2 onward, $\sigma = 1$ and the CPE is zero ($\Delta(0,r^*) = \Delta(0,\theta)$ in that case).

The effect on per capita consumption can be written as:

$$\frac{r}{\hat{c}}\frac{d\hat{c}}{dr} = \frac{r}{\hat{\bar{c}}(0)}\frac{d\hat{\bar{c}}(0)}{dr} - \sigma\frac{r}{\Delta(0,n^*)}\frac{d\Delta(0,n^*)}{dn^*}. \tag{2.46}$$

and is thus also ambiguous in general. The sign of first term on the right-hand side is ambiguous for $\sigma < 1$, because the HWE is negative and the CPE is positive (see (2.45)). For the logarithmic case ($\sigma = 1$), however, the first term must be negative. Since the second term on the right-hand side is positive, it is nevertheless possible for per capita consumption to rise (as is the case in the simulations performed in Section 2.4). Finally, the effect on individual and per capita assets is ambiguous for the general specification of the model.

2.3 Demography

As was stressed by Blanchard (1985, p. 223), exact aggregation of the consumption function is generally impossible because both the propensity to consume (our $1/\Delta(u, r^*)$) and the wealth components (our $\bar{a}(v,t)$ and $\bar{h}(v,t)$) are age dependent. Blanchard cuts this Gordian knot by assuming the mortality rate to be constant, i.e. $m(s) = \mu_0 > 0$ so $M(u) = \mu_0 u$. The advantages of his approach are its simplicity and flexibility—the expected remaining planning horizon is $1/\mu_0$ so, by letting $\mu_0 \to 0$, the infinite-horizon Ramsey model is obtained as a special case. The main disadvantage of the Blanchard approach is that it cannot capture the life-cycle aspect of consumption behaviour. In addition, the perpetual youth assumption is easily refuted empirically as it runs foul of the Gompertz-Makeham law of mortality (Preston et al. (2001) and Section 2.3.1 below).

In the context of a small open economy, however, it is quite feasible to incorporate a realistic demographic structure because the aggregation step is not necessary. Since both the interest rate and the wage rate are exogenous, the macroeconomic equilibrium can be studied directly at the level of individual households (see Sections 2.3.2 and 2.4).

2.3.1 Estimates

In this section we estimate the survival function $(1 - \Phi(\tau - v))$ by using actual demographic data for the Netherlands taken from the Human Mortality Database (2006). We will use these estimates throughout this and the following two chapters. The data are annual and apply to the population cohort born in 1920. Actual mortality figures are available up to 2003, implying that demographic projections have only been used to compute the survival probabilities for the age range 84–105.[17] Denoting the actual surviving fraction up until age u_i of the people born in 1920 by $S(u_i)$, we estimate the parameters of a given parametric distribution function by means of non-linear least squares. Denoting the parameter vector by $\boldsymbol{\mu}$, the model to be estimated is thus:

$$S(u_i) = 1 - \Phi(u_i, \boldsymbol{\mu}) + \varepsilon_i = e^{-M(u_i, \boldsymbol{\mu})} + \varepsilon_i, \tag{2.47}$$

[17] Child mortality was still a real issue in the 1920s—almost 11 percent of the 1920 cohort died during their first year. Since it is not the phenomenon that we wish to focus on, we adjust the mortality figures by assuming the death probability for ages 0–14 to be equal to that of a 15 year old. This takes out the downward sloping segment of the mortality function at the start of life.

where $M(u_i, \boldsymbol{\mu}) = \int_0^{u_i} m(s, \boldsymbol{\mu}) ds$ and ε_i is the stochastic error term. The estimates are reported in Table 2.1 for various specifications of the mortality process. In that table, $\hat{\sigma}$ is the estimated standard error of the regression, $\bar{m}$ is the average mortality rate, the t-statistics are given in round brackets below the estimates, and $1 - \widehat{\Phi(100)}$ represents the estimated proportion of centenarians.

We consider five different functional forms for the instantaneous mortality rate and the associated $M(u_i)$ functions. The Blanchard model based on a *constant* mortality rate (model 1) yields an estimated mortality rate of 0.7% per annum and displays the worst fit of all cases considered–the estimated standard error is 0.22 which far exceeds the standard errors for the other models.

The second and third models are the linear and piecewise linear mortality rates models. The linear model is based on the notion that the mortality rate increases with age. This *linear-in-age* model with the mortality rate defined by $m(u) = \mu_0 + 2\mu_1^2 u$ fits a little better than the constant mortality rate model but it predicts a negative mortality rate for newborns. Constraining the constant to zero gives a standard error of 0.13, better than the Blanchard model, but still quite high. A combination of the Blanchard model and the linear model, the piecewise linear model fits the data much better with $\sigma = 0.024$. According to the piecewise linear model, a human life can be divided in two parts; for young people mortality is constant and low, for old people mortality increases linearly with age. The mortality function can be written as

$$m(u) \equiv \begin{cases} \mu_0 & \text{for } 0 < u < \bar{u} \\ \mu_0 + 2\mu_1^2(u - \bar{u}) & \text{for } u \geq \bar{u} \end{cases} \tag{2.48}$$

For the 1920-generation, the kink in the mortality profile lies around the age of 55. Below this age, the first cause of death dominates, beyond this age, biological wear and tear starts to increase the probability of death.

The fourth model we estimate is the mortality process used by Boucekkine et al. (2002). Their proposed survival law follows

$$S(u) = 1 - \Phi(u) = \frac{e^{-\beta u} - \alpha}{1 - \alpha}, \quad 0 \leq u \leq A \tag{2.49}$$

with $A = -\frac{1}{\beta} \ln \alpha$. Beyond age A no one is supposed to be alive, it is the maximum attainable age. For the estimated parameter values the maximum age is 86 years. This immediately shows the weakness of this model. Although the fit is quite good,

$\hat{\sigma}$ is 0.016, the model predicts that nobody from the 1920 generation will survive this year (2006), however, about 8% is still alive.[18]

Finally, the last model postulates the instantaneous mortality rate to follow the Gompertz Makeham process:

$$m(u) \equiv \mu_0 + \mu_1 e^{\mu_2 u}, \tag{2.50}$$

with $\mu_i > 0$. As Table 2.1 shows, the parameter estimates are all highly significant. The standard deviation is very small and the model features a realistic prediction for the fraction of centenarians (0.1% rather than the unrealistic prediction of almost 32% for the Blanchard model). This model does not suffer from the doubtfull maximum attainable age of the previous model. The G-M model has no maximum age, but mortality rates increase exponentianally, so at very old ages, it is highly unlikely to survive another year. This closely resembles the idea of Gavrilov and Gavrilova (1991) who argued that people die before the age of infinity, not because they cannot pass bounding age, but because the probability of a person avoiding the ever-present risk of death for that long is infinitesimal.

In the top panel of Figure 2.1 we illustrate data points at five-year intervals (stars) as well as the estimated survival functions for the five models. The poor fit of the Blanchard model is confirmed–the surviving fraction is underestimated up to about age 73 and overestimated thereafter. In contrast, the G-M model tracks the data quite well. Another way to visualize the difference between the two models makes use of their predicted mortality rates (middle panel) and expected remaining lifetimes (bottom panel of Figure 2.1). After about age 60, the mortality rate of the G-M model rises exponentially with age. The estimated G-M model thus distinguishes two phases of life, namely 'youth', lasting until about age 60, and 'old age' thereafter. Of course, for the Blanchard model expected remaining lifetime is constant (and equal to 87 years) so the agent enjoys a 'perpetual youth'.

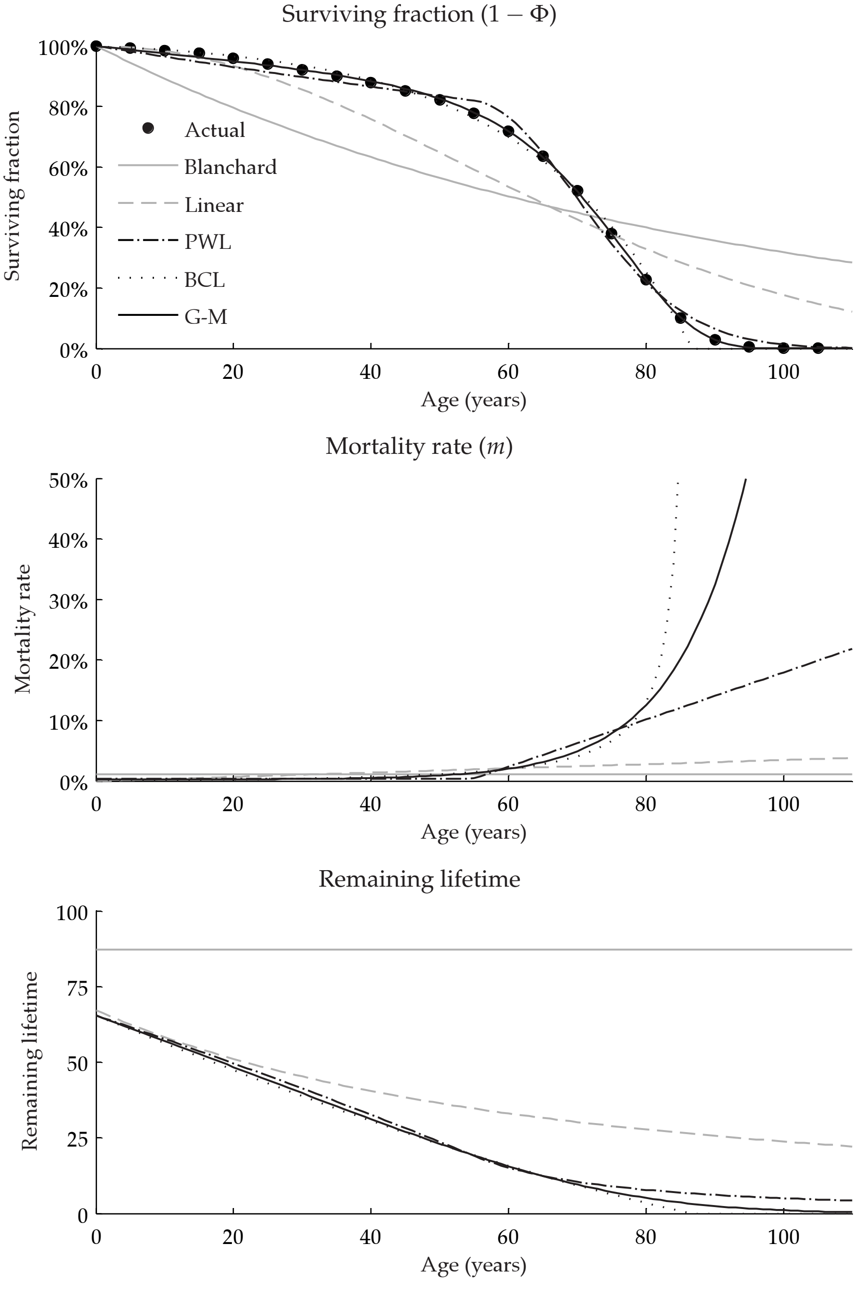
[18] There is a large literature on the maximum length of life. An interesting overview of this literature is given by Kirkwood (2001) who states: 'The truth is that the idea of a fixed limit to human longevity was always a little questionable but it is only now, as understanding of the ageing process improves, that the reason has become apparent. There is no mechanism that measures man's span of time and then activates a destructive process. In fact, quite the reverse is true and nearly every system in the body does its best to preserve life. Even apoptosis is directed mostly at protecting the body by deleting cells that might cause harm. These systems are not perfect, however, and ageing occurs because myriad tiny faults accumulate. Eventually the viability of various organs is compromised, the weakest link is revealed, and so goodbye.'

Table 2.1. Estimated survival functions

1. *Blanchard:*	$M(u) \equiv \mu_0 u$		
$\hat{\sigma} = 0.2213$	$\hat{\mu}_0$		
$\bar{m} = 1.15\%$	0.01147		
$1 - \widehat{\Phi(100)} = 31.8\%$	(14.3)		
2. *Linear:*	$M(u) \equiv \mu_0 u + \mu_1^2 u^2$		
$\hat{\sigma} = 0.1312$	$\hat{\mu}_0$	$\hat{\mu}_1$	
$\bar{m} = 1.11\%$	–	0.0132	
$1 - \widehat{\Phi(100)} = 17.6\%$	–	(42.24)	
3. *Piecewise linear (PWL):*	$M(u) \equiv \begin{cases} \mu_0 u & \text{for } 0 < u < \bar{u} \\ \mu_0 u + \mu_1^2 (u - \bar{u})^2 & \text{for } u \geq \bar{u} \end{cases}$		
$\hat{\sigma} = 0.0243$	$\hat{\mu}_0$	$\hat{\mu}_1$	$\hat{\bar{u}}$
$\bar{m} = 1.04\%$	3.63×10^{-3}	0.0441	54.8
$1 - \widehat{\Phi(100)} = 1.3\%$	(32.14)	(37.74)	(97.61)
4. *Boucekkine et al. (2002):*	$M(u) \equiv \ln\left(\frac{1-\mu_0}{e^{-\mu_1 u} - \mu_0}\right), \quad 0 < u < \ln\left(-\frac{\mu_0}{\mu_1}\right)$		
$\hat{\sigma} = 0.0162$	$\hat{\mu}_0$	$\hat{\mu}_1$	
$\bar{m} = 1.01\%$	41.06	-0.0429	
$1 - \widehat{\Phi(100)} = 0.0\%$	(23.711)	(-78.84)	
5. *Gompertz-Makeham (G-M):*	$M(u) \equiv \mu_0 u + (\mu_1 / \mu_2)[e^{\mu_2 u} - 1]$		
$\hat{\sigma} = 4.852 \times 10^{-3}$	$\hat{\mu}_0$	$\hat{\mu}_1$	$\hat{\mu}_2$
$\bar{m} = 1.02\%$	2.437×10^{-3}	5.52×10^{-5}	0.0964
$1 - \widehat{\Phi(100)} = 0.1\%$	(65.8)	(20.5)	(138.2)

Notes: All function fitted to data for the cohort born in the Netherlands, 1920 (male and female). Observed survival rates for ages 0–85, projected survival rates otherwise. To correct for child mortality, the death probability for ages 0–14 is assumed to be equal to that of a 15 year old. t-statistics between brackets, $\hat{\sigma}$ is the standard deviation, $1 - \widehat{\Phi(100)}$ is the predicted fraction of centenarians. *Source:* Human Mortality Database (2006) and own calculations.

Figure 2.1. Actual and estimated survival rates



Note: All survival functions are fitted to the cohort born in the Netherlands in 1920 (see Table 2.1 for parameter values).

2.3.2 Steady-state profiles

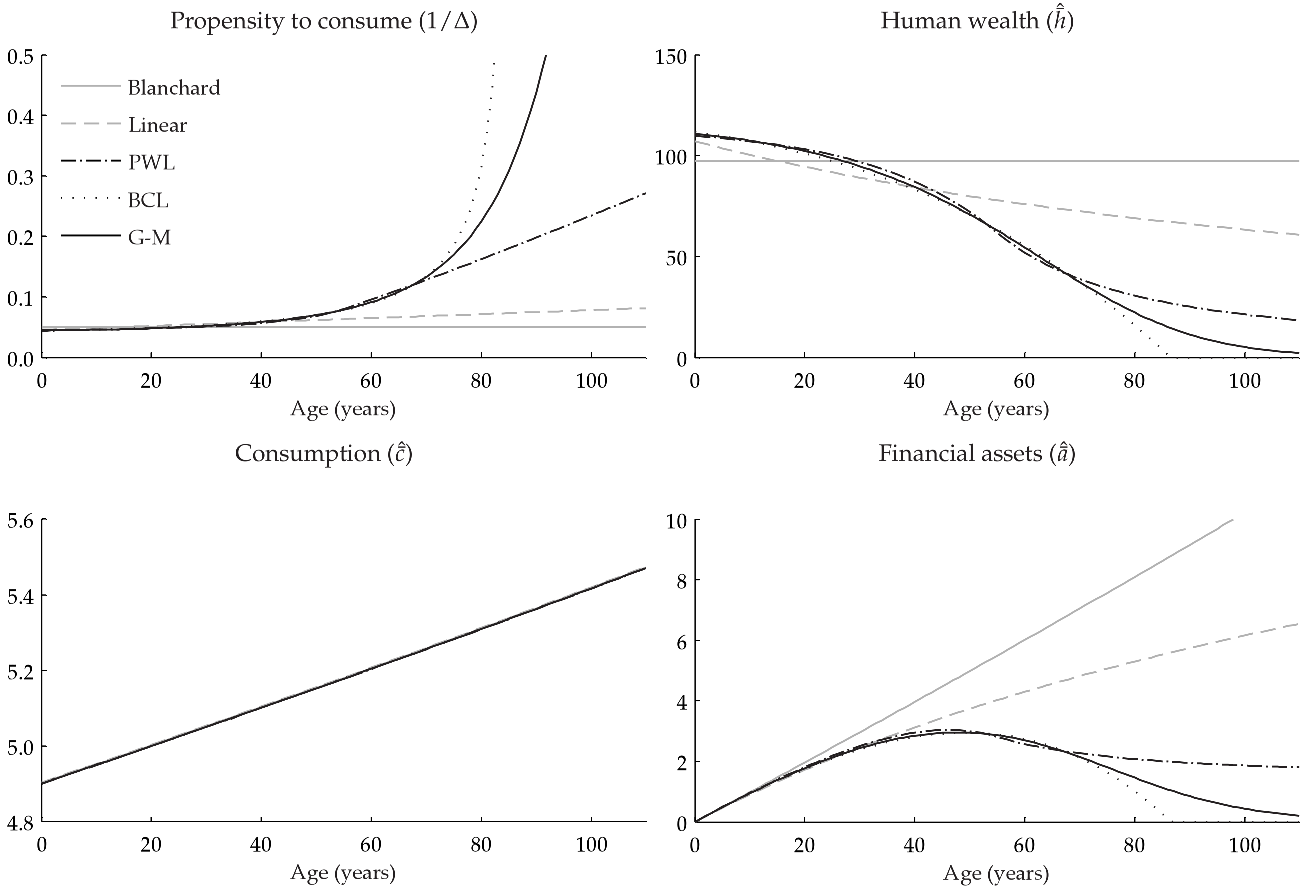
In Figure 2.2 we visualize (for all estimated models) the steady-state age profiles for the propensity to consume, human wealth, consumption, and financial assets. The analytical expressions for these variables are given in, respectively, eqs (2.12), (2.31), (2.32), and (2.33). In order to avoid a taxonomy of cases, we restrict attention to the 'unit-elastic case' in the remainder of this chapter, i.e. we set the intertemporal substitution elasticity equal to one ($\sigma = 1$). This implies that $r^* = \theta$ so that the marginal propensity to consume out of total wealth equals $1/\Delta(u, \theta)$ and is thus independent of the interest rate.

Clearly the $\Delta(u, \lambda)$ function (defined in (2.12)) plays a key role in the model. Fortunately, for all demographic specifications, easily computed closed-form solutions for $\Delta(u, \lambda)$ can be derived. Indeed, for the Blanchard model it reduces to $\Delta(u, \lambda) = 1/(\lambda + \mu_0)$ and is thus independent of the age of the household. We show in Appendix 2.A that the solution for the linear model and the piecewise linear model can be written in terms of the error function and the G-M model can be written in terms of the upper-tailed incomplete gamma function (Kreyszig, 1999, p. A55). Not surprisingly, since all models satisfy the assumptions stated in Proposition 2.1, it follows that the marginal propensity to consume, $1/\Delta(u, \theta)$, increases with age, except for the Blanchard model, for which the marginal propensity to consume is constant (the borderline case in Proposition 2.1(ii) and (iii)). This is confirmed in the top left-hand panel of Figure 2.2.

In the top right-hand panel of Figure 2.2 the age profile for steady-state human wealth is plotted.[19] For the standard Blanchard model the annuity rate of interest is age-independent because the mortality rate is constant. As a result, human wealth is age-independent also. In stark contrast, for the G-M model the annuity rate of interest rises with age so that discounting of after-tax wage income is heavier the older the household is. Human wealth gradually falls with age as a result. Indeed, it follows from (2.31) that $\hat{\bar{h}}(u)$ is proportional to $\Delta(u, r)$ which is downward sloping in u for any demography with a non-decreasing mortality rate (see Proposition 2.1). Exploiting the proportionality between $\hat{\bar{h}}(u)$ and $\Delta(u, r)$, we find that the slope of

[19] Following Gardia (1991, p. 423) we set $r = 0.04$ and $\theta = r^* = 0.039$. We interpret the G-M demography as the truth and choose b such that $n = 0.0134$ (the average population growth rate during the period 1920–1940). This yields a value of $b = 0.0236$ (which falls in between the observed birth rates for 1920 ($= 0.028$) and 1940 ($= 0.02$)). The estimated G-M model yields an expected remaining lifetime at birth of 65.5 years, which is very close to the value used by Cardia ($= 67$). Finally, for the unimportant scaling variables we use $w = 5$ and $z = 0$. The simulation results are quite robust for different parameter values.

Figure 2.2. Steady-state profiles for individuals



Note: All survival functions are fitted to the cohort born in the Netherlands in 1920 (see Table 2.1 for parameter values).

the human wealth profile is given by:

$$\frac{d\hat{\hat{h}}(u)}{du} = (\hat{w} - \hat{z})\big([r + m(u)]\Delta(u,r) - 1\big) < 0, \tag{2.51}$$

where the term in square brackets on the right-hand side of Equation (2.51) is equal to $\partial\Delta(u,r)/\partial u$. During the early phase of life, for the linear, the piecewise linear and the G-M model, the annuity rate $r + m(u)$ is relatively low, $\Delta(u,r)$ is relatively high, and human wealth falls only slightly as young agents are still on the flat part of the mortality curve. At high ages, $r + m(u)$ is high, $\Delta(u,r)$ is low, and $d\hat{\hat{h}}(u)/du$ is again relatively low. These models thus give to an inverse-S-shaped profile for human wealth with a point of inflexion located at the approximate age of 55.

In the bottom left-hand panel of Figure 2.2 the age profile of steady-state consumption is visualized. As follows readily from eq. (2.32), the growth rate of individual consumption is the same for both demographic models. Interestingly, the estimated mortality models all predict very similar steady-state consumption paths (in level terms).

Finally, in the bottom right-hand panel of Figure 2.2 the age profile of steady-state financial assets is visualized. For the Blanchard model financial assets rise with age—see Lemma 2.1(iv). Matters are vastly different for the G-M model. Indeed, for that model financial asset holdings follow the classic life-cycle pattern stressed by Modigliani and co-workers, i.e. households start life with zero assets, then save up until middle age, after which dissaving takes place. Despite the fact that very old agents have hardly any financial assets or human wealth left, the annuity rate of interest is so high for them that a high consumption level can nevertheless be maintained.[20] The results for the piecewise linear model is in between the Blanchard model and the G-M model. For all normal ages, the asset profile shows a humped shape, but after 140 years, assets start to increase again. The asset profile for BCL mortality process hits zero at the maximum age 86 because no one wants to die with assets (we abstract from bequest motives).

The upshot of the discussion so far is as follows. The Blanchard specification tracks the demographic data very poorly and predicts unrealistic age patterns for the consumption propensity, human wealth, and financial wealth. In contrast, the G-M model tracks the data rather well and predicts the relevant life-cycle patterns in these variables. A further theoretical advantage of the G-M model is that it en-

[20] Only the estimated G-M model satisfies the condition stated in Lemma 2.1(iii) so that assets go to zero as the agent gets very old. In addition, the model gives rise to a single peak in the asset profile, a result we have been unable to prove analytically in general.

ables a conceptual distinction between youth and old age (just as is possible in the two-period Diamond (1965) model). The other models have some characteristics of both the Blanchard model and the G-M model, but none of them predicts an actual life-cycle pattern in assets. For this reason, and because the G-M model tracks observed survival rates best, we will focus on these two models from now onward.

2.4 Visualizing shocks with realistic demography

In this section we compute and visualize the effects on the different variables of a number of prototypical shocks affecting a small open economy at time $t = 0$.[21] For reasons mentioned above and to prevent an taxonomy of different models, we will focus on the Blanchard model and the G-M model from now on.

2.4.1 Shocks

Balanced-budget fiscal policy

The first shock consists of an unanticipated and (believed to be) permanent increase in government consumption which is financed by means of lumpsum taxes (i.e. $d\hat{g} = d\hat{z} > 0$). The effects of this shock on individual human wealth and financial assets are illustrated in Figure 2.3. In that figure, the left-hand panels depict the Blanchard case whilst the right-hand panels illustrate the results for the G-M model.

In the Blanchard case, the increase in the lumpsum tax causes a once-off decrease in human wealth which is the same for all existing and future generations. In stark contrast, in the G-M model the fall in human wealth depends both on time and on the generations index. The top right-hand panel of Figure 2.3 shows the effects for two existing households (aged, respectively, 40 and 20 at the time of the shock) and two future households (born respectively one second and 40 years after the shock). As a result of the shock there is a once-off change in the age profile of human wealth. This profile itself does not depend on time because there is no transitional dynamics in after-tax wages (see eq. (2.31) above).

In the bottom two panels of Figure 2.3 the paths for financial assets are illustrated. In the Blanchard case these assets rise monotonically over time for each household. The shock induces a slight kink (at time $t = 0$) in the profile for each generation. For the G-M model in the right-hand panel, the crowding-out effect due to the tax increase is much more visible. The peak in financial asset holdings

[21] These shocks do not have to be infinitesimal as no linearisation techniques have been used.

is higher, the older the existing household is (compare, for example, the 40 and 20 years old households). The profiles for the future households born, respectively, in 0 and 40 years time are identical in shape. Again, this is because of the lack of transitional dynamics in after-tax wages, i.e. in terms of eq. (2.33) the effect operates entirely via steady-state human wealth at birth for post-shock generations.

Temporary tax cut

The second shock consists of a typical Ricardian equivalence experiment. At impact ($t = 0$), the lumpsum tax is reduced, which is financed by issuing bonds. As a result, the stock of government debt gradually increases over time. In order to ensure that government solvency is maintained, the tax is gradually increased over time and ultimately rises to a level higher than in the initial situation. The shock that is administered thus takes the following form:

$$dz(t) = -dz_0 e^{-\chi t} + d\hat{z}\left(1 - e^{-\chi t}\right), \qquad (\text{for } t \geq 0), \tag{2.52}$$

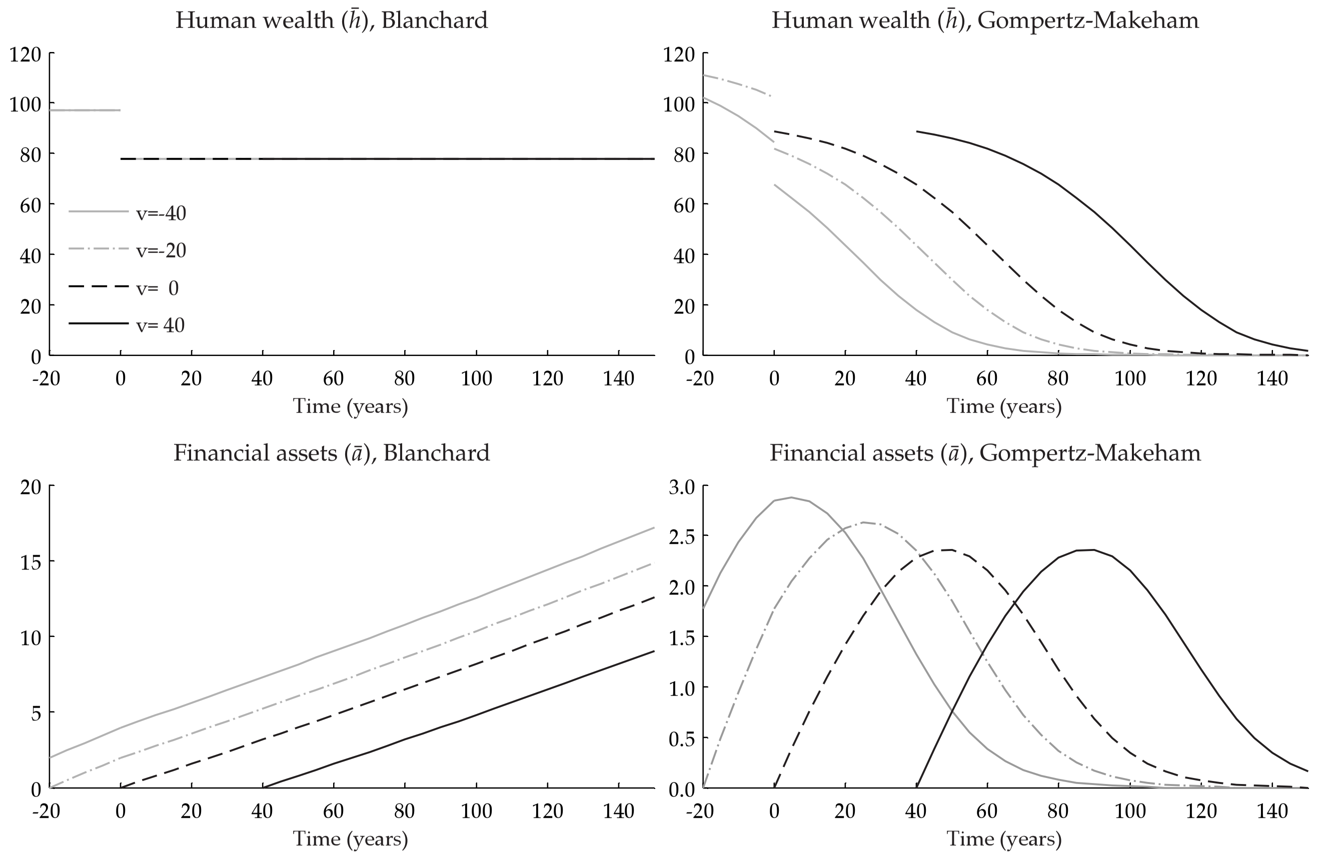
where $0 < \chi \ll \infty$, $dz_0 > 0$, and $d\hat{z} = [(r-n)/\chi]dz_0 > 0$. At impact, the lumpsum tax falls by dz_0 but in the long run it rises by $d\hat{z}$. (The long-run effect on public debt equals $d\hat{d} = dz_0/\chi > 0$.) In the simulations, the persistence parameter is set at $\chi = 0.1$ implying that the tax reaches its pre-shock level after about 15 to 16 years.

The effects on human and financial wealth are illustrated for the two cases in Figure 2.4. In the Blanchard case, human wealth is age-independent. It nevertheless features transitional dynamics because the path of lumpsum taxes is time dependent. Human wealth increases at impact (because of the tax cut), but during transition it gradually falls again (because of the gradual tax increase). In the long run, the permanently higher taxes (needed to finance interest payments on accumulated debt) ensure that human wealth is less than before the shock.

In the G-M model, the effect on human wealth is both time- and age-dependent. At impact, all existing households experience an increase in their human wealth because of the tax cut. For each household, human wealth declines during transition both because of ageing (gradual increase in the annuity rate of interest) and because the tax rises over time. For the future household born 40 years after the shock, the human wealth profile is virtually in the new steady state as most of the shock has worn out by then.

In the bottom panels of Figure 2.4 the profiles for financial assets are illustrated. In the Blanchard case the tax cut causes an acceleration in asset accumulation at

Figure 2.3. Balanced-budget fiscal policy



Note: Both survival functions are fitted to the cohort born in the Netherlands in 1920 (see Table 2.1 for parameter values). Birth rate is 2.36%. At time $t = 0$ lumpsum taxes and government consumption jump from 0 to 1 (20% of gross wage income).

impact. This kink also occurs for the G-M model in the bottom right-hand right panel. The G-M case illustrates quite clearly that the Ricardian equivalence experiment redistributes resources from distant future generations toward near future and existing generations. Especially members of the generation born at the time of the shock react strongly to the tax cut as far as their savings behaviour is concerned. Indeed, their maximum asset holding peaks at a much higher level than that of 40 year old existing generations and generations born 40 years after the shock.

Interest rate shock

The final shock analysed in this chapter consists of an unanticipated and permanent increase in the world interest rate (i.e. $dr > 0$ for $t \geq 0$). The effects of this shock on human and financial wealth are illustrated in Figure 2.5. In the Blanchard case the shock causes a once-off decrease in age-independent human wealth. The higher annuity rate of interest leads to stronger discounting of future after-tax wages. For the G-M model there is a once-off downward shift in the age profile of human wealth. Like the shock itself, this age profile displays no further transitional dynamics over time.

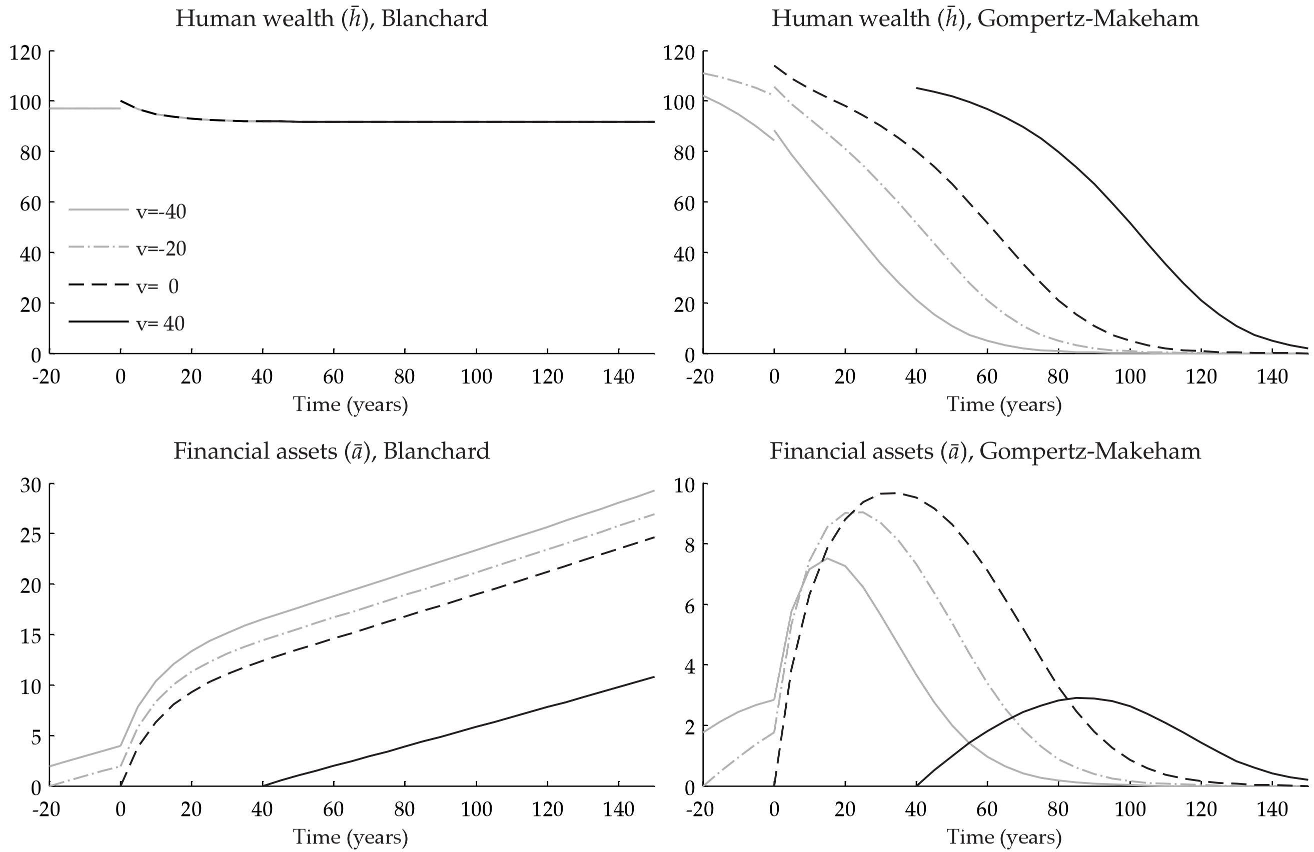
The bottom panels of Figure 2.5 illustrate the effects on financial assets. Whilst the effects for the Blanchard case speak for themselves, those for the G-M model warrant some further comment. For future generations, the age profile of financial assets features a once-off upward shift at impact and displays no further transitional dynamics thereafter. In contrast, for existing generations the time path of assets depends both on their age and on time. This transitional dynamics is caused by the fact that the consumption path for such generations depends on both t and v separately. Existing generations are affected by the interest rate hike both via their human wealth and via their accumulated financial assets which attract a higher rate of return after the shock.

2.4.2 Welfare effects

The Blanchard model is often used to investigate the intergenerational welfare effects of various policy measures.[22] In this section we visualize the intergenerational welfare effects associated with the three shocks studied above. For existing house-

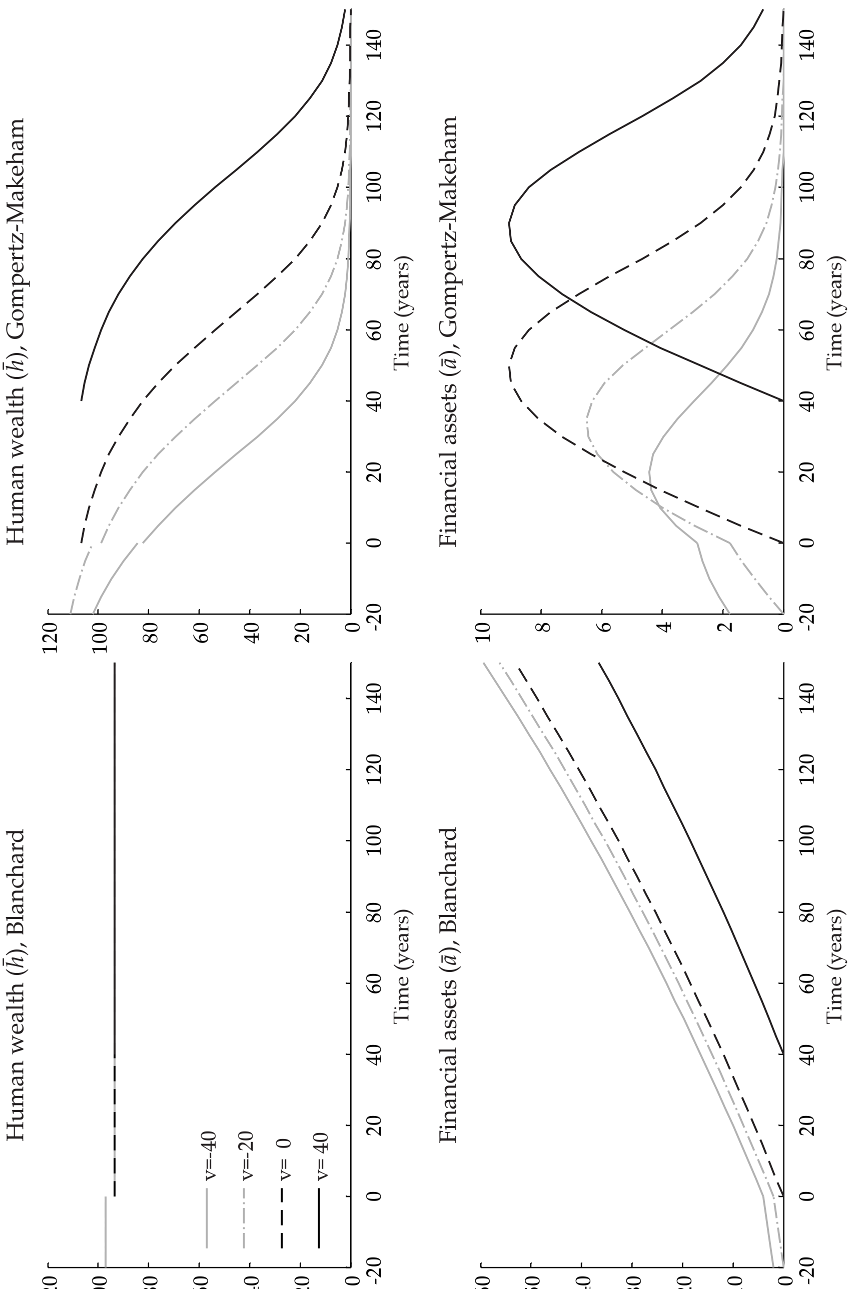
[22] See, for example, Bovenberg (1993, 1994) on capital taxation and investment subsidies, Bettendorf and Heijdra (2001b) and Bettendorf and Heijdra (2001a) on product subsidies and tariffs under monopolistic competition, and Heijdra and Meijdam (2002) on government infrastructure. All these studies are set in the context of a small open economy.

Figure 2.4. Ricardian equivalence experiment: temporary tax cut



Note: Both survival functions are fitted to the cohort born in the Netherlands in 1920 (see Table 2.1 for parameter values). Birth rate is 2.36%. At time $t = 0$ lumpsum taxes drop from 0 to -1 (a subsidy of 20% of gross wage income) and increase exponentially to its new steady state value. v is the date of birth relative to the time of the shock.

Figure 2.5. Increase in the world interest rate



Note: Both survival functions are fitted to the cohort born in the Netherlands in 1920 (see Table 2.1 for parameter values). Birth rate is 2.36%. At time $t = 0$ the exogenous world interest rate jumps from 4% to 4.2%.

holds, the change in welfare from the perspective of the shock period $t = 0$ is evaluated ($d\Lambda(v,0)$ for $v \leq 0$) whereas for future agents the welfare change from the perspective of their birth date is computed ($d\Lambda(v,v)$ for $v > 0$). The welfare effect for existing agents ($v \leq 0$) can be written as:

$$d\Lambda(v,0) = dr \int_0^\infty \tau e^{-\theta\tau - M(\tau - v) + M(-v)} d\tau + \Delta(-v,\theta) \ln \Gamma_E(v), \quad (\text{for } v \leq 0), \tag{2.53}$$

where $\Delta(-v,\theta)$ is defined in eq. (2.12) above and where $\Gamma_E(v)$ is defined as:

$$\Gamma_E(v) \equiv \frac{\hat{\bar{a}}(-v) + \bar{h}(v,0)}{\hat{\bar{a}}(-v) + \hat{\bar{h}}(-v)}, \quad (\text{for } v \leq 0). \tag{2.54}$$

Intuitively, $\Gamma_E(v)$ captures the effect of the impact change in human wealth for existing generations. The welfare effect consists of two separate components. The first term on the right-hand side of (2.53) represents the 'consumption growth effect' and is only relevant for the world interest rate shock (i.e., if $dr > 0$). Individual consumption growth is equal to $r - \theta$ and an increase in r leads to a steeper consumption time profile. The mortality process exerts a non-trivial influence on the consumption growth effect via the utility function. The second term on the right-hand side of (2.53) summarizes the welfare effect of the change in the level of consumption caused by the impact change in human wealth. This 'human wealth effect' is relevant for all shocks and is equal to the product of $\ln \Gamma_E(v)$ (defined in (2.54)) and the inverse propensity to consume of a v-year old agent ($\Delta(-v,\theta)$).

The welfare effect for future generations can be written as:

$$d\Lambda(v,v) = dr \int_0^\infty s e^{-\theta s - M(s)} ds + \Delta(0,\theta) \ln \Gamma_F(v), \quad (\text{for } v > 0), \tag{2.55}$$

where $\Delta(0,\theta)$ is the inverse propensity to consume of a newborn and $\Gamma_F(v)$ is defined as:

$$\Gamma_F(v) \equiv \frac{\bar{h}(v,v)}{\hat{\bar{h}}(0)}, \quad (\text{for } v > 0). \tag{2.56}$$

Here, $\Gamma_F(v)$ represents the effect on the human wealth of a future newborn. Just as for existing generations, the welfare effect for future generations consists of a consumption growth effect (first term on the right-hand side of (2.55)) and a human wealth effect (second term).

The welfare effects of the different shocks are illustrated in Figure 2.6. The left-hand panels present the results for the Blanchard case whilst the right-hand panels visualize those for the G-M model. The welfare effects of balanced-budget fiscal policy are illustrated in the top panels. All present and future generations experience a reduction in human wealth and as a result the welfare effect is negative for all generations. The effect is the same for all future generations because there is no transitional dynamics in human wealth (see above). For existing generations the welfare loss declines with the age of the generation. The human wealth effect decreases with age because both the inverse propensity to consume ($\Delta(-v, \theta)$) and the relative importance of human wealth ($\ln \Gamma_E(v)$ in (2.53) above) decline with age. The Blanchard and G-M models thus give qualitatively similar welfare results for the spending shock. A key difference between the two models concerns the slope of the welfare profile for existing generations. In the G-M model (right-hand panel) the welfare effect is practically zero for all generations older than 100 years. In contrast, for the Blanchard case (left-hand panel) there is still a noticeable welfare effect for 200 year old generations. This low 'generational adjustment speed' of the Blanchard model is also observed for the other shocks. Intuitively, in the Blanchard case, the population share of the old generations is too large because the expected lifetime if too high (see also the top panel of Figure 2.1).

The middle two panels of Figure 2.6 illustrate the welfare effects for the Ricardian tax cut experiment. All existing generations as well as future generations born close to the time of shock benefit at the expense of more distant future generations. For future generations the welfare loss is larger the later they are born. For existing generations the welfare profile is monotonically decreasing in age for the Blanchard case but non-monotonic for the G-M model. In the Blanchard case, $\Delta(-v, \theta) = \Delta(0, \theta) = 1/(\theta + \mu_0)$ is constant and $\ln \Gamma_E(v)$ declines monotonically with age. In contrast, for the G-M model, $\Delta(-v, \theta)$ decreases with age but $\ln \Gamma_E(v)$ is non-monotonic. Indeed, $\ln \Gamma_E(v)$ is increasing in age for all generations up to about 90 years and only decreases in age thereafter. [23] As a result, the welfare profile for existing generations displays a bump around the age of 55 in the middle right-hand panel of Figure 2.6. At that point, the drop in $\Delta(-v, \theta)$ just matches the increase in $\ln \Gamma_E(v)$.

In the bottom two panels of Figure 2.6 the welfare effects for the interest rate shock are illustrated. Since the shock induces no transitional dynamics in the age

[23] Of course, there are virtually no centenarians predicted by the G-M model so the downward sloping part of the $\ln \Gamma_E(v)$ function is practically irrelevant. In contrast, the estimated Blanchard demography predicts that about 32 percent of newborns will still be alive at age 100. See Table 2.1 and the top panel of Figure 2.1.

profile of human wealth for future generations, the welfare effect is the same for all future generations in both models. For existing generations the welfare effect increases with age in the Blanchard model, but is non-monotonic for the G-M model. For an interest rate shock both the consumption growth effect and the human wealth effect are relevant. The shock induces a decrease in $\ln \Gamma_E(v)$ which falls with age in both models. In the Blanchard case, the consumption growth effect is constant (and positive) for all generations. In contrast, for the G-M model, the consumption growth effect is positive and constant for future generations, but falling in age for existing generations. As a result, the total effect on welfare displays a bump around the age of 15 for the G-M model (see the bottom right-hand panel of Figure 2.6).

2.4.3 Aggregate effects

As was pointed out above, Blanchard (1985) assumes a constant mortality rate in order to allow for exact aggregation of the consumption function. With the more general mortality process considered in our model, only numerical aggregation is possible. This section visualizes the aggregate effects on the key variables of the three shocks considered above. To what extent do the aggregate results predicted by the Blanchard and G-M models differ?

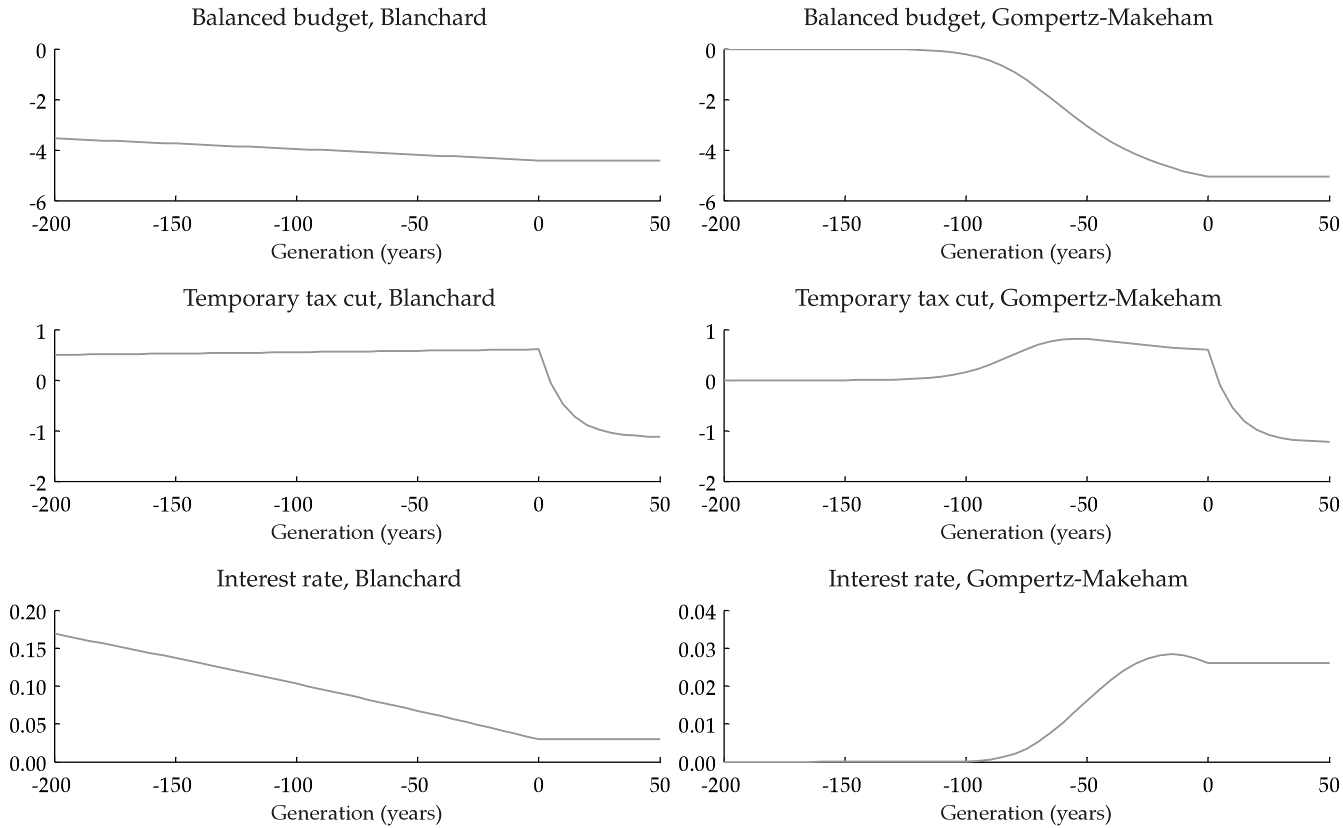
In Figure 2.7 we illustrate the effects on human wealth (first row), consumption (second row), and financial assets (third row) for the spending shock (first column), the Ricardian tax cut (second column), and the interest rate shock (third column). To make it easier to compare the two models, we show the percentage deviations from the steady state for all variables in Figure 2.7, i.e. we show $(h(t) - \hat{h})/\hat{h}$, $(c(t) - \hat{c})/\hat{c}$, and $(a(t) - \hat{a})/\hat{a}$.

For the spending shock, the results for human wealth are identical and those for consumption and financial assets are qualitatively very similar but differ in terms of the speed of adjustment towards the new steady state. The slow speed of convergence is also a feature of the Blanchard results for the other two shocks.

For the Ricardian tax cut, the effects on human wealth are again similar but those on consumption and financial wealth are not. For the G-M model, the impact effect on consumption is much larger, and the slope of the aggregate Euler equation is much steeper during transition, than for the Blanchard model. Similarly, the savings response is much more pronounced for the G-M model.

Finally, for the interest rate shock the effect on human wealth is qualitatively the same for the two models, though the effect is stronger for the Blanchard model.

Figure 2.6. Welfare effects (absolute change



Note: Both survival functions are fitted to the cohort born in the Netherlands in 1920 (see Table 2.1 for parameter values). Birth rate is 2.36%. Balanced budget shock is a jump of lumpsum taxes and government consumption from 0 to 1 (20% of gross wage income). Ricardian Temporary tax cut is a drop of lumpsum taxes from 0 to -1 (a subsidy of 20% of gross wage income) and they increase exponentially to its new steady state value. Interest rate shock is a jump of the exogenous world interest rate jumps from 4% to 4.2%.

The impact reduction in consumption is virtually the same for the two models but transition is much faster for the G-M model. Again, the savings response at impact is stronger for the G-M model.

2.5 Concluding remarks

In this chapter we showed that it is quite feasible to incorporate a realistic demographic structure in an overlapping generation model of a small open economy facing an exogenous world interest rate. At the level of individual households, a realistic description of the mortality process instead of Blanchard's perpetual youth reinstates the classic life-cycle saving insights of Modigliani and co-workers.

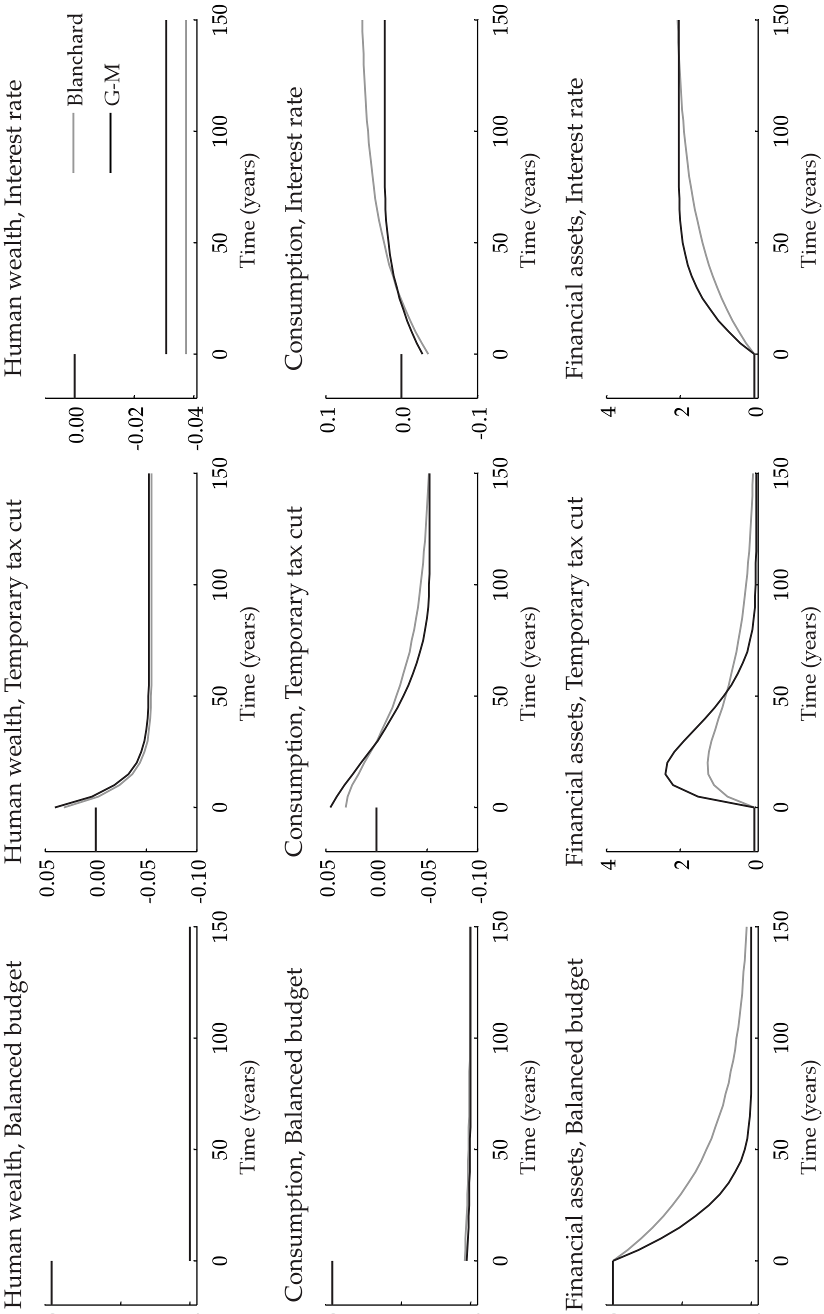
The welfare effects associated with the different shocks are also potentially affected in a non-trivial manner by a more realistic demography. Two key differences stand out between the Blanchard and G-M models. First, the G-M model predicts a much faster (and in our view more realistic) 'generational convergence speed' of the welfare effects than the Blanchard model. Second, the G-M model incorporates more extensive age-dependency and as a result may give rise to a non-monotonic welfare effect on existing generations—something which is impossible in the Blanchard case (for the shocks studied).

Finally, we have demonstrated that the demographic details do not 'wash out' at the aggregate level. The impulse-response functions for the different shocks are quite different for the Blanchard and G-M models, especially the ones for per capita consumption and financial assets.

In some applications of our model, individual behaviour may depend in part on aggregate variables so that knowledge of the latter is crucial. For example, if the revenue of a consumption tax (t_C) is recycled in a lumpsum fashion to households (i.e. $\bar{z}(t) = z(t) = -t_C c(t)$) then individual consumption, human wealth, and financial assets will all depend on the aggregate tax revenue. This complication can be easily dealt with by using an iterative procedure in the simulations. In the first step the initial tax revenue and implied lumpsum transfer are guessed and individual and aggregate consumption levels are computed. In subsequent steps, the aggregate information is used to update the guess for transfers until convergence is achieved. We will use this procedure to solve the extended models in Chapters 3 and 4

The framework developed in this chapter can be extended in a number of directions. First, in order to investigate the effects of demographic change, it is neces-

Figure 2.7. Aggregate effects of the shocks



Note: Both survival functions are fitted to the cohort born in the Netherlands in 1920 (see Table 2.1 for parameter values). Birth rate is 2.36%. Balanced budget shock is a jump of lumpsum taxes and government consumption from 0 to 1 (20% of gross wage income). Ricardian Temporary tax cut is a drop of lumpsum taxes from 0 to -1 (a subsidy of 20% of gross wage income) and they increase exponentially to its new steady state value. Interest rate shock is a jump of the exogenous world interest rate jumps from 4% to 4.2%.

sary to generalize the stochastic distribution for expected remaining lifetimes. Two possibilities can be distinguished. 'Embodied' mortality change can be studied by assuming the instantaneous mortality rate to be generation-specific, i.e. by writing it as $m(v,s)$. An example of embodied mortality change could be the ability to extract and store embryonic stem cells to be used for future organ repairs. In contrast, 'disembodied' demographic change can be modelled by writing the mortality rate $m(t,s)$, i.e. by postulating a time-dependent mortality process. An example of disembodied mortality change would be a comprehensive cure for cancer or heart and vascular diseases.

Second, the age profile for individual consumption could be generalized by introducing shift factors in the utility function. In the current model (with $r > \theta$) consumption is increasing in the age of the household. There are reasons to believe that in reality consumption is hump-shaped, i.e. $\bar{c}(v,t)$ features a rising time profile early on in life followed by a falling profile later on. A simple way to capture this effect is to assume that a household's 'needs' get smaller the older they get. In the diminishing-needs model, lifetime utility is given by:

$$\Lambda(v,t) \equiv e^{M(t-v)} \int_t^\infty \left[\frac{\bar{e}\,(v,\tau)^{1-1/\sigma} - 1}{1-1/\sigma} \right] e^{-\theta\cdot(\tau-t)-M(\tau-v)} d\tau, \tag{2.57}$$

where $\sigma > 0$ is the intertemporal substitution elasticity and $\bar{e}(v,\tau)$ is *effective* consumption:

$$\bar{e}(v,\tau) \equiv \bar{c}(v,\tau) \exp\left\{ \frac{\zeta_0(\tau-v)^{1+\zeta_1}}{1+\zeta_1} \right\}, \tag{2.58}$$

with $\zeta_0 > 0$ and $\zeta_1 > 0$. According to (2.58), a given amount of *actual* consumption, $\bar{c}(v,\tau)$, yields more effective consumption (featuring in the felicity function), the older the household is. Using this specification of preferences, it is straightforward to show that the individual consumption Euler equation is generalized to:

$$\frac{\dot{\bar{c}}(v,\tau)}{\bar{c}(v,\tau)} = \sigma\cdot(r-\theta) - (1-\sigma)\zeta_0(\tau-v)^{\zeta_1}. \tag{2.59}$$

For the empirically relevant case (with $0 < \sigma < 1$), consumption rises during the early phase of life ($\tau - v$ low) and falls during the later stages of life ($\tau - v$ high).

In the next chapter, we present a third extension. We extend the basic framework developed here with an optimal schooling decision. Individuals spend the first years of life at school, accumulating knowledge, which increases their productivity and hence their wage rate later in life. Furthermore, we allow for embodied mortality change and a time varying birth rate and we use the extended model to analyse the effects of embodied mortality changes and baby busts on the aggregate productivity growth rate.

In the final chapter of part I, we endogenise the household's labour supply and retirement decisions. By including leisure hours into the felicity function, the agent has an additional choice variable which determines optimal labour income and lifetime utility. Two approaches can be considered. In the 'divisible labour' case, the agent can freely choose the number of working hours at each instant. In the typical formulation, consumption and leisure are both normal goods so that, as the agent gets older and richer, labour supply gradually declines to its lower bound (of zero). Hence, the agent gradually retires from the labour market. In the 'indivisible labour' case, employment is assumed to be a participation decision, i.e. the agent either works a fixed number of hours (full time) or not at all. In such a setting the retirement decision constitutes a withdrawal from the labour market altogether. In both types of labour supply models, the most interesting shocks that can be studied are ageing shocks and pension reform. In Chapter 4 we focus on the second case and we will analyse how ageing will effect the retirement decision and a nation's retirement system.

2.A Computation of the Δ-function

The model used in this chapter makes extensive use of the Δ-function as defined in Equation (2.12). To be able to solve the model in a reasonable amount of time, we need efficient methods to evaluate this function for the specified mortality processes.

Blanchard: The Δ-function for Blanchard's mortality process can be written as

$$\Delta(u,\lambda) = e^{(\lambda+\mu_0)u} \int_u^\infty e^{-(\lambda+\mu_0)s} ds = \frac{1}{\lambda+\mu_0}$$

Linear and piecewise linear: The demographic discount function for the linear mortality process can be written as:

$$\Delta(u,\lambda) = e^{(\lambda+\mu_0)u+\mu_1^2u^2} \int_u^\infty e^{-(\lambda+\mu_0)s-\mu_1^2s^2} ds$$

We define $\beta(u) = \mu_1 u + \frac{\lambda+\mu_0}{2\mu_1}$ and $t = \beta(s)$. Changing the integrand we obtain:

$$\Delta(u,\lambda) = \frac{\sqrt{\pi}}{2\mu_1} \operatorname{erfcx}\left(\mu_1 u + \frac{\lambda+\mu_0}{2\mu_1}\right)$$

where $\operatorname{erfcx}(x)$ is the scaled complementary error function (defined in general terms as $\operatorname{erfcx}(x) = e^{x^2} \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt$; see Kreyszig, 1999). The scaled complementary error function is well documented and, more importantly, most software packages have very fast routines to calculate it accurately enough for our purposes.

Some tedious, but otherwise straightforward math shows that the expression for $\Delta(u,\theta)$ for the piecewise linear model features two branches, depending on whether the household is still 'young' ($0 < u < \bar{u}$) or has entered 'old age' ($u > \bar{u}$):

$$\Delta(u,\theta) = \begin{cases} \dfrac{1-e^{-(\lambda+\mu_0)(\bar{u}-u)}}{\lambda+\mu_0} \\ \qquad + e^{-(\lambda+\mu_0)(\bar{u}-u)} \dfrac{\sqrt{\pi}}{2\mu_1} \operatorname{erfcx}\left(\dfrac{\lambda+\mu_0}{2\mu_1}\right) & \text{for } 0 \le u < \bar{u} \\ \dfrac{\sqrt{\pi}}{2\mu_1} \operatorname{erfcx}\left(\mu_1 \cdot (u-\bar{u}) + \dfrac{\lambda+\mu_0}{2\mu_1}\right) & \text{for } u \ge \bar{u} \end{cases}$$

Boucekkine et al. (2002): Boucekkine et al. (2002) introduce a survival function

$$S(u) = 1 - \Phi(u) = \frac{e^{-\beta u} - \alpha}{1-\alpha}, \quad 0 \le u \le A$$

with $A = -\frac{1}{\beta}\ln\alpha$. This gives for the mortality rate (only for the relevant domain)

$$m(u) = \frac{\Phi'(u)}{1-\Phi(u)} = \frac{\beta e^{-\beta u}}{e^{-\beta u}-\alpha} \qquad \text{for } 0 \le u < A$$

and integrated mortality rate

$$\begin{aligned} M(u) &= \beta\int_0^u \frac{e^{-\beta s}}{e^{-\beta s}-\alpha}ds = -\int_1^{e^{-\beta u}} \frac{x}{x-\alpha}\frac{1}{x}dx \\ &= \ln|1-\alpha| - \ln|e^{-\beta u}-\alpha| = -\ln S(u) \qquad \text{for } 0 \le u < A \end{aligned}$$

The demographic discount function can be written as (also only for $0 \le u \le A$)

$$\begin{aligned} \Delta(u,\lambda) &= e^{\lambda u + \ln|1-\alpha| - \ln|e^{-\beta u}-\alpha|}\int_u^A e^{-\lambda s - \ln|1-\alpha| + \ln|e^{-\beta s}-\alpha|}ds \\ &= e^{\lambda u}\frac{|1-\alpha|}{|e^{-\beta u}-\alpha|}\int_u^A e^{-\lambda s}\frac{|e^{-\beta s}-\alpha|}{|1-\alpha|}ds \\ &= \frac{1}{e^{-\beta u}-\alpha}\left\{\frac{1}{\lambda+\beta}\left[e^{-\beta u} - e^{-\lambda[A-u]-\beta A}\right] + \frac{\alpha}{\lambda}\left[e^{-\lambda[A-u]}-1\right]\right\} \end{aligned}$$

This expression is easy to evaluate efficiently using any standard mathematical software package. For $\lambda = 0$, we need l'Hopital's rule

$$\Delta(u,0) = \frac{1}{e^{-\beta u}-\alpha}\left\{\frac{1}{\beta}\left[e^{-\beta u} - e^{-\beta A}\right] - \alpha[A-u]\right\}.$$

Gompertz-Makeham: The demographic discount function for the G-M process can be written as:

$$\Delta(u,\lambda) = e^{(\lambda+\mu_0)u + \frac{\mu_1}{\mu_2}e^{\mu_2 u}}\int_u^\infty e^{-(\lambda+\mu_0)s - \frac{\mu_1}{\mu_2}e^{\mu_2 s}}ds.$$

We define $\beta(u) \equiv \frac{\mu_1}{\mu_2}e^{\mu_2 u}$ and $t = \beta(s)$. Changing the integrand we obtain:

$$\Delta(u,\lambda) = \frac{\mu_2^{\alpha-1}}{\mu_1^{\alpha}}e^{(\lambda+\mu_0)u+\beta(u)}\Gamma\big(\alpha,\beta(u)\big),$$

where $\alpha \equiv -(\lambda+\mu_0)/\mu_2$ and $\Gamma\big(\alpha,\beta(u)\big)$ is the upper tailed incomplete gamma function (defined in general terms as $\Gamma(\alpha,x) = \int_x^\infty t^{\alpha-1}e^{-t}dt$; see Kreyszig (1999, p. A55)). The incomplete gamma function is just like the scaled incomplete error function well documented (see e.g. Kreyszig (1999, p. A78)) and software packages have very fast routines to calculate it.

There is one slight complication; the incomplete gamma function is usually only defined for $\alpha \geq 0$, whereas we also need to evaluate it for $\alpha < 0$. We can solve this problem by using the 'functional relation of the incomplete gamma function'. Indeed, by integrating the incomplete gamma function by parts we obtain the following recursion formula:

$$\Gamma(\alpha, x) = \frac{1}{\alpha} e^{-t} t^{\alpha} \Big|_{t=x}^{\infty} + \frac{1}{\alpha} \int_x^{\infty} t^{\alpha} e^{-t} dt = -\frac{1}{\alpha} e^{-x} x^{\alpha} + \frac{1}{\alpha} \Gamma(\alpha + 1, x).$$

Repeated application gives for $k = 0, 1, 2, \ldots$:

$$\Gamma(\alpha, x) = -e^{-x} x^{\alpha} \left[\frac{1}{\alpha} + \frac{1}{\alpha} \frac{1}{\alpha+1} x + \frac{1}{\alpha} \frac{1}{\alpha+1} \frac{1}{\alpha+2} x^2 + \cdots \right.$$
$$\left. + \frac{1}{\alpha} \frac{1}{\alpha+1} \cdots \frac{1}{\alpha+k-1} x^{k-1} \right] + \frac{1}{\alpha} \frac{1}{\alpha+1} \cdots \frac{1}{\alpha+k-1} \Gamma(\alpha+k, x).$$

Hence, by choosing the smallest integer k such that $\alpha + k$ is non-negative, the value of $\Gamma(\alpha, x)$ can be computed in a standard fashion.

Others: For instances of the mortality function that can not be solved explicitly or rewritten in terms of well-documented (and easily calculated) functions, we can evaluate the Δ-function using standard numerical integration techniques. To evaluate the Δ-function for more than one age (u), we can make use of the following algorithm:

1. Initialise; sort the u's, such that $u_1 < u_2 < \ldots < u_n$, calculate $\Delta(u_n, \lambda)$ using any (adaptive) quadrature and set $i = n - 1$.

2. Calculate $\Delta(u_i, \lambda)$ using

$$\Delta(u_i, \lambda) = e^{-\lambda \cdot (u_{i+1} - u_i) - M(u_{i+1}) + M(u_i)} \Delta(u_{i+1}) + \int_{u_i}^{u_{i+1}} e^{-\lambda \cdot (s - u_i) - M(s) + M(u_i)} ds.$$

   The integral can be evaluated using any (adaptive) quadrature.

3. If $i = 1$, then exit, else set $i = i - 1$ and go to step 2.

It is important first to construct the exponents and then to evaluate the e-power, to prevent numerical problems. Furthermore, sorting in descending order ensures that u_{i+1} is always larger than u_i, so the exponential terms are always smaller than 1, which prevents inaccuracies.

Chapter 3

Ageing, Schooling, and Growth

3.1 Introduction

It is a well documented fact that the western world is ageing rapidly. Since the post-war period, the ageing process can be attributed both to increased longevity and reduced fertility (Lee, 2003). For example, in the Netherlands, life expectancy at birth rose from 71.5 years in 1950 to 78.6 years in 2003, whilst the annual (crude) birth rate fell from 2.3% to 1.3% of the population. Because infant mortality stayed relatively constant during that period (at 0.8% of the population), the increase in longevity must be attributed to reduced adult mortality. Not surprisingly, the demographic change has led to a dramatic increase in the population share of elderly people over that period—the old-age dependency ratio (measured as the ratio of the population aged 65 years or over to the population aged 15-64) rose from 12.2% in 1950 to 20.1% in 2002. A similar demographic pattern can be observed for most OECD countries.

The objective of this chapter is to investigate the effects on the economic growth performance of a small open economy of substantial demographic shocks of the type and magnitude mentioned above. We use the Blanchard-Yaari model with a realistic mortality process developed in the previous chapter and extend it with a schooling decision. The finitely-lived agents accumulate both physical and human capital. In this model disconnected generations are born at each instant and individual agents face a positive and age-dependent probability of death at each

This chapter is based on joint work with Ben Heijdra, 'Ageing and Growth in the Small Open Economy', published as CESIfo Working Paper, No. 1740.

moment in time. By making the mortality rate age-dependent, the model can be used to investigate changes in adult mortality.

The other building block of our analysis concerns the engine of growth. Following Lucas (1988), we assume that the purposeful accumulation of human capital forms the core mechanism leading to economic growth. More specifically, like Bils and Klenow (2000), Kalemni-Ozcan et al. (2000), de la Croix and Licandro (1999), and Boucekkine et al. (2002), we assume that individual agents accumulate human capital by engaging in full-time educational activities at the start of life. The start-up education period is chosen optimally and the human capital production function may include an intergenerational external effect of the 'shoulders of giants' variety, as proposed by Azariadis and Drazen (1990).

As we motivate in more detail later on in this chapter, we extend the existing literature in the following directions. First, we generalize Kalemni-Ozcan et al. (2000) by incorporating a realistic (rather than a Blanchard) demographic structure, allowing for non-zero intergenerational spillovers, and by fully characterizing the transitional dynamics. Second, we generalize the analysis by de la Croix and Licandro (1999) and Boucekkine et al. (2002) by incorporating both human and physical capital, by including a concave (rather than linear) felicity function, and by allowing the intergenerational spillover to differ from unity. Third, we generalize the model of Bils and Klenow (2000) by recognising fully-insured-against lifetime uncertainty (rather than a fixed planning horizon), by assuming a more realistic human capital production function, and by characterizing the transitional dynamics. Finally, we generalize all these papers by including an educational subsidy and a labour income tax.

The remainder of this chapter is organized as follows. In Section 3.2 we present the model and demonstrate its main properties. A unique solution for the optimal schooling period is derived which depends on the fiscal parameters and on the mortality process. The mortality process, in combination with the birth rate, also determines a unique path for the population growth rate. For a given initial level of per capita human capital, the model implies a unique time path for all macroeconomic variables. Depending on the strength of the intergenerational external effect, the model either displays exogenous growth (ultimate convergence to constant per capita variables) or endogenous growth (convergence to a constant growth rate).

In Section 3.3 we study the determinants of the optimal schooling decision in detail. An increase in the educational subsidy or the labour income tax leads to an increase in the length of the educational period. Similarly, a reduction in adult mortality also prompts agents to increase the schooling period. In contrast, a reduction in child mortality and a baby bust both leave the optimal schooling period unchanged.

In Section 3.4 we investigate the effects of changes in the birth rate and adult mortality on the population growth rate, both at impact, during transition, and in the long run. A reduction in the birth rate reduces the steady-state population growth rate, whilst an increase in longevity (due to reduced adult mortality) increases this rate because average mortality falls. We use the estimated Gompertz-Makeham mortality process of the previous chapter to illustrate the rather complicated (cyclical) adjustment path resulting from once-off demographic changes. Especially for the embodied mortality shock, convergence toward the new steady state is extremely slow. Indeed, due to the vintage nature of the population, more than 150 years pass until the new demographic steady state is reached.

Section 3.5 deals with the exogenous growth model. In this model there is no or an imperfect intergenerational spillovers and the economy settles at a unique steady state *level* of per capita human capital. We consider this model, on the basis of the empirical evidence, to be the most relevant one. In Section 3.5 we study the effects of fiscal and demographic changes on per capita human capital and the other macroeconomic variables both at impact, in the transition period and in the long run. A positive fiscal impulse leads to an increase in the per capita stock of human capital but leaves the steady-state growth rate of the macro-variables in level terms unchanged (and equal to the steady-state population growth rate). Furthermore, whilst a reduction in the birth rate and an increase in longevity (due to reduced adult mortality) both increase the steady-state per capita human capital stock, the growth effects on level variables are opposite in sign. Again, for both fiscal and demographic shocks, the transitional adjustment is rather slow.

In Section 3.6 we briefly discuss the endogenous growth version of the model. Though this knife-edge case has been studied extensively in the theoretical literature (Azariadis and Drazen, 1990), it is based on an unrealistically strong intergenerational external effect in human capital creation for which very little empirical backing exists. The positive fiscal impulse boosts the steady-state growth rate in per capita human capital due to the scale effect in the growth process. The growth effects of demographic changes are theoretically ambiguous. For a realistic model calibration, however, the asymptotic growth rate is decreasing in the birth rate and in longevity (as measured by life expectancy at birth).

Finally, in Section 3.7 we present some concluding thoughts and give some suggestions for future research. The Appendix contains some key mathematical derivations.

3.2 The model

3.2.1 Households

Individual plans

The core of the model is the same as in Chapter 2. At time t, an individual born at time v ($v \leq t$) has the following (remaining) lifetime utility function:

$$\Lambda(v,t) \equiv e^{M(t-v)} \int_t^{\infty} U[\bar{c}(v,\tau)] e^{-\theta\cdot(\tau-t)-M(\tau-v)} d\tau, \tag{3.1}$$

where $U[\cdot]$ is the felicity function, $\bar{c}(v,\tau)$ is consumption where a bar denotes individual variables and functions as before. θ is the constant pure rate of time preference ($\theta > 0$), and $e^{-M(\tau-v)}$ is the probability that the agent is still alive at time τ. The cumulative mortality rate, $M(\tau - v)$, is defined in Equation (2.3) on page 14 as $M(\tau - v) \equiv \int_0^{\tau-v} m(\alpha)d\alpha$, where $m(\alpha)$ is the instantaneous mortality rate of an agent of age α. As before (see Equation (2.6)), the felicity function is iso-elastic:

$$U[\bar{c}(v,\tau)] = \begin{cases} \dfrac{\bar{c}(v,\tau)^{1-1/\sigma} - 1}{1-1/\sigma} & \text{for } \sigma \neq 1 \\ \ln \bar{c}(v,\tau) & \text{for } \sigma = 1 \end{cases}, \tag{3.2}$$

where σ is the constant intertemporal substitution elasticity ($\sigma \geq 0$).

The budget identity is given by:

$$\dot{\bar{a}}(v,\tau) = [r + m(\tau - v)]\bar{a}(v,\tau) + \bar{w}(v,\tau) - \bar{g}(v,\tau) - \bar{c}(v,\tau), \tag{3.3}$$

where $\bar{a}(v,\tau)$ is real financial wealth, r is the constant world interest rate, $\bar{w}(v,\tau)$ is wage income, and $\bar{g}(v,\tau)$ is total tax payments (see below). As usual, a dot above a variable denotes that variable's time rate of change, e.g. $\dot{\bar{a}}(v,\tau) \equiv d\bar{a}(v,\tau)/d\tau$. As in the previous chapter, we follow Yaari (1965) and Blanchard (1985), by assuming the existence of a perfectly competitive life insurance sector which offers actuarially fair annuity contracts to the agents. Since someone's age is directly observable, the annuity rate of interest faced by an individual of age $\tau - v$ is equal to the sum of the world interest rate and the instantaneous mortality rate of that person. In order to avoid having to deal with a taxonomy of different cases, we again restrict attention to the case of a nation populated by patient agents, i.e. $r > \theta$. Financial wealth can be held in the form of claims on domestic capital, $\bar{v}(v,\tau)$, domestic government bonds, $\bar{d}(v,\tau)$, or net foreign assets, $\bar{f}(v,\tau)$.

$$\bar{a}(v,\tau) \equiv \bar{v}(v,\tau) + \bar{d}(v,\tau) + \bar{f}(v,\tau). \tag{3.4}$$

These assets are perfect substitutes in the agents' investment portfolios and thus attract the same rate of return.

To allow for economic growth, we extend this model by postulating that the agent engages in full time schooling during the early stages of life and works full time thereafter. The production function for human capital is given by:[1]

$$\bar{h}(v,\tau) = \begin{cases} 0 & \text{for } v \leq \tau \leq v + s(v) \\ A_H h(v)^{\phi} s(v) & \text{for } \tau > v + s(v) \end{cases}, \qquad 0 \leq \phi \leq 1, \tag{3.5}$$

where $\bar{h}(v,\tau)$ is the human capital of the agent upon completion of the schooling period, A_H is an exogenous productivity index, $h(v)$ is *per capita* human capital at time v (see below), ϕ is a parameter regulating the strength of the intergenerational external effect in knowledge creation, and $s(v)$ is the length of the schooling period chosen by an agent born at time v. Special cases of (3.5) are used by de la Croix and Licandro (1999, p. 257) and Boucekkine et al. (2002, p. 347), who set $\phi = 1$, and by Kalemni-Ozcan et al. (2000, pp. 5, 10), who set $\phi = 0$.

Available human capital is rented out to competitive producers so that wage income, $\bar{w}(v,\tau)$, can be written as:

$$\bar{w}(v,\tau) = w(\tau)\bar{h}(v,\tau), \tag{3.6}$$

where $w(\tau)$ is the market-determined rental rate of human capital.

The tax system takes the following form. First, all through life, the agent pays a lumpsum tax. Second, during the educational phase, the agent receives a study grant from the government. Third, during working life, the agent faces a labour

[1] This formulation was first proposed in the context of Diamond-Samuelson style overlapping models by Azariadis and Drazen (1990, p. 510) and Tamura (1991, p. 524). Abstracting from their work experience term and using our notation, Bils and Klenow (2000, p. 1161) model the human capital production function as follows:

$$\bar{h}(v,t) = \bar{h}(v-\bar{u},t)^{\phi} e^{\zeta(s)}, \qquad \text{for } t - v > s, \tag{3.5$'$}$$

where $\bar{u}$ is interpreted as the age of the teachers (assumed to be fixed), and $\zeta(s)$ captures the productivity effect of schooling ($\zeta'(s) > 0$). Clearly, for $\zeta(s) \equiv \ln s$ the second term on the right-hand side of (3.5′) is equal to s. In our view, Equation (3.5′) does not adequately capture the notion of an intergenerational externality as the link is only operative between generations v and $v - \bar{u}$, which are locked in a tango through time. In (3.5) the *economy-wide* stock of per capita human capital determines the initial condition facing newborns. Hence, every agent alive at time v exerts an external effect on newborns.

income tax on wage earnings. The tax system is thus given by:

$$\bar{g}(v,\tau) = \begin{cases} [z(\tau)-\rho]w(\tau)A_H h(v)^{\phi} & \text{for } v \le \tau \le v+s(v) \\ [z(\tau)+t_L s(v)]w(\tau)A_H h(v)^{\phi} & \text{for } \tau > v+s(v) \end{cases}, \tag{3.7}$$

where ρ is the *educational subsidy* rate ($\rho > 0$), t_L is the labour income tax rate ($0 \le t_L < 1$), and $z(\tau)$ represents the lumpsum part of the tax. All tax instruments are indexed to the value of marginal schooling productivity to the vintage-v individual (i.e. $A_H h(v)^{\phi}$) to ensure that the tax system continues to play a non-trivial role even in the presence of ongoing economic growth.[2]

From the perspective of the planning date t, the agent chooses remaining time in school ($v+s(v)-t$), and sequences for $\bar{c}(v,\tau)$ and $\bar{a}(v,\tau)$ (for $\tau \in [t,\infty)$) in order to maximize $\Lambda(v,t)$ subject to (3.3)–(3.7), a non-negativity constraint $v+s(v) \ge t$,[3] and a transversality condition. By using this transversality condition as well as Equations (3.3)–(3.7), the lifetime budget constraint for an agent with age $u \equiv t-v$ can be written as follows:

$$e^{M(t-v)} \int_t^{\infty} \bar{c}(v,\tau) e^{-r\cdot(\tau-t)-M(\tau-v)} d\tau = \bar{a}(v,t) + \bar{l}i(v,t), \tag{3.8}$$

where we have used the fact that generations are born without financial assets (i.e. $\bar{a}(v,v)=0$) and where $\bar{l}i(v,t)$ is (remaining) lifetime after-tax wage income of the agent:

$$\begin{aligned} \bar{l}i(v,t) \equiv A_H h(v)^{\phi} e^{M(t-v)} \Bigg[\rho \int_t^{\max\{t,v+s(v)\}} w(\tau) e^{-r\cdot(\tau-t)-M(\tau-v)} d\tau \\ + (1-t_L)s(v) \int_{\max\{t,v+s(v)\}}^{\infty} w(\tau) e^{-r\cdot(\tau-t)-M(\tau-v)} d\tau \\ - \int_t^{\infty} z(\tau) w(\tau) e^{-r\cdot(\tau-t)-M(\tau-v)} d\tau \Bigg]. \end{aligned} \tag{3.9}$$

According to (3.8), the present value of consumption expenditure (left-hand side) must equal total lifetime resources (right-hand side). In the presence of actuarially fair annuity contracts, the annuity rate of interest, $r+m(\tau-v)$, is used for discounting purposes in (3.8)–(3.9).

The following two-stage solution approach can now be used. In the first step,

[2] Alternatively, current gross per capita labour income, $w(\tau)h(\tau)$, could have been used for indexing purposes, but this makes the model intractable.

[3] Older agents have already completed the educational phase ($t-v > s(v)$) and only choose paths for consumption and financial assets. Labour market entry is thus assumed to be an absorbing state.

the agent chooses $s(v)$ in order to maximize lifetime wage income, $\bar{li}(v,t)$. This pushes the lifetime budget constraint out as far as possible and fixes the right-hand side of (3.8). In the second step, the agent chooses the optimal sequence for consumption in order to maximize $\Lambda(v,t)$ subject to (3.8).

Schooling period By using (3.9), the first-order condition for the optimal schooling period, $s^*(v)$, is given by $d\bar{li}(v,t)/ds(v) = 0$ which can be written as:

$$\int_{v+s^*(v)}^{\infty} w(\tau)e^{-r(\tau-v)-M(\tau-v)}d\tau = \left[s^*(v) - \frac{\rho}{1-t_L}\right] w(v+s^*(v))e^{-rs^*(v)-M(s^*(v))}. \tag{3.10}$$

For the case studied in this chapter, the wage rate is constant (see below), and Equation (3.10) reduces to:

$$s^* - \frac{\rho}{1-t_L} = \Delta(s^*, r), \tag{3.11}$$

where $\Delta(u,\lambda)$ is defined in Equation (2.12) on page 16 in the previous chapter. Proposition 2.1 describes the main characteristics of this function.

Equation (3.11) determines the age at which the vintage-v individual completes his education. With a constant mortality process, the optimal schooling period is independent of the agent's date of birth. Since the left-hand side of (3.11) is increasing in s^* and (by Proposition 2.1(ii)) the right-hand side is non-increasing in s^*, it follows that the optimal schooling period is positive and unique.[4] In Section 3.3 below we study changes in the tax parameters and the demographic structure which give rise to once-off changes in the optimal schooling period.

Consumption By using (3.1) and (3.8), the first-order conditions for optimal consumption can be written as $\bar{c}(v,\tau) = e^{\sigma\cdot(r-\theta)(\tau-v)}/\lambda_u$, where λ_u (> 0) is the Lagrange multiplier for the lifetime budget constraint (3.8). Since $r > \theta$, it follows that the agent adopts an upward sloping time profile for its consumption provided the intertemporal substitution elasticity is strictly positive ($\sigma > 0$). The growth rate of individual consumption is thus given by the familiar Euler equation:

$$\frac{\dot{\bar{c}}(v,\tau)}{\bar{c}(v,\tau)} = \sigma\cdot(r-\theta), \qquad \text{for } \tau \in [t,\infty). \tag{3.12}$$

[4] Indeed, for the Blanchard case with a constant death rate, $\Delta(u,\lambda) = 1/(\lambda+\mu_0)$, and (3.11) simplifies even further to $s(v) = \rho/(1-t_L) + 1/(r+\mu_0)$. Apart from the fiscal parameters, this is the expression found in de la Croix and Licandro (1999, p. 258)

By using (3.12) in (3.8) the expression for the consumption level in the planning period is obtained:

$$\Delta(u, r^*)\bar{c}(v,t) = \bar{a}(v,t) + \bar{l}i(v,t), \tag{3.13}$$

where $r^* \equiv r - \sigma\cdot(r-\theta)$ can be interpreted as the effective discount rate facing the agent.

Demography

We allow for non-zero population growth by employing the analytical framework developed by Buiter (1988) which we extended in the previous chapter to include a non-constant mortality rate. To allow for ageing shocks later on, we must extend this model even further. In the previous chapter we assumed that everybody faces the same mortality profile. Here we drop this assumption and assume instead that different cohorts may face different mortality profiles, but that these cohort specific profiles only depend on the individuals age. The instantaneous mortality rate is $m(\alpha, \psi_m(v))$, where $\psi_m(v)$ is a parameter that only depends on the time of birth. We denote the corresponding cumulative mortality by $M(u, \psi_m(v)) = \int_0^u m(\alpha, \psi_m(v))d\alpha$. Wherever possible, we drop the dependency of ψ_m on v or even the dependency of m and M on ψ_m.

The birth rate varies over time, but is still exogenous by assumption. The size of a newborn generation at time v is proportional to the current population at that time, i.e. $L(v,v) = b(v)L(v)$, where $b(v)$ is the – time varying – crude birth rate ($b(v) > 0$), and $L(v)$ is the population size at time v. The size of cohort v at some later time τ is given by:

$$L(v,\tau) = L(v,v)e^{-M(\tau - v, \psi_m(v))} = bL(v)e^{-M(\tau-v,\psi_m)}. \tag{3.14}$$

By definition, the total population at time t satisfies the following expressions:

$$L(t) \equiv \int_{-\infty}^{t} L(v,t)dv, \tag{3.15}$$

$$L(t) \equiv L(v)e^{N(v,t)}, \qquad N(v,t) \equiv \int_v^t n(\tau)d\tau, \tag{3.16}$$

where $n(\tau)$ is the growth rate of the population at time τ. Finally, by combining

(3.14)–(3.16) we obtain:

$$l(v,t) \equiv \frac{L(v,t)}{L(t)} = b(v)e^{-N(v,t)-M(t-v,\psi_m)}, \qquad t \geq v, \tag{3.17}$$

$$\frac{1}{b(v)} = \int_{-\infty}^{t} e^{-N(v,t)-M(t-v,\psi_m)}dv. \tag{3.18}$$

Equation (3.17) shows the population share of the v-cohort at some later time t. It generalizes the corresponding expression (2.23) on page 21 in Chapter 2 to the case of a non-constant population growth rate, $n(t)$. Equation (3.18) implicitly determines $n(t)$ for given demographic parameters (see also Section 3.4). For an economy which has faced the same demographic environment ($b(v) = b$ and $M(t-v,\psi_m) = M(t-v)$) for a long time, the population growth rate is constant ($n(\tau) = n$) and Equation (3.18) reduces to $1/b = \Delta(0,n)$, which is expression (2.23) on page 21.

Per capita plans

Per capita variables are calculated as the integral of the generation-specific values multiplied by the corresponding generation weights, the same as in the previous chapter, section 2.2.1. For example, per capita human capital is defined as:

$$h(t) \equiv \int_{-\infty}^{t} l(v,t)\bar{h}(v,t)dv, \tag{3.19}$$

where $l(v,t)$ and $\bar{h}(v,t)$ are given in, respectively, (3.17) and (3.5) above.

Turning to the wealth components, per capita financial wealth is defined as $a(t) \equiv \int_{-\infty}^{t} l(v,t)\bar{a}(v,t)dv$. By differentiating this expression with respect to time we obtain the dynamic path of per capita financial assets:[5]

$$\dot{a}(t) = [r - n(t)]a(t) + w(t)h(t) - g(t) - c(t), \tag{3.20}$$

where $g(t) \equiv \int_{-\infty}^{t} l(v,t)\bar{g}(v,t)dv$ is per capita tax payments. We assume that the interest rate net of population growth is positive, i.e. $r > n(t)$. As in the standard Blanchard model, annuity payments drop out of the expression for per capita asset accumulation because they constitute transfers (via the life insurance companies) from the deceased to agents who continue to enjoy life.

[5] In deriving (3.20) we have used Equation (3.3) and noted the fact that agents are born without financial assets ($\bar{a}(t,t) = 0$).

3.2.2 Firms

Perfectly competitive firms use physical and human capital to produce a homogeneous commodity, $Y(t)$, that is traded internationally. The technology is represented by the following Cobb-Douglas production function:

$$Y(t) = K(t)^{\varepsilon}[A_Y H(t)]^{1-\varepsilon}, \qquad 0 < \varepsilon < 1, \tag{3.21}$$

where A_Y is a constant index of labour productivity, $K(t) \equiv L(t)k(t)$ is the aggregate stock of physical capital, and $H(t) \equiv L(t)h(t)$ is the aggregate stock of human capital. The cash flow of the representative firm is given by:

$$\Pi(t) \equiv Y(t) - w(t)H(t) - I(t), \tag{3.22}$$

where $w(t)$ is the rental rate on human capital, and $I(t) \equiv \dot{K}(t) + \delta K(t)$ is gross investment, with δ representing the constant depreciation rate. The (fundamental) stock market value of the firm at time t is equal to the present value of cash flows, using the interest rate for discounting, i.e. $V(t) \equiv \int_t^{\infty} \Pi(\tau)e^{r[t-\tau]}d\tau$. The firm chooses paths for $I(\tau)$, $K(\tau)$, $H(\tau)$, and $Y(\tau)$ (for $\tau \in [t, \infty)$) to maximize $V(t)$ subject to the capital accumulation constraint, the production function (3.21) and the definition of cash flows (3.22). Since there are no adjustment costs on investment, the value of the firm equals the replacement value of the capital stock, i.e. $V(t) = K(t)$. In addition, the usual factor demand equations are obtained:

$$r + \delta = \varepsilon \left[\frac{A_Y h(t)}{k(t)}\right]^{1-\varepsilon} = \frac{\partial Y(t)}{\partial K(t)}, \tag{3.23}$$

$$w(\tau) = (1-\varepsilon)A_Y \left[\frac{A_Y h(t)}{k(t)}\right]^{-\varepsilon} = \frac{\partial Y(\tau)}{\partial H(\tau)}. \tag{3.24}$$

For each factor of production, the marginal product is equated to the rental rate. Since the fixed world interest rate pins down the ratio between human and physical capital, it follows from (3.24) that the wage rate is time-invariant, i.e. $w(\tau) = w$,[6]

[6] With labour-augmenting technological change, $\gamma_A \equiv \dot{A}_Y / A_Y$, the wage rate grows exponentially at rate γ_A and Equation (3.11) changes to:

$$s^* - \frac{\rho}{1 - t_L} = \Delta(s^*, r - \gamma_A).$$

It follows from Proposition 1(i) that $\partial s^* / \partial \gamma_A > 0$, i.e. the schooling period depends positively on anticipated wage growth. See also Bils and Klenow (2000, p. 1161) on this issue.

and that physical capital is proportional to human capital at all time:

$$k(t) = A_Y \left[\frac{\varepsilon}{r+\delta} \right]^{1/(1-\varepsilon)} h(t). \tag{3.25}$$

3.2.3 Government and foreign sector

In the absence of government consumption, the government (flow) budget identity in per capita terms is given by:

$$\dot{d}(t) = [r - n(t)]d(t) - g(t), \tag{3.26}$$

where $d(t) \equiv \int_{-\infty}^{t} l(v,t)\bar{d}(v,t)dv$ is per capita government debt. The government solvency condition is $\lim_{\tau\to\infty} d(\tau)e^{r\cdot(t-\tau)+N(t,\tau)} = 0$, so that the intertemporal budget constraint of the government can be written as:

$$d(t) = \int_t^{\infty} g(\tau)e^{r\cdot(t-\tau)+N(t,\tau)}d\tau. \tag{3.27}$$

To the extent that there is outstanding debt (positive left-hand side), it must be exactly matched by the present value of current and future primary surpluses (positive right-hand side), using the net interest rate $(r - n(\tau))$ for discounting purposes.

By using the marginal productivity conditions (3.23)–(3.24) and noting the linear homogeneity of the production function (3.21) and the constancy of factor prices, we find that per capita output, $y(t) \equiv Y(t)/L(t)$, can be written as follows:

$$\begin{aligned} y(t) &= (r+\delta)k(t) + wh(t) \\ &= \left[(r+\delta)^{\varepsilon/(\varepsilon-1)}(\varepsilon A_Y)^{1/(1-\varepsilon)} + w \right] h(t). \end{aligned} \tag{3.28}$$

In going from the first to the second line we have made use of (3.25). It follows from the definition of gross investment that the dynamic evolution of the per capita stock of capital is given by:

$$\dot{k}(t) = i(t) - [\delta + n(t)]k(t), \tag{3.29}$$

where $i(t) \equiv I(t)/L(t)$ is per capita investment. Finally, the current account of the balance of payment, representing the dynamic change in the per capita stock of net foreign assets, $f(t)$, takes the following form:

$$\dot{f}(t) = [r - n(t)]f(t) + y(t) - c(t) - i(t), \tag{3.30}$$

where $f(t) \equiv \int_{-\infty}^{t} l(v,t)\bar{f}(v,t)dv$.[7]

3.2.4 Model solution

The model is recursive and can be solved in three steps. First, for a given demography and with constant tax parameters ρ and t_L, Equation (3.11) determines the optimal schooling period for each agent. Similarly, for a given birth rate, Equation (3.18) can be solved for the population growth rate, $n(t)$. Next, conditional on the optimal value for s^* and the path for $n(t)$, Equation (3.19) can be solved for the equilibrium path of human capital, $h(t)$. Finally, the lumpsum tax z is used to balance the government's intertemporal budget restriction (3.27), after which the values for all remaining variables are fully determined.

In Section 3.3 the effect on the optimal schooling period of both fiscal and demographic shocks are studied. Next, we note that the path for human capital depends critically on the magnitude of the intergenerational externality parameter, ϕ. For values of ϕ in the range $0 \leq \phi < 1$, the model implies a unique steady-state *level* of per capita human capital, i.e. the long-run growth rate in the economy is exogenous (and equal to the population growth rate). This *exogenous growth* case is studied in Section 3.5.

For the knife-edge case with $\phi = 1$, Equation (3.19) gives rise to a unique steady-state *growth rate* in per capita human capital, so that the long-run growth rate is endogenous. This *endogenous growth* model is studied in Section 3.6 below.

3.3 Determinants of schooling

In this section we study the comparative static effect on the optimal schooling period of changes in the fiscal parameters and the demographic process. To keep things simple, only stepwise changes are considered that occur at impact. The time at which the unanticipated and permanent shock occurs is normalised at $t = 0$.

[7] The dynamic expression for per capita assets is given in Equation (3.20), where $a(t) \equiv k(t) + d(t) + f(t)$ (recall that $V(t) = K(t)$). Clearly, total per capita assets $a(t)$ move smoothly over time but its constituting components ($k(t)$, $f(t)$, and $d(t)$) need not. Hence, even in the absence of discrete adjustments in government debt, the capital stock can jump as only $k(t) + f(t)$ moves smoothly over time in that case. A discrete change in $k(t)$ would be engineered by means of an asset swap. Throughout this chapter, however, the world interest rate (r) is held constant so that (via (3.25)) the physical capital stock, $k(t)$, will evolve smoothly because the stock of human capital, $h(t)$, moves smoothly. As a result, the model also gives rise to well-defined current account dynamics—see also Figures 3.4–3.6 below.

3.3.1 Fiscal shocks

The effect of an increase of the educational subsidy on the optimal schooling period have been illustrated in Figure 3.1(a) for the case with a Gompertz-Makeham (G-M) mortality process fitted to actual mortality data for the cohort born in the Netherlands in 1920 (see Chapter 2, Table 2.1 for details).

In terms of Figure 3.1(a), the initial optimum, s_0^*, occurs at the intersection of the line labelled $\Delta + [\rho/(1-t_L)]_0$ and the 45° line. An increase in either ρ or t_L leads to a parallel upward shift in the former line to $\Delta + [\rho/(1-t_L)]_1$ so that the new equilibrium is at s_1^*.

By using (3.11) the comparative static effects of fiscal changes can be computed:

$$\frac{\partial s^*}{\partial \rho} = \frac{1}{(1-t_L)(1-\partial\Delta/\partial s^*)} > 0, \tag{3.31}$$

$$\frac{\partial s^*}{\partial t_L} = \frac{\rho}{(1-t_L)^2(1-\partial\Delta/\partial s^*)} > 0, \tag{3.32}$$

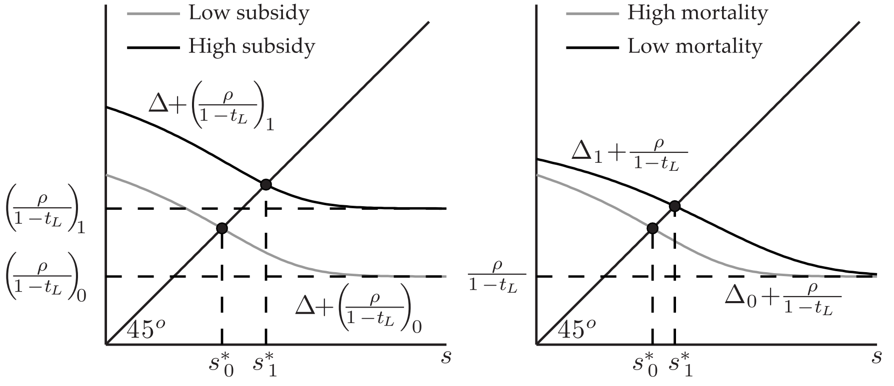
where the signs follow from the fact that $\partial\Delta/\partial s^* \leq 0$ (see Proposition 2.1(ii) on page 17). Not surprisingly, an increase in the educational subsidy leads to a reduction in the opportunity cost of schooling and a longer optimal schooling period. Interestingly, provided the educational subsidy is strictly positive, an increase in the marginal labour income tax also increases the optimal schooling period. Because the educational subsidy is untaxed, the *effective* subsidy affecting the schooling decision is $\rho/(1-t_L)$, which is increasing in t_L.

3.3.2 Demographic shocks

Two types of demographic shocks are considered in our analysis, namely a change in the birth rate and a change in the mortality process. Clearly, in view of (3.11), the birth rate does not affect the optimal schooling period. The mortality process, however, does affect the $\Delta(u,\lambda)$ function and thus the optimal schooling decision. In order to study the effects of changes in the demographic process, we use the notation introduced in section 3.2 and write the instantaneous mortality rate as $m(\alpha, \psi_m)$, where ψ_m is a parameter.[8] In order to investigate the effects of a change in ψ_m we make the following assumptions.

[8] In the Blanchard case, which has only one parameter, μ_0 could be $-\psi_m$ or any decreasing function of ψ_m. For the G-M process, which depends on three parameters (see Table 2.1 on page 31), the parameter *vector* is a function of ψ_m, $(\mu_0, \mu_1, \mu_2) = f(\psi_m)$, and an increase in ψ_m should result in such a change that the G-M mortality function decreases for all ages as ψ_m increases.

Figure 3.1. Effects of educational fiscal shocks (a) and mortality shocks (b) on the optimal schooling period



(a) Fiscal impulse (b) Reduced adult mortality

Assumption 3.1. *The mortality function has the following properties:*

(i) $m(\alpha, \psi_m)$ *is non-negative, continuous, and non-decreasing in age,* $\partial m(\alpha, \psi_m)/\partial\alpha \geq 0$*;*

(ii) $m(\alpha, \psi_m)$ *is convex in age,* $\partial^2 m(\alpha, \psi_m)/\partial\alpha^2 \geq 0$*;*

(iii) $m(\alpha, \psi_m)$ *is non-increasing in* ψ_m *for all ages,* $\partial m(\alpha, \psi_m)/\partial\psi_m \leq 0$*;*

(iv) the effect of ψ_m *on the mortality function is non-decreasing in age,* $\dfrac{\partial^2 m(\alpha, \psi_m)}{\partial\psi_m \partial\alpha} \leq 0.$

An example of a mortality shock satisfying all the requirements of Assumption 3.1 consists of a decrease in μ_1 or μ_2 of the G-M mortality function. In terms of Figure 3.2(a), the shock shifts the mortality function downward, with the reduction in mortality being increasing in age. In panel (b) the function for the surviving fraction of the population shifts to the right. The shock that we consider can thus be interpreted as a reduction in adult mortality. Of course, in view of the terminology of Assumption 1, an increase in ψ_m leads to an increase in the expected remaining lifetime for all ages.

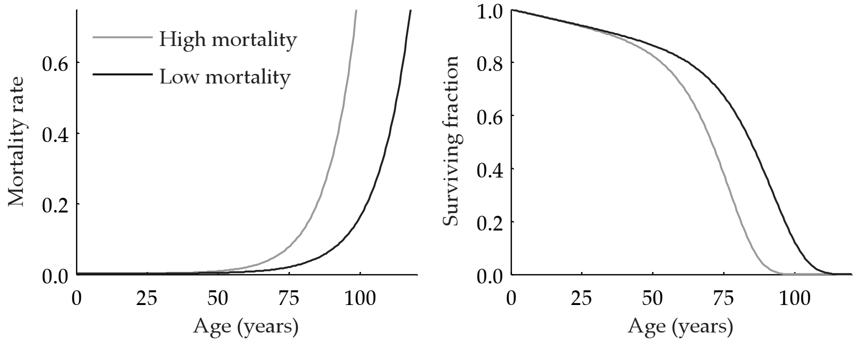
The following results can now be proved.

Proposition 3.1. *Define* $M(u, \psi_m)$ *and* $\Delta(u, \lambda, \psi_m)$ *as:*[9]

$$M(u, \psi_m) \equiv \int_0^u m(\alpha, \psi_m) d\alpha, \tag{3.33}$$

[9] These definitions are generalisations of Equations (2.3) on page 14 and (2.12) on page 16.

Figure 3.2. The effect of reduced adult mortality on the mortality rate (a) and the surviving fraction (b)



(a) Mortality rate, $m(u) = \mu_0 + \mu_1 e^{\mu_2 u}$ (b) Surviving fraction, $S(u) = e^{-M(u)}$

Notes: Mortality process is a Gompertz-Makeham (see Table 2.1 for parameter values). Low mortality correspond to a 50% decrease of μ_1 and 10% decrease of μ_2.

$$\Delta(u, \lambda, \psi_m) \equiv e^{\lambda u + M(u,\psi_m)} \int_u^\infty e^{-\lambda\alpha - M(\alpha,\psi_m)} d\alpha. \tag{3.34}$$

Under Assumption 3.1, the following results can be established.

(i) $$\frac{\partial M(u, \psi_m)}{\partial \psi_m} = \int_0^u \frac{\partial m(\alpha, \psi_m)}{\partial \psi_m} d\alpha \leq 0;$$

(ii) $$\frac{\partial^2 M(u, \psi_m)}{\partial u \partial \psi_m} = \frac{\partial m(u, \psi_m)}{\partial \psi_m} \leq 0;$$

(iii) $$\frac{\partial \Delta(u, \lambda, \psi_m)}{\partial \psi_m} = e^{\lambda u + M(u,\psi_m)} \int_u^\infty \left[\frac{\partial M(u,\psi_m)}{\partial \psi_m} - \frac{\partial M(\alpha,\psi_m)}{\partial \psi_m} \right] e^{-\lambda\alpha - M(\alpha,\psi_m)} d\alpha > 0.$$

Proof. (i) and (ii) follow from simple differentiation and noting assumption 3.1(iii). (iii) follows from differentiation of (3.34) and (i). ■

By using Equation (3.11), and noting the definition (3.34), the comparative static effect on the optimal schooling period of a reduction in adult mortality can be computed:

$$\frac{\partial s^*}{\partial \psi_m} = \frac{\partial \Delta / \partial \psi_m}{1 - \partial \Delta / \partial s^*} > 0, \tag{3.35}$$

where the sign follows from the fact that $\partial \Delta / \partial s^* \leq 0$ (see Proposition 2.1(ii)) and $\partial \Delta / \partial \psi_m > 0$ (see Proposition 3.1(iii)). An increase in longevity prompts agents to increase their human capital investment at the beginning of life. In terms of Figure

3.1(b), the mortality shock shifts the Δ-function to the right, and leads to an increase in the optimal schooling period from s_0^* to s_1^*.

Bils and Klenow argue that a higher life expectancy (as captured in their model by an increase in the exogenous planning horizon) leads to an increase in the optimal schooling period 'since it affords a longer working period over which to reap the wage benefits of schooling' (2000, p. 1164). Similarly, de la Croix and Licandro (1999, p. 258) and Kalemni-Ozcan et al. (2000, p. 11), using the Blanchard demography, show that a decrease in the death probability leads to an increase in the expected planning horizon for all agents and an increase in the optimal schooling period. Our discussion shows that these conclusions are misleading in the presence of lifetime uncertainty *and* age-dependent mortality. In our model, a decrease in child mortality increases expected remaining life time at birth but leaves the optimal schooling period unchanged. In terms of Figure 3.1(b), reduced child mortality flattens the left-hand section of the line $\Delta_0 + \rho/(1 - t_L)$ but the equilibrium solution stays at s_0^*.[10] [11] Of course, with the Blanchard demography one cannot distinguish between child mortality and adult mortality because the death probability is age-independent.

3.4 Demographic shocks and population growth

Demographic changes affect the growth rate of the population, both at impact, during transition, and in the long run. Armed with Proposition 2.1 and 3.1 we can compute the long-run effects of changes in the birth rate and the mortality process. Indeed, since Equation (3.18) reduces in the steady state to $b\Delta(0, \hat{n}, \psi_m) = 1$, it follows that $\hat{n}$ is an implicit function of b and ψ_m, the partial derivatives of which are given by:

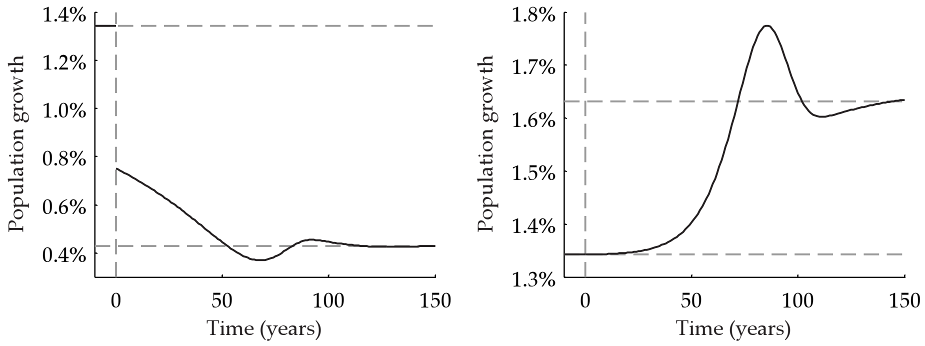
$$\frac{\partial \hat{n}}{\partial b} = -\frac{\Delta(0, \hat{n}, \psi_m)}{b \partial \Delta(0, \hat{n}, \psi_m)/\partial \hat{n}} > 0, \tag{3.36}$$

$$\frac{\partial \hat{n}}{\partial \psi_m} = -\frac{\partial \Delta(0, \hat{n}, \psi_m)/\partial \psi_m}{\partial \Delta(0, \hat{n}, \psi_m)/\partial \hat{n}} > 0, \tag{3.37}$$

[10] Boucekkine et al. also distinguish age-dependent mortality and argue that 'an increase in life expectancy increases the optimal length of schooling' (2002, pp. 352, 370). They thus fail to notice that the mechanism producing this result runs via reduced old-age mortality, not via increased life expectancy in general.

[11] Bils and Klenow (2000, p. 1175) also report that their model implies an unrealistically high sensitivity of the optimal schooling period with respect to life expectancy that is close to unity. In contrast, in the calibrated version of our model, $ds^*/dR(0) = 0.06$ which comes close to the empirical estimate mentioned by Bils and Klenow (2000, p. 1175n27).

Figure 3.3. Population growth rate after a baby bust (a) and an adult mortality shock (b).



(a) Baby bust

(b) Reduced adult mortality

Notes: Mortality process is a Gompertz-Makeham (see Table 2.1 for parameter values), birth rate is 2.36%. Baby bust is a 25% downward jump of the birth rate to 1.78%. Reduced adult mortality is a 50% decrease of μ_1 and 10% decrease of μ_2.

where a hat overstrike designates the steady-state value of a variable, i.e. $\hat{n}$ is the steady-state growth rate of the population. The signs in (3.36)–(3.37) follow from Propositions 2.1(i) and 3.1(iii). Not surprisingly, an increase in the birth rate and an increase in longevity both lead to an increase in the steady-state growth rate of the population.

To compute the transition path for the growth rate of the population we assume that at time $t = 0$ both the mortality process and the birth rate change in a stepwise fashion. The mortality shock is assumed to be *embodied*, i.e. it only affects generations born from time $t = 0$ onwards. Indeed, the mortality process for pre-shock cohorts (with a negative generation index, $v < 0$) is described by $M_0(t - v)$ and $m_0(t - v)$, whereas post-shock cohorts (with $v \geq 0$) face the mortality process described by $M_1(t - v)$ and $m_1(t - v)$. In a similar fashion, the pre-shock and post-shock birth rates are denoted by, respectively, b_0 and b_1. The system is initially in a demographic steady state and the pre-shock population growth rate is denoted by $\hat{n}_0$ (defined implicitly by the condition $1 = b_0\Delta_0(0, \hat{n}_0)$, where $\Delta_0(0, \hat{n}_0)$ is the Δ-function associated with the initial mortality process).

As a consequence of the demographic changes, the path for the population growth rate is implicitly determined by the following expression:

$$1 = b_0 \int_{-\infty}^{0} e^{-M_0(t-v)-N(v,t)} dv + b_1 \int_{0}^{t} e^{-M_1(t-v)-N(v,t)} dv, \tag{3.38}$$

where $N(v,t) \equiv \int_v^t n(\tau)d\tau$ (see also (3.16) above). In Box 3.1 we show that Equation (3.38) can be rewritten in the form of a linear Volterra equation of the second kind with a convolution-type kernel for which efficient numerical solution algorithms are available. In Figure 3.3 we plot the transition path for $n(t)$ for both types of demographic shocks. Panel (a) depicts the path for a baby bust. There is an immediate downward jump at impact ($n(0) = \hat{n}_0 - b_0 + b_1$) followed by gradual cyclical adjustment. Adjustment is rather fast because the birth rate change applies to the entire (pre-shock and post-shock) population alike. Panel (b) of Figure 3.3 depicts the adjustment path following a decrease in adult mortality. Nothing happens at impact and the population growth rate only gradually rises to its long-run steady-state value. Transition is much slower than for the baby bust because the ageing shock is embodied, i.e. the shock only applies to post-shock generations and pre-shock generations only die off gradually during the demographic transition.

BOX 3.1

Population growth after demographic shocks

The transition path for $n(t)$ is determined implicitly by Equation (3.38) in the text, but this equation is useless to compute the path of the population growth rate. We can however rewrite this equation into a so-called Volterra Equation of the second kind, for which there are standard (and efficient) solution algorithms. Start by multiplying both sides of this expression by $e^{N(0,t)}$, and noting that $N(v,t) \equiv N(0,t) - N(0,v)$

$$e^{N(t)} = b_0 \int_{-\infty}^{0} e^{-M_0(t-v)+N(v)} dv + b_1 \int_{0}^{t} e^{-M_1(t-v)+N(v)} dv,$$

where we define $N(t) \equiv N(0,t)$ for notational convenience. Since $N(v) = \hat{n}_0 v$ for $v < 0$ we find

$$e^{N(t)} = b_0 e^{\hat{n}_0 t} \int_{-\infty}^{0} e^{-\hat{n}_0 \cdot (t-v) - M_0(t-v)} dv + b_1 \int_{0}^{t} e^{N(v)-M_1(t-v)} dv$$

$$= b_0 e^{\hat{n}_0 t} \int_t^\infty e^{-\hat{n}_0 u - M_0(u)} du + b_1 \int_0^t e^{N(v) - M_1(t-v)} dv \tag{3.39}$$

$$= b_0 e^{-M_0(t)} \Delta_0(t, \hat{n}_0) + b_1 \int_0^t e^{N(v) - M_1(t-v)} dv, \tag{3.40}$$

where we changed variables from the cohort domain to the age domain in going from the first to the second line, and use the definition of $\Delta_0(t, \hat{n}_0)$ in going from the second to the third line.

Since the long-run population growth rate equals $\hat{n}_1$, it follows that (3.40) can be rewritten in a stationary format by multiplying both sides of the expression by $e^{\hat{n}_1 t}$. We obtain:

$$\xi(t) = \chi(t) + \int_0^t \boldsymbol{K}(t-v)\xi(v)dv, \tag{3.41}$$

where $\xi(t) \equiv e^{N(t) - \hat{n}_1 t}$, $\chi(t) \equiv b_0 e^{-M_0(t) - \hat{n}_1 t} \Delta_0(t, \hat{n}_1)$, and $\boldsymbol{K}(t-v) \equiv b_1 e^{-M_1(t-v) - \hat{n}_1(t-v)}$. Equation (3.41) is a so-called a *renewal equation*, i.e. a linear Volterra equation of the second kind with a convolution type kernel—see *inter alia* Linz (1985, p. 14) and Bellman and Cooke (1963, ch. 7). We use the standard solution algorithm proposed by Linz (1985, p 98) which generates approximations for $\xi(t)$ on a grid with constant step size. If we denote the step size by ϵ, we can approximate $\xi(t)$ at each gridpoint $t = i\epsilon$, $i = 1, 2 \ldots$

$$\xi(i\epsilon) \doteq \chi(i\epsilon) + \epsilon \left[\frac{1}{2}\boldsymbol{K}(i\epsilon)\xi(0) + \sum_{j=1}^{i-1} \boldsymbol{K}([i-j]\epsilon)\xi(j\epsilon) + \frac{1}{2}\boldsymbol{K}(0)\xi(i\epsilon) \right].$$

where we used the simple trapezoidal rule to numerically evaluate the integral. Solving this equation for $\xi(i\epsilon)$ we get

$$\xi(i\epsilon) \doteq \frac{\chi(i\epsilon) + \epsilon \left[\frac{1}{2}\boldsymbol{K}(i\epsilon)\xi(0) + \sum_{j=1}^{i-1} \boldsymbol{K}([i-j]\epsilon)\xi(j\epsilon) \right]}{1 - \epsilon K(0)/2}.$$

The value of $\xi(0)$ is known, $\xi(0) = \chi(0)$. From this we can calculate the value of $\xi(\epsilon)$ and keep on jumping forward in time until we have all the required points. From the path of $\xi(t)$ it is easy to derive path for $n(t)$ by noting that $n(t) \equiv \tilde{\xi}(t) - \hat{n}_1$, where $\tilde{\xi}(t) \equiv d \ln \xi(t)/dt$ can be computed easily with the aid of finite difference methods. The only problem is to determine the initial growth rate, $n(0)$, because the population growth might not be continuous at

$t = 0$ so finite difference methods do not work. By differentiating (3.39)) with respect to time and evaluating the result for $t = 0$ we find that $n(0) = \hat{n}_0 - b_0 + b_1$. In deriving this result we make use of the fact that $M_0(0) = M_1(0) = 0$, $N(0) = 0$, and $b_0\Delta_0(0, \hat{n}_0) = 1$. For our purposes, setting $\epsilon = 1$ is accurately enough, but if more accuracy is required, we can decrease the stepsize ϵ.

3.5 Exogenous growth

In Section 3.3 it was shown that both fiscal and demographic shocks lead to a change in the optimal schooling period, s^*. In this section we study the resulting transitional and long-run effects on human capital formation for the *exogenous growth* case, i.e. we assume that the intergenerational knowledge transfer incorporated in the human capital production function (3.5) is either absent ($\phi = 0$) or subject to diminishing returns ($0 < \phi < 1$). First, in Section 3.5.1 we analytically characterize the steady-state and study its sensitivity with respect to fiscal and demographic shocks. Next, in Section 3.5.2 we visualise the rather complicated transitional dynamics associated with the various shocks for a plausibly parametrized model which incorporates the estimated G-M process introduced above.

3.5.1 Long-run effects

In the long-run equilibrium, Equation (3.19) gives rise to the following expression for the steady-state stock of per capita human capital, $\hat{h}$:

$$\hat{h}^{1-\phi} = A_H s^* b \int_{s^*}^{\infty} e^{-\hat{n}u - M(u,\psi_m)} du. \tag{3.42}$$

Equation (3.42) clearly shows the various mechanisms affecting $\hat{h}$, namely (i) the birth rate, (ii) the optimal schooling decision of agents, s^*, which itself depends on the fiscal and mortality parameters (ρ, t_L, ψ_m), (iii) the population growth rate, $\hat{n}$, which depends on (b, ψ_m), and (iv) the cumulative mortality factor, $M(u, \psi_m)$, which depends on the mortality parameter ψ_m.

Pure schooling shock In order to facilitate the interpretation of our results, we first study the effects of a change in the schooling period in isolation. By differentiating Equation (3.42) with respect to s^* and simplifying we obtain:

$$\begin{aligned}\frac{\partial \hat{h}^{1-\phi}}{\partial s^*} &= A_H b e^{-\hat{n}s^* - M(s^*,\psi_m)} [\Delta(s^*,\hat{n}) - s^*] \\ &= A_H b e^{-\hat{n}s^* - M(s^*,\psi_m)} \left[\Delta(s^*,\hat{n}) - \Delta(s^*,r) - \frac{\rho}{1-t_L}\right], \end{aligned} \tag{3.43}$$

where we have used (3.11) to arrive at the second expression. In the absence of an educational subsidy ($\rho = 0$), a pure schooling shock unambiguously leads to an increase in the per capita stock of human capital. Indeed, since by assumption the interest rate exceeds the steady-state growth rate of the population ($r > \hat{n}$), it follows from Proposition 2.1(i) that $\Delta(s^*,\hat{n}) > \Delta(s^*,r)$ so that $\partial\hat{h}^{1-\phi}/\partial s^* > 0$ in that case. With a non-zero educational subsidy, Equation (3.43) shows that the effect on $\hat{h}$ of a pure schooling shock is no longer unambiguous because a sufficiently high effective educational subsidy will render the term in square brackets negative even for the case with $r > \hat{n}$. Intuitively, in such a case the economy is 'over-educated', i.e. agents study for too long a period and thus have too short a career as productive workers. Because in actual economies r is much greater than $\hat{n}$ and educational subsidies are typically quite low, we make the following assumption which rules out over-education and ensures that $\partial\hat{h}^{1-\phi}/\partial s^*$ is positive.

Assumption 3.2. *The steady-state net interest rate $r - \hat{n}$ is sufficiently positive to ensure that $\Delta(s^*,\hat{n}) > \Delta(s^*,r) + \rho/(1-t_L)$.*

Fiscal shock A fiscal shock, consisting of an increase in either ρ or t_L, affects the steady-state per capita human capital stock according to:

$$\frac{\partial \hat{h}^{1-\phi}}{\partial\left[\rho/(1-t_L)\right]} = \frac{\partial \hat{h}^{1-\phi}}{\partial s^*}\frac{\partial s^*}{\partial\left[\rho/(1-t_L)\right]} > 0, \tag{3.44}$$

where the sign follows from (3.31)–(3.32) above. The fiscal shock leads to an increase in the optimal schooling period which, in view of Assumption 3.2, leads to an increase in $\hat{h}$.

Birth rate shock A change in the birth rate affects steady-state per capita human capital both directly and via its effect on the steady- state population growth rate. By differentiating Equation (3.42) with respect to b and simplifying we obtain:

$$\frac{\partial \hat{h}^{1-\phi}}{\partial b} = A_H s^* \left[\int_{s^*}^{\infty} e^{-\hat{n}u - M(u,\psi_m)} du - b\frac{\partial \hat{n}}{\partial b}\int_{s^*}^{\infty} u e^{-\hat{n}u - M(u,\psi_m)} du\right] < 0, \tag{3.45}$$

where the sign follows from Lemma 3.1 in Appendix 3.B. Intuitively, a higher birth rate leads to an upward shift in the steady-state path of the human capital stock in *level* terms, but also induces an increase in the population growth rate. The latter effect dominates the former so that *per capita* human capital declines in the steady state.

Mortality shock The mortality change is by far the most complicated shock under consideration because it affects the schooling period, s^*, the population growth rate, $\hat{n}$, and the cumulative mortality factor, $M(u, \psi_m)$. By differentiating (3.42) with respect to ψ_m we obtain:

$$\frac{\partial \hat{h}^{1-\phi}}{\partial \psi_m} = \frac{\partial \hat{h}^{1-\phi}}{\partial s^*}\frac{\partial s^*}{\partial \psi_m} + A_H s^* b \frac{\partial}{\partial \psi_m}\int_{s^*}^{\infty} e^{-\hat{n}u - M(u,\psi_m)} du > 0, \tag{3.46}$$

where the sign follows from (3.35), (3.43), and Lemma 3.2 in Appendix 3.B. The first composite term on the right-hand side is straightforward: increased longevity boosts the optimal schooling period which in turn increases per capita human capital in the steady state. The second term on the right-hand side represents the joint effect of increased longevity on the integral appearing on the right-hand side of (3.42). An increase in ψ_m has two effects on the discounting factor of that integral. First, the population growth rate is increased ($\partial \hat{n}/\partial \psi_m > 0$) leading to heavier discounting and a lower value for the integral. Higher population growth constitutes a higher drag on human capital as the cake must be shared over ever more people. This effect leads to a decrease in per capita human capital. Second, the cumulative mortality factor is decreased for higher age levels ($\partial M(u, \psi_m)/\partial \psi_m < 0$) leading to reduced discounting and a higher integral. Educated people live longer as a result of the shock and per capita human capital increases as a result. Lemma 3.2 in Appendix 3.B shows that, under our set of assumptions regarding mortality change, the first effect is dominated by the second and, ceteris paribus the schooling period, *human-capital deepening* occurs as a result of increased longevity, i.e. the second composite term on the right-hand side of (3.46) is positive.

Balanced growth Up to this point attention has been restricted to steady-state per capita human capital. This focus is warranted because all remaining variables are uniquely related to $\hat{h}$. Indeed, it follows directly from, respectively, (3.25) and (3.28), that $\hat{k}$ and $\hat{y}$ are both proportional to $\hat{h}$. The level of per capita human capital determines individual human capita, which fixes the individual lifetime income profile. From this lifetime income profile and the propensity to consume follow consumption and individual assets. The other per capita variables follow from these individual profiles and the generational weights given in (3.17). In the steady state all per capita aggregate variables are constant, so their levels grow at the steady state population growth rate.

3.5.2 Transitional dynamics

In this subsection we compute and visualise the transitional effects of fiscal and demographic shocks using a plausibly calibrated version of the model.[12] The world interest rate is $r = 0.055$, the pure rate of time preference is $\theta = 0.03$, the intertemporal substitution elasticity is $\sigma = 1$, the capital depreciation rate is $\delta = 0.07$, and the efficiency parameter for physical capital is $\varepsilon = 0.3$.

The human capital externality parameter is set at $\phi = 0.3$. We rationalize this choice as follows. In a recent paper, de la Fuente and Doménech (2006, p. 12) formulate an aggregate production function of the form:

$$\ln y_i(t) = \ln TFP_i(t) + \alpha_1 \ln k_i(t) + \alpha_2' \ln s_i(t), \tag{3.47}$$

where i is the country index, TFP_i is total factor productivity, k_i is capital per worker, and s_i measures education attainment, i.e. the average years of education of *employed* workers. Since their data on educational attainment refers to the total (rather than the employed) population, they postulate the relationship $\ln s_i(t) = \beta_1 \ln \bar{s}_i(t) - \beta_2 \ln PR_i(t)$, where $\bar{s}_i$ measures population average education attainment (i.e. average years of schooling in the adult population), and PR_i is the participation rate (i.e. the proportion of employed adults). Substituting this expression into (3.47) they derive the equation to be estimated:

$$\ln y_i(t) = \ln TFP_i(t) + \alpha_1 \ln k_i(t) + \alpha_2 \ln \bar{s}_i(t) + \alpha_3 \ln PR_i(t), \tag{3.48}$$

where $\alpha_2 \equiv \alpha_2' \beta_1$ and $\alpha_3 \equiv -\alpha_2' \beta_2$. They present panel data estimates for the parameters, using different specifications for $\ln TFP_i(t)$, and find large and highly significant values for α_2 ranging from 0.378 to 0.958 (de la Fuente and Doménech, 2006, p. 14). They argue on the basis of meta-estimation that the lower bound for the key parameter of interest, α_2', lies in the range of 0.752 to 0.844 for the fixed-effect regressions. They conclude that '...investment in human capital is an important

[12] Kalemni-Ozcan et al. (2000) restrict attention to the steady state. Boucekkine et al. (2002, pp. 363-365) only show the adjustment path in the endogenous growth rate following a drop in the birth rate.

growth factor whose effect on productivity has been underestimated in previous studies because of poor data quality' (de la Fuente and Doménech, 2006, p. 28).

What does this say about our ϕ parameter? In the steady state our model implies the following relationship:

$$\ln \hat{y} = \alpha_0 + \varepsilon \ln \hat{k} + \frac{1-\varepsilon}{1-\phi} \ln s^*, \tag{3.49}$$

where $\alpha_0 \equiv (1-\varepsilon)\ln A_Y + \frac{1-\varepsilon}{1-\phi}\ln(bA_H \int_{s^*}^{\infty} e^{-\hat{n}u - M(u,\psi_m)}du)$. Ignoring the fact that in Equation (3.49) the constant term itself depends negatively on s^*, we find that $\hat{\alpha}_1$ is an estimate of ε and $\hat{\alpha}_2'$ is an estimate of $(1-\varepsilon)/(1-\phi)$. de la Fuente and Doménech find estimates for $\hat{\alpha}_1$ in the range 0.448 to 0.491, so that the implied estimate for ϕ is given by $\hat{\phi} \equiv 1 + (\hat{\alpha}_1 - 1)/\hat{\alpha}_2'$ which ranges from 0.266 to 0.397.[13] Our chosen value of ϕ falls within this range.

On the demographic side, we use the same specification as in section 2.4 in the previous chapter. We interpret the estimated G-M demography as the truth and choose the birth rate, b, such that $\hat{n} = 0.0134$ (the average population growth rate during the period 1920-1940). This yields a value of $b = 0.0237$. The estimated G-M model yields an expected remaining lifetime at birth of 65.5 years (see Table 2.1 on page 31 for details). We compute the implied wage rate from the factor price frontier and find $w = 1.019$. The initial lumpsum tax follows from the government solvency condition for an initial debt level of $\hat{d}_0 = -2.112$ and fiscal parameters $\rho = 4.915$ and $t_L = 0.15$. The implied value for the lumpsum tax is $z_0 = 0.2645$. Finally, for the scaling variables we use $A_H = A_Y = 1$. The initial age at which agents leave school and enter the labour market is $s_0^* = 21.82$ years. The initial steady state has the following main features: $\hat{a}_0 = 7.8$, $\widehat{li}_0 = 647.2$, $\hat{h}_0 = 36.1$, $\hat{y}_0 = 52.6$, $\hat{c}_0 = 37.2$, $\hat{i}_0 = 10.5$, $\hat{k}_0 = 126.2$, and $\hat{f}_0 = -116.2$. The output shares of consumption, investment, and net exports are, respectively, 0.71, 0.20, and 0.09.

The economy is initially in a steady-state equilibrium, the stepwise shock occurs at time $t = 0$, and we refer to pre-shock ($v < 0$) and post-shock agents ($v \geq 0$). In the interest of brevity, we focus the discussion on the transition path of per capita human capital. As is seen readily from (3.25) and (3.28), the time paths for $k(t)$ and $y(t)$ are proportional to that of $h(t)$. The remaining variables of the model (such as $d(t)$, $i(t)$, $f(t)$, $li(t)$, $a(t)$, and $c(t)$) feature more complicated dynamic adjustment paths but are of less interest for the main purpose of this chapter. Where

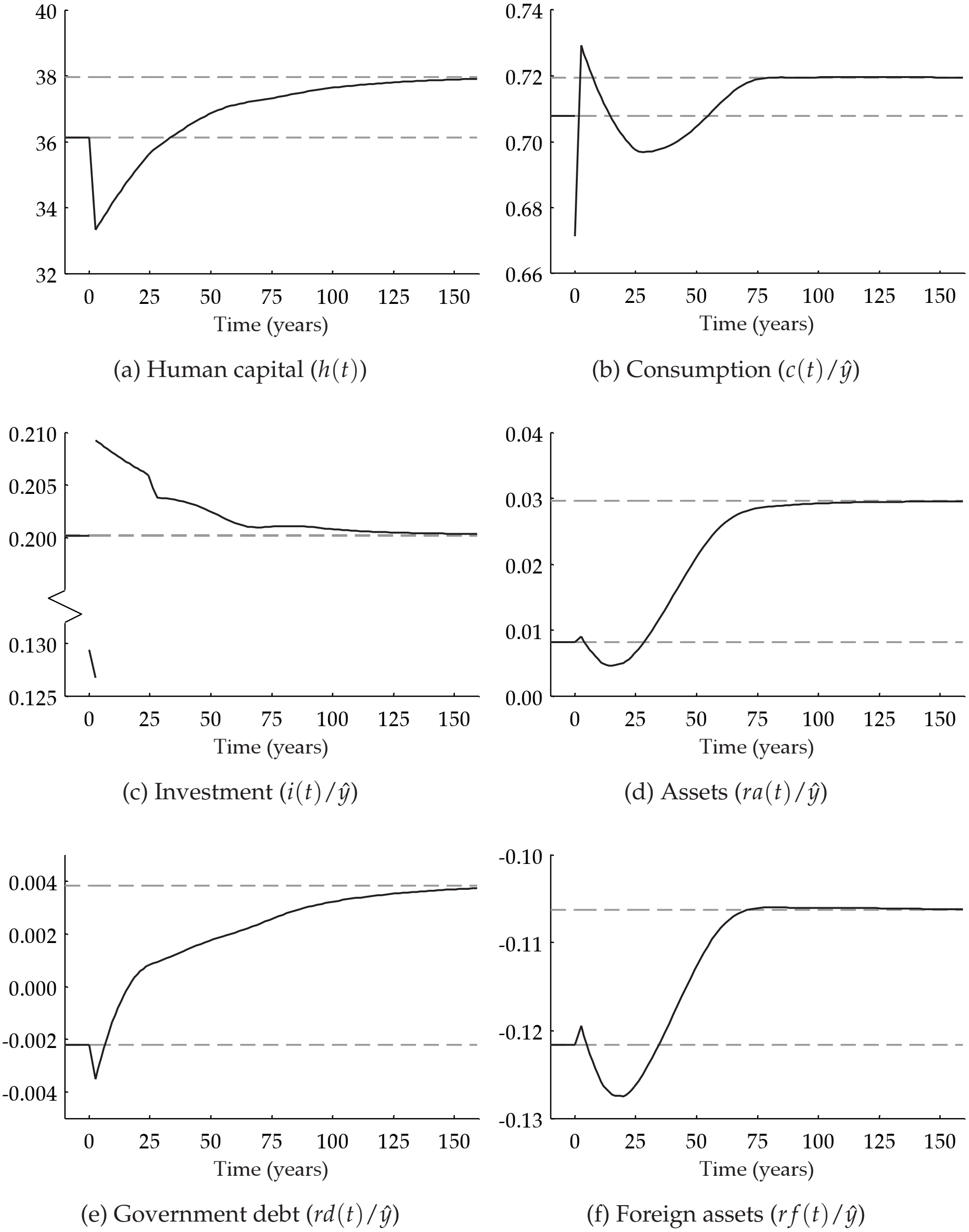
[13] Of course, this is only a very tentative estimate for ϕ for at least two reasons. First, the data may not represent observations for the steady state. Second, the procedure ignores the fact that α_0 itself also depends on s^*. This may lead to an under-estimate for ϕ.

no confusion can arise we drop the 'per capita' adjective in the intuitive discussion of our results.

Fiscal shock In Figure 3.4 we illustrate the transitional dynamics associated with a fiscal education impulse, consisting of a 50% increase in the educational subsidy, from $\rho_0 = 4.915$ to $\rho_1 = 7.372$. There is no effect on the demography so the population growth rate is unchanged ($n(t) = \hat{n}_0$). The human capital of pre-shock workers is unaffected because labour market entry is an *absorbing state*, i.e. workers cannot go back to school by assumption. Pre-shock students, however, react to the improved fiscal incentives by extending their schooling period from $s_0^* = 21.8$ to $s_1^* = 22.9$. As a result, in the time interval $0 \le t < s_1^* - s_0^*$ there are no new labour market entrants and human capital declines sharply as a result of the mortality process—see Figure 3.4(a). Labour market entry resumes for $t \ge s_1^* - s_0^*$ and the entrants have a higher level of education, so human capital starts to rise as a result. During the interval $s_1^* - s_0^* \le t < s_1^*$ entry consists entirely of pre-shock students, whereas for $t \ge s_1^*$ only post-shock cohorts enter the labour market. Since these cohorts choose the same schooling period s_1^*, adjustment in human capital is monotonic. For $t \to \infty$, the system reaches a new steady-state which features a higher stock of human capital (see also (3.44) above).

Panels (b)–(f) of Figure 3.4 illustrate the adjustment paths of the other macroeconomic variables. In panel (b) consumption falls at impact due to the once-off increase in the lumpsum tax needed to finance the increase in the educational subsidy. During transition, however, consumption increases non-monotonically as a result of the increase in lifetime income caused by the increase in human capital. In panel (e) the path for government debt is illustrated. Debt fluctuates during transition because the government engages in tax smoothing with respect to the lumpsum tax, z. The current account dynamics is illustrated in panel (f). At impact, the reduction in consumption and investment dominates the reduction in output, so that net exports increase and the stock of net foreign assets rises sharply. During transition, however, net foreign assets gradually fall during the first two decades of adjustment after which they rise to a permanently higher level. In a similar fashion, the path for total assets is non-monotonic due to the population heterogeneity that exists during transition. Indeed, during transition three broad cohort types coexist, namely pre-shock workers (who base their savings decisions on the pre-shock schooling choice s_0^*), pre-shock students (who switched from s_0^* to s_1^* at time $t = 0$ and changed their savings plans accordingly), and post-shock cohorts (who

Figure 3.4. Aggregate effect of a fiscal education impulse.



(a) Human capital ($h(t)$)

(b) Consumption ($c(t)/\hat{y}$)

(c) Investment ($i(t)/\hat{y}$)

(d) Assets ($ra(t)/\hat{y}$)

(e) Government debt ($rd(t)/\hat{y}$)

(f) Foreign assets ($rf(t)/\hat{y}$)

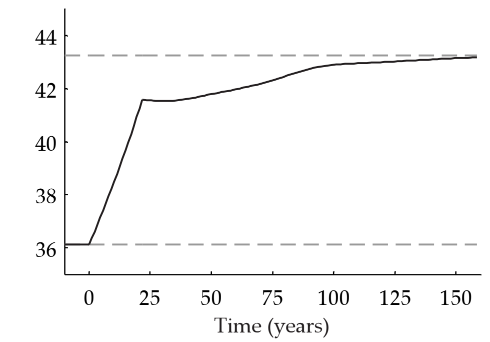
Notes: Mortality process is a Gompertz-Makeham (see Table 2.1 for parameter values), birth rate is 2.36%. Fiscal education shock is a 50% increase in the educational subsidy, from $\rho_0 = 4.915$ to $\rho_1 = 7.372$. Results are absolute differences relative to the old steady state values.

all choose s_1^* and, provided $\phi > 0$, face changing initial conditions because human capital changes over time).

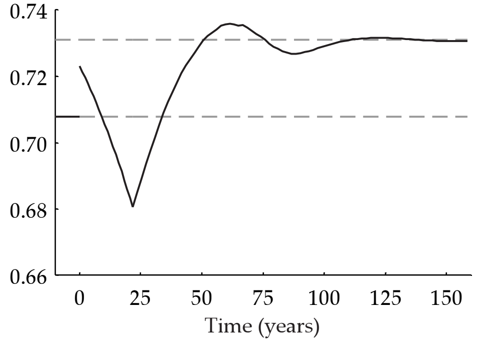
Birth rate shock In Figure 3.5 we illustrate the transitional dynamics associated with a baby bust, that is the birth rate drops once and for all by 25% from $b_0 = 2.37\%$ to $b_1 = 1.78\%$. Nothing happens to the optimal schooling choice, but the population growth rate falls in a non-monotonic fashion from $\hat{n}_0 = 1.34\%$ to $\hat{n}_1 = 0.43\%$ as is illustrated in Figure 3.3(a). The sharp increase in human capital in Figure 3.5(a) is entirely attributable to the fast reduction in $n(t)$ during the early phase of transition. At time $t = s_0^*$, the population growth rate is close to its new steady state and the slope of the per capita human capital stocks flattens out. This is because the flow of labour market entrants is smaller than before as it consists entirely of post-shock newborns. In the new steady state, per capita human capital increases as a result of the baby bust (see also (3.45) above). For completeness sake, the paths for the remaining macroeconomic variables are also illustrated in panels (b)–(f) of Figure 3.5.

Mortality shocks In Figure 3.6 we illustrate the transitional dynamics associated with an adult mortality shock leading to increased longevity. The μ_1-parameter of the G-M process is reduced by 50% and the μ_2 parameter by 10%, leading to an increase of the expected lifetime at birth from $R_0(0) = 65.45$ to $R_1(0) = 77.57$ years. In the face of increased longevity, post-shock cohorts choose a longer schooling period ($s_1^* = 22.5$ instead of $s_0^* = 21.8$). Furthermore, the shock perturbs the demographic steady-state and causes a rather slow non-monotonic increase in the population growth rate, from $\hat{n}_0 = 1.34\%$ to $\hat{n}_1 = 1.63\%$ as is illustrated in Figure 3.3(b). The transition in human capital passes through the following phases. During the interval $0 \leq t < s_0^*$ nothing happens to human capital because only pre-shock students (facing an unchanged mortality process) enter the labour market and the mortality process for pre-shock workers has not changed. For $s_0^* \leq t < s_1^*$ there are no new labour market entrants at all because post-shock students choose a schooling period s_1^*. Human capital declines sharply because (a) pre-shock cohorts die off at the rate implied by the pre-shock mortality process, and (b) the population growth rate increases. For $t \geq s_1^*$ post-shock cohorts enter the labour market. The closer the birth rate of such cohorts is to s_1^*, the worse are their initial conditions in the human capital formation process. Indeed, the cohort born at time $t = s_1^*$ faces low schooling productivity because $h(s_1^*)$ is quite low. As is clear from Figure 3.6(a), human capital increases in a non-monotonic fashion after $t = s_1^*$, where the

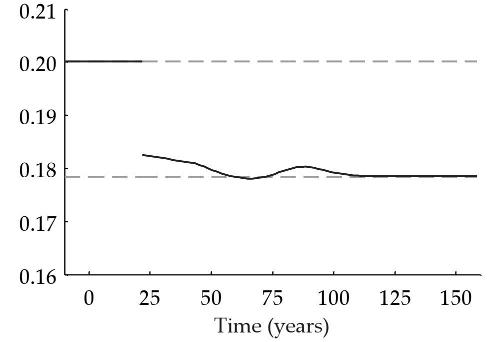
Figure 3.5. Aggregate effect of a baby bust



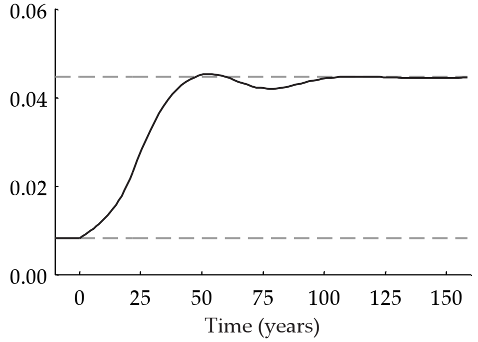
(a) Human capital ($h(t)$)



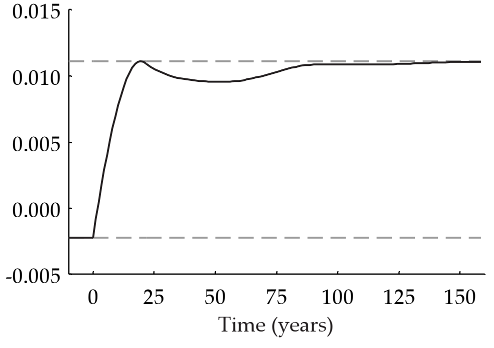
(b) Consumption ($c(t)/\hat{y}$)



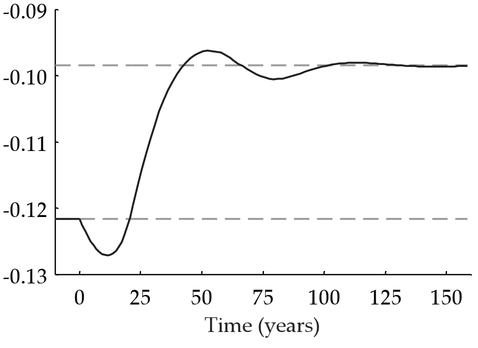
(c) Investment ($i(t)/\hat{y}$)



(d) Assets ($ra(t)/\hat{y}$)



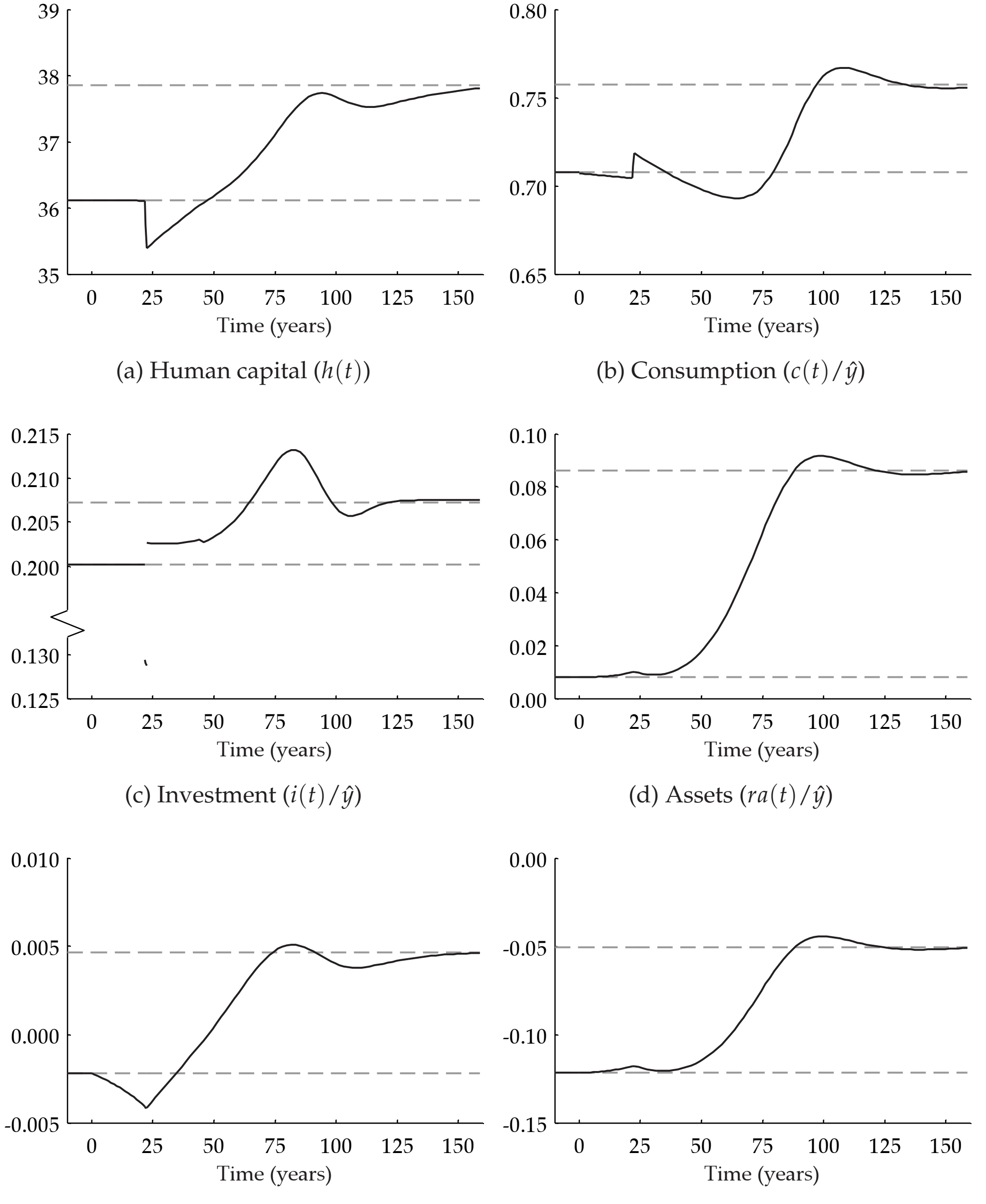
(e) Government debt ($rd(t)/\hat{y}$)



(f) Foreign assets ($rf(t)/\hat{y}$)

Notes: Mortality process is a Gompertz-Makeham (see Table 2.1 for parameter values), birth rate is 2.36%. Baby bust is a 25% downward jump of the birth rate to 1.78%. Results are absolute differences relative to the old steady state values.

Figure 3.6. Aggregate effect of reduced adult mortality



(a) Human capital ($h(t)$)

(b) Consumption ($c(t)/\hat{y}$)

(c) Investment ($i(t)/\hat{y}$)

(d) Assets ($ra(t)/\hat{y}$)

(e) Government debt ($rd(t)/\hat{y}$)

(f) Foreign assets ($rf(t)/\hat{y}$)

Notes: Mortality process is a Gompertz-Makeham (see Table 2.1 for parameter values, birth rate is 2.36%. Reduced adult mortality is a 50% decrease of μ_1 and 10% decrease of μ_2. Results are absolute differences relative to the old steady state values.

bump after about 95 years is due to the corresponding maximum in the population growth rate at that time—see Figure 3.3(b).

3.5.3 Discussion

The main findings of this section are as follows. Provided the intergenerational externality parameter is below the knife-edge value of unity, the stock of per capita human capital settles at a constant level in the long run. Balanced growth in consumption, investment, output, employment, and human and physical capital is thus entirely due to population growth as in the celebrated Solow-Swan model. Fiscal incentives, though causing permanent level effects, only produce temporary growth effects. In contrast, demographic shocks change both levels and the population growth rate in the long run. In particular, the baby bust reduces long-run growth whilst increased longevity—due to reduced adult mortality—increases it. It is thus an empirical issue whether ageing countries, experiencing the combined demographic shock mentioned in the Introduction, will ultimately converge to a lower or a higher long-run rate of economic growth. Since convergence is extremely slow, time series tests for the exogenous growth model will be hard to conduct given the paucity of data.

3.6 Endogenous growth

Up to this point we have restricted attention to the case for which the intergenerational knowledge externality is relatively weak (i.e. $0 \leq \phi < 1$) and the system reaches a steady state in terms of per capita levels. In this section we study the knife-edge case for which the intergenerational knowledge transfer is very strong and subject to constant returns ($\phi = 1$). This case has been studied extensively in the literature; see among others Azariadis and Drazen (1990) and Boucekkine et al. (2002). However, as we argued in the previous section (p. 75), we do believe that the intergenerational knowledge externality (ϕ) is not unity.

3.6.1 Long-run effects

The steady-state growth path for per capita human capital can be written as follows:

$$\hat{h}(t) = \int_{-\infty}^{t-s^*} l(t-v)\hat{\hat{h}}(t-v)dv$$

$$= A_H s^* b \int_{-\infty}^{t-s^*} e^{-\hat{n}\cdot(t-v)-M(t-v,\psi_m)} \hat{h}(v) dv, \tag{3.50}$$

where we have used (3.5) and (3.17) to arrive at the second expression. In Appendix 3.C we show that there is a unique steady state growth rate of per capita variables. Moreover, after any shock, the growth rate of all per capita variables (most importantly human capital) converge to this unique value.

Denoting the steady-state growth rate by $\hat{\gamma}$, it follows that along the balanced growth path we have $\hat{h}(v) = \hat{h}(t)e^{-\hat{\gamma}\cdot(t-v)}$. By using this result in (3.50) and simplifying we obtain the implicit definition for $\hat{\gamma}$:

$$1 = A_H s^* b \int_{s^*}^{\infty} e^{-(\hat{\gamma}+\hat{n})u - M(u,\psi_m)} du. \tag{3.51}$$

Clearly, the model implies a *scale effect* in the growth process, i.e. a productivity improvement in the human capital production function gives rise to an increase in the steady-state growth rate ($\partial\hat{\gamma}/\partial A_H > 0$). Equation (3.51) can also be used to compute the effect on the asymptotic growth rate of the fiscal and demographic shocks.

Pure schooling shock Just as in Subsection 3.5.1 above, the interpretation of our results is facilitated by first considering a pure schooling shock. By differentiating (3.51) with respect to $\hat{\gamma}$ and s^*, and gathering terms we find:

$$\begin{aligned} \frac{\partial\hat{\gamma}}{\partial s^*} &= \frac{e^{-(\hat{\gamma}+\hat{n})s^* - M(s^*,\psi_m)} \left[\Delta(s^*, \hat{\gamma}+\hat{n}) - s^*\right]}{s^* \int_{s^*}^{\infty} u e^{-(\hat{\gamma}+\hat{n})u - M(u,\psi_m)} du} \\ &= \frac{e^{-(\hat{\gamma}+\hat{n})s^* - M(s^*,\psi_m)}}{s^* \int_{s^*}^{\infty} u e^{-(\hat{\gamma}+\hat{n})u - M(u,\psi_m)} du} \left[\Delta(s^*, \hat{\gamma}+\hat{n}) - \Delta(s^*, r) - \frac{\rho}{1-t_L}\right] > 0, \end{aligned} \tag{3.52}$$

where we have used quation (3.11) to arrive at the final expression. The sign of $\partial\hat{\gamma}/\partial s^*$ is determined by the term in square brackets on the right-hand side of (3.52). By appealing to the endogenous- growth counterpart of Assumption 3.2 (with $\hat{n}$ replaced by $\hat{n}+\hat{\gamma}$) we find that the steady-state growth rate increases as a result of the pure schooling shock.

Fiscal shock An increase in the educational subsidy or the labour income tax affects the steady-state growth rate via its positive effect on the schooling period.

Indeed, we deduce from (3.31)–(3.32) and (3.52) that:

$$\frac{\partial \hat{\gamma}}{\partial \left[\rho/(1-t_L)\right]} = \frac{\partial \hat{\gamma}}{\partial s^*} \frac{\partial s^*}{\partial \left[\rho/(1-t_L)\right]} > 0. \tag{3.53}$$

Birth rate shock The growth effects of a birth rate change are computed most readily by restating the shock in terms of the steady-state population growth rate, $\hat{n}$, and noting the monotonic relationship between $\hat{n}$ and b stated in (3.36) above. Indeed, by substituting the steady-state version of (3.18) into (3.51) we find an alternative implicit expression for $\hat{\gamma}$:

$$\int_0^\infty e^{-\hat{n}u-M(u,\psi_m)}du = A_H s^* \int_{s^*}^\infty e^{-(\hat{\gamma}+\hat{n})u-M(u,\psi_m)}du. \tag{3.54}$$

Since the birth rate shock leaves the schooling period unchanged, it follows from (3.54) that:

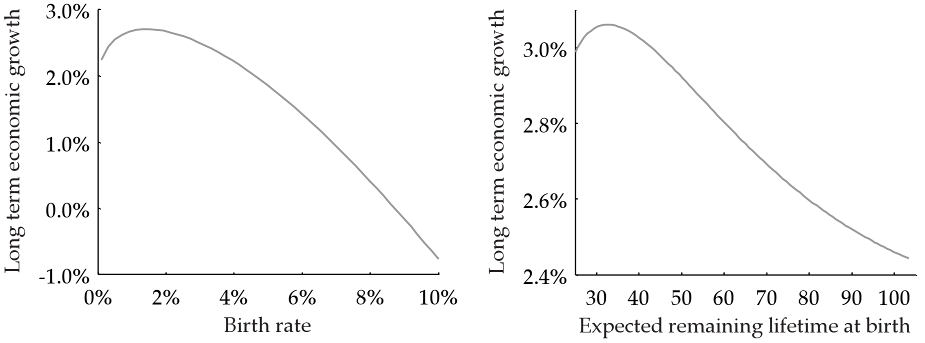
$$\frac{\partial \hat{\gamma}}{\partial b} = \frac{\partial \hat{\gamma}}{\partial \hat{n}} \frac{\partial \hat{n}}{\partial b} = \frac{\partial \hat{n}}{\partial b} \left[\frac{\int_0^\infty u e^{-\hat{n}u-M(u,\psi_m)}du}{A_H s^* \int_{s^*}^\infty u e^{-(\hat{\gamma}+\hat{n})u-M(u,\psi_m)}du} - 1 \right] \gtreqless 0. \tag{3.55}$$

Despite the fact that $\partial \hat{n}/\partial b > 0$, the growth effect of a birth rate change is ambiguously because the term in square brackets on the right-hand side of (3.55) cannot be signed a priori. Indeed, using the calibrated version of the model, we find that the relationship between $\hat{\gamma}$ and b is hump-shaped. As is illustrated in Figure 3.7(a), the growth rate rises with the birth rate for low birth rates, but is decreasing for higher birth rates. For the calibrated model, the maximum growth rate is attained at a birth rate of 1.25% per annum.

Mortality shock Just as in the exogenous growth model, increased longevity constitutes by far the most complicated shock studied here. Indeed, as can be seen from Equation (3.51) above, a mortality shock affects three distinct items featuring in the implicit expression for the steady-state growth rate, $\hat{\gamma}$, namely (a) the optimal schooling period, s^*, (b) the steady-state growth rate of the population, $\hat{n}$, and (c) the cumulative mortality factor, $M(u,\psi_m)$. By differentiating (3.51) with respect to $\hat{\gamma}$ and ψ_m (and recognising the dependence of s^* and $\hat{n}$ on ψ_m) we find after some steps:

$$\frac{\partial \hat{\gamma}}{\partial \psi_m} = \frac{\partial \hat{\gamma}}{\partial s^*} \frac{\partial s^*}{\partial \psi_m} - \frac{\partial \hat{n}}{\partial \psi_m} + \frac{\int_{s^*}^\infty -\frac{\partial M(u,\psi_m)}{\partial \psi_m} e^{-(\hat{\gamma}+\hat{n})u-M(u,\psi_m)}du}{\int_{s^*}^\infty u e^{-(\hat{\gamma}+\hat{n})u-M(u,\psi_m)}du} \gtreqless 0. \tag{3.56}$$

Figure 3.7. The effect of the birth rate (a) and mortality (b) on long-term growth



(a) Growth and the birth rate (b) Growth and life expectancy at birth

The overall growth effect of increased longevity is ambiguous. The first composite term on the right-hand side of (3.56) represents the schooling effect, which is positive (see (3.35) and (3.52)). The third term on the right-hand side represents the cumulative mortality effect and is also positive (given Proposition 3.1(i)). The ambiguity thus arises because the second term on the right-hand side exerts a negative influence on growth, i.e. increased longevity boosts the steady-state population growth rate (see (3.37) above) which in turn slows down growth.

In Figure 3.7(b) we use the calibrated version of the model to plot the relationship between the steady-state growth rate and a measure of longevity, namely life expectancy at birth, $R(0, \psi_m) \equiv \Delta(0, 0, \psi_m)$. Except for very low values of $R(0, \psi_m)$, there is negative relationship between long-term growth and longevity.

3.6.2 Transitional dynamics

In this subsection we visualise the transitional effects of fiscal and demographic shocks in the endogenous growth model. We restrict attention to the growth rate of per capita human wealth, $\gamma(t) \equiv \dot{h}(t)/h(t)$, since this variable drives all other macroeconomic variables. Except for ϕ and A_H, we use the same calibration values as before (see Subsection 3.5.2). Because the model contains a scale effect, we set $A_H = 0.13$ and obtain a realistic steady-state growth rate, $\hat{\gamma}_0 = 1.096\%$. The discussion here can be quite brief because, following a shock, the transition proceeds along the same phases as in the exogenous growth model.

Fiscal shock Figure 3.8(a) illustrates the path for $\gamma(t)$ following a 20% increase in the educational subsidy. For $0 \leq t < s_1^* - s_0^*$ there are no new labour market entrants and the growth rate collapses. Then, for $s_1^* - s_0^* \leq t < s_1^*$ pre-shock students enter the labour market and the growth rate jumps above its initial steady-state level. Finally, for $t \geq s_1^*$ the growth rate converges in a non-monotonic fashion to its long-run value, i.e. $\lim_{t\to\infty} \gamma(t) = \hat{\gamma}_1 = 1.13\%$, where $\hat{\gamma}_1$ exceeds the initial steady-state growth rate $\hat{\gamma}_0$ (see quation (3.53) above).

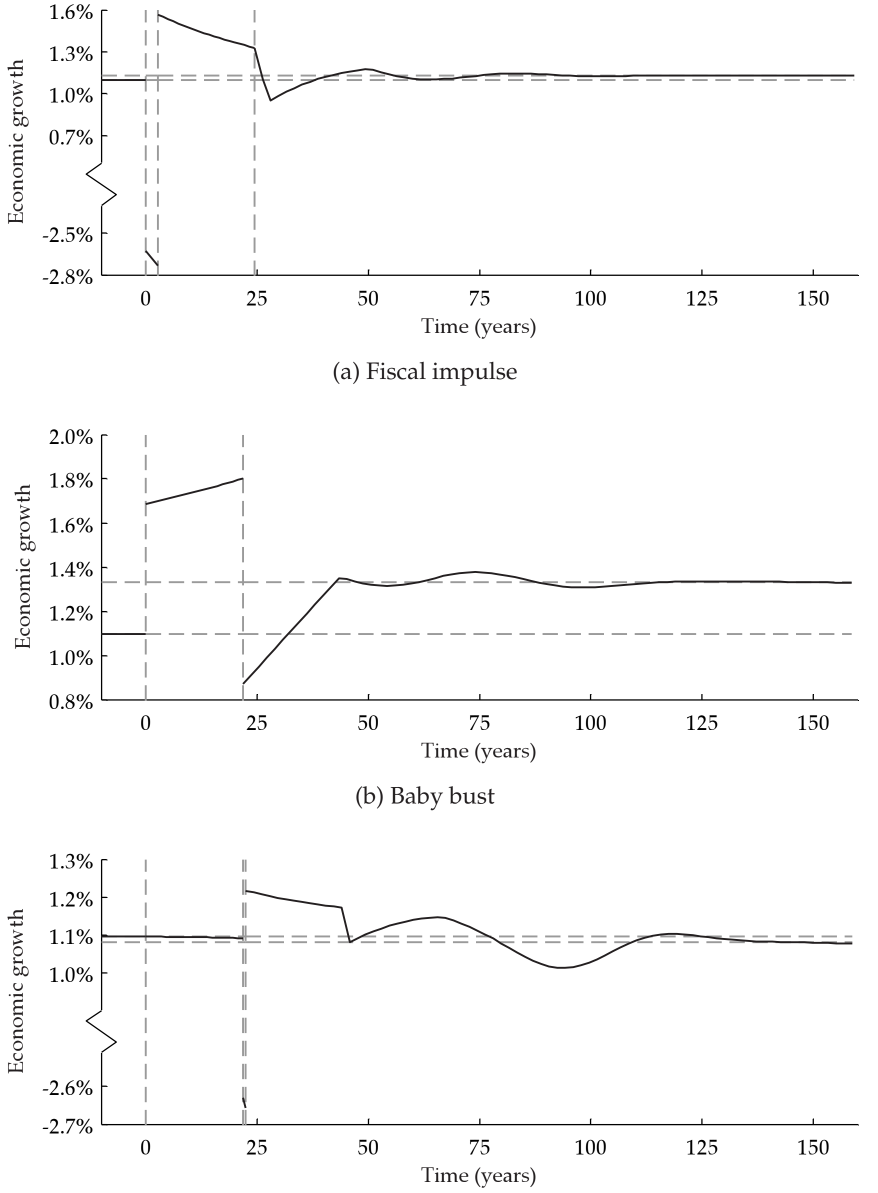
Birth rate shock In Figure 3.8(b) the transitional effects of a baby bust are illustrated. There is no effect on the optimal schooling period but the population growth rate falls from $\hat{n}_0$ to $\hat{n}_1$—see Figure 3.3(a). Growth jumps sharply due to the fast reduction in $n(t)$ that occurs at impact and immediately hereafter. Intuitively, pre-shock students enter the labour market but their human capital is spread out over fewer people than before the shock so that growth in per capita terms increases sharply. About twenty-two years after the shock, $n(t) \approx \hat{n}_1$ and there is a sharp decline in growth. This is because the post-shock students start to enter the labour market. Despite the fact that they have higher human capital than existing workers, as a group they are not large enough to maintain the previous growth in per capita human capital. Thereafter, the growth rate converges in a non-monotonic fashion to its long-run level $\hat{\gamma}_1 = 1.33\%$, which is higher than the initial steady-state growth rate, i.e. $\hat{\gamma}_1 > \hat{\gamma}_0$. Given our calibration, the economy lies to the right of the peak in the curve for $\hat{\gamma}$ in Figure 3.7(a) so that a baby bust increases long-run growth.

Mortality shock In Figure 3.8(c) the effect on the growth rate of increased longevity of generations born after time $t = 0$ is illustrated. Just as for the exogenous growth model, nothing happens to growth for the period $0 \leq t < s_0^*$ because only pre-shock agents enter the labour market and the same type of agents die off. For $s_0^* \leq t < s_1^*$ there are no new labour market entrants and the growth rate collapses. At time $t = s_1^*$ the oldest of the post- shock cohorts enter the labour market and as a result growth is boosted again. For $t > s_1^*$, the growth rate converges non-monotonically towards the new steady-state growth rate $\hat{\gamma}_1 = 1.09\% < \hat{\gamma}_0$. In terms of Figure 3.7(b), the calibration places the economy on the downward sloping segment of the $\hat{\gamma}$ curve so increased longevity reduces the long-run growth rate.

3.6.3 Discussion

The main findings of this section are as follows. For the calibrated model, the long-run growth rate in per capita human capital increases as a result of a positive fiscal

Figure 3.8. Per capita human capital growth



(a) Fiscal impulse

(b) Baby bust

(c) Reduced adult mortality

Notes: Mortality process is a Gompertz-Makeham (see Table 2.1 for parameter values), birth rate is 2.36%. Fiscal education shock is a 50% increase in the educational subsidy, from $\rho_0 = 4.915$ to $\rho_1 = 7.372$. Baby bust is a 25% downward jump of the birth rate to 1.78%. Reduced adult mortality is a 50% decrease of μ_1 and 10% decrease of μ_2. Results are absolute growth rates.

impulse or a fall in the birth rate. Increased longevity, however, reduces this long-run growth rate. The transition path in the growth rate is cyclical and rather complex for all shocks considered, and the new equilibrium is reached only very slowly.

3.7 Conclusion

In this chapter we extended the basic model of Chapter 2 with a schooling decision and we have studied how fiscal incentives and demographic shocks affect the growth performance of a small open economy populated by disconnected generations of finitely-lived agents facing age-dependent mortality and constant factor prices. Our analysis shows that only for a unrealistically strong intergenerational knowledge spillover, policy changes and demographic shocks lead to a permanent higher (or lower) growth rate. Moreover, if the intergenerational spillover is unrealistically large, the link between longevity and economic growth is non-monotonic. For realistic parameter values a higher life expectancy at birth causes a lower long-run growth rate in most developed countries.

This chapter highlights the crucial role played by the strength of the intergenerational external effect in the production of human capital. Also, the vintage nature of the model gives rise to very slow and rather complicated dynamic adjustment. This feature of the model may help explain why robust empirical results linking education and growth have been so hard to come by.

This chapter focused on the decision an individual faces at the beginning his economic life. In the next chapter we focus on the economic decision at the end of an individual's life, the retirement decision. Many OECD countries have experienced an increase in the old-age dependency ratio over the last half century. This has important implications for the feasibility of existing pay-as-you-go pension schemes, a phenomenon that has been abstracted from in this chapter. In Chapter 4 we drop the schooling decision, but we endogenise the agent's labour force participation decision in the presence of a stylized public pension system including realistic institutional features such as the early retirement age and the mandatory retirement age (Gruber and Wise, 1999). With this extended model we hope to contribute to the literature on pension reform in an ageing society.

3.A Optimal schooling period

Consider the household that is still in school. By differentiating (3.9) with respect to $s(v)$ we get:

$$\begin{aligned}\frac{d\bar{l}i(v,t)}{ds(v)} \equiv A_H h(v)^{\phi} e^{M(t-v)} \Bigg[& \rho(v+s(v))w(v+s(v))e^{-r(v+s(v)-t)-M(s(v))} \\ & + \int_{v+s(v)}^{\infty} [1-t_L(\tau)]w(\tau)e^{-r\cdot(\tau-t)-M(\tau-v)}d\tau \\ & - [1-t_L(v+s(v))]s(v)w(v+s(v))e^{-r\cdot(v+s(v)-t)-M(s(v))}\Bigg].\end{aligned}$$

By simplifying and setting $\frac{d\bar{l}i(v,t)}{ds(v)} = 0$ we obtain (for time-invariant ρ and t_L):

$$\begin{aligned}(1-t_L)\int_{v+s(v)}^{\infty} & w(\tau)e^{-r\cdot(\tau-v)-M(\tau-v)}d\tau \\ & = [(1-t_L)s^*(v)-\rho]w(v+s^*(v))e^{-rs^*(v)-M(s(v))}.\end{aligned}$$

For the time-invariant wage rate we obtain:

$$\begin{aligned}s^*(v) - \frac{\rho}{1-t_L} &= e^{rs^*(v)+M(s^*(v))}\int_{v+s^*(v)}^{\infty} e^{-r(\tau-v)-M(\tau-v)}d\tau \\ &\equiv \Delta(s^*(v),r).\end{aligned}$$

3.B Useful Lemmas

Birth rate shock To determine effect of a birth rate shock on the level of human capital in Equation (3.45) we need the following lemma

Lemma 3.1. *By using* (3.36) *in* (3.45) *we obtain:*

$$\frac{\partial \hat{h}^{1-\phi}}{\partial b} = \frac{A_H s^*}{b}\psi_m(s^*),$$

where $\psi_m(s)$ is defined as:

$$\psi_m(s) \equiv \frac{\int_s^{\infty} e^{-\hat{n}u-M(u,\psi_m)}du}{\int_0^{\infty} e^{-\hat{n}u-M(u,\psi_m)}du} - \frac{\int_s^{\infty} ue^{-\hat{n}u-M(u,\psi_m)}du}{\int_0^{\infty} ue^{-\hat{n}u-M(u,\psi_m)}du},$$

with $\hat{n} > 0$ and $M(u,\psi_m)$ as defined in equation (2′). The following results can be estab-

lished: (i) $\psi_m(s) \leq 0$ *for all* $s \geq 0$*, (ii)* $\psi_m(0) = 0$*, (iii)* $\lim_{s\to\infty} \psi_m(s) = 0$.

Proof. Results (ii) and (iii) follow directly from the definition of $\psi_m(s)$. Differentiation with respect to s gives

$$\frac{\partial \psi_m}{\partial s} = e^{-\hat{n}s-M(s,\psi_m)} \left[\frac{s}{\int_0^\infty e^{-\hat{n}u-M(u,\psi_m)}du} - \frac{1}{\int_0^\infty ue^{-\hat{n}u-M(u,\psi_m)}du} \right],$$

which is continuous in s and has only one root. The second derivative is positive in this unique stationary point, so it is a global minimum. Together with (ii) and (iii) this implies result (i). ■

Mortality shock To determine the effect of a mortality shock on the level of human capital in the long run (Equation (3.46), we need the following lemma

Lemma 3.2. *Define* $\Xi(s,\psi_m)$ *for* $s \geq 0$ *as:*

$$\Xi(s,\psi_m) = \int_s^\infty e^{-\hat{n}u-M(u,\psi_m)}du.$$

Then $\dfrac{\partial \Xi(s,\psi_m)}{\partial \psi_m} \geq 0$ *for all* $s > 0$*, where the equality holds if and only if* $\dfrac{\partial^2 m(u,\psi_m)}{\partial u \partial \psi_m} = 0$.

Proof. For the sake of readability define

$$\begin{aligned}\Xi_{\psi_m}(s,\psi_m) &\equiv \frac{\partial \Xi(s,\psi_m)}{\partial \psi_m} \\ &= \int_s^\infty \frac{\partial M(u,\psi_m)}{\partial \psi_m} e^{-\hat{n}u-M(u,\psi_m)}du - \frac{\partial \hat{n}}{\partial \psi_m} \int_s^\infty ue^{-\hat{n}u-M(u,\psi_m)}du,\end{aligned} \tag{3.B.1}$$

Note that $\lim_{s\to\infty} \Xi_{\psi_m}(s,\psi_m) = 0$ and:

$$\frac{\partial \hat{n}}{\partial \psi_m} = \frac{\int_0^\infty \frac{\partial M(u,\psi_m)}{\partial \psi_m} e^{-\hat{n}u-M(u,\psi_m)}du}{\int_0^\infty ue^{-\hat{n}u-M(u,\psi_m)}du}. \tag{3.B.2}$$

By substituting (3.B.2)) into (3.B.1)) we find that $\Xi_{\psi_m}(0,\psi_m) = 0$. The stationary points of $\Xi_{\psi_m}(s,\psi_m)$ with respect to s are determined by the roots of:

$$\frac{\partial \Xi_{\psi_m}(s,\psi_m)}{\partial s} = e^{-\hat{n}u-M(u,\psi_m)} \left[\frac{\partial M(s,\psi_m)}{\partial \psi_m} - s\frac{\partial \hat{n}}{\partial \psi_m} \right]. \tag{3.B.3}$$

From Proposition 3.1 we know that $\frac{\partial M(s,\psi_m)}{\partial \psi_m}$ is non-positive, non-increasing and concave in s. This implies together with $\frac{\partial \Xi_{\psi_m}(0,\psi_m)}{\partial s} = 0$ that (3.B.3) has at most two roots (one at $s = 0$) or is 0 everywhere (if $\frac{\partial \Xi_{\psi_m}(0,\psi_m)}{\partial s} = 0$ on the interval $[0, s^*]$, $0 \leq s^* \ll \infty$, then $\lim_{s\to\infty} \Xi_{\psi_m}(s, \psi_m) = 0$ does not hold). If $\frac{\partial \Xi_{\psi_m}(s,\psi_m)}{\partial s} = 0$ for all $s \geq 0$, then $\Xi_{\psi_m}(s, \psi_m) = 0$ for all $s \geq 0$. This last situation only occurs if $\frac{\partial M(s,\psi_m)}{\partial \psi_m}$ is linear in s, i.e. if $\frac{\partial^2 m(u,\psi_m)}{\partial u \partial \psi_m} = 0$.

If $\frac{\partial^2 m(u,\psi_m)}{\partial u \partial \psi_m} < 0$ for some $s \geq 0$, then $\Xi_{\psi_m}(s, \psi_m)$ has exactly two stationary points for a given ψ_m, one at $s = 0$ and one at $s = s^* > 0$. Concavity of $\frac{\partial M(s,\psi_m)}{\partial \psi_m}$ implies that the stationary point at $s = s^*$ is a maximum. Since $\frac{\partial \Xi(s,\psi_m)}{\partial \psi_m}$ goes to 0 as $s \to \infty$ and is continuous, $\frac{\partial \Xi(s,\psi_m)}{\partial \psi_m}$ must be positive for all $s > 0$, otherwise there would be a minimum somewhere at $s > s^*$. This completes the proof. ■

3.C Convergence of the endogenous growth model

We have already demonstrated that, following a demographic shock, the population growth rate converges to a constant value, $\hat{n}$. In this section we assume for simplicity that $n(t) = \hat{n}$ and consider the stability of the growth rate in per capita human capital for the case with $\phi = 1$, i.e. we prove that $\gamma(t) \equiv \dot{h}(t)/h(t)$ converges to $\hat{\gamma}$ as t gets large. Taking the past as given and focusing on $t > s^*$, we can rewrite the first expression in Equation (3.50) in the form of a normal integral equation. Define:

$$\mathbf{K}(u) = \begin{cases} 0 & \text{for } 0 \leq u \leq s^* \\ A_H s^* l(u) & \text{for } u > s^* \end{cases} \quad \text{and} \quad \chi(t) = \int_{-\infty}^{0} \mathbf{K}(t-v)h(v)dv \quad \text{for } t > s.$$

With these definitions, we can write $h(t)$ for $t > s$ in the form of a renewal equation:

$$h(t) = \chi(t) + \int_0^t \mathbf{K}(t-v)h(v)dv,$$

where the exogenous function $\chi(t)$ is called the forcing equation, and $\mathbf{K}(t-v)$ is the kernel of the integral operator. By definition, human capital in the past ($h(t)$, for $t < 0$) is bounded and continuous. This makes the forcing equation continuous and monotonically decreasing.

We want to show that no matter what the path of human capital was before $t = 0$, human capital growth always converges to a constant. To show this, we closely follow Bellman and Cooke (1963, ch. 7). The integral is the convolution of

$h(t)$ and $\mathbf{K}(t)$, so Laplace techniques are a logical choice to analyse the behaviour of $h(t)$. Taking the Laplace transform of $h(t)$ and using the convolution theorem we obtain

$$\mathcal{L}(h) = \mathcal{L}(\chi) + \mathcal{L}(\mathbf{K})\,\mathcal{L}(h) \quad \Rightarrow \quad \mathcal{L}(h) = \frac{\mathcal{L}(\chi)}{1 - \mathcal{L}(\mathbf{K})}$$

Using the complex inversion formula we find that the solution of $h(t)$ is given by the contour integral

$$h(t) = \int_{(\mathbf{b})} \frac{\mathcal{L}(\chi)(\varsigma)}{1 - \mathcal{L}(\mathbf{K})(\varsigma)} e^{\varsigma t} d\varsigma$$

with $\int_{(\mathbf{b})}$ as in Bellman and Cooke (1963, p. 233):

$$\int_{(\mathbf{b})} F(s)ds = \lim_{T\to\infty} \frac{1}{2\pi i} \int_{\mathbf{b}-iT}^{\mathbf{b}+iT} F(s)ds = \lim_{T\to\infty} \frac{1}{2\pi} \int_{-T}^{T} F(\mathbf{b}+it)dt,$$

and $\mathbf{b}$ such that all the singularities of the function under the integral sign lie to the left of the line $\mathrm{Re}(z) < \mathbf{b}$ in the complex plane. That is, $h(t)$ is the contour integral taken over the vertical line in the complex plane defined by $\mathrm{Re}(z) = \mathbf{b}$.

Note that all singularities of the function under the integral sign are determined by the roots of $1 - \mathcal{L}(\mathbf{K})$. As we will see later, it is sufficient to take any $\mathbf{b}$ such that $\mathbf{b} > \hat{\gamma}$, where $\hat{\gamma}$ is implicitly defined as the real solution of:

$$1 = \mathcal{L}(\mathbf{K})(\hat{\gamma}) = A_H s^* \int_{s^*}^{\infty} l(u) e^{-\hat{\gamma}u} du = A_H s^* b \int_{s^*}^{\infty} e^{-[\hat{\gamma}+\hat{n}]u - M(u)} du$$

Following Bellman and Cooke (1963, par. 7.11), it is quite simple to show that $1 - \mathcal{L}(\mathbf{K})$ has only one real root. Denote this real root by $\hat{\gamma}$. Existence and uniqueness of $\hat{\gamma}$ follows directly from continuity and monotonicity of $\mathcal{L}(\mathbf{K})(\varsigma)$ and:

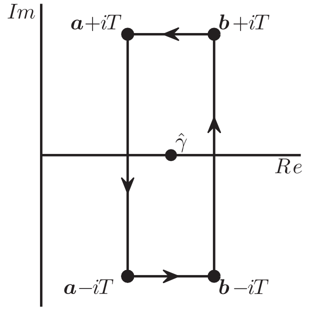
$$\lim_{\varsigma\to-\infty} \mathcal{L}(\mathbf{K})(\varsigma) = -\infty \quad \text{and} \quad \lim_{\varsigma\to\infty} \mathcal{L}(\mathbf{K})(\varsigma) = 0.$$

Note that $\hat{\gamma}$ is not necessarily positive. To prove that $\hat{\gamma}$ is the root with the largest positive real part, suppose there is an other (complex) root $\varsigma = x + i\tau$.

$$1 = \left| \int_{s}^{\infty} e^{-xu} e^{-i\tau u} l(u) du \right| < \int_{s}^{\infty} e^{-xu} l(u) du$$

The only way the term on the right can be larger than one is if $x < \hat{\gamma}$ which means $\mathrm{Re}(\varsigma) < \hat{\gamma}$.

Figure 3.C.1. Contour shifting



To analyse the behaviour of $h(t)$ as $t \to \infty$ we shift the contour $(\mathbf{b} - iT, \mathbf{b} + iT)$ in

$$h(t) = \lim_{T\to\infty} \frac{1}{2\pi i} \int_{\mathbf{b}-iT}^{\mathbf{b}+iT} \frac{\mathcal{L}(\chi)(\varsigma)}{1-\mathcal{L}(\mathbf{K})(\varsigma)} e^{\varsigma t} d\varsigma$$

to the left such that we pick up the pole at $\hat{\gamma}$ and we can (hopefully) write $h(t)$ as

$$h(t) = \beta e^{\hat{\gamma} t} + \int_{(\mathbf{a})} \frac{\mathcal{L}(\chi)(\varsigma)}{1-\mathcal{L}(\mathbf{K})(\varsigma)} e^{\varsigma t} d\varsigma.$$

That is, we write $h(t)$ as the sum of the residue at $\hat{\gamma}$ and the contour integral taken over the line $\mathrm{Re}(z) = \mathbf{a}$ with $\mathbf{a}$ smaller than $\hat{\gamma}$, but larger than the other poles of $1 - \mathcal{L}(\mathbf{K})$. If for $t \to \infty$ the exponential term dominates the integral, then human capital growth converges to $\hat{\gamma}$.

In Figure 3.C.1, the contour is shifted from the line $(\mathbf{b} - iT, \mathbf{b} + iT)$ to $(\mathbf{a} - iT, \mathbf{a} + iT)$. The residue theorem tells us that the contour integral taken over the square equals the sum of the residues of the singular points within this square. Since $\mathbf{b}$ and $\mathbf{a}$ are chosen in such a way that $\hat{\gamma}$ is the only singular point in this region, the contour integral over the square equals the residue at $\hat{\gamma}$ which is:

$$\lim_{\varsigma\to\hat{\gamma}} \frac{\varsigma - \hat{\gamma}}{1-\mathcal{L}(\mathbf{K})(\varsigma)} \mathcal{L}(\chi)(\varsigma) e^{\varsigma t} = \frac{\int_0^\infty e^{-\hat{\gamma} t}\chi(t)dt}{\int_0^\infty t e^{-\hat{\gamma} t}\mathbf{K}(t)dt} e^{\hat{\gamma} t}$$

This residue is a function of t and grows at rate $\hat{\gamma}$.

Next we will show that the contribution of the two horizontal contours vanishes as $t \to \infty$. To show that

$$\lim_{T\to\infty} \frac{1}{2\pi i} \int_{\mathbf{b}\pm iT}^{\mathbf{a}\pm iT} \frac{\mathcal{L}(\chi)(\varsigma)}{1-\mathcal{L}(\mathbf{K})(\varsigma)} e^{\varsigma t} d\varsigma = 0, \tag{3.C.1}$$

it is sufficient to show that for $\mathbf{a} \le x \le \mathbf{b}$

$$\lim_{T\to\infty} \mathcal{L}(\chi)(x \pm iT) = \lim_{T\to\infty} \mathcal{L}(\mathbf{K})(x \pm iT) = 0.$$

Integration by parts of $\mathcal{L}(\chi)$ on the contour $\varsigma = x \pm iT$ gives

$$\mathcal{L}(\chi) = \frac{1}{T}\left[e^{-xt} f(t) \frac{e^{\pm iT}}{i}\bigg|_{t=0}^{\infty} \pm \frac{x}{i}\int_0^\infty e^{-xt}\chi(t)e^{\pm iT}dt \pm \frac{1}{i}\int_0^\infty e^{-xt}\chi'(t)e^{\pm iT}dt \right].$$

Using the definition of $\chi(t)$ it is easy to see that on $\mathbf{a} \le x \le \mathbf{b}$

$$\lim_{t\to\infty} e^{-xt}\chi(t) = 0 \quad \text{and} \quad \int_0^\infty e^{-xt}\chi'(t)dt \ll \infty.$$

This means that the whole term within square brackets is a constant and we can write

$$|\,\mathcal{L}(\chi)| = O(1/T).$$

A similar result holds for $\mathcal{L}(\mathbf{K})$. This means that numerator in the integrand in equation (3.C.1) goes to zero as T goes to infinity and the denominator goes to 1 and the integral vanishes.

The final part we have to show is

$$\lim_{T\to\infty} \frac{1}{2\pi i} \int_{\mathbf{a}-iT}^{\mathbf{a}+iT} \frac{\mathcal{L}(\chi)(\varsigma)}{1-\mathcal{L}(\mathbf{K})(\varsigma)} e^{\varsigma t} d\varsigma = O(e^{\mathbf{a}t}),$$

which follows directly from the the fact that all singularities in $\mathcal{L}(\chi)$ are cancelled by singularities in $\mathcal{L}(\mathbf{K})$ so the fraction $\mathcal{L}(\chi)/\,\mathcal{L}(\mathbf{K})$ remains bounded.

Finally we can write $h(t) = \beta e^{\hat{\gamma} t} + O(e^{\mathbf{a}t})$ as $t \to \infty$, with

$$\beta = \frac{\int_0^\infty e^{-\hat{\gamma}t}\chi(t)dt}{\int_0^\infty te^{-\hat{\gamma}t}\mathbf{K}(t)dt} \quad \text{and } \mathbf{a} < \hat{\gamma}$$

This implies that no matter what the initial path of human capital was in the past, the growth of human capital will always converge towards $\hat{\gamma}$.

Chapter 4

Ageing, Pensions, and Retirement

4.1 Introduction

Population ageing is playing havoc with the public pension schemes of many western countries. In a celebrated sequence of international comparative studies, Gruber and Wise (1999, 2004, 2005) and their collaborators have established a number of stylized facts pertaining to a subset of OECD countries. These facts are:

(SF1) For most developed countries, the pay-as-you-go social security system includes promises that cannot be kept without significant system reforms. In the absence of reform, the current systems are *fiscally unsustainable*.

(SF2) Over the last four decades, the trend is for older people to leave the labour force at ever younger ages. For most countries there is a clear downward trend in the labour force participation of pension-age males. Retirement is a *normal good* in the sense that the demand for years of retirement rises as agents' income rises (Barr and Diamond, 2006, p. 27).

(SF3) Only a very small fraction of the labour force retires before the earliest age at which public retirement benefits are available, the so-called *early eligibility age* (EEA hereafter). For males, the EEA in 2003 typically falls in the range of

This chapter is based on joint work with Ben Heijdra, 'Retirement, Pensions, and Ageing', mimeo

60-62 years of age. Similarly, only very few people work until the *normal retirement age* (NRA hereafter), which is typically 65 for most countries (Duval, 2003, p. 35) Together this implies that most people retire either at the EEA or somewhere in between the EEA and the NRA.

(SF4) Most social security programs contain strong incentives for older workers to leave the labour force. In most countries it simply does not pay to work beyond the EEA because *actuarial adjustments* are less than fair. As Gruber and Wise put it, 'once benefits are available, a person who continues to work for an additional year will typically receive less in social security benefits over his lifetime than if he quit work and started to receive benefits at the first opportunity' (2005, p. 5). The present value of expected social security benefits declines with the retirement age, and there is a high *implicit tax* on working beyond the EEA.

(SF5) In many European countries disability programs and age-related unemployment provisions essentially provide early retirement benefits, even before the EEA.

A formal analysis of issues surrounding ageing, retirement, and pensions should accommodate at least some, but preferably all, of these stylized facts. In this chapter we study the consumption, saving, and retirement decisions of individual agents facing lifetime uncertainty, or longevity risk. In addition, we also determine the macroeconomic consequences of individual behaviour and policy changes. We use the extended-Blanchard-Yaari framework developed in Chapter 2. We maintain the assumption that the country is small in world capital markets and thus faces an exogenous world interest rate which we take to be constant. To analyse the effects of ageing societies, we use the generalised demographic framework of Chapter 3. By allowing the mortality rate to depend on age and time of birth, the model can be used to investigate the micro- and macroeconomic effects of a reduction in *adult mortality*. We still assume that finitely lived agents fully insure against the adverse affects of lifetime uncertainty by purchasing actuarially fair annuities.

The second building block of our analysis concerns the labour market participation decision of individual agents. Following the seminal contribution by Sheshinski (1978) and much of the subsequent literature, we assume that labour is indivisible (the agent either works full time or not at all), that the retirement decision is irreversible, and that the felicity function is additively separable in consumption and leisure. All agents are blessed with perfect foresight and maximize an intertemporal utility function subject to a lifetime budget constraint. Workers choose the optimal retirement age, taking as given the time- and age profiles of wages, the fiscal system, and the public pension system. Not surprisingly, like Mitchell and Fields and many others we find that 'the optimal retirement age ... equates the marginal utility of income from an additional year of work with the marginal utility of one more year of leisure' (1984, p. 87).

The two papers most closely related to the analysis in this chapter are Sheshinski (1978) and Boucekkine et al. (2002).[1] We extend the analysis of Sheshinski (1978) in two directions. First, as was already mentioned above, we incorporate a realistically modelled lifetime uncertainty process, rather than a fixed planning horizon. Second, we embed Sheshinski's *microeconomic* model in the context of a small open economy and are thus able to study the *macroeconomic* repercussions of ageing and pension reform. We generalize the analysis of Boucekkine et al. (2002) by including a concave, rather than linear, felicity function, and by modelling a public pension system with realistic features such as an EEA which differs from the NRA and non-zero implicit tax rates. Furthermore, we conduct our analysis with a general description of the demographic process, whereas they use a specific functional form for this process throughout their paper.

The remainder of this chapter is organized as follows. In Section 4.2 we present the model and demonstrate its main properties. Consumption is proportional to total wealth, consisting of financial and human wealth. With a realistic demography, the marginal propensity to consume out of wealth is increasing in the agent's age because the planning horizon shortens as one grows older and the agent does not wish to leave any bequests. We derive the first-order condition for the optimal retirement age and show that it depends not only on the mortality process but also on the features of the fiscal and pension systems. The mortality process, in combination with the birth rate, also determines a unique path for the population growth rate.

In Section 4.3 we abstract from the public pension system and study the comparative static effects on the optimal retirement age of various age related shocks. A reduction in the disutility of working leads to an increase in the optimal retirement age. In contrast, an upward shift in the age profile of wages causes a negative wealth effect but a positive substitution effect, rendering the total effect on the optimal retirement age ambiguous. A reduction in adult mortality increases the

[1] In the interest of brevity, we refer the interested reader to the literature surveys on retirement and ageing by Lazear (1999); Hurd (1990, 1997) and D. N. Weil (1997). For a recent literature survey on pension reform, see Lindbeck and Persson (2003).

expected remaining lifetime for everyone, though more so for older agents. The effect of increased longevity on the optimal retirement age is ambiguous in general because the lifetime-income effect cannot be signed a priori. For realistic scenarios, however, the increased longevity only starts to matter quantitatively at ages exceeding the NRA so that the lifetime-income effect works in the direction of increasing the optimal retirement age.

Section 4.3 also presents the graphical apparatus that we use throughout the chapter. We demonstrate that the optimal retirement decision is best studied in terms of its consequences for lifetime income and the *transformed retirement age*. The transformed retirement age is a monotonically increasing transformation of the calender age of retirement and captures the notion of an agent's economic (rather than biological) age. Our graphical apparatus has the attractive feature that indifference curves are convex and that, with an age invariant wage rate, actuarially fair adjustment leads to a linear budget constraint. We believe that our graphical representation is more intuitive than the conventional one based on biological years.

In Section 4.4 we re-introduce the public pension system (including disability programs and age-related unemployment provisions, see SF5) and determine its likely consequences for the retirement decision of individual agents. Using data from Gruber and Wise (1999) for nine OECD countries, we compute conservative estimates for standardized lifetime income profiles and find that these profiles are concave in the transformed age domain. For at least six of these countries, the lifetime income profile features a kink at the EEA as a result of non-trivial implicit tax rates. Combined with convex indifference curves, it is not surprising that many agents choose to retire at the EEA, conform stylized facts (SF3) and (SF4).

In Section 4.5 we take the concavity of lifetime income profiles for granted and discuss the comparative static effects on the optimal steady-state retirement age of various changes in the tax system or the public pension system. We restrict attention to interior solutions because an optimum occurring at the kink in the lifetime income profile is insensitive to infinitesimal changes. An increase in the lumpsum[2] tax leads to a reduction in lifetime income and an increase in the optimal retirement age. Retirement is thus a normal good in our model, conform stylized fact (SF2). Not surprisingly, an increase in the labour income tax has an ambiguous effect on the retirement age because the substitution effect is negative and the wealth effect is positive. Holding constant the slope of the pension benefit curve, an increase in its level unambiguously leads to a decrease in the retirement age—the wealth effect

[2] In the public economics literature, a lumpsum tax is usually named 'poll' tax. Since we used the term 'lumpsum' in the previous chapters, we will continue using it here.

and the substitution effect operate in the same direction. In contrast, an increase in the slope of the benefit curve, holding constant its level, leads to an increase in the optimal retirement age as a result of the positive substitution effect.

In Section 4.6 we compute and visualize the general equilibrium effects of various large demographic shocks and several assumed policy reform measures. Conform stylized fact (SF1), we postulate that in the initial steady state individuals are stuck at the early retirement kink. Because both the shocks and the policy reform measures are infra marginal, we simulate a plausibly calibrated version of our model to compute the impact-, transitional-, and long-run effects on the macroeconomy.

Finally, in Section 4.7 we present some concluding thoughts and give some suggestions for future research. The chapter also contains a brief Appendix containing some additional material on the retirement age transformation as well as data on replacement rates and implicit taxes for nine OECD countries.

4.2 The model

4.2.1 Households

An individual values both consumption and leisure. The (remaining) lifetime utility function at time t for an agent born at time $v(v \leq t)$ is written as:

$$\Lambda(v,t) \equiv e^{M(u)} \int_t^{\infty} \left[U(\bar{c}(v,\tau)) - \mathbf{I}(\tau - v, R(v))D(\tau - v)\right] e^{-\theta \cdot (\tau - t) - M(\tau - v)} d\tau, \quad (4.1)$$

where $u \equiv t - v$ is the agent's age in the planning period and $\mathbf{I}\left(\tau - v, R(v)\right)$ is an indicator function capturing the agent's labour market status:

$$\mathbf{I}(\tau - v, R(v)) = \begin{cases} 1 & \text{for } 0 < \tau - v < R(v) \\ 0 & \text{for } \tau - v \geq R(v) \end{cases}. \quad (4.2)$$

In Equation (4.1), $U(\cdot)$ is a concave consumption-felicity function (to be discussed below), $\bar{c}(v,\tau)$ is goods consumption, $D(\cdot)$ is the age-dependent disutility of working, $R(v)$ is the retirement age (see below), θ is the constant pure rate of time preference ($\theta > 0$), and $e^{-M(\tau - v)}$ is the probability that the agent is still alive at time τ. The cumulative mortality rate is defined as $M(\tau - v) \equiv \int_0^{\tau - v} m\left(s\right) ds$, where $m(s)$ is the instantaneous mortality rate of a household of age s (see the discussion in Section 2.3 for details). Two features of the lifetime utility function are worth noting.

First, following the standard convention in the literature, the instantaneous utility function is assumed to be additively separable in goods consumption and labour supply.[3] Previous to retirement the agent works full time, and inelastically supplies its unitary time endowment to the labour market. After retirement the agent does not work at all. Hence, we model the labour market participation decision (rather than an hours-of-work decision). Leaving the labour force is assumed to constitute an irreversible decision.[4] As a result, the age at which the agent chooses to withdraw from the labour market, which we denote by $R(v)$, can be interpreted as the *voluntary retirement age*. Second, the disutility of working is non-decreasing in age, i.e. $D'(\tau - v) > 0$. This captures the notion that working becomes more burdensome as one grows older (cf. Boucekkine et al. (2002, p. 346)).

The budget identity is given by:

$$\begin{aligned}\dot{\bar{a}}(v,\tau) = {} & [r + m(\tau - v)]\bar{a}(v,\tau) + \mathbf{I}(\tau - v, R(v))\bar{w}(\tau - v)[1 - t_L(\tau)] \\ & + [1 - \mathbf{I}(\tau - v, R(v))]\bar{p}(v,\tau,R(v)) - \bar{c}(v,\tau) - \bar{z}(\tau), \quad (4.3)\end{aligned}$$

where $\bar{a}(v,\tau)$ is real financial wealth, r is the exogenously given (constant) world rate of interest, $\bar{w}(\tau - v)$ is the age-dependent before-tax wage rate, t_L is the labour income tax, $\bar{p}(\cdot)$ is the public pension benefit, and $\bar{z}$ is the lumpsum tax (see below). As usual, a dot above a variable denotes that variable's time rate of change, e.g. $\dot{\bar{a}}(v,\tau) \equiv d\bar{a}(v,\tau)/d\tau$. Like in the previous chapter, we follow Yaari (1965) and Blanchard (1985), and postulate the existence of a perfectly competitive life insurance sector which offers actuarially fair annuity contracts. As a result, the annuity rate of interest facing an agent of age $\tau - v$ is given by $r + m(\tau - v)$.

The public pension system is modelled as follows. The government cannot force people to work, i.e. the voluntary retirement age, $R(v)$, is chosen freely by each individual agent. However, there exists an *early eligibility age* (EEA hereafter), which we denote by R_E. The EEA represents the earliest age at which social retirement benefits can be claimed. An agent who chooses to retire before reaching the EEA ($R(v) < R_E$) will only get a public pension benefit from age R_E onward, i.e. this agent will have no non-asset income during the age interval $[R(v), R_E]$. The pension

[3] See, for example, Sheshinski (1978); Burbidge and Robb (1980); Mitchell and Fields (1984); Kingston (2000); Boucekkine et al. (2002); Kalemni-Ozcan and Weil (2002), and d'Albis and Augeraud-Véron (2005)

[4] Apart from lifetime uncertainty there are no other stochastic shocks in our model and agents are blessed with perfect foresight. The empirical literature models retirement under uncertainty using the option-value approach. See, for example, Stock and Wise (1990b, 1990a); Lumsdaine et al. (1992) and the recent survey by Lumsdaine and Mitchell (1999).

benefits someone ultimately receives depends solely on that person's retirement age:[5]

$$\bar{p}(v,\tau,R(v)) = \begin{cases} 0 & \text{if } \tau - v < R_E \\ B(R(v)) & \text{if } \tau - v \geq R_E \end{cases} \tag{4.4}$$

where $B(R(v))$ is non-decreasing in the retirement age, i.e. $B'(R(v)) \geq 0$. Note that $B(R(v))$ might be discontinuous at some retirement ages, but if it exists such a jump is positive by assumption.

Lifetime income (or human wealth) is defined as the present value of after-tax non-asset income using the annuity rate of interest for discounting. For a working individual, whose age in the planning period falls short of the desired retirement age ($t - v < R(v)$), lifetime income is given by:

$$\bar{l}i(v,t,R(v)) \equiv e^{ru+M(u)} \left[\int_u^{R(v)} \bar{w}(s) e^{-rs-M(s)} ds - \int_u^{\infty} \bar{z}(v+s) e^{-rs-M(s)} ds \right] + SSW(v,t,R(v)), \tag{4.5}$$

where $SSW(v,t,R(v))$ represents the value of *social security wealth*:

$$SSW(v,t,R(v)) = e^{ru+M(u)} \left[B(R(v)) \int_{\max\{R_E,R(v)\}}^{\infty} e^{-rs-M(s)} ds - \int_u^{R(v)} t_L(v+s)\bar{w}(s) e^{-rs-M(s)} ds \right]. \tag{4.6}$$

Intuitively, social security wealth represents the present value of retirement benefits minus contributions, again using the annuity rate of interest for discounting. Lifetime income is an important wealth component for each agent. Indeed, by integrating the budget identity (4.3) for $\tau \in [t,\infty)$ and imposing the No-Ponzi-Game (NPG) condition, we obtain the lifetime budget constraint:

$$e^{ru+M(u)} \int_t^{\infty} \bar{c}(v,t) e^{-r\cdot(\tau-v)-M(\tau-v)} d\tau = \bar{a}(v,t) + \bar{l}i(v,t,R(v)). \tag{4.7}$$

The present value of current and future consumption is equated to total wealth,

[5] We thus assume a *pure defined benefit system*, i.e. previous payments into to the pension system do not influence the benefit. Sheshinski (1978, p. 353) assumes that pension benefits also depend on characteristics of the worker's wage profile before retirement, e.g. the arithmetic average wage, $\bar{w}_R \equiv (1/R)\int_0^R \bar{w}(s)ds$, or the maximum earned wage, $\bar{w}_R \equiv \max\{\bar{w}(s)\}$ for $0 \leq s \leq R$. We have abstracted from this dependency to keep the analysis a simple as possible.

which equals the sum of financial wealth and human wealth.

The agent of vintage v chooses a time path for consumption $\bar{c}(v,\tau)$ (for $\tau \in [t,\infty)$) and a retirement age $R(v)$ in order to maximize lifetime utility (4.1) subject to the lifetime budget constraint (4.7). Unfortunately, the optimisation procedure used in Chapter 3 does not work here. In the previous chapter, an individual had to choose an optimal schooling period and picked that schooling period that maximised lifetime income. Here the lifetime income maximising retirement age does not maximise lifetime utility, since leisure increases utility. However, due to the separability of preferences, the optimization problem can be solved in two steps. In the first step, we solve for optimal consumption conditional on total wealth. As before we postulate an iso-elastic consumption-felicity function (see Equation (2.6) on page 15):

$$U(\bar{c}(v,\tau)) \equiv \begin{cases} \dfrac{\bar{c}(v,\tau)^{1-1/\sigma}-1}{1-1/\sigma} & \text{for } \sigma \neq 1 \\ \ln \bar{c}(v,\tau) & \text{for } \sigma = 1 \end{cases} \tag{4.8}$$

where σ is the intertemporal substitution elasticity ($\sigma > 0$). The level and time profile for consumption are given by:

$$\bar{c}(v,t) = \frac{\bar{a}(v,t)+\bar{l}i(v,t,R(v))}{\Delta(u,r^*)}, \tag{4.9}$$

$$\bar{c}(v,\tau) = \bar{c}(v,t)e^{\sigma\cdot(r-\theta)(\tau-t)}, \quad \text{for } \tau \geq t, \tag{4.10}$$

where $r^* \equiv r - \sigma\cdot(r-\theta)$.[6] The Δ-function is defined in (2.12) and the important properties of this function are stated in Proposition 2.1.

Equation (4.9) shows that consumption in the planning period is proportional to total wealth, with $1/\Delta(u,r^*)$ representing the marginal propensity to consume. It follows from Proposition 2.1(v) that the consumption propensity is an increasing function of the individual's age in the planning period. Old agents face a relatively short expected remaining lifetime, due to increasing mortality rates, and thus consume a larger fraction of their wealth in each period. Equation (4.10) states the time path for consumption. As in the previous chapters, we assume throughout this chapter that $r > \theta$, i.e. we study a small nation populated by patient agents. It follows from (4.10) that the desired consumption profile is exponentially increasing over time.

[6] The derivation of Equations (4.9)–(4.10) is explained in detail in Chapter 2

In the second step of the maximization problem the optimal retirement age is chosen. This in turn determines optimal lifetime income. The retirement decision is only relevant for a working individual, because labour market exit is an absorbing state. By substituting (4.9)–4.10 into (4.1) we obtain the expression for lifetime utility of a working individual:

$$\bar{\Lambda}(v,t) \equiv e^{\theta u+M(u)} \int_u^\infty \left[U\left(\frac{\bar{a}(v,t)+\bar{l}i(v,t,R(v))}{\Delta(u,r^*)} e^{\sigma\cdot(r-\theta)(s-u)} \right) e^{-\theta s+M(s)} ds \right. \\ \left. - \int_u^{R(v)} D(s) e^{-\theta s - M(s)} ds \right], \qquad \text{for } u < R(v). \quad (4.11)$$

Borrowing terminology from econometrics, we refer to $\bar{\Lambda}(v,t)$ as the *concentrated* utility function, i.e. it is a transformation of the original lifetime utility function with the maximized solution for the consumption path incorporated in it. As a result, the concentrated utility function only depends on total wealth (including lifetime income) and on the retirement age. Every working individual maximizes (4.11) by choosing $\bar{l}i(v,t,R(v))$ and $R(v)$ subject to the definition of lifetime income (4.5), taking as given the stock of financial assets in the planning period.[7] This is a simple two-dimensional optimization problem with a single constraint. The optimal retirement age, $R^*(v)$, is the implicit solution to the following first-order condition:[8]

$$D(R(v))e^{-\theta R(v)-M(R(v))} = \left[\frac{\bar{a}(v,t)+\bar{l}i(v,t,R(v))}{\Delta(u,r^*)} \right]^{-1/\sigma} \frac{d\bar{l}i(v,t,R(v))}{dR(v)}. \quad (4.12)$$

The comparative static effects of the optimal retirement age with respect to ageing and pension shocks are studied in detail in Sections 4.3 and 4.5 below. One important property of the solution is immediately apparent from (4.12): no rational agent will choose a retirement age at which lifetime income is downward sloping. Because the disutility of working is strictly positive, the optimal solution must be situated on the upward sloping part of the $\bar{l}i(v,t,R(v))$ function. A direct corollary to this argument is as follows. If there exists a lifetime-income maximizing retirement age, say R_I, then this age is an upper bound for the utility-maximizing retirement age, i.e. it is never optimal to retire after age R_I.[9]

[7] After retirement, $R(v)$ is fixed and lifetime income is no longer a choice variable. Each individual simply chooses consumption such that the lifetime budget constraint is just satisfied.

[8] Similar expressions can be found in Sheshinski (1978, p. 354) and Burbidge and Robb (1980, p. 424)). Our expression differs from theirs because we allow for lifetime uncertainty, whereas they assume that agents have fixed lifetimes.

[9] See also Footnote 14 below. As is pointed out by Kingston (2000, p. 834f5), Lazear (1979) assumes that

4.2.2 Demography

We use the demographic framework presented in Section 3.2.1 on page 60. The instantaneous mortality rate is $m(\alpha, \psi_m(v))$, and the corresponding cumulative mortality by $M(u, \psi_m(v)) = \int_0^u m(\alpha, \psi_m(v))d\alpha$, where $\psi_m(v)$ is a parameter that only depends on the time of birth.

The birth rate varies over time, but is still exogenous by assumption. In Section 3.2.1 we showed that the size of a newborn generation at time v is proportional to the current population at that time, i.e. $L(v, v) = b(v)L(v)$, with $b(v)$ the birth rate and $L(v)$ is the population size at time v. The size of cohort v at some later time τ is given by:

$$L(v, \tau) = L(v, v)e^{-M(\tau - v, \psi_m(v))} = bL(v)e^{-M(\tau - v, \psi_m)}. \tag{4.13}$$

The population shares are given by

$$l(v, t) \equiv \frac{L(v, t)}{L(t)} = b(v)e^{-N(v,t) - M(t - v, \psi_m)}, \qquad t \geq v, \tag{4.14}$$

and the population growth rate is implicitly defined by

$$\frac{1}{b(v)} = \int_{-\infty}^{t} e^{-N(v,t) - M(t - v, \psi_m)} dv. \tag{4.15}$$

Box 3.1 shows how to calculate the population growth rate from Equation (4.15) by rewriting this equation as a Volterra equation of the second kind.

4.2.3 Firms

Perfectly competitive firms use physical capital and efficiency units of labour to produce a homogeneous commodity, $Y(t)$, that is traded internationally. The technology is represented by the following Cobb-Douglas production function:

$$Y(t) = K(t)^{\varepsilon}[A_Y H(t)]^{1-\varepsilon}, \qquad 0 < \varepsilon < 1, \tag{4.16}$$

where A_Y is a constant index of labour-augmenting technological change, $K(t)$ is the aggregate stock of physical capital, and $H(t)$ is employment in efficiency units. Following Blanchard (1985, p. 235) and Gomme et al. (2005, p. 431) we assume

the disutility of labour is zero, so that retirement occurs at the point where lifetime income is maximized. Since this typically occurs late in life, Lazear uses this result to rationalize the existence of mandatory retirement.

that labour productivity is age dependent, i.e. a surviving worker of age $\tau - v$ is assumed to supply one unit of 'raw' labour and $E(\tau - v)$ efficiency units of labour. The efficiency profile is exogenous.[10] Aggregate employment in efficiency units is thus given by:

$$H(t) \equiv \int_{-\infty}^{t} L(v,t)E(t-v)\mathbf{I}(t-v,R(v))dv. \tag{4.17}$$

Following the same steps as in Section 3.2.2 we obtain the usual factor demand equations

$$r + \delta = \varepsilon \left(\frac{A_Y h(t)}{k(t)} \right)^{1-\varepsilon} = \frac{\partial Y(t)}{\partial K(t)}, \tag{4.18}$$

$$w(t) = (1-\varepsilon) A_Y \left(\frac{A_Y h(t)}{k(t)} \right)^{-\varepsilon} = \frac{\partial Y(t)}{\partial H(t)}, \tag{4.19}$$

where $h(t) \equiv H(t)/L(t)$ and $k(t) \equiv K(t)/L(t)$. For each factor of production, the marginal product is equated to the rental rate. Since the fixed world interest rate pins down the ratio between $h(t)$ and $k(t)$, it follows from (4.19) that the rental rate on efficiency units of labour is time-invariant, i.e. $w(\tau) = w$. Hence, both physical capital and output are proportional to employment at all time:

$$k(t) = A_Y \left(\frac{\varepsilon}{r+\delta} \right)^{1/(1-\varepsilon)} h(t), \tag{4.20}$$

$$y(t) = A_Y \left(\frac{\varepsilon}{r+\delta} \right)^{\varepsilon/(1-\varepsilon)} h(t), \tag{4.21}$$

where $y(t) \equiv Y(t)/L(t)$. Finally, since efficiency units of labour are perfectly substitutable in production, cost minimization of the firm implies that the wage rate for a worker of age u is equal to:

$$\bar{w}(u) = wE(u). \tag{4.22}$$

Despite the fact that w is constant, the wage facing individual workers is age-dependent because individual labour productivity is.

[10] The comparative static effects of changes in the $E(\tau - v)$ function on the retirement decision are studied in Section 4.3 below. Note that there exists a large literature on life-cycle labour supply and human capital accumulation. See, for example, Ben-Porath (1967), Razin (1972), Weiss (1972), Heckman (1976), Driffill (1980), Gustman and Steinmeier (1986) , Heckman et al. (1998), and Mulligan (1999).

4.3 Retirement and ageing in the absence of pensions

In this section we study the comparative static effect on the optimal retirement age of various ageing shocks. In order to build intuition, we abstract from a public pension system and restrict attention to a comparison of steady states. A supplementary aim of this section is to introduce the graphical apparatus with which the effects of pensions and ageing can be visualized in an intuitive manner.

4.3.1 The retirement decision

In the steady state, we have $t_L(s) = t_L$, $\bar{z}(s) = \bar{z}$, $\bar{a}(v,t) = \bar{a}(u)$, $R(v) = R$, $\bar{li}(v,t,R(v)) = \bar{li}(u,R)$, and the concentrated lifetime utility function and the expression for lifetime income can both be written in terms of the individual's actual age, u, and the planned retirement age, R:

$$\bar{\Lambda}(u) \equiv e^{\theta u + M(u)} \left[\int_u^\infty U\left(\frac{\bar{a}(u) + \bar{li}(u,R)}{\Delta(u,r^*)} e^{\sigma\cdot(r-\theta)(s-u)} \right) e^{-\theta s - M(s)} ds - \int_u^R D(s) e^{-\theta s - M(s)} ds \right], \quad (4.23)$$

$$\bar{li}(u,R) = e^{ru+M(u)} \int_u^R \bar{w}(s) e^{-rs-M(s)} ds - \bar{z}\Delta(u,r), \quad (4.24)$$

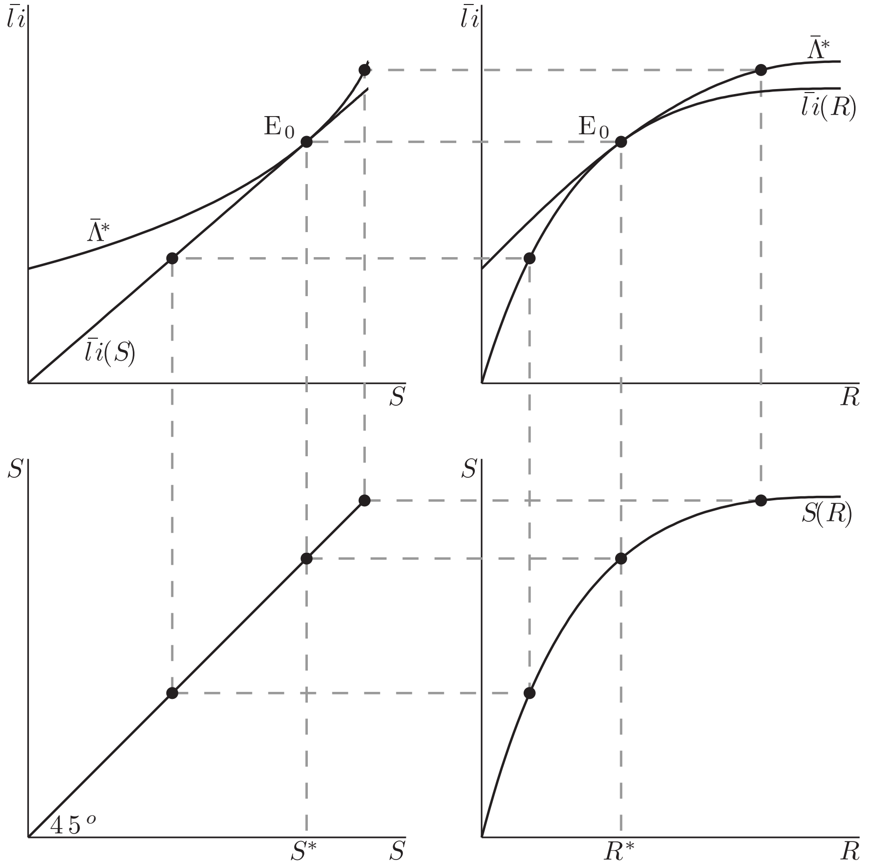
where $\bar{z}\Delta(u,r)$ represents the present value of lumpsum tax payments for an agent of age u.

In principle, it is possible to analyse the steady-state optimization problem directly in $(\bar{li}, R)$-space, but the solution is not easy to visualize because both indifference curves and the budget constraint are not well behaving, e.g. indifference curves are S-shaped or concave—see Appendix 4.A. This is not a problem, in and of itself, because it can be shown that, under mild restrictions, the budget constraint is always more curved in an interior solution than the indifference curves are. However, for the sake of simplicity and to facilitate the graphical exposition, it is more convenient to use a monotonic transformation of the retirement age (rather than R itself) as the retirement choice variable. In particular, we define the auxiliary variable S, which we refer to as the *transformed retirement age*, as follows:

$$S(u,R) = e^{ru+M(u)} \int_0^R e^{-rs-M(s)} ds, \qquad \text{for } 0 \le u \le R. \quad (4.25)$$

Clearly, S is a continuous, monotonically increasing transformation of R for a given

Figure 4.1. Optimal retirement and the transformed retirement age



age u, which ensures that the inverse function, $R = R(u, S)$, also exists. In the lower right panel of Figure 4.1 the transformation from R to S for a newborn (i.e. $S(0, R)$) is illustrated, using the Gompertz-Makeham mortality process fitted to the cohort born in the Netherlands in 1920 as in Chapters 2 and 3. See table 2.1 for details and parameters values of the Gompertz-Makeham mortality function. The concave shape of the transformation stretches the S intervals for young ages and compacts these intervals for old ages.

For a general demography, such as the Gompertz-Makeham process, the inverse function, $R(u, S)$, is only defined implicitly by Equation (4.25).[11] The derivative of

[11] For the Blanchard (1985) case the instantaneous mortality rate is constant and equal to μ_0. Equation

this inverse function is given by

$$\frac{\partial R}{\partial S} = e^{-ru-M(u)} e^{rR(u,S)+M(R(u,S))} > 0. \tag{4.26}$$

Where no confusion arises we drop the dependency of R on S and u from here on. For future reference we note that the EEA, utility-maximizing, and lifetime-income maximizing values for S are given by, respectively, $S_E = S(u, R_E)$, $S^* = S(u, R^*)$, and $S_I = S(u, R_I)$.

The slope and curvature of the indifference curves in $(\bar{l}i, S)$-space are obtained by implicit differentiation of Equation (4.23):

$$\left.\frac{d\bar{l}i}{dS}\right|_{\bar{\Lambda}_0} \equiv -\frac{\partial\bar{\Lambda}/\partial R}{\partial\bar{\Lambda}/\partial\bar{l}i} \times \frac{\partial R}{\partial S} = e^{(r-\theta)(R-u)} D(R) \left[\frac{\bar{a}(u) + \bar{l}i}{\Delta(u, r^*)}\right]^{1/\sigma} > 0, \tag{4.27}$$

$$\left.\frac{d^2\bar{l}i}{dS^2}\right|_{\bar{\Lambda}_0} = \left[\frac{1}{\sigma\cdot(\bar{a}(u) + \bar{l}i)} \left.\frac{d\bar{l}i}{dS}\right|_{\bar{\Lambda}_0} + \left(\frac{D'(R)}{D(R)} + r - \theta\right)\frac{\partial R}{\partial S}\right] \left.\frac{d\bar{l}i}{dS}\right|_{\bar{\Lambda}_0} > 0. \tag{4.28}$$

The indifference curves are upward sloping, since postponing retirement causes additional disutility of labour which must be compensated with a higher lifetime income. By assumption $D'(R) \geq 0$ and $r > \theta$, so the indifference curves are convex. In the upper left panel of Figure 4.1 an indifference curve for a newborn is illustrated—see the curve labelled $\bar{\Lambda}^*$.

By differentiating (4.24), noting (4.22) and (4.26), we find that the slope and curvature of the $\bar{l}i(u, S)$ curve are given by:

$$\frac{d\bar{l}i}{dS} = \bar{w}(R) = wE(R) > 0, \tag{4.29}$$

$$\frac{d^2\bar{l}i}{dS^2} = \bar{w}'(R)\frac{\partial R}{\partial S} = wE'(R)\frac{\partial R}{\partial S} \lesseqgtr 0. \tag{4.30}$$

By increasing the (transformed) retirement age slightly, lifetime income is increased by an amount equal to the wage rate facing an agent of age R. Depending on the age profile of wages, the budget constraint may contain convex segments (for

(4.25) simplifies to:

$$S(u, R) \equiv \frac{e^{(r+\mu_0)u}}{r + \mu_0}\left[1 - e^{-(r+\mu_0)R}\right], \qquad \text{for } R \geq 0,$$

and the $R(u, S)$ function is given by:

$$R(u, S) \equiv -\frac{1}{r + \mu_0} \ln\left[1 - (r + \mu_0)Se^{-(r+\mu_0)u}\right], \qquad \text{for } 0 \leq S < \frac{e^{(r+\mu_0)u}}{r + \mu_0}.$$

$\bar{w}'(R) > 0$), linear segments (for $\bar{w}'(R) = 0$), and concave segments (for $\bar{w}'(R) < 0$). The relevant case, however, appears to be that the wage is either constant or declining with age around the optimal age of retirement—see OECD (1998, p. 133) for empirical evidence on OECD countries. To streamline the discussion, we adopt the following assumption.

Assumption 4.1. *The wage schedule is non-increasing around the optimal retirement age and beyond, i.e.* $\bar{w}'(R) \leq 0$ *around and beyond* R^*.

In the upper left panel of Figure 4.1 we illustrate the linear budget constraint that results for the special case of an age-invariant wage rate ($\bar{w}'(R) = 0$ for all R). The optimum is located at point E_0, where there exists a tangency between the lifetime budget line and an indifference curve. The upper right panel shows the same equilibrium in $(\bar{l}i, R)$-space.

4.3.2 Ageing effects

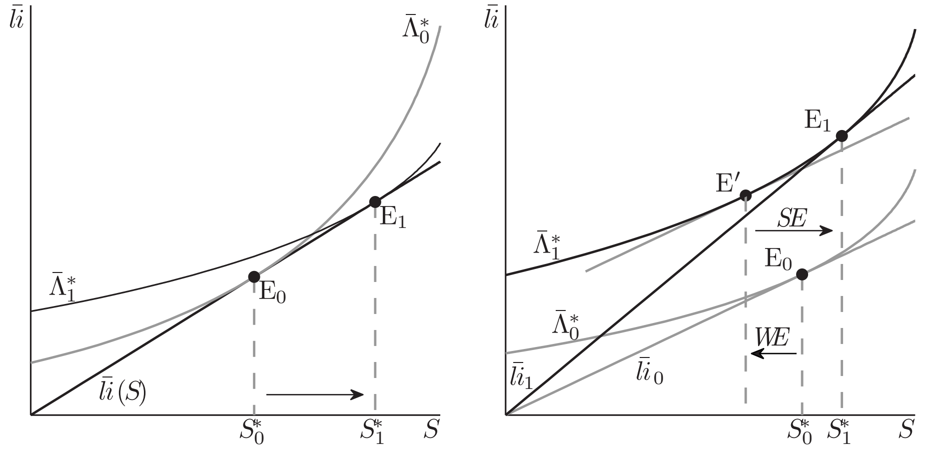
Our model distinguishes both *biological* and *economic* age dependencies. A biological ageing effect involves changes in the mortality structure, as captured by the mortality function $M(u, \psi_m)$, where ψ_m is the shift parameter introduced in Chapter 3. Economic ageing, on the other hand, refers to changes in the disutility of working or in the efficiency of labour over the life cycle, as captured by the functions $D(u, \psi_d)$ and $E(u, \psi_e)$, respectively, where ψ_d and ψ_e are the associated shift parameters. In the remainder of this section we focus on the retirement decision of a newborn, i.e. we set $u = \bar{a}(u) = 0$ in Equations (4.23)–(4.24). This entails no loss of generality because the agent's plans are dynamically consistent, i.e. the optimal retirement age is age-invariant.

Economic ageing

In Figure 4.2(a) we illustrate the effect on lifetime income and the optimal retirement age of a decrease in the disutility of labour, i.e. $\partial D(u, \psi_u)/\partial \psi_d \leq 0$ for all u, with strict inequality around $u = R^*$. Such a preference shock leaves the budget constraint unchanged, but changes the slope of the indifference curves. Indeed, it follows from (4.27) that:

$$\frac{\partial}{\partial \psi_d} \left[\frac{d\bar{l}i}{dS} \right]_{\bar{\Lambda}_0} = e^{r-\theta)(R-u)} \left[\frac{\bar{l}i}{\Delta(0, r^*)} \right]^{1/\sigma} \frac{D(R, \psi_d)}{\partial \psi_d} < 0. \tag{4.31}$$

Figure 4.2. Optimal retirement and economic ageing shocks



(a) Reduced disutility of working

(b) Increase in labour productivity

The indifference curves become flatter and the agent chooses a higher retirement age as a result—see the move from E$_0$ to E$_1$ in Figure 4.2(a).

In Figure 4.2(b) we depict the comparative static effect of an increase in the age profile of labour efficiency, i.e. $\partial E(u,\psi_e)/\partial\psi_e \geq 0$ with strict inequality for $u = R^*$. Indifference curves are not affected by this shock but the budget constraint is. Indeed, the effects of such a shock are complicated because there are offsetting wealth- and substitution effects. It follows from (4.24) that the budget constraint shifts up:

$$\frac{\partial \bar{li}}{\partial \psi_s} = w \int_0^R \frac{\partial E(s,\psi_e)}{\partial \psi_e} e^{-rs-M(s)} ds > 0, \tag{4.32}$$

and from (4.29) that it becomes steeper:

$$\frac{\partial}{\partial \psi_s}\left[\frac{d\bar{li}}{dS}\right] = w\frac{\partial E(R,\psi_e)}{\partial \psi_e} > 0. \tag{4.33}$$

In Figure 4.2(b) we illustrate the case for which the optimal retirement increases. The total effect is the change from E$_0$ to E$_1$, consisting of a negative wealth effect (from E$_0$ to E$'$) and a positive substitution effect (from E$'$ to E$_1$).

Biological ageing

Two types of demographic shocks are considered in our analysis, namely a change in the birth rate and a change in the mortality process. Clearly, in view of (4.23)–(4.24), the birth rate does not directly affect the retirement choice of individual agents.[12] The mortality process, however, affects the $\Delta(u, \lambda)$ function and thus the optimal retirement choice. In this chapter we will make the same assumptions regarding the effect of a change of ψ_m on the mortality process as in Chapter 3 (see Assumption 3.1 on page 65). These assumptions assure that the expected remaining lifetime increases for all ages and that the mortality profile shifts more downward for old ages than for young ages, it is a so-called adult mortality shock.

In order to compute the effect of increased longevity on retirement, we write the first-order condition for the optimal transformed retirement age, $d\bar{\Lambda}/dS = 0$, as follows:[13]

$$\Gamma(R^*, \psi_m) \equiv \bar{w}(R^*) - D(R^*)e^{(r-\theta)R^*} \left[\frac{\bar{l}i(0, R^*, \psi_m)}{\Delta(0, r^*, \psi_m)} \right]^{1/\sigma} = 0, \tag{4.34}$$

where the second-order condition for utility maximization implies that $\partial\Gamma/\partial R^* < 0$, and where $\bar{l}i(0, R^*)$ is given by:

$$\bar{l}i(0, R^*, \psi_m) \equiv \int_0^{R^*} \bar{w}(s)e^{-rs-M(s,\psi_m)}ds - \bar{z}\Delta(0, r, \psi_m). \tag{4.35}$$

In Equation (4.35), lifetime income depends on the mortality parameter ψ_m because both wage income and lumpsum taxes are discounted using the actuarially fair annuity rate of interest, $r + m(s, \psi_m)$. In addition, in Equation (4.34) the mortality parameter affects the marginal propensity to consume out of total wealth.

It follows from (4.34) that:

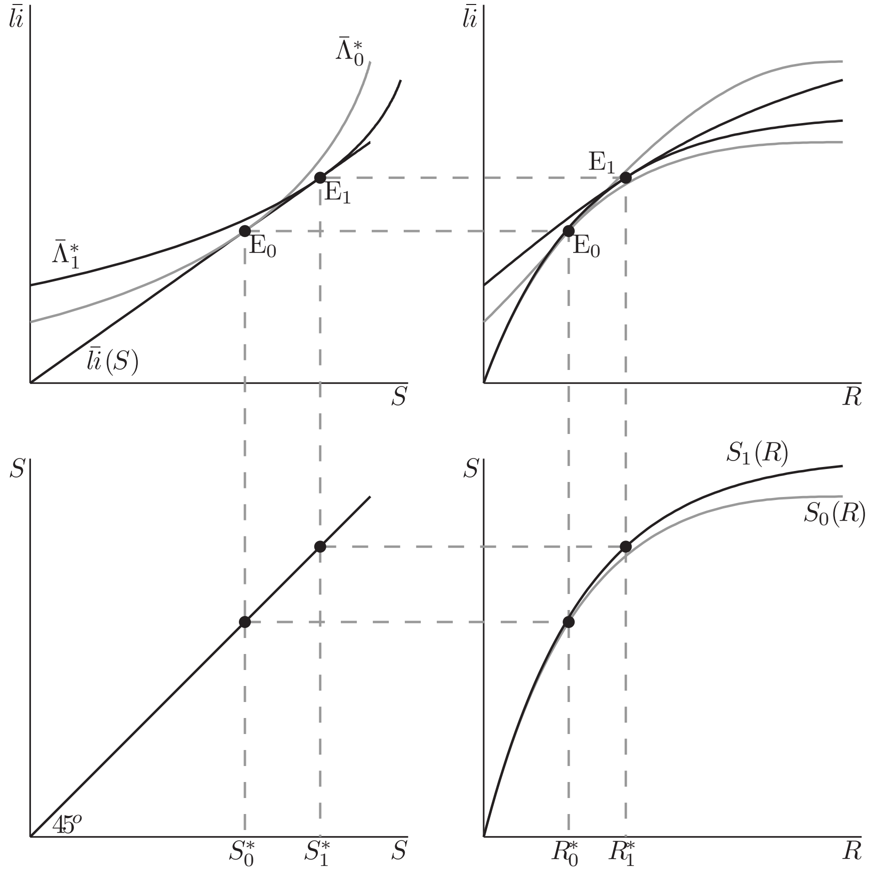
$$\frac{dR^*}{d\psi_m} = \frac{\partial\Gamma/\partial\psi_m}{|\partial\Gamma/\partial R|} = \frac{\bar{w}(R^*)}{\sigma|\partial\Gamma/\partial R^*|} \left[\frac{\partial\Delta(0, r^*, \psi_m)/\partial\psi_m}{\Delta(0, r^*, \psi_m)} - \frac{\partial\bar{l}i(0, R^*, \psi_m)/\partial\psi_m}{\bar{l}i(0, R^*, \psi_m)} \right] \gtreqless 0. \tag{4.36}$$

Clearly, the sign of the comparative static effect is determined by the term in square brackets on the right-hand side of (4.36). Using Proposition 3.1(iii) we find that $\partial\Delta(0, r^*, \psi_m)/\partial\psi_m > 0$ so the *propensity effect* operates in the direction of increasing

[12] Of course, in general equilibrium the birth rate may affect the retirement choice via the fiscal system. See Section 4.6 for a further analysis.

[13] This expression is obtained by combining Equations (4.27) and (4.29) and setting $u = \bar{a}(u) = 0$.

Figure 4.3. Optimal retirement and increased longevity



the retirement date. We find from (4.35) that the *lifetime-income effect* is ambiguous in general:

$$\frac{\partial \bar{li}(0,R^*,\psi_m)}{\partial \psi_m} = -\int_0^{R^*} \bar{w}(s)\frac{\partial M(s,\psi_m)}{\partial \psi_m} e^{-rs-M(s,\psi_m)}ds - \bar{z}\frac{\partial \Delta(0,r,\psi_m)}{\partial \psi_m} \gtreqless 0. \quad (4.37)$$

The first term on the right-hand side is positive (see Proposition 3.1(i)), i.e. as a result of reduced discounting of wage income, lifetime income increases. But lighter discounting also increases the lifetime burden of the lumpsum tax, i.e. the second term on the right-hand side is also positive. As a result, the wage effect moves in the opposite direction of the tax effect and the net effect of ageing on lifetime income cannot be signed a priori. Of course, in the absence of lumpsum taxes, the lifetime-income effect is positive and thus works in the direction of decreasing the retirement age. There is a strong presumption, however, that the first term on the right-hand side of (4.37) is rather small. Indeed, as can be gleaned from Figure 3.2(a) on page 67, an adult mortality shock starts to matter quantitatively for age levels at which most agents have already retired in advanced countries. Hence, the retirement age is likely to increase as longevity increases because the tax effect is dominant, i.e. $dR^*/d\psi_m > 0$ in realistic scenarios.

In Figure 4.3 we illustrate the comparative static effects of increased longevity. In panel (d), the mortality shock increases the transformed retirement age at all values of R, though more so for higher ages. Intuitively, by making the transformation curve steeper, a post-shock octogenarian is 'younger' than his/her pre-shock counterpart. As a result, the indifference curves in panel (a) flatten out so that, with a linear budget constraint (with $\bar{w}'(R) = 0$), the equilibrium shifts from E_0 to E_1. In panel (b) the same comparative static effect is shown in $(\bar{l}i, R)$-space.

4.4 Realistic pension system

In this section we re-introduce the public pension system and investigate its likely consequences for the trade-offs facing workers in advanced economies. As in the previous section, we continue to assume that the pension system is in a steady state. As a result, social security wealth (4.6) can be written as follows:

$$SSW(u,R) = e^{ru+M(u)}\left[B(R)\int_{\max\{R,R_E\}}^{\infty} e^{-rs-M(s)}ds - t_L\int_u^R \bar{w}(s)e^{-rs-M(s)}ds\right]. \tag{4.38}$$

By incorporating social security wealth into the steady-state budget constraint (4.24) and differentiating with respect to the transformed retirement age we obtain:

$$\frac{d\bar{l}i}{dS} = \begin{cases} (1-t_L)\bar{w}(R) + B'(R)\Pi(R,R_E,\infty,r) > 0 & \text{for } S < S_E \\ (1-t_L)\bar{w}(R) + B'(R)\Delta(R,r) - B(R) \geq 0 & \text{for } S_E \leq S \leq S_I \end{cases} \tag{4.39}$$

where R_E and R_I (S_E and S_I) are, respectively, the (transformed) EEA and lifetime-income maximizing retirement age.[14] The Π-function appearing in the upper branch

[14] In the presence of a public pension system, R_I is defined implicitly by $\Delta(R_I,r) = B(R_I) - (1 - t_L)w(R_I)$. Since $B'(R_I) \geq 0$, $\bar{w}'(R_I) \leq 0$ (Assumption 4.1) and $\partial\Delta(R_I,r)/\partial R_I < 0$ (Proposition 2.1(v)), it follows that there exists a unique value for R_I.

of (4.39) is defined in general terms as:

$$\Pi(u, \underline{u}, \bar{u}, \lambda) = e^{\lambda u + M(u)} \int_{\underline{u}}^{\bar{u}} e^{-\lambda s - M(s)} ds. \tag{4.40}$$

$\Pi(u, \underline{u}, \bar{u}, \lambda)$ is the present value of an annuity that one receives during the age interval $(\underline{u}, \bar{u})$, evaluated at age u, using the discount rate λ. The 'regular' Δ-function, is a special case of the Π-function, with $\underline{u} = u$ and $\bar{u} = \infty$.

As is evident from (4.39), the shape, slope, and curvature of the budget constraint is complicated by the existence of the EEA. If $B(R)$ and $B'(R)$ are both continuous at $R = R_E$, then the budget constraint is continuous but features a kink at that point equal to $-B(R_E)$. The kink represents the retirement benefit that is foregone by not retiring at R_E but at some later age.

The curvature of the lifetime income function is ambiguous in general, i.e. it cannot be inferred from theoretical first principles whether or not it is concave in the relevant region. Our reading of the empirical comparative-institutional literature for OECD countries, however, gives us enough confidence to formulate the following assumption.

Assumption 4.2. *In the relevant calender age domain of 55 to 70, the lifetime income function is concave in the transformed retirement age S. It may feature a single kink at the EEA.*

Our defence for this assumption takes up the remainder of this section and proceeds as follows. In the literature (e.g. Gruber and Wise (1999, 2004), and OECD (2005)) retirement incentives are typically summarized with the EEA, the NRA, the replacement rate, the pattern of benefit accrual, and the implicit tax rate. Using these incentive indicators, it is possible to derive the shape and slope of the lifetime income function.

The *replacement rate* is defined as the ratio of the retirement benefit to net wages. In terms of our theoretical framework, the replacement rate RR for someone retiring at or after the EEA is given by:

$$RR(R) \equiv \frac{B(R)}{(1 - t_L)\bar{w}(R)} \qquad \text{for } R \geq R_E. \tag{4.41}$$

This replacement rate differs greatly between countries, but also between ages. As can be seen from Table 4.B.1 in Appendix 4.B, the replacement rate for France starts at 92% at the EEA (59 years) and slowly increases to 96% thereafter. In contrast, in

Canada the replacement rate starts at 18% at age 59, after which it increases to 91% for a 69 year old.

The *benefit accrual* is the nominal change in social security wealth if one postpones retirement by one year (i.e., it is $\partial SSW/\partial R$ in terms of our model). The benefit accrual depends on the age of the individual. To compare accrual levels, we can either hold constant the age at which social security wealth is evaluated or evaluate social security wealth at the actual retirement age. Both methods have their advantages and drawbacks. The first makes it easier to track social security wealth over time, the second allows for easier comparison of retirement incentives at the retirement age. In this chapter we will use the second definition because it allows for easier mathematics in the transformed retirement age S-space. By differentiating Equation (4.38) with respect to R and evaluating at age $u = R$ we obtain:

$$ACC(R) \equiv \left.\frac{\partial SSW(u,R)}{\partial R}\right|_{u=R} = \begin{cases} B'(R)\Pi(R, R_E, \infty, r) - t_L\bar{w}(R) & \text{for } R < R_E \\ B'(R)\Delta(R,r) - t_L\bar{w}(R) - B(R) & \text{for } R > R_E \end{cases} \tag{4.42}$$

The level of benefit accrual is closely connected to the slope of the lifetime income function (as a function of S). Indeed, by combining (4.39) and (4.42) we obtain:

$$\frac{d\bar{l}i}{dS} = ACC(R) + \bar{w}(R). \tag{4.43}$$

In this context, actuarial adjustment of the pension benefit is called *fair* if and only if $ACC(R) = 0$ for all R.

The benefit accrual depends on the monetary units in which social retirement benefits are measured and the age at which the social security wealth is evaluated. Most studies standardize the benefit accrual either with the level of social security wealth or with the present value of net wages. The first measure is the accrual rate, the second measure is the implicit subsidy.

The negative of the implicit subsidy is the *implicit tax rate* (IT), measuring the additional tax rate one 'implicitly' faces over and above the normal taxes. A negative accrual is an additional tax on labour and discourages work. Conversely, a positive accrual is an implicit subsidy on labour and encourages the individual to work an additional year. Since we evaluate the accrual level at the retirement age, we should not discount the net wage rate, so the implicit tax can be written in terms of our model as:[15]

$$IT(R) \equiv -\frac{ACC(R)}{(1-t_L)\bar{w}(R)}. \tag{4.44}$$

By substituting this expression for the implicit tax rate into Equation (4.43), we can write the slope of the lifetime income function as:

$$\frac{d\bar{li}}{dS} = (1-t_L)\bar{w}(R)\left[\frac{1}{1-t_L} - IT(R)\right]. \tag{4.45}$$

Under the maintained assumption that gross wages are constant for higher ages (typically in the range 55–70), Equation (4.45) can be used to compute the shape of the lifetime income function. Dividing lifetime income by net wages gives a 'standardized' measure of lifetime income which is more easily comparable between countries. The only caveat is that we do not have data on the relevant labour income tax, so we cannot estimate $1/(1-t_L)$. This is not a problem, however, because we are only interested in the curvature of the lifetime income function (its convexity or concavity). The term $1/(1-t_L)$ only influences the slope of the lifetime income profile, but it has no effect on its curvature. To get an idea of the shape of the lifetime income profile we proceed as if there are no labour income taxes. Since the fraction $1/(1-t_L)$ has a lower bound of 1 (because in reality taxes are positive), we thus obtain a conservative estimate for lifetime income.

Figure 4.4 shows the lifetime income profiles for nine OECD countries, as we computed them using the implicit tax rates published in Gruber and Wise (1999). For convenience these tax rates have also been reported in Table 4.B.2 in Appendix 4.B. The lifetime income profiles are normalized at age 54 to enable comparison between countries. The graphs contain two horizontal axes. The main (lower) horizontal axis measures the transformed retirement age, S, whilst the secondary (upper) axis shows the corresponding values for the untransformed retirement age, R. The effect of the non-monotonic scaling is clearly visible.

Figure 4.4(a) characterizes the retirement systems in continental Europe. Lifetime income profiles are increasing in the retirement age, more or less concave and usually have a clear kink at the EEA (which is 60 years in most countries, but only 55 in Italy) or at the NRA (65). A notable exception is formed by the Netherlands. Its profile has a sharp spike at age 59 and decreases until the NRA of 65. The pen-

[15] Some contributors to Gruber and Wise (1999) do not provide information concerning the age at which they evaluate the present value of social security wealth. This is not a problem, however, provided we do not use the accrual levels, but the accrual rates or implicit tax rates.

sion system in the Netherlands is such that there exists an implicit tax of more than 141% of net earnings. The pension benefits someone receives hardly increases if someone retires after age 59, but one still has to pay contributions to the pension system. Moreover, replacement rates are very high due to the usually generous, but mandatory, company pension systems. It is not surprising that most employees in the Netherlands retire at age 60.[16]

Figure 4.4(b) characterizes the retirement systems in Canada, Japan, the United Kingdom, and the United States. A feature of these systems is the rather low implicit tax rates. A low implicit tax is a symptom of either (i) a near-actuarially fair system, or (ii) a rather poorly developed pension system. The replacement rates in Table 4.B.1 indicate that the former is the case in Canada and the United States, whereas the latter is relevant for the United Kingdom, leaving Japan as a somewhat mixed case. As a result of the small implicit tax rates, the wage effect in the lifetime income function (4.45) is dominant and the standardized lifetime income profiles are roughly the same in these four countries.

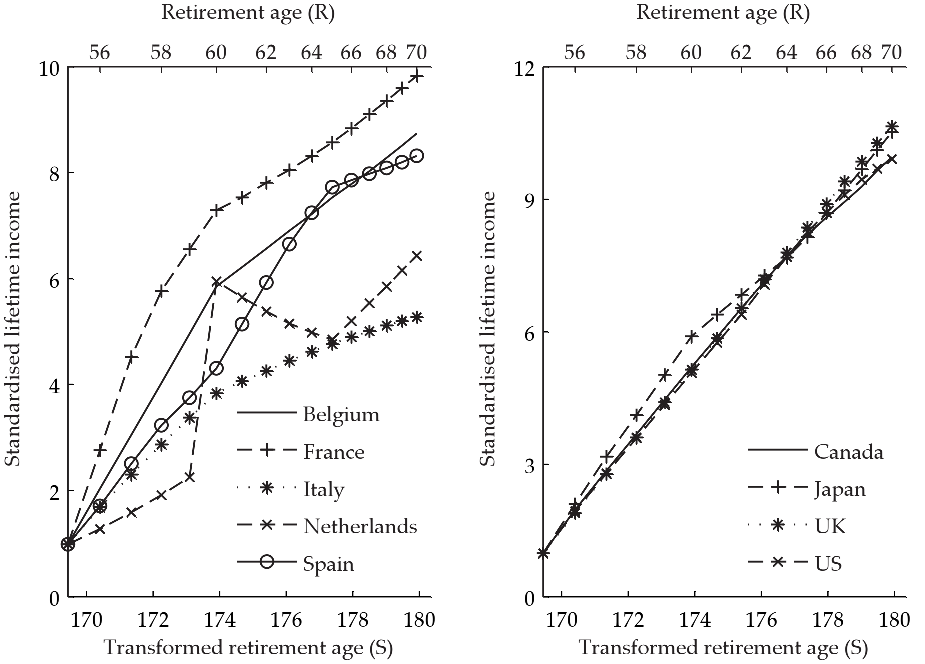
Although Figure 4.4 only shows conservative estimates for the lifetime income profiles, it does give an accurate picture concerning the shape of these profiles. Apart from Spain and the Netherlands, the income profiles are concave and may feature a kink at the EEA. Even though the profile for the Netherlands is not concave, there is a pronounced kink at age 60 which precludes individuals from working beyond that age.

4.5 Ageing and pension shocks

In this section we study the comparative static effects on the optimal steady-state retirement age of various marginal changes in the tax system or the public pension scheme. In view of Assumption 4.2 and because indifference curves are convex in $(\bar{l}i, S)$-space, the optimum retirement age is unique. If there is no kink in the lifetime income profile, then there will be an interior solution. In the presence of a single kink, however, there are three possible outcomes. First, if the agent's disutility of labour is high, and indifference curves are relatively steep, then the interior optimum occurs to the left of the kink, i.e. the agent chooses $R^* < R_E$, contra stylized fact (SF3). Second, if labour disutility is moderate, then indifference curves are relatively flat and there will be a corner solution at the kink, i.e. $R^* = R_E$. Third,

[16] The graph is based on retirement schemes as they existed in the late 1980s. More recent figures published in Gruber and Wise (2004) provide a qualitatively similar picture.

Figure 4.4. Lifetime income profiles for nine OECD countries (lower bound)



(a) Continental Europe (b) Other countries

Source: Gruber and Wise (1999) and own calculations.

if labour disutility is very low then there will be an interior solution to the right of the EEA, i.e. $R^* > R_E$. The second and third cases are not inconsistent with reality.

In this section we focus on the interior solutions and we assume that the retirement age is strictly larger than the EEA ($R^* > R_E$). By combining Equations (4.27) and (4.39) and setting $u = \bar{a}(u) = 0$ we obtain the first-order condition which implicitly defines a unique solution for R^*:

$$\Gamma(R^*) \equiv (1-t_L)\bar{w}(R^*) + B'(R^*)\Delta(R^*,r) \\ - B(R^*) - e^{(r-\theta)R^*} D(R^*) \left[\frac{\bar{li}(0,R^*)}{\Delta(0,r^*)}\right]^{1/\sigma} = 0, \quad (4.46)$$

where $\bar{li}(0,R^*)$ is obtained by adding $SSW(0,R^*)$ to the right-hand side of Equation (4.35) above. The second-order condition of utility maximization implies that

$\partial\Gamma/\partial R^* < 0$. For future reference we define the following partial derivative:

$$\left|\frac{\partial\Gamma(R^*)}{\partial\bar{l}i}\right| \equiv e^{(r-\theta)R^*}\frac{D(R^*)}{\sigma\Delta(0,r^*)}\left[\frac{\bar{l}i(0,R^*)}{\Delta(0,r^*)}\right]^{(1-\sigma)/\sigma} > 0. \tag{4.47}$$

Changes in the tax system affect the optimal retirement age in the following way. First, an increase in the lumpsum tax leads to a reduction in lifetime income and an increase in the retirement age:

$$\frac{dR^*}{d\bar{z}} = \frac{\partial\Gamma/\partial\bar{z}}{|\partial\Gamma/\partial R^*|} = \Delta(0,r)\left|\frac{\partial\Gamma(R^*)}{\partial\bar{l}i}\right| > 0. \tag{4.48}$$

Intuitively, the tax change induces a pure wealth effect. Because consumption and leisure are both normal goods, labour supply is increased, i.e. the agent retires later in life. Second, a change in the labour income tax rate has an ambiguous effect:

$$\frac{dR^*}{dt_L} = \frac{\partial\Gamma/\partial t_L}{|\partial\Gamma/\partial R^*|} \equiv -\frac{\bar{w}(R^*)}{|\partial\Gamma/\partial R^*|} + \left|\frac{\partial\Gamma(R^*)}{\partial\bar{l}i}\right|\frac{\int_0^{R^*}\bar{w}(s)e^{-rs-M(s)}ds}{|\partial\Gamma/\partial R^*|} \lesseqgtr 0. \tag{4.49}$$

The first term on the right-hand side of (4.49) represents the substitution effect, which is negative. A higher tax discourages working and thus encourages retiring earlier in life via that effect. The second term is the positive wealth effect. The tax increase makes the agent poorer and thus provides incentives to retire later in life. In summary, the labour income tax increase operates qualitatively like a decrease in labour efficiency (see Equations (4.32)–(4.33) and Figure 4.2(b)).

Changes in the pension system affect the retirement decision as follows. First, holding constant the slope of the retirement benefit curve, the effect of a change in its *level* is negative:

$$\frac{dR^*}{dB(R)} = \frac{\partial\Gamma/\partial B(R)}{|\partial\Gamma/\partial R^*|} = -\frac{1}{|\partial\Gamma/\partial R^*|} - \left|\frac{\partial\Gamma(R^*)}{\partial\bar{l}i}\right|\frac{\int_{R^*}^{\infty}e^{-rs-M(s)}ds}{|\partial\Gamma/\partial R^*|} < 0. \tag{4.50}$$

In this case the wealth- and substitution effects operate in the same direction. The first term on the right-hand side of (4.50) is the negative substitution effect: by increasing the public retirement benefit the rewards to working longer are reduced, i.e. the relevant branch of the budget constraint (4.39) is rotated in a clockwise fashion. The second term on the right-hand side of (4.50) is the negative wealth effect. The benefit increase boosts lifetime income and thus induces agents to work less and to retire earlier in life. In graphical terms, the wealth effect leads to an upward shift of the lifetime budget constraint.

Second, ceteris paribus the level of the benefit function, a change in its *slope* $B'(R)$ causes a positive substitution effect:

$$\frac{dR^*}{dB'(R)} = \frac{\partial\Gamma/\partial B'(R)}{|\partial\Gamma/\partial R^*|} = \frac{\Delta(R^*,r)}{|\partial\Gamma/\partial R^*|} > 0. \tag{4.51}$$

Intuitively, the steeper slope of the benefit function induces agents to postpone retirement somewhat. In graphical terms, the budget constraint rotates counter-clockwise and the optimal retirement age shifts to the right.[17]

4.6 Demographic change and policy reform

In this section we compute and visualize the general equilibrium computational results of various demographic shocks and their assumed fiscal reform measures. We restrict attention on measures characterizing the aggregate economy. The per capita aggregate variables are calculated as in the previous chapters. Per capita consumption, for example, is computed as $c(t) \equiv \int_{-\infty}^{t} l(v,t)\bar{c}(v,t)dv$, where the relative cohort weight, $l(v,t)$, is defined in Equation (4.14) above, and individual consumption, $\bar{c}(v,t)$, is given in (4.9).

In accordance with stylized fact (SF4), we calibrate the model in such a way that the initial optimum retirement age is at the EEA, i.e. the budget constraint features a kink at the EEA and individual agent are 'stuck' in this corner solution. The main demographic and economic features of the calibrated model are as follows. The mortality process is as in the previous chapters. It represents the fitted G-M process for the cohort born in 1920 in the Netherlands. Life expectancy at birth for this cohort is 65.5 years. The crude birth rate is set at $b = 0.0237$, a value that lies in between the observed birth rates of 1920 and 1940. In combination, the demographic parameters imply an initial steady-state population growth rate equal to $\hat{n}_0 = 1.34\%$. For households we assume in this chapter that the world interest rate facing them equals $r = 0.05$, the rate of time preference is $\theta = 0.03$, and the intertemporal substitution elasticity is $\sigma = 0.8$. In combination, these parameter values imply an annual consumption growth for individuals of $\dot{\bar{c}}(v,t)/\bar{c}(v,t) = \sigma\cdot(r-\theta) = 0.016$. Disutility of labour and labour efficiency are both age-invariant and set at, respectively, $D(u) = 0.15$ and $E(u) = 10$. On the production side, we set the share of capital in the production function at $\varepsilon = 0.4$,

[17] Following Sheshinski (1978, pp. 357-8), we write the pension benefit as $B(R,\psi_b)$, where ψ_b is a shift parameter. The first pension shock assumes $\partial B/\partial\psi_b > 0$ and $\partial^2 B/\partial\psi_b\partial R = 0$. The second shock sets $\partial B/\partial\psi_b = 0$ and $\partial^2 B/\partial\psi_b\partial R > 0$.

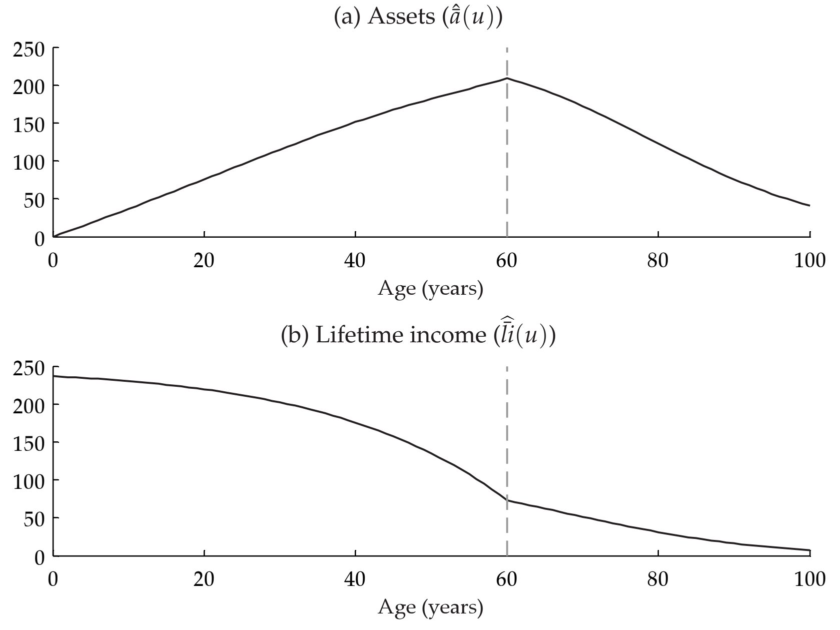
the technology index is $A_Y = 1$, and the depreciation rate of capital is $\delta = 0.06$. For the policy parameters we use the following values. The labour income tax is $t_L = 0.1$, the lumpsum tax is $\bar{z}_0 = -0.166$, the initial debt level is $\hat{d}_0 = 10$, and the EEA is set at $R_E = 60$. For somebody who retires before the EEA, pension benefits are zero until the EEA and equal to $\beta_0 = 7.094$ from age EEA onward (this value for β_0 amounts to 50% of an agent's gross wage). For somebody retiring at or after the EEA, pension benefits are zero until actual retirement, and equal to $\beta_0 + \beta_1(R - R_E)$ from age R onward, where $\beta_1 = 0.05$.

Figure 4.5 shows the steady-state age profiles of financial assets and lifetime income. Panel (a) shows that the profile for assets is inverse U-shaped and reaches a peak at $u = R_E$. After retirement, the agent slowly decumulates its assets. Panel (b) shows that there is a kink in the profile for lifetime income at $u = R_E$. The initial steady state has the following aggregate features: $\hat{w} = 14.2$, $\hat{a}_0 = 100.5$, $\hat{li}_0 = 184.4$, $\hat{h}_0 = 9.0$, $\hat{y}_0 = 21.2$, $\hat{c}_0 = 16.0$, $\hat{i}_0 = 5.7$, $\hat{k}_0 = 77.0$, and $\hat{f}_0 = 13.5$. The output shares of consumption, investment, net exports, and the government primary surplus are, respectively, 75.6%, 26.7%, -2.3% and 1.73%. We summarize the initial steady state in column (1) of Table 4.1. Our model economy is clearly not a banana republic. It is a wealthy country ruled by a fiscally responsible government (that is running a primary surplus), and populated by long-lived and patient citizens (who as a group hold a net claim on the rest of the world).

The comparative dynamic exercises performed throughout this section take the following form. Starting from the initial steady state, the economy is hit by one of two types of demographic change occurring at time $t = 0$, namely a baby bust or an increase in longevity (reduced adult mortality). The demographic shocks are the same as in the previous chapter, a 24% drop in the birth rate and a longevity shock that increases expected remaining lifetime from 65.5 years to 77.6 years. In both cases, the demographic shock renders the public pension system fiscally unsustainable in the long run, conform stylized fact (SF1). At time $t = 0$, however, the policy maker announces a policy reform—to be implemented at some later date, $T_R > 0$—which restores fiscal sustainability. The announcement is believed by individual agents as the policy maker has been credible in the past.

We study the effects of three types of policy reform. In Section 4.6.1 we assume that the policy maker engineers a once-off change in the lumpsum tax, $\bar{z}$, at time $t = T_R$ which restores government solvency. The policy response is the same for the two types of demographic change. In keeping the lumpsum tax time-invariant, both before and after the reform, the government engages in tax smoothing.

Figure 4.5. Individual steady-state wealth profiles with optimal retirement



In Section 4.6.2 we assume that the policy maker uses different instruments to address the two types of demographic change. For the baby bust, the policy response consists of a once-off increase in the labour income tax rate, t_L, occurring at time $t = T_R$. This is again a tax smoothing scenario as t_L is time-invariant both before and after the shock. For the longevity shock, the policy response consists of a permanent increase in the EEA, occurring at time $t = T_R$, which restores solvability without any further tax changes.

BOX 4.1

The government budget constraint

The key to solving the model is to determine the policy instrument that keeps the government finances sustainable. It is impossible to determine government debt in the infinite future, we therefore cut off the problem by postulating that the government debt must be stable at some finite future date, T. Government debt at date T is determined by the outstanding debt at time t and the budget deficits in between $G(\tau)$, which is determined by the policy instrument mix.

$$d(T) = d(t)e^{r\cdot(T-t)-N(T)+N(t)} + \int_t^T G(\tau)e^{r\cdot(T-\tau)-N(T)+N(\tau)}d\tau, \tag{4.52}$$

$$G(\tau) = \int_{-\infty}^{\tau} b(v)Q(v,\tau)e^{N(v)-N(\tau)-M(v,\tau)}dv \tag{4.53}$$

where $Q(v,\tau)$ are the net payments to the system at time τ of someone born at time v,

$$Q(v,\tau) = \begin{cases} t_L(\tau)\bar{w}(\tau-v) + \bar{z}(\tau) & \text{if working, } \tau - v \leq R(v) \\ \bar{z}(\tau) - \bar{p}(v,\tau,R(v)) & \text{if retired, } \tau - v > R(v) \end{cases} \tag{4.54}$$

Equation (4.52) shows the first problem, any approximation errors in the initial debt level $d(t)$ and the government deficits $G(\cdot)$ explode because of the multiplication with exploding exponential terms. To solve this we solve the problem backwards in time. Multiply both sides with $e^{-r\cdot(T-t)+N(T)-N(t)}$ and rearrange

$$d(t) = d(T)e^{-r\cdot(T-t)+N(T)-N(t)} - \int_t^T G(\tau)e^{r\cdot(t-\tau)-N(t)+N(\tau)}d\tau. \tag{4.55}$$

If we pick T large enough that debt is in the new steady state, we determine the stable debt level implied by the government deficit, that is fully determined by the policy instruments. From this new steady state level and the deficits, we can calculate the implied government debt at time $t = 0$. The whole problem translates into a simple one dimensional root finding problem: find for level of the policy instrument for which $d(0)$ as defined in (4.55) equals the predetermined debt level. Any root finding algorithm works.

There is however one more caveat. We must numerically evaluate both integrals in Equations (4.52) and (4.53). The problem is that the net payments $Q(v,\tau)$ are not continuous in v, $Q(\cdot)$ jumps at the retirement age. Before retirement, the agent has to pay the lumpsum tax and labour tax, after retirement, he only has to pay the lumpsum tax and if the agent passed the EEA, he also receives a retirement benefit. The usual solution to this problem is to split the integral, such that the integrand is continuous on each part.

Although we faced the same problem in the previous chapter, the problem here is slightly more complicated. In the previous chapter, the policy instrument was a neutral lumpsum tax that did not affect the schooling decision. The discontinuity in net payments in Chapter 3 always occurred at $\tau = v + s(v)$ (s in the previous chapter is the schooling period) and we knew where to split the integral. Here we do not know this beforehand. In principle, we could first determine the retirement ages for the specific policy instrument combination, then split the integral in parts, and continue, but this is too time consuming.

Much faster and ultimately simpler to implement is to write the integral in (4.55) in terms of the generational accounts (see Auerbach et al., 1994)

$$\begin{aligned} GA(v,t) &= e^{r\cdot(t-v)+M(t-v)} \int_t^\infty Q(v,\tau) e^{-r\cdot(\tau-v)-M(\tau-v)} d\tau \\ &= SSW(v,t) + e^{r\cdot(t-v)+M(t-v)} \int_{t-v}^\infty z(v+u) e^{-ru-M(u)} du \end{aligned}$$

The generational accounts are the present value of net payments to the system and are easy to calculate for stepwise policy changes in the system. For sake of readability write the integral in (4.55) as

$$D_A(t) = \int_t^T \int_{-\infty}^\tau b(v) Q(v,\tau) e^{r\cdot(t-\tau)+N(v)-N(t)-M(v,\tau)} dv d\tau$$

Now split the inner integral at $v = t$ and change the order of integration

$$\begin{aligned} D_A(t) = &\int_{-\infty}^t b(v) e^{N(v)-N(t)} \int_t^T Q(v,\tau) e^{r\cdot(t-\tau)-M(v,\tau)} d\tau dv \\ &+ \int_t^T b(v) e^{N(v)-N(t)} \int_v^T Q(v,\tau) e^{r\cdot(t-\tau)-M(v,\tau)} d\tau dv \end{aligned}$$

Some tedious, but otherwise straightforward math shows that the inner integrals can be written in terms of the generational accounts

$$\begin{aligned} D_A(t) = &\int_{-\infty}^t b(v) e^{N(v)-N(t)} \left[e^{-M(v,t)} GA(v,t) - e^{-r\cdot(T-t)-M(v,T)} GA(v,T) \right] dv \\ &+ \int_t^T b(v) e^{N(v)-N(t)} \left[e^{r\cdot(t-v)} GA(v,v) - e^{-r\cdot(T-t)-M(v,T)} GA(v,T) \right] dv \end{aligned}$$

The integrand in the first term has a discontinuity at $v = 0$ caused by the a possible demographic shock, so we should split that one. This leaves us with three separate integrals, all with a continuous integrand, that we can evaluate with any numerical integration method. With an efficient method to calculate $D_A(0)$, we can solve the whole model within a number of seconds.

4.6.1 Tax reform

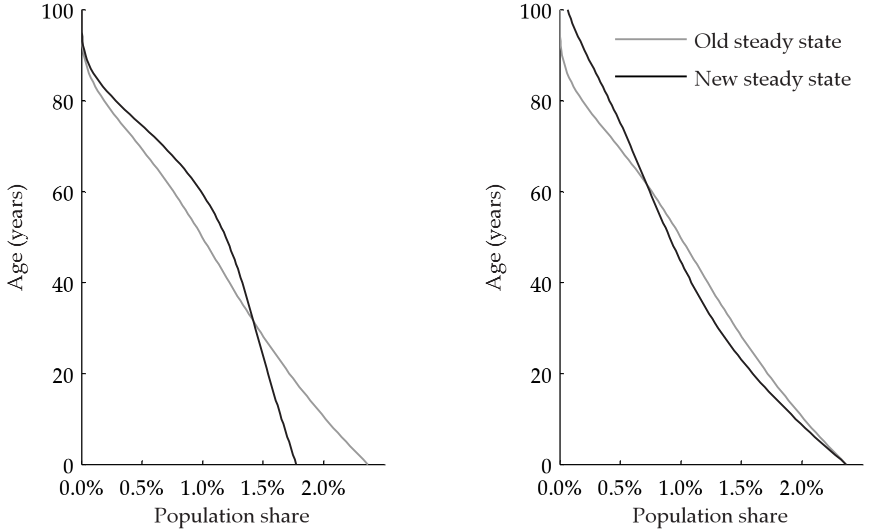
Throughout this subsection the announced policy reform consists of a once-off change in the lumpsum tax which makes government finances healthy again.

Baby bust The effects of a once-off decrease in the birth rate occurring at time $t = 0$ are visualized in Figures 4.6(a), 4.8 and 4.7(a). The baby bust causes a twenty-five percent decrease in the birth rate, from $b_0 = 0.0237$ to $b_1 = 0.0177$. It is assumed that policy reform is implemented twenty years after the baby bust, i.e. $T_R = 20$ in these figures. Since this reform has no effect on the kink in the lifetime income profile, individuals continue to retire at the EEA. Figure 4.6(a) depicts the change in the steady-state age composition of the population. The mass of the distribution is moved from younger to older ages. Figure 4.8(a) shows the demographic transition due the baby bust. There is an immediate drop in the population growth rate because the arrival rate of new agents has decreased permanently, i.e. $n(0) - \hat{n}_0 = b_1 - b_0$ (see the last paragraph in Box 3.1). Following the initial jump, $n(t)$ adjusts in a non-monotonic fashion to the new demographic equilibrium at $\hat{n}_1 = 0.43\%$.

Figure 4.8(b) illustrates the transition path for per capita employment in efficiency units (Recall that the paths for per capita output and physical capital are both proportional to $h(t)$—see Equations (4.20) and (4.21) above). Employment declines in a non-monotonic fashion, from $\hat{h}_0 = 9.0$ to $\hat{h}_1 = 8.5$. There is a gradual decline in $h(t)$ from $t = 0$ until $t = 60$ both because fewer workers enter the labour force than before the shock and because the larger pre-shock cohorts retire. At time $t = 60$, the path for $h(t)$ starts to rise again because the flow of retirees consists entirely of relatively small post-shock cohorts. Beyond $t = 60$, the path for employment converges in a cyclical fashion to the new steady state.

Figure 4.8(c) depicts the adjustment path for per capita consumption. At impact, consumption falls because all pre-shock generations anticipate the future lumpsum tax increase and cut their consumption level accordingly. During the first half cen-

Figure 4.6. Steady state population composition before and after a birth rate shock (a) and a longevity shock (b)



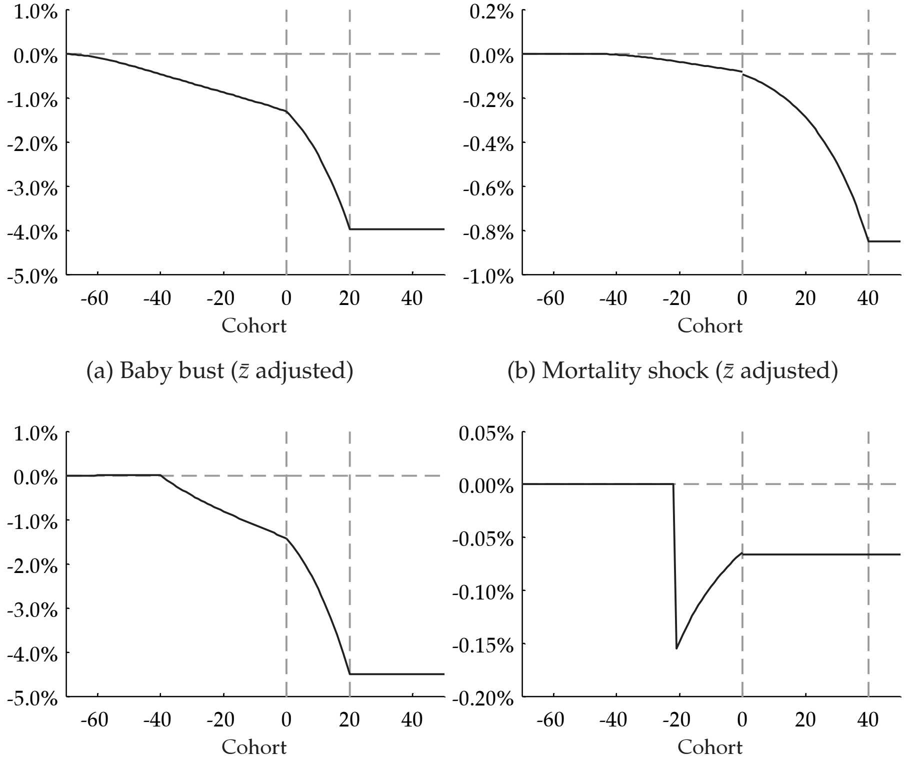
(a) Baby bust (b) Reduced adult mortality

Notes: Mortality process is a Gompertz-Makeham (see Table 2.1 for parameter values), birth rate is 2.36%. Baby bust is a 25% downward jump of the birth rate to 1.78%. Reduced adult mortality is a 50% decrease of μ_1 and 10% decrease of μ_2.

tury following the shock consumption rises due to a strong numerator effect caused by the reduction in the population growth rate. Consumption reaches a peak at the point where the weight of the relatively rich pre-shock cohorts starts to dwindle as a result of mortality. Consumption declines thereafter because post-shock generations have a lower consumption level due to the heavier lumpsum tax burden they are faced with during their lifetimes. The path of asset income, depicted in Figure 4.8(d) shows the strong savings response that occurs during the time period $0 < t < T_R$. Agents anticipate the higher taxes from T_R onward and save more than before the shock. At time T_R, the slope of the asset path is reduced because the tax increase is implemented. Eventually, the last of the relatively large pre-shock cohorts enter retirement and start to dissave so that aggregate assets fall somewhat. The long-run effect of the baby bust is an increase in per capita assets from $\hat{a}_0 = 100.5$ to $\hat{a}_1 = 107.9$.

Figure 4.8(e) illustrates the path of per capita government debt. The baby bust destabilizes the public pension system and leads to a gradual build up of gov-

Figure 4.7. Welfare effects of demographic shocks and the policy reforms



(a) Baby bust ($\bar{z}$ adjusted) (b) Mortality shock ($\bar{z}$ adjusted)

(c) Baby bust (t_L adjusted) (d) Mortality shock (EEA adjusted)

Notes: Mortality process is a Gompertz-Makeham (see Table 2.1 for parameter values), birth rate is 2.36%. Baby bust is a 25% downward jump of the birth rate to 1.78%. Mortality shock is a 50% decrease of μ_1 and 10% decrease of μ_2.

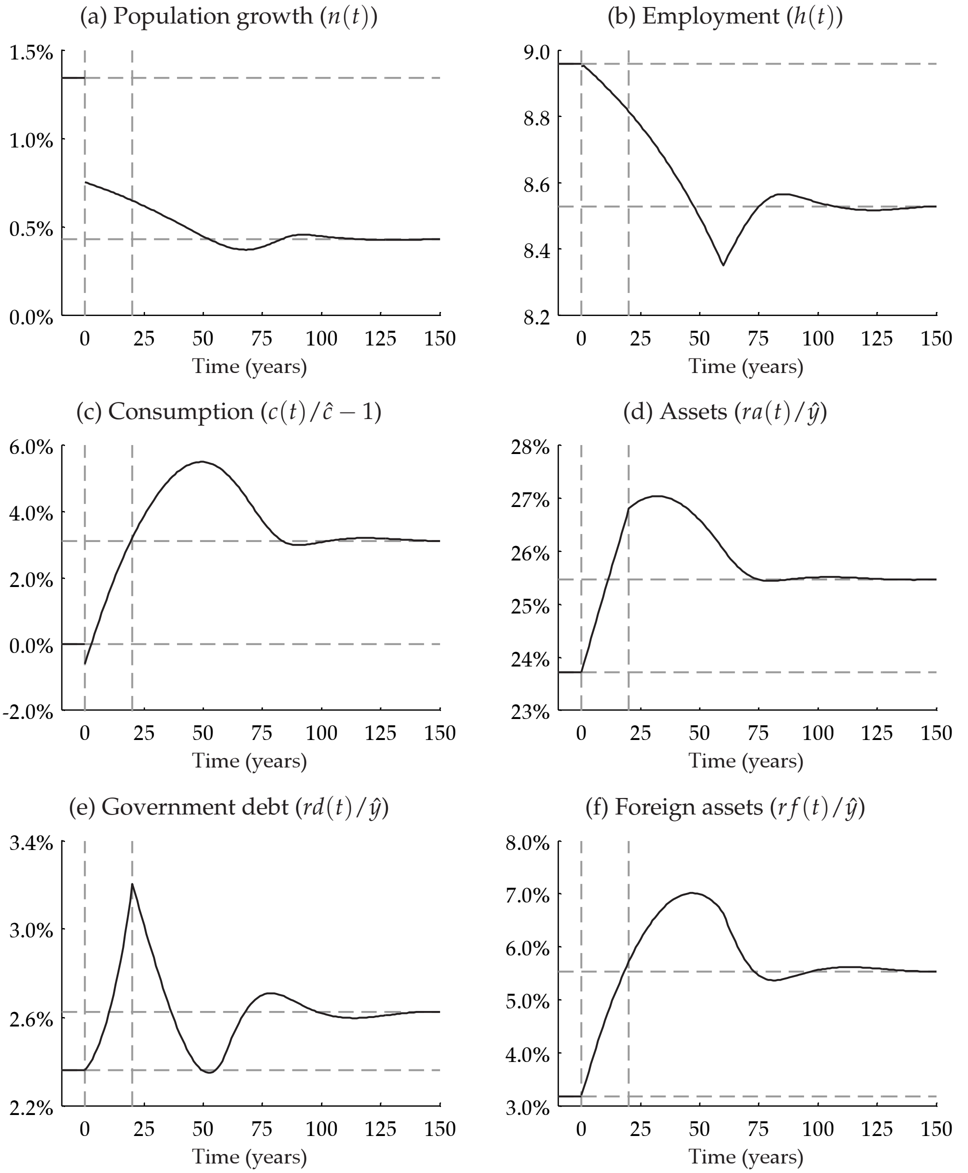
ernment debt in the pre-reform period, $0 < t < T_R$. At time T_R, the lumpsum tax is increased, solvency is restored, and the government can redeem some of its outstanding debt obligations. Interestingly, the post-reform transition path is non-monotonic because the relatively large pre-shock cohorts die and thus stop paying taxes. In the long run, the baby bust leads to an increase in per capita debt from $\hat{d}_0 = 10$ to $\hat{d}_1 = 11.1$.

Finally, in Figure 4.8(e) we plot the adjustment path for net foreign assets. Obviously, since $a(t) = k(t) + d(t) + f(t)$, the path for net foreign assets mirrors that of total assets, the capital stock, and government debt. During the first half century of adjustment, agent's strong savings response (panel (d)) coincides with the accumulation of net foreign assets. Note that at time T_R the government starts to redeem public debt, i.e. both $k(t)$ and $d(t)$ are falling immediately after T_R. Total assets are still rising, however, so it follows that foreign asset accumulation continues quite vigorously even after the tax reform has taken place. The long-run effect of the baby bust consists of an increase in net foreign assets from $\hat{f}_0 = 13.5$ to $\hat{f}_1 = 23.5$. For convenience we summarize the quantitative results of the baby bust in column (2) of Table 4.1.

In Figure 4.7(a) we illustrate the change in welfare experienced by the different generations. To facilitate the interpretation of the effects, we present equivalent-variation (EV) measures expressed in terms of initial wealth level. For pre-shock generations ($v \leq 0$) we compute the change in lifetime utility from the perspective of the shock period ($t = 0$), i.e. we plot the EV-value of $d\bar{\Lambda}(v, 0)$ for $v \leq 0$. In contrast, for post-shock generations ($v > 0$), we compute the welfare change from the perspective of their birth date, i.e. we plot the EV-value of $d\bar{\Lambda}(v, v\$v > 0$ in Figure 4.7. The welfare effects of the baby bust are straightforward. All generations lose out as a result of the lumpsum tax increase. For old pre-shock generations the welfare effect is small. These generations have a very short time horizon and for them the tax increase that will occur only at time $T_R = 20$ hardly poses any burden at all. The younger the pre-shock generations are, the heavier the burden of the anticipated tax increase become. Similarly, for post-shock generations the welfare loss becomes larger the closer they are born to the time at which the tax increase takes place. Worst off are those generations born at or after T_R: the welfare loss is about 4 percent of initial wealth for them.

Increased longevity The effect of an embodied longevity shock occurring at time $t = 0$ are visualized in Figures 4.6(b), 4.9 and 4.7(b) (welfare effects). The effect on the mortality rate itself is illustrated Figure 3.2(a) on page 67. The μ_1-parameter of the G-M process is reduced by 50% and the μ_2 parameter by 10%, leading to an increase of the expected lifetime at birth from $R_0(0) = 65.45$ to $R_1(0) = 77.57$ years. Figure 4.6(b) depicts the long-run effect on the age composition of the population. The population pyramid is squeezed for ages up to about 62, but is thickened for higher ages. Figure 4.9(a) shows that the demographic transition, following an embodied longevity shock, is rather slow. Indeed, even 30 years after the shock the population growth rate is virtually at its initial steady-state level. For that reason we assume that the policy reform is implemented 40 years after the longevity shock, i.e. $T_R = 40$ in Figures 4.9 and 4.7(b). Just as for the baby bust, the tax reform has

Figure 4.8. Aggregate effect of a baby bust ($\bar{z}$ balances the budget)



Notes: Mortality process is a Gompertz-Makeham (see Table 2.1 for parameter values), birth rate is 2.36%. Baby bust is a 25% downward jump of the birth rate to 1.78%. Policy change is announced at $t = 0$, implementation is at $t = 20$.

Table 4.1. Initial steady state and long-run effects of demographic shocks

	(1) Initial st. st.	(2) Baby bust $\bar{z}$ adjusts	(3) Mortality $\bar{z}$ adjusts	(4) Baby bust t_L adjusts	(5) Mortality R_E adjusts
μ_1	5.61×10^{-5}	5.61×10^{-5}	2.80×10^{-5}	5.61×10^{-5}	2.80×10^{-5}
μ_2	0.09616	0.09616	**0.0867**	0.09616	**0.0867**
$\Delta(0,0)$	65.5	65.5	81.6	65.5	81.6
b	0.02365	**0.01774**	0.02365	**0.01774**	0.02365
R^*	60	60	60	60	62.13
$\hat{n}$	0.0134	0.0043	0.0163	0.0043	0.0163
T_R		20	40	20	40
t_L	0.100	0.100	0.100	*0.142*	0.100
R_E	60	60	60	60	*61.7*
$\bar{z}$	−0.166	*0.343*	*−0.0573*	−0.166	−0.166
$\hat{d}$	10.0	11.1	−1.4	10.9	−1.2
$\hat{a}$	100.5	107.9	134.9	105.8	134.6
$\widehat{li}$	184.4	165.4	188.1	166.0	191.6
$\hat{h}$	9.0	8.5	8.5	8.5	8.6
$\hat{y}$	21.2	20.2	20.0	20.2	20.3
$\hat{c}$	16.0	16.5	16.4	16.4	16.6
$\hat{\imath}$	5.7	4.7	5.5	4.7	5.6
$\hat{f}$	13.5	23.5	60.75	21.5	59.7
$\hat{k}$	77.0	73.3	72.7	73.3	73.7
$\hat{c}/\hat{y}$	0.756	0.820	0.820	0.815	0.812
$\hat{\imath}/\hat{y}$	0.267	0.278	0.280	0.234	0.278

Notes: Exogenous shocks are indicated by bold text, policy instruments by italic text

no effect on the retirement choice, i.e. pre-shock and post-shock agents all retire at the EEA. It follows that post-shock agents expect a much longer retirement period than pre-shock agents do.

The quantitative long-run effects of the longevity shock have been reported in column (3) of Table 4.1. The key features of the transition paths in Figure 4.9 are as follows. In Figure 4.9(b), employment is virtually constant until the tax reform takes place and rises slightly thereafter. People live longer so the inflow into the labour market exceeds the outflow. For $R_E < t < 90$ there is a sharp decrease in employment because the post-shock cohorts start to retire. Because their longevity is higher than for the pre-shock cohorts, the retiring cohorts are relatively large and the outflow from the labour market is huge. In the new steady state, employment is permanently lower because the weight of retirees is larger than before. People live longer but they do not work for a longer period of time. As a result, per capita employment falls.

Figure 4.9(c) depicts the adjustment path for consumption. For $t < R_E$, per capita consumption falls because post-shock newborns consume less than pre-shock newborns, i.e. the negative horizon effect dominates the positive lifetime-income effect. Consumption rises again for $R_E < t < 90$. Pre-shock generations have all passed away but post-shock generations—who live longer lives—have a relatively high consumption level later on in life. In the new steady state per capita consumption is higher as a result. Figure 4.9(d) shows that per capita assets rise during the transition. As is shown in Figure 4.5(a), the individual age profile for assets is increasing up to age $u = R_E$. The longevity shock implies that larger population fractions ultimately reach the EEA and beyond. As result, per capita assets increase.

Figure 4.9(e) shows that public debt is virtually constant for $0 < t < T_R$. This is because the longevity shock takes a long time before it starts to seriously affect the government finances. Were the government to do nothing, debt would ultimately explode, conform stylized fact (SF1). However, our fiscally responsible government slightly increases thelumpsuml tax from T_R onward, thus making room for higher future outlays on pension payments. Figure 4.9(f) shows that net foreign assets rise during the transition.

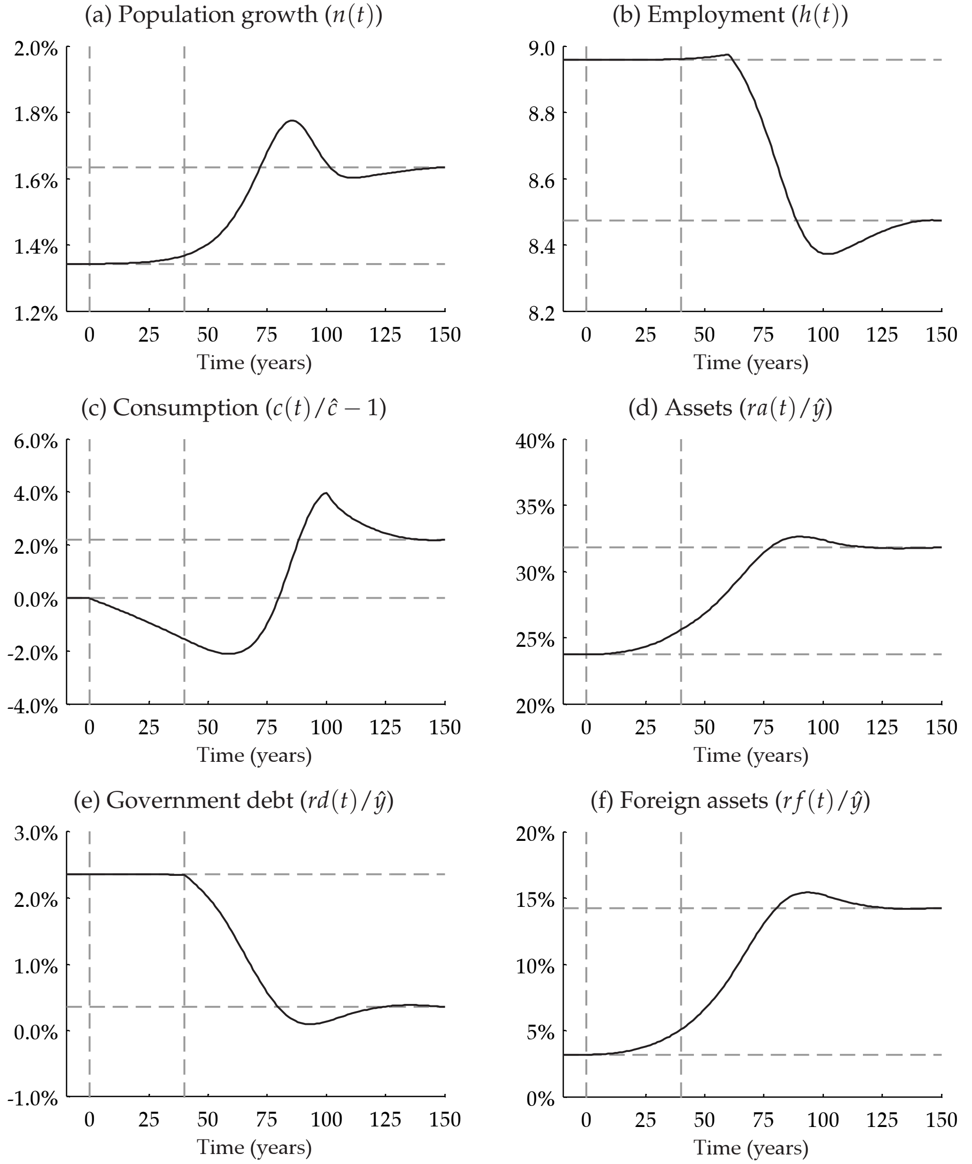
The welfare effects of the longevity shock are visualized in Figure 4.7(b). Just as for the baby bust, (a) all generations lose out as a result of the lumpsum tax increase and (b) welfare losses are increasing in the generations index, v. Because the tax increase is much smaller than for the baby bust scenario, the welfare losses are smaller for all generations.

4.6.2 Pension reform

In this subsection the announced pension reform is assumed to be specific to the type of demographic shock hitting the economy. Indeed, we assume that t_L is increased following a baby bust, whereas the EEA is increased in reaction to increased longevity.

Baby bust The quantitative long-run effects of the baby bust have been reported in column (4) of Table 4.1. A crucial feature of the solution is that the increase in the

Figure 4.9. Aggregate effect of reduced adult mortality ($\bar{z}$ balances the budget)



Notes: Mortality process is a Gompertz-Makeham (see Table 2.1 for parameter values), birth rate is 2.36%. Reduced adult mortality is a 50% decrease of μ_1 and 10% decrease of μ_2. Policy change is announced at $t = 0$, implementation is at $t = 40$.

labour income tax is not sufficiently large to induce individuals to retire at an age beyond the EEA. Indeed, both pre-shock and post-shock agents continue to retire at the EEA, and as a result the labour income tax operates just like a lumpsum tax. The only difference between the two scenarios is that retirees do not have to pay the labour income tax, whereas they do pay the lumpsum tax. For this reason, the welfare profiles are slightly different for the two scenarios. Comparing Figures 4.7(a) and (c) we find that the welfare loss is zero for all pre-shock cohorts older than $R_E - T_R$ in the labour tax scenario. These generations will be retired from the labour force by the time the tax reform is implemented.

Increased longevity In column (5) of Table 4.1, and Figures 4.10 and 4.7(d) we characterize the effects of the longevity shock. The EEA is increased at time T_R in such a way that the government maintains solvency. This implies that the EEA rises from $R_{E0} = 60$ to $R_{E1} = 61.7$. For $0 < t < T_R$ agents continue to retire at age R_{E0} but thereafter agents retire almost a year later in life, at R_{E1}. Comparing Figures 4.9 and 4.10 we find that the main difference between thelumpsuml and EEA scenarios is found in the adjustment path for employment (panel (b) in these figures). In Figure 4.10(b) there is a sharp increase in employment at time T_R because nobody retires at that time. Some pre-shock generations delay their retirement by 0.9 years. Since new cohorts continue to enter the labour market, employment rises sharply. The remainder of the adjustment path is similar as for the lumpsum tax case: there is a sharp decline at $t = R_{E1}$ as the first of the post-shock cohorts retire.

Comparing Figures 4.7(b) and (d) we find that the welfare effects are rather different for the two scenarios. Five groups of cohorts can be identified in Figure 4.7(d). Group 1 consists of cohorts whose generations index satisfies $v < T_R - R_{E1}$. These cohorts have either already retired at the time of the shock ($t = 0$) or will be just old enough at the time of the policy reform (T_R) to retire at that time and receive benefits immediately. This means that at time $t = T_R$ such agents must be at least R_{E1} years of age. For these generations there is no welfare loss as a result of the anticipated EEA perform. They continue to retire at age R_{E0}.

Groups 2 and 3 are cohorts for which $T_R - R_{E1} < v < T_R - R_{E0}$. Agents in this group face a choice. Option 1: they can either retire early at age R_{E0} (the old EEA) and be without income for a brief period of time because they retire too early under the new regime. Option 2: they can adjust their planned retirement age from R_{E0} to R_{E1}. It turns out that the oldest generations will choose option 1 whereas the youngest generations will choose option 2, with the pivotal generation index being at $v^* = -20.5$. Agents in both groups experience a welfare loss as a result of the reform. Interestingly, the welfare loss is increasing in v for $T_R - R_{E1} < v < v^*$ but decreasing in v for $v^* < v < T_R - R_{E0}$.

Group 4 consists of cohorts for which $T_R - R_{E0} < v < 0$. People in this group did not have any real choice. At time T_R they are too young to retire with benefits under the under the old regime and thus have to retire at age R_{E1}. Their delayed pension is compensated partially by higher a level of lifetime income because they have a longer working life. The welfare loss for agents in this group is decreasing in v.

Finally, group 5 consists of post-shock cohorts for which $v > 0$. Agents in this group are all affected equally. They all choose the retirement age R_{E1} and they all face the same initial conditions in life.

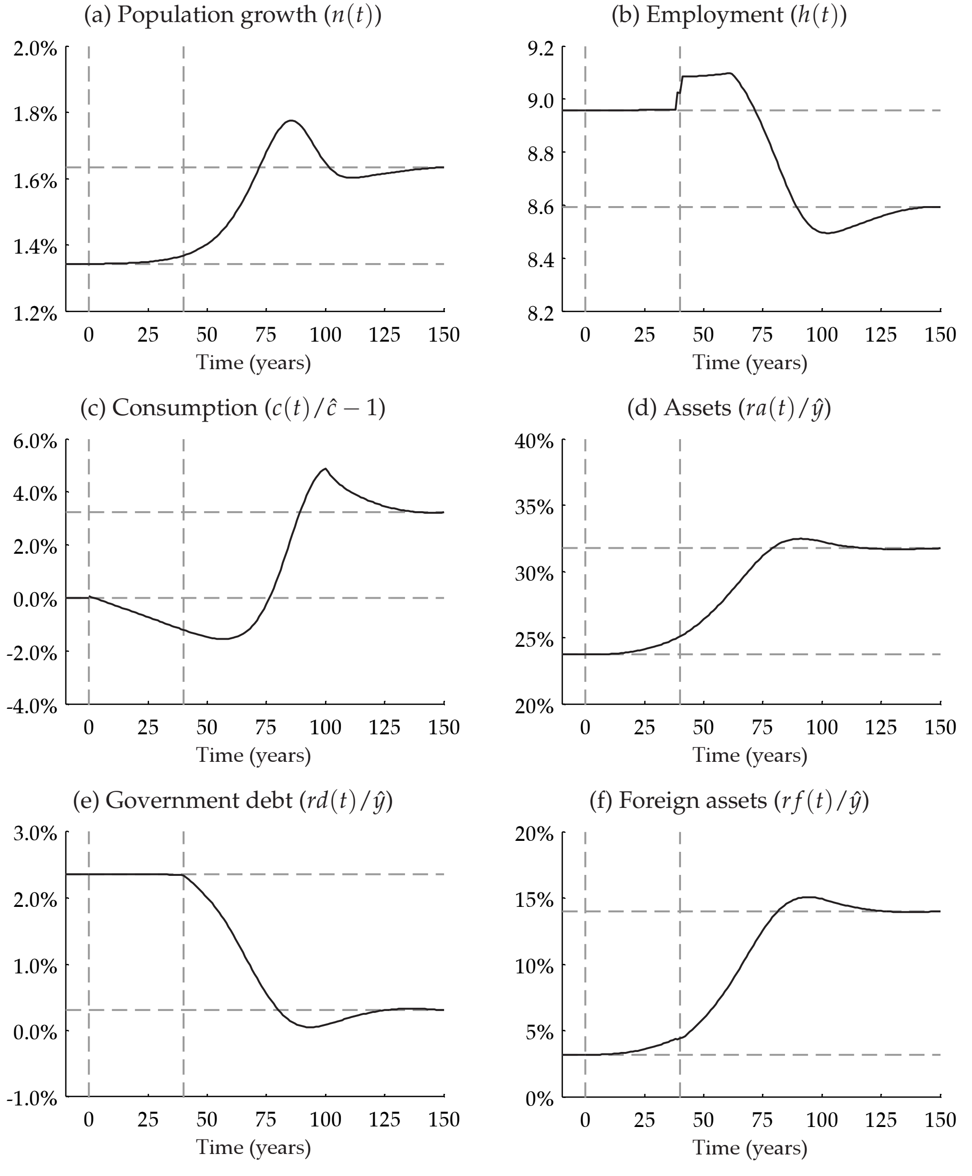
4.6.3 Discussion

The key findings of this section are as follows. First, although both a baby bust and a longevity boost have an adverse effect on the government's budget, there is a striking difference in the speed with which such effects become apparent. Indeed, for the baby bust the adverse effects show up immediately. Government debt starts to rise sharply immediately after the shock because the flow of tax payers dwindles. In contrast, for the longevity shock it takes a very long time before any effect on the government's balances can be observed.

Second, even though we simulated very large demographic changes, wealth effects are simply too weak to get agents to move out of the kink and to postpone retirement beyond the EEA. For a realistic calibration, the implicit tax rates are rather high, ranging from 11.1% until age 60, jumping to 62.8% at that age, and subsequently rising to 67% at age 70. The kink in the lifetime income profile acts as a kind of early retirement trap. Changes in the lumpsum tax or the labour income tax are insufficiently powerful instruments to get agents out of the trap. The welfare costs of the tax increase are non-trivial. Indeed, our simulations show that post-shock agents experience a welfare loss that is the equivalent of more than 4% of initial wealth!

Third, an increase in the EEA itself constitutes a rather good policy measure. By increasing the EEA, the kink in the lifetime income profile is shifted to right, and agents retire later on in life despite the existence of high implicit tax rates. We show that the welfare effects of the EEA increase are tiny: post-shock agents experience a

Figure 4.10. Aggregate effect of reduced adult mortality (EEA balances the budget)



Notes: Mortality process is a Gompertz-Makeham (see Table 2.1 for parameter values), birth rate is 2.36%. Reduced adult mortality is a 50% decrease of μ_1 and 10% decrease of μ_2. Policy change is announced at $t = 0$, implementation is at $t = 40$.

welfare loss that is the equivalent of less than 0.1% of initial wealth as a result of the EEA increase! Agents not only work longer but they also get a higher consumption level as a result.

4.7 Conclusions

In this chapter we have studied the microeconomic and macroeconomic effects of ageing in the context of a small open economy populated by disconnected generations of finitely-lived agents facing age-dependent mortality and constant factor prices. From a policy perspective, our main finding is as follows. Most actual pension systems induce a kink in the lifetime income function which acts as an early retirement trap. Fiscal changes are not potent enough to get individuals out of the trap. Increasing the early entitlement age appears to be a low cost policy measure to counteract the adverse effects of the various demographic shocks.

Our analysis is subject to a number of potentially important limitations, some of which we wish to address in the near future. First, in this chapter the age profile of labour efficiency is exogenous, i.e. there is no endogenous human capital accumulation decision. A possible solution is to combine the models of this chapter Chapter 3. This is quite feasible, but it will blur the two results and this makes the interpretation of effects of ageing on various macroeconomic variables harder.

Second, we only consider once-off changes in the demographic processes. In reality, demographic changes occur only gradually over time. The main complication lies in the calculation of the population dynamics, i.e. the population growth rate. The macroeconomic block of the model (individual optimisation, production, saving) remains the same since the interest rate is constant in the small open economy. The main difference with the current stepwise shock is that the transition periods are longer and the costs are spread out over more generations.

Third, we have focused attention of mortality and have ignored the equally important issue of morbidity. One of the main functions of a retirement system is to support people that are not capable of working due to old age related diseases. Identification problems arise if health is not perfectly observable.The risk exists that either the retirement system becomes too expensive because too many people use it, while they are perfectly capable of working, or that people that cannot work are kept out of the system.

4.A Years-of-retirement transformation

Burbidge and Robb (1980, p. 424) use a linear space transformation. Instead of using the retirement age directly, they reformulate their model in terms of years or retirement, $T - R$, where T is the fixed planning horizon. Two things are worth noting. First, their linear transformation does not solve the problem of non-convex indifference curves—see below. Second, in our model, T is a stochastic variable and the expected planning horizon at birth is given by $\Delta(0,0)$, where $\Delta(u,\lambda)$ is defined in (2.12) in Chapter 2. Transforming our model in terms of expected years of retirement (from the perspective of birth), $\Delta(0,0) - R$ suffers from the same defects.

The basic point is that a linear transformation does not guarantee well-behaved indifference curves. This can easily be demonstrated in the context of our model. The steady-state concentrated utility function is given by (4.23) in the text. We write it as $\bar{\Lambda}(u, \bar{l}i, R)$ but hold u constant. To determine the slope and curvature of the indifference curves, we need the following building blocks:

$$
\begin{aligned}
\bar{\Lambda}_{\bar{l}i} &\equiv \frac{\partial \bar{\Lambda}}{\partial \bar{l}i} = \left[\frac{\bar{a}(u) + \bar{l}i}{\Delta(u, r^*)} \right]^{-1/\sigma} > 0, \\
\bar{\Lambda}_R &\equiv \frac{\partial \bar{\Lambda}}{\partial R} = -D(R) e^{\theta \cdot (u-R) + M(u) - M(R)} < 0, \\
\bar{\Lambda}_{\bar{l}i, \bar{l}i} &\equiv \frac{\partial^2 \bar{\Lambda}}{\partial \bar{l}i^2} = -\frac{1}{\sigma \cdot \left[\bar{a}(u) + \bar{l}i \right]} \bar{\Lambda}_{li} < 0, \\
\bar{\Lambda}_{\bar{l}i, R} &\equiv \frac{\partial^2 \bar{\Lambda}}{\partial \bar{l}i \partial R} = 0, \\
\bar{\Lambda}_{R,R} &\equiv \frac{\partial^2 \bar{\Lambda}}{\partial R^2} = \left[\frac{D'(R)}{D(R)} - \theta - m(R) \right] \bar{\Lambda}_R \gtrless 0.
\end{aligned}
$$

The slope of the indifference curve in $(R, \bar{l}i)$-space is

$$
\left. \frac{d\bar{l}i}{dR} \right|_{\bar{\Lambda}_0} \equiv -\frac{\bar{\Lambda}_R}{\bar{\Lambda}_{\bar{l}i}} = D(R) e^{\theta \cdot (u-R) + M(u) - M(R)} \left[\frac{\bar{a}(u) + \bar{l}i}{\Delta(u, r^*)} \right]^{1/\sigma} > 0.
$$

Hence, the indifference curves are always upward sloping.

To compute the curvature of the indifference curve in $(R, \bar{l}i)$-space we must take into account the dependency of $\bar{l}i$ on R along a given indifference curve. After some manipulation, we find:

$$
\left. \frac{d^2\bar{l}i}{dR^2} \right|_{\bar{\Lambda}_0} \equiv -\frac{d\left(\bar{\Lambda}_R / \bar{\Lambda}_{\bar{l}i} \right)}{dR} = -\frac{1}{\bar{\Lambda}_{\bar{l}i}^2} \left[\bar{\Lambda}_{\bar{l}i} \bar{\Lambda}_{R,R} - \bar{\Lambda}_R \bar{\Lambda}_{\bar{l}i, \bar{l}i} \left. \frac{d\bar{l}i}{dR} \right|_{\bar{\Lambda}_0} \right]
$$

$$= \left.\frac{d\bar{l}i}{dR}\right|_{\bar{\Lambda}_0} \left[\frac{1}{\sigma \cdot [\bar{a}(u) + \bar{l}i]} \left.\frac{d\bar{l}i}{dR}\right|_{\bar{\Lambda}_0} + \frac{D'(R)}{D(R)} - \theta - m(R) \right] \gtrless 0.$$

Equation (4.7) is a rather intractable expression, and the sign is ambiguous in general. However, numerical simulations reveal that for realistic parameter values the indifference curves are either concave in R or S-shaped (i.e., convex for small R and concave for large R). Similar results can be derived for the specification used by Burbidge and Robb (1980, p. 425), so their assumption that the indifference curves are convex in the relevant region is problematic.

The key point to note is that a linear transformation of the retirement age is unhelpful. Hence, transforming our model in terms of expected years of retirement, $\Delta(0,0) - R$, is not useful either.

4.B Data

In Section 4.4 of we use data on replacement rates and implicit tax rates that were gathered from the various chapters in Gruber and Wise (1999). For convenience we present an overview of these data here. The figures refer to data taken from Tables 1.4, 2.2, 3.5, 5.4, 6.1, 7.1, 8.7, 9.2, 10.4, and 11.1. Note that we report the retirement age in the first column of Tables 4.B.1 and 4.B.2. In contrast, Gruber and Wise (1999) report the last age of active employment. Our entries for age 60 are thus equivalent to their entries for age 59.

Table 4.B.1. Replacement rates in nine OECD countries

Age	Belgium	Canada	France	Italy	Japan	Neth's	Spain	UK	US
55				0.726					
56				0.744					
57				0.761					
58				0.780					
59	0.749	0.182	0.920	0.798	0.552	0.910	0.590		
60	0.771	0.202	0.910	0.799	0.800	0.906	0.661		
61	0.794	0.217	0.920	0.804	0.799	0.900	0.730		0.403
62	0.817	0.245	0.910	0.805	0.802	0.902	0.816		0.440
63	0.839	0.270	0.920	0.805	0.801	0.892	0.895		0.476
64	0.863	0.508	0.920	0.809	0.438	0.909	0.996		0.703
65	0.874	0.518	0.930	0.809	0.549	0.909	0.998	0.464	0.749
66	0.882	0.527	0.940	0.809	0.547	0.909	0.996	0.491	0.798
67	0.890	0.850	0.950	0.809	0.716	0.909	0.988	0.519	0.845
68	0.898	0.881	0.960	0.809	0.608	0.909	0.981	0.549	0.872
69	0.905	0.914	0.960	0.809	0.607	0.909	0.973	0.581	0.898

Source: Gruber and Wise (1999)

Table 4.B.2. Implicit tax rates in nine OECD countries

Age	Belgium	Canada	France	Italy	Japan	Neth's	Spain	UK	US
55	-0.129	-0.049	-0.910	0.245	-0.195	0.687	0.216	0.020	-0.022
56	-0.134	0.003	-0.970	0.308	-0.202	0.650	0.108	0.010	0.046
57	-0.145	0.037	-0.460	0.338	-0.106	0.612	0.153	0.030	0.060
58	-0.148	0.038	0.040	0.372	-0.112	0.578	0.362	0.030	0.069
59	-0.157	0.040	0.050	0.401	-0.138	-3.777	0.286	0.030	0.072
60	0.496	0.063	0.670	0.697	0.338	1.410	-0.149	0.030	0.071
61	0.497	0.066	0.600	0.711	0.340	1.384	-0.120	0.020	0.064
62	0.491	0.064	0.630	0.718	0.342	1.339	-0.112	0.020	-0.028
63	0.489	0.071	0.560	0.729	0.340	1.280	0.046	0.020	-0.005
64	0.473	0.169	0.560	0.746	0.204	1.222	0.160	0.020	0.031
65	0.529	0.285	0.520	0.756	0.000	0.357	0.757	0.010	0.188
66	0.519	0.323	0.480	0.772	0.000	0.347	0.767	0.020	0.225
67	0.476	0.259	0.460	0.787	0.000	0.337	0.777	0.030	0.269
68	0.463	0.203	0.450	0.803	0.000	0.327	0.741	0.050	0.439
69	0.440	0.229	0.430	0.818	0.000	0.315	0.705	0.070	0.455

Source: Gruber and Wise (1999)

Part II

Public Capital and Economic Growth

Chapter 5

Public Capital and economic growth: A Survey

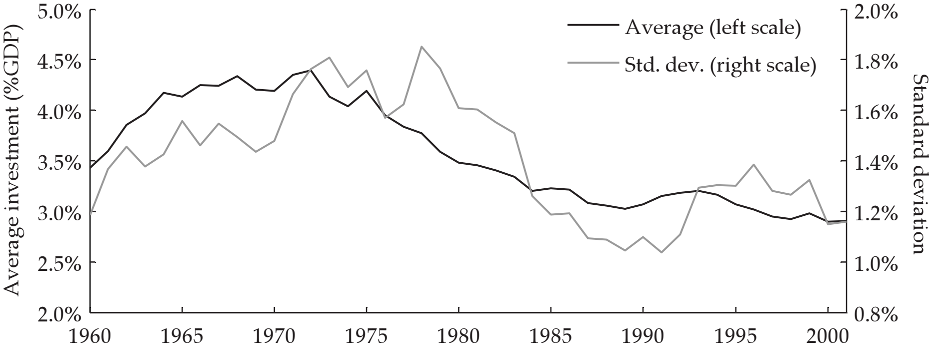
5.1 Introduction

Public capital, and especially infrastructure, is central to the activities of households and firms. According to the World Bank (1994), public capital represents the 'wheels' – if not the engine – of economic activity. Input-output tables show, for example, that telecommunications, electricity, and water are used in the production process of nearly every sector, while transport is an input for every commodity. However, the World Bank (1994, p. 19) also concludes that 'infrastructure investment is not sufficient on its own to generate sustained increases in economic growth'.

In recent years, a substantial research effort focused on estimating the contribution of public capital to the productivity of private factors of production and to economic growth. This research was motivated by two factors (Aschauer, 2000). First, for many years the ratio of public capital investment to gross domestic product (GDP) declined in the OECD area. Figure 5.1 shows average government investment spending as a percentage of GDP for 22 OECD countries over the period 1963–2001 (left-hand side scale) and its standard deviation (right-hand side scale). The data relate to consolidated general government and are based on the Standardised National Accounts compiled and published by the OECD. Figure 5.1 shows

This chapter is based on joint work with Jakob de Haan, 'Public capital and economic growth: a critical survey', EIB Papers, Vol. 10 (2005), No. 1, pp. 40–73.

Figure 5.1. Government investment in 22 OECD countries, 1961–2001, average (%GDP) and standard deviation



Source: Kamps (2006)

that public capital spending as a share of GDP declined between 1971 and 1990 and slightly recovered afterwards.[1] Another conclusion that can be drawn from Figure 5.1 is that government investment spending varies considerably across countries. As Table 5.1 in the Annex shows, in 2000–01, government capital spending ranged between 1.6 percent of GDP in the United Kingdom and 6.9 percent in Japan.

Second, various authors claim that the decline in public non-military capital spending in the United States contributed to the productivity slowdown of the 1970s and 1980s. The early empirical work in this area, conducted largely at the national level, reported a significant and large impact of public capital on productivity. For instance, using a production-function approach for the US between 1949 and 1985, Aschauer (1989) found that a 10-percent rise in the public capital stock would raise multifactor productivity by almost 4 percent. Other studies using aggregate data also reported large effects of public capital spending. At a time when the slowdown in productivity growth was a widespread concern, these findings suggested that a decline in the rate of public capital accumulation was 'a potential new culprit' (Munnell, 1990b, p. 3).

However, several economists questioned Aschauer's estimates on the grounds that they were implausibly high (see, for instance, Gramlich, 1994). Furthermore, the early studies were fraught with methodological and econometric difficulties.

[1] According to Oxley and Martin (1991, p.161) the decline of government investment reflected 'the political reality that it is easier to cut-back or post-pone investment spending than it is to cut current expenditures.' de Haan et al. (1996) report evidence that during large scale fiscal contractions government capital spending is indeed reduced more than other categories of government spending

Issues ranking high on the list of potential problems include reverse causation from productivity to public capital and a spurious correlation due to non-stationarity of the data.

Perhaps the most important concern is the direction of causality between public capital and aggregate output: while public capital may affect productivity and output, economic growth can also shape the demand and supply of public capital services, which is likely to cause an upward bias in the estimated returns to public capital if endogeneity is not addressed.[2] The recent literature on the economic growth effects of public capital suggests various ways of solving this problem.

Some of the earlier studies have also been criticised for not taking the stationarity of the data properly into account (see, for instance, Sturm and de Haan, 1995). Unit root tests often suggest that output and public capital contain a unit root. However, it is well known that unit root tests have low power to discriminate between unit root and near unit root processes. This problem is especially pronounced for small samples. One way to alleviate the small-sample problem that has become popular in recent research is to make use of the cross-sectional dimension of the data and to apply panel data techniques.

In some of the earlier studies unit roots in GDP and capital stock were removed by taking first differences. But this may ignore evidence of a long-run relationship in the data if the series are cointegrated (Munnell, 1992). Indeed, various recent studies report evidence for such a cointegrating relationship between public capital (or infrastructure) and output. By exploiting this cointegrating relationship, these studies estimate the long-run effect of public capital (or infrastructure) on GDP per capita. However, the existence of a cointegrating relationship in itself does not necessarily imply that causality runs from infrastructure to long-run growth (Canning and Pedroni, 1999).

In their survey of the earlier literature, Sturm et al. (1998) show that the literature contained a relatively wide range of estimates, with a marginal product of public capital that is much higher than that of private capital (e.g., Aschauer, 1989), roughly equal to that of private capital (e.g., Munnell, 1990a), well below that of private capital (e.g., Eberts, 1986) and, in some cases, even negative (e.g., Hulten and Schwab, 1991). The wide range of estimates makes the results of these older

[2] The problem not only occurs in studies like that of Aschauer (1989), but also in studies based on panel data, like Munnell (1990a), who found positive elasticities of output to public capital using panel data at the US state level. According to Holtz-Eakin (1994, p. 13), '[b]ecause more prosperous states are likely to spend more on public capital, there will be a positive correlation between the state-specific effects and public sector capital. This should not be confused, however, with the notion that greater public capital leads a state to be more productive'.

studies almost useless from a policy perspective.

However, more recent studies generally suggest that public capital may, under specific circumstances, raise income per capita. The purpose of this chapter is to review this literature, thereby providing an update of the survey of Sturm et al. (1998). We focus on two important questions. First, does an increase in public capital spur economic growth? Second, to what extent do conclusions on the effect of more infrastructure change once it is taken into account that infrastructure construction diverts resources from other uses?

The remainder of the chapter is organised as follows. Before we start reviewing the literature in some detail, Section 5.2 zooms in on our central questions and some other general considerations. Section 5.3 reviews studies belonging to the production-function approach in which the public capital stock is considered as an additional input factor in a production function. The next sections review three other approaches that have been applied to assess the impact of public capital on economic growth: the cost-function approach (Section 5.4), vector autoregressions (Section 5.5), and cross-country models (Section 5.6). In Section 5.7, we discuss the issue of the optimal capital stock. Section 8 offers some concluding comments.

5.2 Key questions concerning the link between public capital and economic growth

5.2.1 What do we want to know?

Empirical research on the relationship between public capital and growth should provide answers to two important questions. First, does an increase in the public capital stock foster economic growth?[3] Second, the policy relevant question for infrastructure investment is not what is the effect of extra infrastructure, holding everything else constant, but what is the net effect of more infrastructure given that infrastructure construction diverts resources from other uses (Canning and Pedroni, 1999). In other words, is the existing stock of capital optimal?

Of course, the possibility of a long-run impact of infrastructure on income very much depends on whether the data are generated by a neoclassical exogenous growth model or an endogenous growth model. In the exogenous growth model,

[3] The impact of public investment on economic growth is also relevant from a regional policy perspective. Governments can influence the rate at which regions accumulate various productive factors, particularly infrastructure. If these factors affect productivity and the location of mobile private production factors, there will be room for supply-side policies to influence the regional dispersion of income (de la Fuente and Vives, 1995).

in which technical progress drives long-run growth, shocks to the infrastructure stock can only have transitory effects. In an endogenous growth model, shocks to infrastructure can raise the steady-state income per capita. For instance, in the endogenous growth model with constant returns to aggregate capital of Canning and Pedroni (1999), positive shocks to infrastructure stocks raise long-run income per capita when the economy is below the efficient infrastructure level.

Apart from the growth model selected, the existing capital stock matters for the marginal productivity of public capital. This is clear from a network perspective: a new network may yield a one-time increase in productivity rather than a continuing path to prosperity (Fernald, 1999). Furthermore, according to the law of diminishing returns, an increment to the public capital stock would have a small (large) output effect if the capital stock in the previous period was large (small). There is evidence that countries with a small public capital stock have the highest marginal productivity of public capital (Demetriades and Mamuneas, 2000). Many empirical studies focus on the average, as opposed to the marginal, productivity of public capital and can therefore not be used to assess whether the existing capital stock is optimal. Kamps (2005) adopts the methodology proposed by Aschauer (2000) in order to investigate whether there is a lack of public capital in European Union countries.

In addressing the second question, it comes natural to take a government budget perspective and to look at how additional public investment is financed. The effect of public investment on growth is likely to depend on how the increased spending is financed. Increases in taxes are widely considered to reduce the rate of economic growth. An increase in public capital stimulates economic growth only if the productivity impact of public capital exceeds the adverse impact of higher taxes. If cutting other government spending finances an increase in capital spending, there is still no guarantee that growth will be enhanced. Hulten (1996) argues, for instance, that new infrastructure construction may have a perverse effect if it draws scarce government resources away from maintenance and operation of the existing capital stock.

Sections 5.3 to 5.6 will focus on the growth-enhancing effects of public capital spending while Section 5.7 will turn to the issue of the optimality of the public capital stock. But first we review why public capital may affect growth and how the stock of public capital can be measured.

5.2.2 Why does public capital matter for economic growth?

How does public capital affect economic growth? This issue has received only scant attention in the literature on the relationship between public capital spending and economic growth. As Holtz-Eakin and Lovely (1996, p.106) note, 'A somewhat surprising feature of this literature is the noticeable absence of formal economic models of the productivity effects of infrastructure'.

In the earlier literature it is generally assumed that public capital forms an element in the aggregate production function. The stock of public capital (G_t) may enter the production function in two ways: directly, as a third input, or it may influence multifactor productivity (A)

$$Q_t = A(G_t)f(K_t, L_t, G_t), \tag{5.1}$$

where Q_t is real aggregate output of the private sector, L_t is (aggregate hours worked by) the labour force and K_t is the aggregate non-residential stock of private fixed capital.

Although is it pretty common to model the growth effects on government capital by adding a third factor in the production function, on second thoughts it is questionable whether it makes much sense. After all, government roads as such do not produce anything. Implicitly, it is assumed that the services of public capital are a pure, non-rival public good, with services proportional to the stock of capital. However, as pointed out by the World Bank (1994), many infrastructure services are almost (although not perfectly) private goods. Private goods can be defined as both rival (i.e., consumption by one user reduces available supply to others) and excludable (i.e., a user can be prevented from consuming them).

Furthermore, public capital is treated symmetric to labour and private capital. According to Duggal et al. (1999), this goes against standard marginal productivity theory in assuming that a market determined per unit cost of infrastructure is known to the individual firms and can be used in calculating total cost. However, since public investment is financed through general tax revenues or government debt, per unit costs of public capital are not market determined. Moreover, there is no guarantee that the total cost of infrastructure to the firm is related to the amount it uses. Aaron (1990) argues that this absence of a market test, coupled with possible government pricing inefficiencies, makes it impossible to assume that public capital as a factor input would be remunerated in line with its marginal product.

An alternative would be to incorporate public capital into the production function as part of the technological constraint that determines total factor productivity (see Duggal et al., 1999). Rather than acting as a discretionary factor input, public investment increases total productivity by lowering production costs. By increasing the technological index, additional public capital shifts the production function upward, and thus enhances the marginal products of the factor inputs. However, as pointed out by Sturm et al. (1998), in a Cobb-Douglas function (estimated in log levels) it does not make any difference whether public capital is treated as a third production factor or as influencing output through the factor representing technology. Both ways of modelling the influence of public capital yield similar equations to be estimated, so that the direct and indirect impact of public capital cannot be disentangled.

A better way to model the growth effect of public capital is by focusing explicitly on the services provided by the assets. For instance, Fernald (1999) assumes that for each industry i, production depends, apart from L_i and K_i, on transport services (T_i) produced within that particular sector. These services, in turn, depend upon the flow of services provided by the aggregated stock of government capital (roads) G and the stock of vehicles in the sector V_i. Output also depends on the Hicks neutral level of technology U_i. This yields

$$Q_t = U_i F^i\big(K_i, L_i, T(V_i, G)\big). \tag{5.2}$$

This way of modelling the growth effects of public capital also makes it possible to introduce the effects of congestion and network externalities. Many services provided by the stock of public capital may be subject to congestion: more vehicles on a road lower the productivity of this road. More roads will reduce congestion, and therefore, improve productivity. Above a certain threshold, however, marginal increments will no longer affect output since they no longer cause a decline in congestion (Sanchez-Robles, 1998). So congestion will give rise to non-linearities in the relationship between public capital and economic growth.

Public capital, notably infrastructure, is often distinguished from other types of capital because several market imperfections make accumulating and operating those assets prone to extensive government interventions and give rise to a special role for institutional characteristics. Economies of scale due to network externalities are a widely recognised imperfection in infrastructure services (World Bank, 1994). An important characteristic of modern infrastructure is the supply of services through a networked delivery system designed to serve a multitude of users. This interconnectedness means that the benefits from investment at one point in the network will generally depend on capacities at other points. The network character also has important consequences for the relationship between public capital and economic growth. Once the basic parts of a network are established, opportunities for highly productive investment diminish. In line with this argument, Fernald (1999) reports that once the highway system in the US was roughly completed, after 1973, the hypothesis that the marginal productivity of roads is zero cannot be rejected. In other words, road building gave a boost to productivity growth in the years before 1973, but post-1973 investment did not yield the same benefits at the margin.

There is broad consensus among economists and politicians that public infrastructure investment is an important aspect of a competitive location policy.[4] Often it is argued that infrastructure lowers fixed costs, attracting companies and factors of production and, thereby, raising production (see e.g., Haughwout, 2002 and Egger and Falkinger, 2003). This does not necessarily imply higher growth at the national level, however, since production in other regions might go down. A common result in this type of models is that, under certain assumptions, the resulting stock of capital without coordination between regions or countries is sub-optimal. Since more infrastructure in the 'home' region attracts production factors out of the 'foreign' region, there is a risk of the infrastructure being too high in both regions compared to the situation in which they coordinate their actions. That said, spillover effects of infrastructure could lead to the opposite outcome: because the investing region only gets part of the benefits, both regions end up with too little infrastructure.

The size of spillover effects will depend on the size of the country or region concerned and its openness. One simple way to model these spillovers has been suggested by Cohen and Morrison Paul (2004). Their model for a cost function of the manufacturing sectors in US states not only includes the public capital stock in the state concerned, but also the public capital stock in geographically connected states.[5] In a similar way, the public capital stock of a neighbouring state (G_j) can be included in a production function, which gives

$$Q_i = A_i K_i^{\alpha} L_i^{\beta} G_i^{\gamma} G_j^{\eta}. \tag{5.3}$$

A somewhat different reason why public capital may affect economic growth is suggested by the new economic geography (e.g., Krugman, 1991, Holtz-Eakin and

[4] The member countries of the European Union, for example, agreed upon a benchmark method to determine the competitiveness of the EU economies in which infrastructure plays a prominent role.

[5] Also Holtz-Eakin and Schwartz (1995) consider interstate spillovers.

Lovely, 1996, Fujita et al., 1999), which considers transport costs to be a central determinant of the location and scale of economic activity and of the pattern of trade. More transport infrastructure has a profound impact on the size of the market, so producers can cluster together in one central region. This clustering of activities leads to specialisation and economies of scale. In these theoretical models it is common to model transport costs as 'iceberg costs' (Krugman, 1991, Bougheas et al., 1999). The producer of a particular good sells a certain quantity and during transport a fraction of the shipped quantity 'melts' away. The longer the distance, the larger the fraction that melts and the higher are the transport costs. The buyer has to pay for more goods than he actually receives. This bypasses the need to model the transport sector separately. However, the concept of iceberg costs implicitly assumes that the transport sector's production function is equal to the production function of transported products, which is a rather strong assumption.

De la Fuente and Vives (1995) offer another nice and simple way to model transportation costs. They assume that final output Q in region i depends positively on intermediate production Y_i and negatively on transportation costs C_i. Transportation costs rise with the land area S of the region (as a proxy for distance) and decrease with the region's public capital stock G. de la Fuente and Vives further assume that Q_i exhibits constant returns to scale with respect to Y and C and that there is perfect private capital mobility across regions (so: $Q_i = Y_i^c G_i^\gamma S_i^{1-c-\gamma}$ where $c < 1 < c + \gamma$ so that transportation costs increase with land area). For intermediate production they assume a Cobb-Douglas production function with private capital and labour. Substitution results in

$$Q_i = A_i K_i^\alpha L_i^\beta G_i^\gamma S_i^{1-\alpha-\beta-\gamma}. \tag{5.4}$$

Even though the theoretical reasoning is different, the specification of de la Fuente and Vives is remarkably similar to Equation (5.1), suggesting observational equivalence.

Finally, the effects of government capital spending on growth will also crucially depend on the extent to which private and public capital are substitutes. The literature generally assumes that public and private capital spending are complements. However, public investment might also be a substitute for private investment. For instance, firms might build a road on their own, thereby allowing the government to withhold from this investment.

5.2.3 How to define public capital?

Most people probably think about roads and other infrastructure – such as electricity generating plants and water and sewage systems – when they refer to the public capital stock. However, it is important to point out that this does not fully correspond to the concept of public sector investment expenditure as defined in national accounts statistics, which are typically used to construct data on public capital stock. First, only spending by various government sectors is included. That implies that spending by the private sector (including public utility firms concerned with electricity generation, gas distribution, and water supply) is excluded. Secondly, public investment includes spending on various items (public buildings and swimming pools, for instance), which may not add anything to the productive capacity of an economy.

In calculating the stock of public capital on the basis of investment flow data, researchers typically use the sum of past investments, adjusted for depreciation. In applying the so-called perpetual inventory method, the researcher has to make certain assumptions about the assets' lifespan and depreciation. Furthermore, one needs an initial level for the capital stock. Especially with infrastructure these assumptions are far from trivial. There is a huge variation in the economic lifespan of different types of infrastructure; the lifespan of a railway bridge cannot be compared with the lifespan of an electricity transmission line. Usually, the initial stock is calculated by assuming that the real investments were constant at the level of the first observed investment level and that the capital stock was at its steady state at the start of the observed time series. With very low depreciation rates, the rate of convergence towards the steady-state level is very low, which requires a very long time of constant investment.

To calculate the public capital stock one needs long-term time-series data on public investment. Long-term national account time-series data on government investment spending are available for most OECD countries. However, for many developing countries the availability of long-term data is more of a problem, so that the public capital stock cannot be constructed for these countries. Therefore various studies use government investment or some physical measure of infrastructure instead of the government capital stock. A drawback of the use of government investment spending (as share of GDP) as regressor – which is a fairly common approach in studies based on cross-country growth regressions and in some vector autoregression studies – is the implicit assumption that the effects of public investment are independent of the level of the corresponding capital stock. Economic theory suggests that this assumption is dubious (Kamps, 2006). Also the use of some physical measure of infrastructure, like the number of kilometres of paved roads, has certain advantages and disadvantages (see below).

Pritchett (1996) points to some serious problems with using monetary values to calculate the stock of public capital. Prices for infrastructure capital vary widely across countries. Furthermore, the level of expenditure may say little about the efficiency in implementing the investment project. Especially if the investment project is carried out by the public sector, actual and economic costs (defined as the minimum of possible costs given available technology) may deviate. So, monetary investment in infrastructure may be a poor guide to the amount of infrastructure capital produced because government investment may be very inefficient. According to Pritchett (1996), this is probably true, in particular, in developing countries. He estimates that only slightly more than half the money invested in investment projects will have a positive impact on the public capital stock.[6] This implies that public capital stock series constructed on the basis of investment series will tend to be overvalued.

Also from a network perspective, the monetary value as obtained by the perpetual inventory method of measuring capital stock is not appropriate. In particular, the internal composition of the stock matters since the marginal productivity of one link depends on the capacity and configuration of all links in the network. Using measures of the total stock may thus allow estimating the average marginal product of, say, roads in the past, but these estimates may not be appropriate for considering the marginal product of additional roads today (Fernald, 1999).

Given these problems, many recent studies have employed some physical measure of infrastructure in analysing its impact on economic growth. Studies have used, in particular, the number of kilometres of paved roads, kilowatts of electricity generating capacity, and the number of telephones (see, for instance, Canning and Pedroni, 1999, Sanchez-Robles, 1998, and Esfahani and Ramíres, 2003).[7] As these physical measures are available for many countries for long time spans, they are ideal for estimating panel models. An advantage of using physical measures of infrastructure is that they do not rely on the concept of public investment as em-

[6] How the project is financed may affect these figures; the stronger the incentives for the government to minimise costs, the higher the contribution to the public capital stock of an investment project.

[7] Canning (1998) describes an annual database of physical infrastructure stocks for 152 countries for 1950-95. The database contains six measures: kilometres of roads, kilometres of paved roads, kilometres of railway lines, number of telephones, number of telephone main lines, and kilowatts of electricity generating capacity.

ployed in the national accounts. For instance, by whom electricity is generated does not matter. However, simple physical measures do not correct for quality. Furthermore, some of the measures do not necessarily refer to (the results of) government spending.

Initially research on the impact of public capital on economic growth focused on the United States. Only few of the earlier studies investigated the productivity of government capital for a group of OECD countries (see, for instance, Ford and Poret, 1991 and Evans and Karras, 1994). These authors drew their data from the OECD that assembled capital stock series for 12 countries over the period 1970–1996, provided directly by the national authorities. However, these data were not internationally comparable because estimation methods differed widely across countries. This was one of the reasons why the OECD suspended the publication of the capital stock series after 1997. Recently, Kamps (2006) has provided internationally comparable annual capital stock estimates for 22 OECD countries for the period 1960–2001.

Whereas Aschauer (1989) and many subsequent studies employed national data for the United States, other studies used regional data again with mixed findings (see Sturm et al., 1998). For the US, data at the state level are only available after 1970. Also for some European countries (Spain, France, Germany, and Italy) regional public capital stock data are available. Using regional data increases data variation, which may make the estimates more reliable.

To summarise, this section has set out the main research questions addressed by the literature on the relationship between public capital and growth, explained the meaning of public capital and its link to infrastructure, and sketched theoretical insights about the role of public capital for economic growth. The following sections elaborate on alternative empirical research strategies used to learn more about the role of public capital for economic growth.

5.3 Production function approach

Let us start with a description of the theoretical framework underlying the empirical studies that follow the production-function approach. In this type of analysis, the production function as given in equation (5.1) is generally written as an aggregated Cobb-Douglas production function in which the public capital stock (or the monetary value of the stock of infrastructure),G_t, is added as an additional input factor,

$$Q_t = A_t K_t^{\alpha} L_t^{\beta} G_t^{\gamma}. \tag{5.5}$$

Writing Equation (5.5) in per capita terms, taking the natural logarithm, and assuming constant returns to scale across all inputs ($\alpha + \beta + \gamma = 1$), gives

$$\ln \frac{Q_t}{L_t} = \ln A_t + \beta \ln \frac{K_t}{L_t} + \gamma \ln \frac{G_t}{L_t}. \tag{5.6}$$

The parameter γ gives the elasticity of infrastructure. To assess γ, a straightforward procedure is to estimate the production function in log-level or, alternatively, in first-difference or growth. This is indeed common practice in the initial attempts at measuring the role of infrastructure. Aschauer (1989) introduces a constant and a trend variable as a proxy for $\ln A_t$. The capacity utilization rate is added to control for the influence of the business cycle. Many subsequent papers have used this or a similar specification.[8] A drawback of the estimated production functions is that labour and capital are exogenous; it is implicitly assumed that both factors are paid according to their marginal productivity. Some studies have used a translog function, which is more general than the Cobb-Douglas function (e.g. Canning and Bennathan, 2000, Albala-Bertrand and Mamatzakis, 2004, Everaert and Heylen, 2004, and Charlot and Schmitt, 1999).

A major problem in estimating a production function is the potential for reverse causation. If capital investments ($I_t = \Delta K_t$) depend on income (for example, through a savings function S_t) we can write

$$\Delta K_t = sY_t - \delta K_t, \tag{5.7}$$

[8] Various authors have taken issue with the specification of Aschauer's model. Tatom (1991), for instance, uses another specification with energy prices included and capacity utilization entered multiplicatively to both the private and public capital stock and finds little evidence that the public capital stock raises productivity. However, Duggal et al. (1999) criticize Tatom's approach arguing that the relative price of energy is a market cost factor that would be included in the firm's cost function and therefore also in the factor input demand functions.

where Y_t is total income and δ is the depreciation rate. This gives the steady state relationship

$$K_t = \frac{sY_t}{\delta}. \tag{5.8}$$

This implies a feedback from income to the capital stock, making it difficult to identify the results of regressions such as equation (5.6) as a production function relationship. There is also a potential feedback from income to a demand for infrastructure. Dealing with this problem has been at the heart of the controversy over the infrastructure-growth relationship.

Various approaches have been followed in the literature to deal with the problem of causality. One is to derive an appropriate test in such a way that it is clear how the causality runs. Other approaches that have been followed are: estimating panel models, estimating simultaneous equation models, and using instrumental variables.

Fernald (1999) is a good example of the first approach. Using data for 29 sectors in the US economy for the years 1953-89, he finds that changes in road growth are associated with larger changes in productivity growth in industries that are more vehicle intensive. Fernald argues that if roads were endogenous, one would not expect any particular relationship between an industry's vehicle intensity and its relative productivity performance when road growth changes. According to Fernald, his results suggest that the massive road building in the US of the 1950s and 1960s offered a one-time boast to the level of productivity. His results have important policy implications: building an interstate highway network may be very productive, but building a second network may not.

Another highly relevant study that belongs to the first approach is Canning and Pedroni (1999). They derive a reduced form of a model in which public and private capital are financed out of available savings so that there is a growth-maximising level of public capital. The nature of the long-run relationship and the short-run dynamics may vary across countries. Since they find that in each country the physical stock of infrastructure and per capita income are individually non-stationary but cointegrated, they can represent the series in the form of a dynamic error-correction model. By testing restrictions in this model, they can decide on the direction of causality. It appears that causality runs in both directions. For balanced panels of different countries they find that, on average, telephones and paved roads are supplied at around the growth-maximising level, but some countries have too few, others too many. Canning and Pedroni also find that long-run effects of investment in electricity generating capacity are positive in many countries, with negative effects being found in only a few.

Canning and Bennathan (2000) argue that the causality problem may be solved by using a panel data approach. If the cointegrating Equation (5.4) in a panel setting is a homogeneous relationship, while Equation (5.5) differs across countries, pooling the data across countries allows identifying the long-run production-function relationship. For two infrastructure stock variables (electricity generating capacity and the length of paved roads) they find higher rates of returns than for other types of capital, although there is some heterogeneity in their sample.

The most intuitive way to solve the causality problem is to develop a simultaneous-equations model, consisting of two equations. The first equation links production to public capital, the second equation links public capital to production. The main question is the functional form for the second equation. Demetriades and Mamuneas (2000) estimate a system of equations that is derived from an intertemporal profit maximisation framework.[9] The estimates refer to a pooled model for 12 OECD countries over 1972-91. In the short run, the output effect of public capital varies from 0.36 percent in the UK to 2.06 percent in Norway. Also for the intermediate to long run, Demetriades and Mamuneas find diverging rates of return across countries. In their theoretical model, producers take at each point in time the publicly provided inputs as given and maximise the present value of future profits to determine their output, variable inputs, and quasi-fixed factor demands. In the first stage, firms decide on the optimal output and variable input demands, conditional on the private and public capital stocks. In the second stage, firms choose the optimal sequence of capital inputs. The authors claim that 'by taking into account the optimising behaviour of firms we avoid the simultaneity problem typical of the production-function approach' (pp. 688-89). Although this may be true for the private capital stock, it is not true for the public capital stock, which is simply assumed to be exogenous.

A better attempt to estimate a simultaneous-equations model is the cross-country growth study by Esfahani and Ramíres (2003), who develop a structural growth model that helps discern the reciprocal effects of infrastructure and the rest of the economy. The model specifies the ways in which country characteristics and policies enter the infrastructure GDP interactions and lead to heterogeneity of outcomes across situations. The authors distinguish heterogeneity in the steady state and in the rate of convergence towards a steady state. They derive the infrastruc-

[9] This paper belongs to the cost-function approach as discussed in the next section, but is taken up here since it is a good example of the simultaneous equations approach.

ture-output interactions as a recursive system that can be estimated simultaneously while solving the identification problem. The relationships between infrastructure and income are formulated as error-correction processes to account for the simultaneous effects of infrastructure innovations and responses to deviations from the steady state. Esfahani and Ramíres find that the contribution of infrastructure services to GDP is substantial and, in general, exceeds the cost of providing these services. The findings of Esfahani and Ramíres also shed light on the factors that shape a country's response to its infrastructure needs. An interesting result in this respect is that private ownership of infrastructure and government credibility (low risk of contract repudiation) matter for infrastructure growth, but mainly in speeding up the rate of adjustment rather than the steady-state infrastructure-income ratios

Cadot et al. (2002) also endogenise public capital formation by focusing on the decision making process of public capital spending. The policy equation explicitly models the political decision process, including lobbying from different regions. Estimating the model for 21 regions in France over the period 1985–91, Cadot et al. (2002) find an elasticity of output with respect to public capital of 0.101 for France as a whole. This is very close to their simple single equation OLS estimates of 0.099, which suggests that the simultaneous-equation bias is only moderate. Interestingly, they find evidence that roads and railways are not built to reduce traffic jams: they are built essentially to get politicians re-elected. The number of large companies in a region seems to be an important determinant in explaining the total public investment allocated to that region.

Kemmerling and Stephan (2002) also focus on the political decision-making process on public investment. Using panel data for 87 German cities for the years 1980, 1986, and 1988 in a simultaneous equations model, they estimate the relationship between infrastructure investments, investment grants, local manufacturing output, policy and lobbying variables. Their main findings are that political affiliation, measured by the coincidence of party colour between state and local government, is decisive in explaining the distribution of investment grants across cities, and that cities with 'marginal voters' neither spend more on public infrastructure nor receive more investment grants from higher-tier governments. Interestingly, they also conclude that efficiency considerations do not seem to determine the observed intergovernmental grant allocation across cities.[10]

[10] These studies point to an interesting area for future research, i.e. the explanation of differences in public investment spending across regions/countries and over time. So far, most of the theoretical literature assumes that desicion-making on public capital spending is only based on efficiency considerations; the evidence presented by Cadot et al. (2002) and Kemmerling and Stephan (2002) suggest that this assumption is highly unrealistic.

Finally, some instrumental variable approach may be used. Some of the older studies already applied the Generalized Method of Moments (GMM) estimator, which resembles an instrumental-variables procedure and therefore avoids the possible reverse-causation bias (Finn, 1993; Ai and Cassou, 1995).[11] A more recent study is Calderón and Servén (2002). They chose for the instrumental variable method since this is easier to carry out than the simultaneous-equations model. These authors estimate a per capita Cobb-Douglas production function (in log-levels) for a panel of 101 countries for the period 1960–97. To solve the causality problem they use lagged values of the explanatory variables. Because of non-stationary data, they estimate a per capita Cobb-Douglas production function in first differences. Allowing for country-specific effects by a 'within' estimator they find an average elasticity of 0.16 for different types of infrastructure.

Table 5.2 summarises key features and results of the papers reviewed above and other studies based on the production-function approach. The table is an update of Table 1 in the survey of Sturm et al. (1998) and has a similar set-up. The first column presents the study, the second to fourth columns show the aggregation level, the sample, the specification, and the way public capital has been measured, respectively, while the final column summarises the study's main findings. Although not all studies find a growth-enhancing impact of public capital, it is worth noting that – compared to the results surveyed by Sturm et al. (1998) – there is more consensus that public capital furthers economic growth. Another interesting result is that the impact as reported in recent studies is substantially less than suggested in earlier studies.

5.4 The cost function approach

A key shortcoming of the production-function approach is that it violates standard marginal productivity theory. Some studies have tried to get around the violation by focusing on the cost function and assuming that public capital is externally provided by the government as a free input. These studies specify a cost function for the private sector, with firms being assumed to aim at producing a given level of output at minimum private cost (C). Because the input prices (p_i) are exogenously determined, the instruments of the firm are the quantities of the private inputs (q_i). Alternatively, firms are assumed to maximize their profits (Π) given the output (p^Q)

[11] Finn (1993) reports a significant elasticity of the stock of public highways in the US of 0.16. The elasticity estimates of Ai and Cassou (1995) for the total stock of public capital in the US range between 0.15 and 0.26.

and input prices.

$$C\left(p_t^i, q_t^i, A_t, G_t\right) = \min \sum p_t^i q_t^i \qquad \text{s.t.} \quad Q_t = f(q_t^i, A_t, G_t) \tag{5.9}$$

$$\Pi\left(p_t^Q, p_t^i, q_t^i, A_t, G_t\right) = \max p_t^Q Q_t - \sum p_t^i q_t^i \qquad \text{s.t.} \quad Q_t = f(q_t^i, A_t, G_t) \tag{5.10}$$

When firms optimise, they take into account the environment in which they operate. One of these environmental variables is the state of technical knowledge (A). Another is the amount of public infrastructure capital available (G). The public capital stock enters the cost or profit function as an unpaid fixed input. Although the stock of infrastructure is considered externally given in the cost-function approach, each individual firm must still decide the amount it wants to use. This implies that a firm's use of the infrastructure is part of its optimisation problem, which, in turn, leads to the need of a demand function for infrastructure that must satisfy the conditions of standard marginal productivity theory (Duggal et al., 1999). To make this approach comparable with the production-function approach, various authors (e.g., Demetriades and Mamuneas, 2000) use Hotelling's Lemma to obtain supply functions, which can be used to calculate output elasticities of public capital.

Sturm et al. (1998) note that many authors estimating a cost or profit function adjust the stock of public capital by an index, such as the capacity utilisation rate, to reflect its use by the private sector. Two reasons have been advocated for adjusting the stock of public capital. First, public capital is a collective input that a firm must share with the rest of the economy. However, since most types of public capital are subject to congestion, the amount of public capital that one firm may employ will be less than the amount supplied. Moreover, the extent to which a capacity utilisation index measures congestion is dubious. Second, firms might have some control over the use of the existing public capital stock. For example, a firm may have no influence on the highways provided by the government, but can vary its use of existing highways by choosing routes. Therefore, there are significant swings in the intensity with which public capital is used.

As pointed out by Sturm et al. (1998), an important advantage of the cost-function approach is that it is less restrictive than the production-function approach. The use of a flexible functional form hardly enforces any restrictions on the production structure. For example, *a priori* restrictions placed on the substitutability of production factors, as in the production-function approach, do not apply. Apart from the focus on the direct effects in the production-function approach, public capital might also have indirect effects. Firms might adjust their demand for private inputs if public capital is a substitute or a complement to these other production factors. It seems very plausible that, for instance, a larger stock of infrastructure raises the quantity of private capital used and therefore indirectly raises production.

By using a flexible functional form, the influence of public capital through private inputs can be determined. A flexible function not only consists of many parameters that need to be estimated, but also of many second-order terms which are cross products of the inputs. These second-order variables can create multicollinearity problems. Therefore, the data set not only has to be relatively large, but must also contain enough variability so that multicollinearity can be dealt with. In other words, the most appealing feature of the cost-function approach also induces the biggest problem, i.e., the flexibility of the functional form requires considerable information to be included in the data. Most cost-function studies therefore use panel data, which combine a time dimension with either a regional dimension or a sectoral dimension.

Interestingly, whereas Sturm et al. (1998) found that the cost-function approach was used in many studies they reviewed, we have found only a few studies that rest on the cost-function approach. Table 5.3 summarizes these studies, thereby updating Table 2 of Sturm et al. (1998). We discuss two of these studies – probably the most interesting – in some detail.

Moreno et al. (2003) estimated cost functions for 12 manufacturing sectors in Spanish regions during the period 1980–91. They conclude that the average cost elasticity of public capital is only -0.022. However, there is wide variety in the effect across regions and industries; in fact, the range of values (-0.062 to 0.033) is wide enough to suggest the possibility that some regions and sectors did not benefit from public capital in some years. Costs in industries such as electric machinery, food and drinks, and textiles seem to have been most sensitive to a rise in infrastructure, while the opposite applies to sectors such as metallic and non-metallic minerals and chemistry. Among the regions with higher-than-average cost-infrastructure sensitivities are some of the least and most developed regions in Spain.

Cohen and Morrison Paul (2004) estimated a cost-function model by maximum likelihood techniques; they used data for 48 US states on prices and quantities of aggregate manufacturing output and inputs (specifically: capital, production and non-production labour, and materials) and on public highway infrastructure; their analysis covers the period 1982–96. They assume that manufacturing firms minimise short-run costs by choosing a combination of inputs for a given level of input prices, demand (output), and capacity (capital) and for given (external) technological and environmental conditions. The model also distinguishes between intra- and interstate effects of public infrastructure and accounts for interaction between the two. More specifically, for a given state, the model includes not only the public infrastructure of that state but also the infrastructure in neighbouring states. Cohen and Morrison Paul find a significant contribution of public infrastructure investment to lowering manufacturing cost - an effect enhanced by spillover effects across states. If the stock of infrastructure of a neighbouring state is not included, as in most of this literature, the elasticity is around -0.15, which is comparable to those found in other studies. However, taking spillovers into account raises the average elasticity to -0.23. So recognising spatial linkages increases the estimated effects of intrastate infrastructure investment. They also find that the intra- and interstate effects of public capital increase over time.[12]

In conclusion, the results of the cost-function studies reviewed in this section are broadly in line with those of studies using the production-function approach: public capital reduces cost, but there is much heterogeneity across regions and/or industries.

5.5 Vector autoregression models

Various recent studies use vector autoregression (VAR) models, which – unlike the production function and cost-function approaches – do not impose causal links among the variables under investigation.[13] In a VAR model, all variables are jointly determined with no a priori assumptions about causality. So VAR models allow to test whether the causal relationship assumed in other approaches is valid or whether there are feedback effects from output to public capital. Furthermore, the VAR approach allows testing for indirect effects between the variables of the model. An unrestricted VAR model can be simply estimated by standard ordinary least squares (OLS). OLS will yield consistent and asymptotically normally distributed estimates, even if variables are integrated and possibly cointegrated (Sims et al., 1990).

However, even in a simple VAR model some choices with respect to the spe-

[12] The results of Cohen and Morrison Paul are also interesting from the viewpoint of the causality issue. To test for the potential endogeneity of infrastructure, they conducted a Hausman test and found that they could not reject the null hypothesis of infrastructure exogeneity, which they argue is 'consistent with our a priori conjectures that manufacturing sector activity is unlikely to drive policy decisions across states (or even within a state), due to the small share of manufacturing production in states' overall GSP' (p. 555).

[13] This section heavily draws on Kamps (2004).

cification of the model have to be made, and all of them may affect the estimated responses and, thus, alter the conclusions about the link between public investment and economic growth. For instance, to simulate the cumulative response functions, restrictions with regard to ordering are imposed. These restrictions are rationalised by invoking assumptions of exogeneity and/or pre-determinedness, both of which can only be derived from theoretical considerations. In the absence of ordering assumptions, the non-structural VAR model can be used to characterise the data, but it cannot be used to spell out causation. Furthermore, Phillips (1998) shows that impulse responses and forecast error variance decompositions based on unrestricted VAR models are inconsistent at long-run horizons in the presence of non-stationary data. In contrast, Vector Error Correction Models (VECMs) yield consistent estimates of impulse responses and of forecast error decompositions if the number of cointegrating relationships is estimated consistently.

Table 5.4 summarizes VAR studies, updating Table 3 of Sturm et al. (1998). The following conclusions can be drawn. First, only few studies (for example Mittnik and Neumann, 2001, and Kamps, 2004) refer to a group of OECD countries; the rest focuses on one or two countries only. Second, most studies consist of a four variables VAR with output, employment, private capital, and public capital. Third, there is a wide variety of model specifications. Some studies specify VAR models in first differences, without testing for cointegration, while others explicitly test for cointegration. Some studies specify VAR models in levels, following the argument of Sims et al. (1990) that OLS estimates of VAR coefficients are consistent even if the variables are non-stationary and possibly cointegrated. Fourth, in the majority of studies the long-run response of output to public capital shock is positive.[14] However, as pointed out by Kamps (2004), most studies fail to provide any measure of uncertainty surrounding the impulse response estimates so that it is impossible to judge the statistical significance of the results. Kamps (2004) employs bootstrapping techniques to provide confidence intervals. Fifth, many VAR studies report evidence for reverse causality, i.e. feedback from output to public capital. Finally, some studies (e.g. Everaert, 2003) report that public capital has less impact on economic growth than reported by Aschauer (1989).

[14] Voss (2002) gives no conclusions regarding output effects of infrastructure as he focuses on possible crowding in effects found by Aschauer (1989). These 'crowding in' effects enforce the positive effects of public investment, but using cointegrating techniques to correct for non-stationarity in the data, Voss does not find evidence for these effects in both the US and Canada. Only Ghali (1998) finds negative effects on growth, but these can easily be explained by the very structure of the Tunisian economy where 'highly subsidized and inefficient state owned enterprises [] have often reduced the possibilities for private investment'.

5.6 Cross-section studies

Since the mid-1980s, the study of economic growth and its policy implications has vigorously re-entered the research agenda. Various studies tried to explain, theoretically and empirically, why differences in income over time and across countries did not disappear as the neoclassical models of growth predicted. The idea that emerged from this literature is that economic growth is endogenous. That is, economic growth is influenced by decisions of economic agents, and is not merely the outcome of an exogenous process. Endogenous growth theory assigns a central role to capital formation, where capital is not just confined to physical capital, but includes human capital, infrastructure and knowledge capital.

Initially, the econometric work on growth was dominated by cross-country regressions, in which growth of real per capita GDP is estimated by a catch-up variable, human capital, investment, and population factors like fertility. Some of these studies add government investment as an explanatory variable. The equations estimated in various studies can be summarised as

$$\Delta \ln \left(\frac{Y}{L}\right)_{0,T} = \alpha + \beta \left(\frac{Y}{L}\right)_{0} + \gamma \left(\frac{I^G}{L}\right)_{0,T} + \delta, \tag{5.11}$$

where $(Y/L)_{0,T}$ is the average per capita GDP over a period $[0, T]$, $(Y/L)_0$ is the initial level of real per capita GDP, and $(I^G/Y)_{0,T}$ is the average rate of public investment (as percentage of GDP) over a period $[0, T]$. The variable δ captures a set of conditional variables such as private investment (as percentage of GDP) and primary and/or secondary enrolment (as a proxy for human capital). The parameter γ measures the effect of public investment on growth and is not the same as the marginal productivity of public capital.

Unfortunately, most empirical economic growth studies do not distinguish between public and private investment, instead relying on an aggregate measure of total investment. However, the services from public investment projects are likely to differ from those of private investment projects for a number of reasons, and this suggests that an aggregate investment measure is inappropriate (Milbourne et al., 2003). Table 5.5 at the end of this chapter, which updates Table 4 in Sturm et al. (1998), provides a summary of cross-country growth models that include public investment.

Probably the first study that included public capital in an empirical growth model is Easterly and Rebelo (1993), who ran pooled regressions (using decade averages for the 1960s, 1970s and 1980s) of per capita growth on (sectoral) public investment and conditional variables (see Sturm et al. (1998) for a summary). They found that the share of public investment in transport and communication infrastructure is correlated with growth. Likewise, Gwartney et al. (2006) find a significant positive effect of public investment, although its coefficient is always smaller than that of private investment.

However, other studies using the public investment share of GDP as regressor report different results. For instance, Sanchez-Robles (1998) finds a negative growth impact of infrastructure expenditure in a sample of 76 countries. Devarajan et al. (2000) report evidence for 43 developing countries, indicating that the share of total government expenditure (consumption plus investment) has no significant effect on economic growth. However, the authors find an important composition effect of government expenditure: increases in the share of consumption expenditure have a significant positive impact on economic growth whereas increases in the share of public investment expenditure have a significant negative effect. Devarajan et al. attribute their results to the fact that excessive amounts of transport and communication expenditures in those countries make them unproductive. Pritchett (1996) suggested another explanation, arguing that public investment in developing countries is often used for unproductive projects. As a consequence, the share of public investment in GDP can be a poor measure of the actual increase in economically productive public capital.

Milbourne et al. (2003) report that for the steady-state model, there is no significant effect from public investment on the level of output per worker. Using standard ordinary least squares (OLS) methods for the transition model, they find that public investment has a significant effect on economic growth. However, when instrumental variables methods are used, the associated standard errors are much larger and the contribution of public investment is statistically insignificant.

The only study in this category that we are aware of that has used physical indicators of infrastructure instead of public investment spending is Sanchez-Robles (1998). When she includes indicators of physical units of infrastructure, she finds they are positively and significantly correlated with growth in a sample of 76 countries.

There are two important general problems in the cross-country growth regressions: one is model uncertainty and the other is outliers and parameter heterogeneity (Temple, 2000; Sturm and de Haan, 2005). Model uncertainty has been discussed extensively in the literature. The main issue here is that several models may all seem reasonable given the data, but yield different conclusions about the parameters of interest. In these circumstances, presenting only the results of the model preferred by the author can be misleading (Temple, 2000). Unfortunately, economic theory does not provide enough guidance to properly specify the empirical model. For instance, Sala-i-Martin (1997) identifies around 60 variables supposedly correlated with economic growth. The so-called extreme bound analysis (EBA) of Leamer (1983) and Levine and Renelt (1992) is therefore often used to examine how 'robust' the economic growth effect of a certain variable is. The key idea of EBA is to report an upper and lower bound for parameter estimates, thereby indicating the sensitivity to the choice of model specification. The upper and lower bounds are based on a set of regressions using different subsets of the set of explanatory variables. If the upper and lower bounds have a different sign, the relation is not robust.

The second problem – the role of outliers and parameter heterogeneity – has been largely ignored by the empirical growth literature . Although economists engaged in estimating cross-country growth models often test the residuals of their regressions for heteroskedasticity and structural change, they hardly ever test for unusual observations. Still, their data sets may frequently contain unusual observations. In particular, less developed countries tend to have a lot of measurement error in national accounts and other data. This may have affected the conclusions of cross-country growth models.

Unfortunately, none of the studies reviewed in this section takes the issues of model uncertainty and outliers and parameter heterogeneity seriously into account, which casts considerable doubt on their findings. With this somewhat sober remark we finish the review of different empirical strategies to estimate the link between public capital and economic growth, and we move on to a brief discussion of what could constitute an optimal capital stock.

5.7 Optimal capital stock

In estimating the optimal stock of public capital, the assumption on the public good character of infrastructure is crucial. For pure public goods, one could define total marginal benefits of public capital as the sum of the shadow values over all firms plus the sum of corresponding marginal benefits over all final consumers, yielding what might be called the social or total marginal benefit of public capital. Alternatively, if there is no congestion in the consumption of public goods, the total marginal benefit could be the largest benefit accruing to any one or set of consumers and producers rather than the sum over all consumers and producers. The simplest rule to determine the optimal provision of public capital is to calculate the amount of infrastructure for which social marginal benefits just equal marginal costs.

The difficulty in the empirical implementation of this rule lies in approximating the marginal costs of public capital. Sturm et al. (1998) found only a few studies that estimated the optimal amount of public capital and compared it with the actual stock of public capital. These studies use some measure for the cost of borrowing, such as the government bond yield, to approximate the marginal costs of public capital. Adopting this approach, Berndt and Hansson (2004), for instance, report excess public capital in the United States, which has declined over time, however. Alternatively, Conrad and Seitz (1994) interpret the case in which the social marginal benefit of public capital is greater than the price of private capital as a shortage of public capital, whereas the reverse indicates over-investment in public capital. These authors find that during 1961-79 the social marginal benefit of public capital in Germany was larger than the user cost of private capital, whereas in the 1980–88 period the opposite was true.

The more recent literature has taken other ways of modelling the optimal public capital stock. Canning and Pedroni (1999) develop a model in which public investment spending lowers investment in other types of capital because they all need to be financed out of savings. In this approach, there is a certain level of public capital that maximises economic growth, and if there is too much infrastructure, it diverts investment away from other productive uses to the point where income growth falls. In this setting, the effect of an increase in public investment on economic growth depends on the relative marginal productivity of private versus public capital. In other words, we need to know not only whether public capital is productive but also whether it is productive enough to boost economic growth. An interesting finding of this study is that the assumption of parameter homogeneity can clearly be rejected. In other words, there is much heterogeneity among countries with regard to the optimal level of public capital.

Aschauer (2000) has developed a non-linear theoretical relationship between public capital and economic growth in order to obtain estimates of the growth-maximising ratio of public to private capital. Permanent increases in the public capital ratio bring forth permanent increases in growth – but only if the marginal product of public capital exceeds the after-tax marginal product of private capital. Using data for 48 US states over the period 1970–90, Aschauer finds that for most of the United States the actual levels of public capital were below the growth-maximising level. Kamps (2005) is the first study to use the methodology of Aschauer (2000) in the European context to assess the gap between actual and optimal public capital stocks. The empirical results suggest that there currently is no lack of public capital in most 'old' EU countries. However, current fiscal policies imply that in the long run the public capital to GDP ratio will be significantly lower than its growth-maximizing level in 4 out of 14 EU countries considered if the government investment to GDP ratio in these countries is not raised.

5.8 Concluding comments

Our review of recent studies that examine the relationship between public capital and economic growth suggests the following main results. First, although not all studies find a growth-enhancing effect of public capital, there is more consensus in the recent literature than in the older literature as summarised by Sturm et al. (1998). Second, according to most studies, the impact is much lower than found by Aschauer (1989), which is generally considered to be the starting point of this line of research. Third, many studies report that there is heterogeneity: the effect of public investment differs across countries, regions, and sectors. This is perhaps not a surprising result. After all, the effects of new investment spending will depend on the quantity and quality of the capital stock in place. In general, the larger the stock and the better its quality, the lower will be the impact of additions to this stock. The network character of public capital, notably infrastructure, causes non-linearities. The effect of new capital will crucially depend on the extent to which investment spending aims at alleviating bottlenecks in the existing network. Some studies also suggest that the effect of public investment spending may also depend on institutional and policy factors.

In concluding, we would like to mention a few issues we believe have not been well researched. First, attempts at explaining existing differences in capital stocks are only in their infancy. Second, only a few of the enormous bulk of studies on the output effects of infrastructure base their estimates on solid theoretical models. But to understand non-linearities and heterogeneity, we must understand the channels through which infrastructure affects economic growth. After all, government roads as such do not produce anything, and to include infrastructure or public capital as a separate input in a production function neglects the usually complex links. Third, most of the literature has focused on the importance of additional public investment spending, while maintenance of the existing stock is as important, if not more important, as additions to the stock. As pointed out by the World Bank (1994), inadequate maintenance imposes large and recurrent capital costs. For instance, paved roads will deteriorate fast without regular maintenance. Likewise, insufficient maintenance of a railway system will lower its reliability, causing delays for travellers when parts of the system break down. Unfortunately, policy makers have a perverse incentive: given their higher visibility, new public investment projects are politically more attractive than economically crucial, but politically less rewarding spending on infrastructure maintenance.

Table 5.1. Government investment in 22 OECD countries as percentage of GDP, 1960-2001

Country	1960–69	1970–79	1980–89	1990–99	2000–01
Australia	3.77	3.61	2.59	2.56	2.76
Austria	5.03	5.50	3.74	2.72	1.37
Belgium	2.06	3.44	3.15	1.74	1.62
Canada	3.40	2.65	2.36	2.59	2.48
Denmark	5.15	4.42	2.07	1.74	1.86
Finland	2.82	3.40	3.34	3.11	2.49
France	4.02[1]	3.55	2.97	3.23	2.99
Germany	4.05	3.86	2.61	2.37	1.95
Greece	3.90	3.34	2.78	3.12	3.86
Iceland	4.21	4.29	3.23	3.48	3.48
Ireland	5.65	6.24	4.56	2.29	3.01
Italy	3.31	2.88	3.15	2.58	2.39
Japan	7.50	9.32	7.47	7.68	6.91
Netherlands	6.21	4.88	3.18	2.96	3.27
New Zealand	5.65[2]	6.42	5.37	3.21	3.02
Norway	3.31	4.13	3.25	3.48	3.13
Portugal	2.37	2.08	2.60	3.69	3.92
Spain	2.82	2.54	2.98	3.86	3.14
Sweden	2.72	2.65	2.15	2.63	2.19
Switzerland	2.55	3.29	2.90	3.17	2.99
United Kingdom	3.96	3.52	1.85	1.99	1.57
United States	4.51	2.99	3.14	3.37	3.41

[1] 1963–1969, [2] 1962–1969
Source: Kamps (2006)

Table 5.2. Studies using some kind of production function approach

Study	Countries	Sample	Specification	Public capital variable	Conclusion
Albala-Bertrand & Mamatzakis (2004)	Chile	1960–98	Translog PF	Infrastructure capital stock (transportation, communication, general purpose)	Infrastructure capital growth appears to reduce productivity slightly up to 1971. From 1972 onwards, the reverse seems true
Albala-Bertrand (2004)	Chile and Mexico, regions	1950-2000	Gap approach using a Leontief PF (with private and public capital as inputs)	Infrastructure capital stock (transportation, communication, general purpose)	In Chile potential output is mostly constrained by shortages of normal capital, in Mexico infrastructure is the binding factor
Batina (1999)	U.S.	1948–1993	Cobb-Douglas production function with public capital as separate factor	Various proxies for state level public capital	The productivity of public capital depends on the proxies used for private and public capital
Bonaglia et al. (2000)	Italy, regions	1970-94	Cobb-Douglas PF with public capital as separate factor	Public capital stock	Elasticity is 0.05 (insignificant) for Italy as a whole, large variation between regions
Cadot et al. (2002)	France, regions	1985-92	Cobb-Douglas PF combined with policy equation for transport infrastructure	Infrastructure capital stock (transportation)	Elasticity is 0.08
Calderón & Servén (2002)	101 countries	1960-97	Cobb-Douglas PF with different types of infrastructure as separate factor	Infrastructure capital stock (transportation, communication, general purpose)	Elasticity is 0.16
Canning (1999)	57 countries	1960-90	Cobb-Douglas PF with different types of infrastructure as separate factor	Number of telephones, electricity generating capacity and kilometres of paved roads and railways	Electricity and transportation routes have normal capitals rate of return, telephone above normal
Canning & Pedroni (1999)	Panel of countries, different length	1950-92	Tests whether infrastructure has long-run effect on growth based on dynamic error-correction model	Number of telephones, electricity generating capacity and kilometres of paved roads and railways	Evidence of long-run effects running from infrastructure to growth, but results differ across countries and type of infrastructure

table continues on next page...

Table 5.2. (continued)

Study	Countries	Sample	Specification	Public capital variable	Conclusion
Canning & Bennathan (2000)	62 countries	1960-90	Cobb-Douglas and translog PF with different types of infrastructure as separate factor	Number of telephones, electricity generating capacity and kilometres of paved roads and railroads	On average, only the low- and middle-income countries benefit from more infrastructure
Charlot & Schmitt (1999)	France, regions	1982-93	Cobb-Douglas and translog PF with public capital as separate factor	Public capital stock	Elasticity is 0.3 (Cobb-Douglas), 0.4 (translog), but very sensitive to region and period
Delgado Rodriguez & Álvarez Ayuso (2000)	Spain, regions	1980-95	Cobb-Douglas Productive capital stock (Factor model)	Km of roads, km of railway, no. of telephone lines lines, and so on	The results indicate that productive infrastructure encourages private investment and can be considered to be essential for economic growth.
Duggal et al. (1999)	USA, national	1960-89	PF, technology index is non-linear function of infrastructure and time trend	Public capital stock	Elasticity for infrastructure is 0.27
Everaert & Heylen (2004)	Belgian regions	1965-96	Translog PF. Using a general equilibrium model, they analyse labour market effects of public investment. As a by-product they estimate the output elasticity.	Public investments	Elasticity is 0.31
La Ferrara & Marcellino (2000)	Italy, regions	1970-94	Cobb-Douglas PF with physical capital stocks as separate input	Public capital stock	
Holtz-Eakin & Schwartz (1995)	US states	1971-86	Neo-classical growth model that separates adjustment effects from steady state effects	Infrastructure capital (transportation and communications) and public capital stock	Infrastructure has a negligible effect on output nowadays
Kamps (2006)	22 OECD countries	1960-2001	Aschauer (1989) model for individual countries and panel	Public capital stock	Elasticity is 0.22 in panel, but much higher in time-series models

table continues on next page…

Table 5.2. (continued)

Study	Countries	Sample	Specification	Public capital variable	Conclusion
Kemmerling & Stephan (2002)	87 large German cities	1980, 1986 and 1988	Cobb-Douglas PF combined with policy equation for transport infrastructure and investment function for private capital		Rate of return on infrastructure is 16%. Political colour is important determinant for receiving grants
Ligthart (2002)	Portugal	1965-95	Cobb-Douglas PF, with and without CRS	Public capital stock	Positive and significant output effects of public capital
Seung & Kraybill (2001)	Ohio	Calibrated on 1990	Computable general equilibrium model with congestion adjusted infrastructure as third factor in Cobb-Douglas PF	Public capital stock	Welfare effects of infrastructure are non-linear
Shioji (2001)	US states and Japanese regions	US: 1963-93 & Japan: 1955-95, 5 year interval	Computable general equilibrium model with public capital in the technology term of a Cobb-Douglas PF	Public capital stock	Elasticity between 0.10 and 0.15
Stephan (2000)	W-German and French regions	Germany: 1970-95, France: 1978-92	Cobb-Douglas PF with public capital as separate factor and translog PF	Infrastructure capital stock (transportation)	Cobb Douglas gives elasticity of 0.11. Translog specification runs into multicolinearity problems.
Stephan (2003)	West-German regions (11)	1970-96	Cobb-Douglas PF with public capital as separate factor	Infrastructure capital (transportation and communications)	Elasticity between 0.38 (first differences) and 0.65 (loglevels)
Vijverberg et al. (1997)	US, time series	1958-89	Cobb-Douglas and semi-translog	Net stock of non-military equipment in the hands of the government	

Table 5.3. Studies using some kind of cost/profit function approach

Study	Countries	Sample	Specification	Public capital variable	Conclusion
Bonaglia et al. (2000)	Italy, regions	1970-94	Cobb-Douglass variable cost function	Public capital stock	Inconclusive, no good measure of the social user cost of public capital available
Boscá et al. (2000)	Spain, regions	1980-93	Generalized Leontief	Infrastructure capital stock (transportation, communication, general purpose)	Elasticity is 0.08
Canaleta et al. (2002)	Spain, regions	1964-91	Flexible cost function	Infrastructure capital (transportation) and public capital stock	Public capital reduces private production costs, public and private capital factors are complementary. Spillovers exist in Spain
Cohen & Morrison Paul (2004)	US, States	1982-96	Generalized Leontief	Public highway stock constructed using perpetual inventory method	Infrastructure investment reduces own costs and increases cost reducing effect of adjacent states
Demetriades & Mamuneas (2000)	12 OECD countries	1972-91	Quadratic cost function	Public capital stock	Output elasticity varies from 2.06 (Norway) to 0.36 (UK)
La Ferrara & Marcellino (2000)	Italy, regions	1970-94	Cobb-Douglas and generalized Leontief with physical capital stocks as separate input	Public capital stock	
Mamatzakis (1999a)	Two digit Greek industries (20)	1959-90	Translog cost function	Infrastructure capital stock (transportation, communication)	Cost saving impact of public infrastructure ranges from 0.02% in food manufacturing to 0.78% in wood and cork
Moreno et al. (2003)	Spain, regions and sectors	1980-91	Translog cost function	Infrastructure capital stock (transportation, communication, general purpose)	Public and private investments increase efficiency
Vijverberg et al. (1997)	US	1958-89	Translog cost and profit functions	Net stock of non-military equipment in the hands of the government	

Table 5.4. Summary of VAR/VECM studies

Study	Countries	Sample[1]	Specification	Variables[2]	Public capital variable	Conclusion
Agénor et al. (2005)	Egypt, Jordan, Tunesia	1965–2002(A)	VAR	I^G/Y, I^K/Y, growth Y, private sector credit/Y, growth r, G	Public capital stocks	There is a weak effect, short-lived and usually insignificant effect of public capital on private capital
Batina (1998)	US	1948–93 (A)	VAR and VECM	Y, L, different types of G and K	Public capital stock	Public capital has long-lasting effects on output and vice-versa
Belloc & Vertova (2006)	7 highly indebted countries	1970–99, some sub–sample (A)	VECM	Y, I^G, I^K	Public investment	In 6 of the 7 cases there is a positive effect of public investment on output
Crowder & Himarios (1997)	US	1947–89 (A)	VECM	Y, K, G, L, E	Public capital stock	Public capital is at the margin slightly more productive or as productive as private capital
Everaert (2003)	Belgian regions	1953–96 (A)	VECM	Y, K, G	Public capital stock	Output elasticity of public capital is 0.14, which is only a fraction (0.4) of output elasticity of private capital
Flores de Frutos et al. (1998)	Spain	1964–92 (A)	VARMA (first dif's log levels)	Y,K,G,L	Infrastructure capital (transport and communications)	Transitory increase of public capital growth implies a permanent increase of output, private capital and employment
Ghali (1998)	Tunesia	1963–93	VECM	Y, I^G, I^K	Public investment	Public investment has a negative effect on growth
Kamps (2004)	22 OECD countries	1960–2001 (A)	VECM	Y, K, G, L	Public capital stock	For majority of countries there is a positive and significant effect on growth
Ligthart (2002)	Portugal	1965–95 (A)	VAR, log levels	Y, K, G, L	Public capital stock	Positive output effects of public capital

table continues on next page. . .

Table 5.4. (continued)

Study	Countries	Sample	Specification	Variables	Public capital variable	Conclusion
Mamatzakis (1999b)	Greece	1959–93	VECM	Y, K, G, L	Public capital stock	Positive effect of public capital on productivity, no reverse effect
Mittnik & Neumann (2001)	Canada, France, UK, Japan, Netherlands, and Germany	Different periods per country (Q)	VECM	Y, I^G, C^G, I^K	Public investment	Weak positive output effect of infrastructure, public investment induces private investment; no reverse causation from GDP to public capital
Pereira (2000)	US	1956–97 (A)	VAR, first dif's log levels	Y, I^G, I^P, L	Public investment (different types)	Positive effect through crowding in of private investment
Pereira (2001)	US	1956–97 (A)	VAR, first dif's log levels	Y, I^G, I^P, L	Public investment (different types)	All types of public investment are productive, but core infrastructure displays the highest rate of return
Pereira & Andraz (2003)	US (sectoral and national)	1956–97 (A)	VAR, first dif's log levels	Y, I^G, I^P, L	Public investment	Public investment positively affects private investment, employment and output
Pereira & Flores de Frutos (1999)	US	1956–89 (A)	VAR, first dif's log levels	Y, K, G, L	Public capital stock	Public capital is productive, but substantially less than suggested by Aschauer (1989)
Pereira & Roca-Sagales (1999)	Spain (regional and national)	1970–89 (A)	VAR, first dif's log levels	Y, K, G, L	Infrastructure capital (transport and communications)	Positive and significant long-run effects on output, employment and private capital
Pereira & Roca-Sagales (2001)	Spain (sectoral and national)	1970–93 (A)	VAR, first dif's log levels	Y, K, G, L	Infrastructure capital (transport and communications)	Positive and significant long-run effects on output, employment and private capital
Pereira & Roca-Sagales (2003)	Spain (regional and national)	1970–95 (A)	VAR, first dif's log levels	Y, K, G, L	Infrastructure capital (transport and communications)	Positive and significant long-run effects on output, employment and private capital

table continues on next page. . .

Table 5.4. (continued)

Study	Countries	Sample	Specification	Variables	Public capital variable	Conclusion
Pina & St. Aubyn (2005)	Portugal	1960–2001 (A)	VAR, first dif's log levels	Y, K, G, H	Public capital stock	Without feedback effects, the rates of return on public investment are higher then on private investment
(Pina and St. Aubyn, 2006)	US	1956–2001 (A)	VAR, first dif's log levels	Y, K, G, L	Public capital stock	Taking crowding out of private investment into account lowers rate of return on public investment from 7.3% to 4%
Sturm et al. (1999)	Netherlands	1853–1913	VAR, levels	Y, I^G, I^K	Public investment	Positive and significant short-run effect; no long-run effect
Voss (2002)	US and Canada	US: 1947–88 (Q) Canada: 1947–96 (Q)	VAR, 11 lags first dif's	$Y, p^G, p^K, r, I^G/Q, I^K/Q$	Public investment as share of output	Public investment tends to crowd out private investment

[1] A = Annual, Q = Quarterly.

[2] Y = real GDP (output), K = private capital stock, G = public capital stock, L = number of employed persons, p^G = relative price of public investment, p^K = relative price of private investment, r = real interest rate, I^G = public investment, I^K = private investment, C^G = public consumption, E = energy price.

Table 5.5. Summary of cross section studies

Study	Countries	Sample	Public capital variable	Conclusion
Devarajan et al. (2000)	43 LDCs		Transportation and communication expenditure	Significant negative effect
Esfahani & Ramíres (2003)	75 countries	1965–95 (three decades)	Growth rates of telephones and power production per capita	Significant positive impact
Gwartney et al. (2006)	86 countries of which 66 LDCs	1980–2000	Public investment/GDP	Significant positive effect, but coefficient is less than coefficient of private investment
Milbourne et al. (2003)	74 countries	1960–85	Public investment as share of GDP, total and disaggregated into 6 sectors	Not significantly different from zero in steady state model; in transition model with IV also not significantly different from zero
Miller & Tsoukis (2001)	Varying per decade from 30 to 91	1960–89 (three decades)	Public investment/GDP	Government investment appears to be an important factor in growth
Sanchez-Robles (1998)	57 countries and 19 Latin American countries	1970–85 and 1980–92 (large sample); 1970–85 (small sample)	Index of physical units of infrastructure at beginning of the sample	Significant positive effect

Chapter 6

Public Capital and Private Productivity: The Long-Run Effect

6.1 Introduction

The previous chapter showed that estimates of the impact of infrastructure investment on economic growth differ substantially depending on the countries and time period covered, the level of aggregation and the econometric methodology employed. This chapter attempts to arrive at a set of robust estimates of the output elasticity of public capital using internationally comparable aggregate and industry data for a considerable number of developed countries and a substantial number of years. State-of-the-art econometric methods are used to counter many of the criticisms raised against earlier studies. In particular, the pooled mean group estimator (PMG) that is used allows for the identification of the long-run effect of infrastructure on productivity.

As discussed in the previous chapter, many studies use growth rates to identify the effect of infrastructure on productivity, thereby destroying the long-run relationship, while infrastructure investment mostly consists of projects with long dur-

This chapter is based on joint work with Robert Inklaar and Jan Egbert Sturm, 'Public Capital and Private Productivity: In search of the long-run effect', mimeo.

ations and long-run effects.[1] The PMG estimator proposed by Pesaran et al. (1999) avoids this problem by identifying the long-run relationship between variables in an error-correction framework. In cross-country and cross-industry estimates, efficiency gains are possible by restricting the parameter of interest to be equal across countries and/or industries. However, the PMG estimator only restricts the long-run parameter to be the same across countries or industries, while allowing for heterogeneity in the adjustment to this long run. In recent years, this estimator has been successfully applied in a number of studies on various issues, such as the effect of information and communication technologies (ICT) on growth (O'Mahony and Vecchi, 2005), the influences of human capital on growth (Bassanini and Scarpetta, 2002), and the impact of foreign direct investment (FDI) on growth (Ruschinski and Sturm, 2004).

The PMG estimator is first used to test the effect of infrastructure on productivity in an aggregate dataset for 21 OECD countries. However, one concern is that heterogeneity in the short-run adjustment may not be sufficient to arrive at robust elasticity estimates because of differences in the effect of infrastructure across industries. Indeed, a number of recent industry studies find such differences.[2] Furthermore, the impact of infrastructure may vary systematically across industries depending on the amount of transport equipment used by an industry. Fernald (1999) was the first to use this idea by postulating an industry production function with transport services as an input, rather than public capital. Transport services in turn depend on both the amount of transport equipment owned and the amount of public capital. His specification allows for heterogeneity across industries and at the same time tests whether more vehicle intensive sectors benefit more from extra public investment then less vehicle intensive sectors.

This chapter contains estimates of the impact of infrastructure based both on aggregate and industry data. Even though data and econometric methodology are state-of-the-art and counter many, if not all, criticisms raised in this literature, stable output elasticity estimates are elusive. Indeed, the estimated parameters vary wildly between equally plausible econometric specifications and range between -2 and 2. The aggregate estimates tend to be more stable, but even here, output elasticities range between 0.04 and 1.13. While it is hard to discount cross-country variation, the cross-specification variation we find suggests extreme sens-

[1] An exception is the work by Canning and Pedroni (1999). They find that imposing a common long-run effect across industries and types of public capital is not justified. However, this is not surprising given their sample of both developed and developing countries.

[2] See e.g. Fernald (1999), Mamatzakis (1999a), Pereira and Andraz (2003) and Pereira and Roca-Sagales (2001).

itivity to conceptually innocuous specification choices. Overall, this suggests that production function estimates of the impact of infrastructure are not well-suited to be used for infrastructure policy recommendations.

This chapter continuous as follows. In the next section we describe our theoretical framework and the estimation procedure in more detail. The data are discussed in Section 6.3 and Section 6.4 summarises the empirical results. Section 6.5 offers some concluding remarks.

6.2 Methodology

6.2.1 Basic production function framework

The starting point of our analysis is an industry specific long-run production function with public capital (G) as a third input besides private capital (K) and labour (L) and Hicks neutral technology (U)

$$Y_i = U_i F^i \left(K_i, L_i, G\right). \tag{6.1}$$

Assuming a Cobb-Douglas production function and taking natural logarithms gives

$$y_i = u_i + \alpha_{Ki} k_i + \alpha_{Li} l_i + \alpha_{Gi} g + \epsilon_i, \tag{6.2}$$

where lowercase variables denote the logs. The data is given for time periods $t = 1 \ldots T$ and industries $i = 1 \ldots N$ (time subscripts are omitted). The parameter u_{it} is a proxy for the state of technology. The parameters α_K, α_L and α_G are the elasticities of the input factors, and ϵ_{it} is a time-sector specific error term. If we postulate constant returns to private inputs (capital and labour), i.e. $\alpha_K + \alpha_L = 1$, we get the well-known equation (see Equation (5.6) in Section 5.3 on page 155)

$$y_i - k_i = u_i + \alpha_{Li}(l_i - k_i) + \alpha_{Gi} g + \epsilon_i \tag{6.3}$$

We use a constant and a time trend as a proxy for technology. In this chapter we are only interested in the long-run effects of public capital, so we interpret this equation as a long-run production function. The estimation procedure that we use, the pooled mean group estimator, is specifically developed for this purpose so capacity utilisation correction terms as used by e.g. Aschauer (1989) are not necessary. Equation (6.3) is the most common model in the literature; various authors (e.g. Aschauer, 1989 and Kamps, 2006) have estimated this equation at a macroeconomic or state level. However, to our knowledge, we are the first to use this particular specification at the industry level.

If we assume cost minimisation and price taking in factor markets, we can replace the output elasticity of labour (α_L) by the observed input share of labour, s_L

$$y_i - (1 - s_{Li})k_i - s_{Li}l_i = u_i + \alpha_{Gi}g + \epsilon_i. \tag{6.4}$$

As explained in the previous chapter, estimating the output effect of public capital using Equation (6.3) or (6.4) is problematic, in particular due to the non-stationarity of the dependent and independent variables, forcing most authors to estimate in first differences. However, the PMG estimator also generates consistent estimates if the series are I(1). This allows us to estimate (6.3) and (6.4) in levels.

6.2.2 Inter-industry heterogeneity

Equations (6.3) and (6.4) suffer from one major drawback. To use the cross section dimension of the data to gain efficiency, we have to assume equal long-run public capital parameters; that is, the α_G parameter is equal for all countries and/or industries. This restriction is unlikely to hold. For example the transport and storage sector will probably benefit more from extra public capital than an industry like financial intermediation.[3] To allow for heterogeneity between sectors we use the production framework proposed by Fernald (1999). This method allows for different output elasticities of infrastructure across industries, but assumes that this variation is perfectly correlated with the amount of transport equipment owned by each industry. This restriction allows us to obtain a more efficient estimator of the sector-specific output elasticity of public capital without unduly restricting the output elasticity.

In Fernald's model, public capital does not enter the production function directly as a separate input, but via transport services (T). Transport services are produced within the sector and depend on the stock of vehicles (V) and available public capital. Again we assume that technological process is Hicks-neutral, so we can write the long-term production function as

$$Y_i = U_i F^i \left(K_i^N, L_i, T(V_i, G) \right), \tag{6.5}$$

[3] In this chapter, public capital and infrastructure capital are used interchangeably as most public capital consists of infrastructure capital (see Section 5.2.3, page 152). Results based on different capital measures are discussed below.

with K^N non-vehicle capital. As before, price taking in factor markets and cost minimization are assumed as well as constant returns to private production factors. Under these assumptions, the output elasticity of, for example, vehicles is equal to the cost share of vehicles

$$s_{v,i} = \frac{dY}{dV_i}\frac{V_i}{Y_i} = \frac{dY}{dT}\frac{dT}{dV_i}\frac{V_i}{Y_i}. \tag{6.6}$$

Ultimately we are interested in the output elasticity of transport capital G. Using the separability in the production function allows us to rewrite this elasticity as

$$\frac{dY}{dG}\frac{G}{Y_i} = \frac{dY}{dT}\frac{dT}{dG}\frac{G}{Y_i} = \left[\frac{dT/dG}{dT/dV}\frac{G}{V_i}\right]\left[\frac{dY}{dT}\frac{dT}{dV}\frac{V_i}{Y_i}\right] = \phi_i s_{v,i}. \tag{6.7}$$

The parameter ϕ is the output elasticity of transport capital relative to the output elasticity of vehicles. This parameter links the unobserved output elasticity of public capital to the observed cost share of vehicles. We assume that the production function that transforms vehicles and infrastructure capital into transport services is of a Cobb-Douglas type with the same coefficients for all sectors

$$T(V_i, G) = A_i V_i^{\beta_V} G^{\beta_G}. \tag{6.8}$$

Substituting the first derivatives of T with respect to V and G in (6.7) shows that ϕ can be written as

$$\phi_i = \frac{dT/dG}{dT/dV} GV = \frac{\beta_G}{\beta_V}, \tag{6.9}$$

which implies that ϕ is the same in all sectors. Note that in Equation (6.9) only the output elasticities had to be assumed constant across sectors, while the transformation technology A may differ across industries.

Returning to the original production function and postulating a Cobb-Douglas specification with industry-specific technology, the production function in log levels becomes

$$y_i = u_i + \alpha_{Ki} k_i^N + \alpha_{Li} l_i + (\alpha_{Ti}\beta_V) v_i + (\alpha_{Ti}\beta_G) g + \epsilon_i. \tag{6.10}$$

Equation (6.10) closely resembles Equation (6.2), except that (private) capital is split between vehicle (v) and non-vehicle capital (k^N). However, if we replace output elasticities of the private inputs by observed, industry-specific shares that are constant over time and use Equation (6.9) we get

$$y_i - s_{Ki}k_i^N - s_{Li}l_i - s_{Vi}v_i = u_i + \phi(s_{Vi}g) + \epsilon_i. \tag{6.11}$$

The term within brackets on the right hand side is fully observable and the transport services technology A is absorbed into the general technology term due to the Cobb-Douglas production function specification. The key point to note about Equation (6.11) is that the estimated coefficient ϕ is constant over industries, allowing us to pool all series and estimate the one coefficient efficiently. However, a constant ϕ implies a different impact of public capital across industries, but this effect is assumed to vary depending on the amount of vehicle capital. Fernald (1999) originally proposed this identification scheme, although he estimated Equation (6.11) in first differences. Denoting the left hand side of (6.11) by *mfp* (multifactor productivity) and its first difference by Δmfp we get

$$\Delta mfp_i = \phi(s_{Vi}\Delta g) + \epsilon_i. \tag{6.12}$$

Equation (6.12) links sector specific multifactor productivity growth to public investment. There is a causality problem since higher productivity growth possibly leads to higher public investment. A simple correction (Fernald (1999, pp 622-3) for details) is to de-mean the series, i.e. to estimate

$$\Delta mfp_i - \overline{\Delta mfp} = \phi(s_{Vi} - \bar{s}_V)\Delta g + \epsilon_i. \tag{6.13}$$

Under the assumption that only higher average productivity growth has a potential effect on public investment, the estimator of ϕ using Equation (6.13) will be consistent.[4]

6.2.3 Estimation procedure

There are various ways to estimate the panel equations (6.3), (6.4), (6.11) and (6.13). One way, used by Fernald (1999), is the Seemingly Unrelated Regression (SUR) proposed by Zellner (1962). SUR allows the contemporaneous error covariances to be freely estimated, but it neglects further similarities between sectors. The main drawback, however, is that the number of series (N) must be smaller than the length

[4] Equation (6.13) is a simplified version of Fernald's (1999) estimated equation. His more complex equation also allows for different cyclicality of the industries, i.e. it allows industries to be more responsive to aggregate productivity shocks. However, our more restrictive representation does capture the main idea and does not change the interpretation and sign of the key parameter ϕ.

of each series (T). In our industry level analysis we have 22 years of data on 24 industries per country and 31 years of data for 21 countries on the macro level. Clearly we cannot use SUR estimation directly. Fernald solves this problem by grouping comparable sectors or by selecting only a subset, thereby reducing the number of series. While this is likely to improve estimation efficiency, it reduces cross-section variation. Reducing the cross-sectional dimension and thus variation is strange if one considers that identification in Fernald's model depends on this variation.

Another approach is to estimate a fixed effects panel data model. The main problem with this approach is that these models only allow for different intercepts while all other coefficients are assumed to be constant across all series, both in the short run as in the long run. Especially in Equations (6.3), (6.4) this assumption is troublesome, because it is violated on theoretical grounds since the coefficients of the inputs are all industry-specific.

As a third option we could estimate a separate equation for each industry and examine the distribution of the estimated coefficients across series (countries or industries). The mean of the estimates, often called the Mean Group (MG) estimator will then be of particular interest. However, this procedure does not take into account that some of the parameters might the same across groups and one could obtain a more efficient estimator by using this extra information.

We will use an intermediate procedure proposed by Pesaran et al. (1999) and referred to as the Pooled Mean Group (PMG) estimator. This estimator, based on a maximum likelihood approach, constrains the long-run coefficients to be identical, but allows the short-run coefficients and error variances to differ across industries, thereby allowing for short-run heterogeneity between sectors. As long as the variables are I(0) or I(1), the PMG estimators are consistent and the asymptotic distribution of the PMG estimator can be derived.

To illustrate the PMG method, we start with the long-run relationship in Equation (6.4). If all variables are I(1)[5] and cointegrated, then the error term in the estimating equations is stationary. Instead of estimating the long-run relationships directly, the PMG estimates the error equation of the autoregressive distributed lag (ARDL) representation. For notational convenience we derive the equations here with one lag in the dependent and explanatory variables for all series. In the actual

[5] Kamps (2006) investigates whether these series are actually I(1) on a macro level for 22 OECD series. If each series is tested individually, he comes to the conclusion that this is probably not true. However, unit root tests do suffer from low power to discriminate between unit root and near unit root processes, especially for small samples. Panel unit root tests do not reject the null hypothesis that the variables are I(1).

estimation procedures, we allow the lags to differ between series and between the dependent and explanatory variables (see Pesaran et al., 1999 for a more general discussion). The ARDL equation is

$$mfp_{it} = \lambda_1 mfp_{i,t-1} + \delta_{i,0} g_{it} + \delta_{i,1} g_{i,t-1} + \beta_i + \gamma t + \epsilon_{it}, \tag{6.14}$$

with the left hand side

$$mfp_i = y_i - (1 - s_{Li})k_i - s_{Li} l_i \tag{6.15}$$

and where we have imposed a sector specific intercept β and a time trend γt. Rewriting gives the equivalent, but notationally more convenient error correction equation

$$\Delta mfp_{it} = (1 - \lambda_1) mfp_{i,t-1} + (\delta_{i,0} + \delta_{i,1}) g_{it} - \delta_{i,1} \Delta g_{it} + \beta_i + \gamma t + \epsilon_{it}, \tag{6.16}$$

This error correction equation implies the following long-run relation

$$mfp_{it} = \frac{\delta_{i,0} + \delta_{i,1}}{1 - \lambda_1} g_{it} + \beta_i' + \gamma' t + \epsilon_{it}. \tag{6.17}$$

Imposing equal parameters in the long run gives

$$\frac{\delta_{i,0} + \delta_{i,1}}{1 - \lambda_1} = \phi, \tag{6.18}$$

so the ratio of the error correction parameters are restricted to be equal across series. The PMG estimator is the likelihood maximising value of θ. Estimating a restricted version of (6.16) instead of the long-run Equation (6.3), (6.4), (6.11) and (6.13) directly allows the parameters to differ across series in the short run. Another advantage of (6.16) over the non-dynamic estimating equations is that including lagged variables solves problems of possible autocorrelation.

Tests of homogeneity of error variances and short-run or long-run slope coefficients can be easily carried out using Likelihood Ratio tests, since – like fixed effects estimators – the PMG estimator is a restricted version of the set of industry-specific estimators. Although it is common to use pooled estimators without testing the implied restrictions, in case of cross country studies, the Likelihood Ratio tests has a tendency to reject the restrictions at conventional significance levels.[6] In

[6] Pesaran et al. (1999) explain this by pointing at possibly omitted group specific variables that are correlated with the regressors causing the group specific estimates to vary wildly. When estimating

our industry-level application, the LR test also rejects the null hypothesis of equal long-run coefficients.

Pesaran et al. (1999) suggest using a Hausman (1978) type test instead of a Likelihood ratio test. The MG estimator provides consistent estimates of the mean of the long-run coefficients, though these will be inefficient if slope heterogeneity holds. Under long-run slope homogeneity, the pooled estimators are consistent and efficient. Therefore, the effect of heterogeneity on the means of the coefficients can be determined by a Hausman-type test applied to the difference between the MG and the PMG or the fixed effects estimators.

6.3 Data

Our data are annual. The macro level data cover the period 1960–2001, the industry level data start in 1979. The national public and private capital data are taken from Kamps (2006). He provides international comparable capital stock estimates for 22 OECD countries (see the first column of Table 6.1 for a complete list).[7] The capital stock estimates are calculated using the perpetual inventory method, assuming geometric depreciation. Kamps' concept of public capital includes all capital, not just infrastructure. This dataset has the main advantage that the capital estimates are constructed in a consistent manner, so diverging results for different countries are not caused by merely different concepts of public capital. To be certain that this broad definition of public capital does not influence the results, we also collected data specifically on infrastructure investment for the six countries for which we have industry data. For this set of countries (except Australia), we also gathered data on road length as an admittedly crude, but very straightforward measure of infrastructure.[8]

For data on output, and industry specific capital and labour, we rely on the GGDC database (Inklaar et al., 2006). The second column of Table 6.1 presents the 24 industries included in our analysis.[9] Unfortunately, the industry level data is not

equations for a large number of groups it is not possible to include all group specific variables. However, if the coefficients are really the same and the bias inducing correlations are not systematic, then pooled estimation will be appropriate.

[7] We excluded Switzerland from the original 22 country dataset due to data availability, leaving 21 countries.

[8] This data is available upon request.

[9] The original dataset covers 26 industries. We exclude two of them; petroleum and coal products and non-market services. The first because value added of this sector is mainly driven by the exogenous oil price, the second because of measurement problems with the output of this sector and because public capital is included in the estimation of output in this sector by the statistical offices.

available for all countries in the macroeconomic dataset for the full sample. Due to limited data availability, we had to restrict our industry level analysis to the time period 1979–2001 and to a subset of six countries, i.e. the Netherlands, Germany, France, the UK, Australia and the United States. We used hours worked as a proxy for labour input and value added as the measure of output. The capital data is divided into growth rates for six types of capital assets, one of which is transport equipment. These six specific capital growth rates are aggregated using the two-period average share of each asset type in total nominal capital compensation (see Inklaar et al., 2006 for details).

$$\Delta \ln K_t = \sum_j \bar{V}^K_{j,t} \Delta \ln K_{j,t},$$

where $\bar{V}^K_{j,t}$ is the two-period average share of asset type j in total nominal capital compensation. This aggregate capital growth rate is finally used to construct a private capital services index.

Of particular interest is the transport capital input or vehicle share, since this is used to calculate the output elasticity of public capital in equations (6.11) and (6.13). Table 6.2 shows this vehicle share per industry per country. The average input share is 2.7%, and similar for all countries, except Australia (4.23%). Australia also has a different distribution over industries. For example, the transport and storage sector, in most other countries the most vehicle intensive sector by far, has only a vehicle share of 5%. In contrast, the construction sector in Australia spends 13.2% of inputs on transport capital, far more than the same sector in the other countries.

6.4 Results

For all estimations we use a three-stage estimation procedure. In the first stage, we determine the optimal number of lags for each series chosen by the Schwarz-Bayesian information criterium (SBC) subject to a maximum lag order of three. All results are robust for alternative selection criteria like Akaike's Information Criterium (AIC) and the Hanna-Quinn Information Criterium (HQ). In the second, stage we search for a suitable initial value of the public capital coefficient. If any output elasticities of other inputs are included in the estimating equation, we fix them temporarily to their observed cost shares. With the other output elasticities fixed, we use a grid search over the relevant domain to obtain an initial value for the public capital parameter in the optimisation algorithm. For the basic estimating

Table 6.1. Countries and industries included in the empirical analysis

Countries	**Industries**
Australia[1]	Agriculture, forestry & fishing
Austria	Mining & quarrying
Belgium	Food products
Canada	Textiles, clothing & leather
Denmark	Wood products
Finland	Paper, printing & publishing
France[1]	Chemical products
Germany[1]	Rubber & plastics
Greece	Non-metallic mineral products
Iceland	Metal products
Ireland	Machinery
Italy	Electrical & optical equipment
Japan	Transport equipment
Netherlands[1]	Furniture & misc. manufacturing
New Zealand	Electricity, gas & water
Norway	Construction
Portugal	Wholesale trade
Spain	Retail trade
Sweden	Hotels and restaurants
United Kingdom[1]	Transport & storage
United States[1]	Communications
	Financial intermediation
	Business services
	Social & personal services

[1] Included in the industry level analysis

equations (6.3) and (6.4) we search over the interval from -0.10 to 0.50; this covers most point estimates of the public capital output elasticity reported in other studies (see Table 5.2 on page 171). For the extended estimating equations (6.11) and (6.13) the interpretation of the public capital coefficient is different, the product of this coefficient and the vehicle share gives the output elasticity of public capital. The average vehicle share is 2.72% so we use the interval from -4 to 18.[10]

[10] Other studies using the PMG estimator (e.g. Pesaran et al., 1999; O'Mahony and Vecchi, 2005) use the mean group estimator as initial values in the optimisation step. This has the drawback that the mean group estimator might be very far from the likelihood maximising value and that the optimisation algorithm (usually of the Gauss-Newton type) might get stuck at a local maximum. Our grid search prevents this.

Table 6.2. Transport share per industry (%).

Industries	All countries	Netherlands	France	Germany	UK	Australia	USA
Agri., forestry & fishing	6.99	2.56	0.74	0.60	2.90	14.37	9.96
Mining & quarrying	3.00	0.26	0.67	0.71	1.54	3.17	3.91
Food products	1.60	1.72	0.77	1.40	1.75	3.60	1.61
Textiles, clot. & leather	0.60	0.67	0.44	0.65	1.37	1.96	0.43
Wood products	1.54	1.19	1.74	1.04	3.54	3.67	1.30
Paper, print. & publishing	1.43	0.85	1.43	1.30	1.14	3.70	1.39
Chemical products	1.28	0.42	3.34	0.96	1.12	4.66	1.05
Rubber & plastics	0.99	0.62	2.82	1.22	0.91	4.04	0.40
Non-metallic mineral prod's	1.75	1.94	1.20	1.55	2.18	3.83	1.67
Metal products	0.94	0.89	1.79	0.99	0.67	3.69	0.57
Machinery	0.97	1.01	0.93	0.95	2.21	2.36	0.78
Electrical & optical equip.	0.48	0.23	0.85	0.69	1.17	0.47	0.28
Transport equipment	0.68	1.74	0.65	1.07	0.22	1.71	0.53
Furniture & misc. man.	0.78	0.68	1.01	1.10	2.88	1.85	0.41
Electricity, gas & water	1.89	0.30	2.19	1.44	0.76	2.38	2.09
Construction	6.87	3.39	4.46	4.63	8.99	13.17	7.42
Wholesale trade	4.37	5.91	3.12	4.12	8.79	4.35	4.13
Retail trade	1.39	2.29	0.36	1.07	2.42	2.98	1.39
Hotels and restaurants	1.00	1.41	0.22	0.63	1.18	3.24	1.02
Transport & storage	10.85	17.34	10.89	4.45	12.99	5.06	12.57
Communications	3.28	0.89	1.10	1.20	0.94	3.48	4.11
Financial intermediation	1.47	1.01	0.57	1.08	3.64	4.92	1.22
Business services	7.20	9.96	8.62	15.59	4.85	4.53	4.97
Social & personal services	1.58	0.87	0.38	1.41	2.80	3.26	1.61
Average	2.19	2.42	2.09	2.08	2.96	4.19	2.70

Note: the table shows the share of transport equipment capital compensation in value added, averaged over the period 1979-2001.

In the final stage we use a Newton optimisation algorithm to maximise the likelihood function, using the cost shares of the private inputs and the likelihood maximising coefficient of public capital from the second stage as initial values. Naturally we do not restrict the estimated coefficients in any way. In particular, we do not restrict the public capital coefficient to the interval used in stage two.

6.4.1 Output elasticities: country estimates

Table 6.3 shows the country-level estimates of the output elasticities of public capital. The first specification is Equation (6.3). It is the most restrictive, assuming constant returns to private inputs and a common output elasticity of private capital, labour and public capital for all countries. The first row in each block shows the estimation results with all 21 countries included. The estimated output elasticity of public capital is 0.56, with an asymptotic t-value of 5.87 (between brackets). The pooled error correction is -0.11 and also highly significant (t=-5.07). The mean group elasticity estimate is higher, 0.90, with a much higher variance, but still significantly different from zero at the usual significance levels. The last column gives the p-value of the Hausman test. It clearly shows that we have to reject the null hypothesis of equal long-run coefficients. In this specification it is mostly due to wildly varying private input coefficients (not shown).

The second row in the first block shows the estimates if we only include relatively small countries (all countries except the US, UK, France, Germany and Japan). The pooled mean group estimate is lower than the estimate for all countries; however, the mean group estimate is higher, 1.28. Again, the Hausman test shows that we must reject the null hypothesis of equal long-run parameters. The third and fourth row show the estimation results for a pool with only the five large countries and for the 15 European countries. Finally, the last row for each specification shows the estimation results for the six countries that are included in the industry-level analysis.

The second specification shows the same Equation (6.3), but we allow private input coefficients to vary by country. The output elasticity estimates vary between 0.64 and 1.13, although this might seem ridiculously high from an economic viewpoint, they are within the range found in other studies. The Hausman test gives no reason to reject the null hypothesis, mainly because the coefficient of labour is not restricted to be equal among countries in the long run.

Finally, the third specification shows the estimation results for Equation (6.4). Here we fix the private input coefficients to their observed cost shares. The estimated output elasticities all lie within the expected range, varying from 0.04 for the large countries to 0.15 for the small countries. All estimates, except for the large country group, differ from zero at a 5% significance level. As with the previous specification, none of the Hausman tests give reason to reject the null hypothesis of equal coefficients.

The estimated output elasticities in Table 6.3 vary considerably, especially between the various specifications, but also across the subsets of countries. Figure 6.1 shows this variation, the markers indicating the point estimates and the lines the 95% confidence intervals. As mentioned before, specifications 1 and 2 both generate relatively high output elasticities. Given the fairly stringent assumptions about output elasticities of private inputs, this might point to a misspecification. The vertical lines also indicate that the high estimates have high standard errors; the lower, more realistic estimates have smaller confidence intervals. Overall, the country-level estimates cover the entire range of estimates reported in other studies. Furthermore, the estimates based on specification 3 cover the range of estimates found in most modern, sophisticated studies. However, this range of 5-15 percent is still fairly wide and there is no independent information to verify whether the cross-country variation is reasonable.

6.4.2 Output elasticities: industry estimates

Basic estimation results

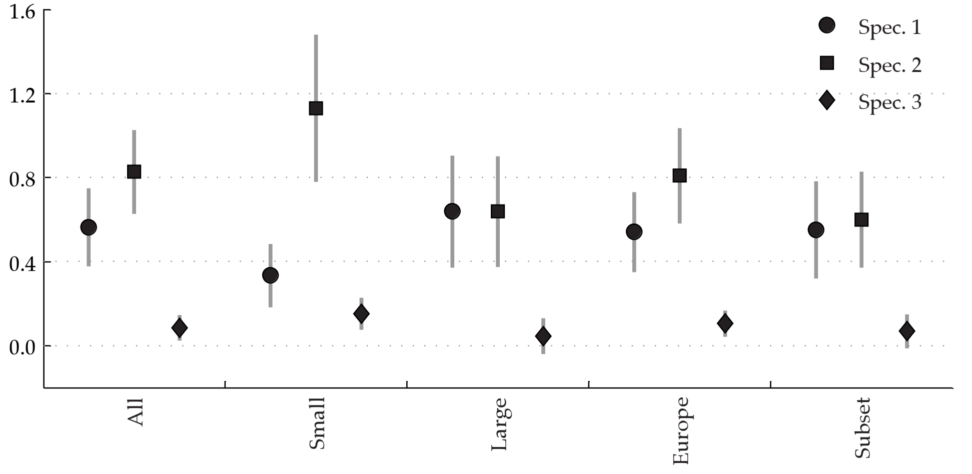
Table 6.4 shows the estimation results based on the same specifications as above. The first line for each specification shows the results for the complete industry panel. With all 24 industries in the six countries included, this estimation is more or less comparable to the 'Subset' group from Table 6.3. Comparing these lines from Table 6.3 and Table 6.4 shows positive estimates in Table 6.3 and negative coefficients in Table 6.4, raising questions about the robustness of the results. Furthermore, the country results in Table 6.4 show a range of output elasticity estimates of -2.38 to 1.53, depending on the country and the specification. This spread is very substantial compared to the spread for the country-level estimates from Table 6.3 of 0.04 to 1.13. A possible explanation is that within a country, some sectors benefit considerably more from public capital than others and restricting the public capital output elasticity to be equal across industries is not justified. However, the joint Hausman tests suggest this is not the case. Only five of the 18 tests reject the null hypothesis of equal long-run parameters and two of them are rejected not because

Table 6.3. Country-level estimates of the output elasticity of public capital

	Pooled Mean Group		Mean Group		Hausman
	Output elasticity	Error correction	Output elasticity	Error correction	p-val
Specification 1 (common labour elasticity)					
All	0.56 (5.87)	-0.11 (-5.07)	0.90 (2.94)	-0.22 (-7.05)	2.9%
Small	0.33 (4.33)	-0.19 (-5.57)	1.28 (3.20)	-0.24 (-5.68)	2.9%
Large	0.64 (4.67)	-0.16 (-2.04)	0.59 (1.35)	-0.18 (-2.27)	0.4%
Europe	0.54 (5.53)	-0.13 (-4.31)	1.16 (3.56)	-0.25 (-6.80)	0.3%
Subset	0.55 (4.66)	-0.17 (-2.48)	0.82 (4.77)	-0.24 (-4.52)	10.2%
Specification 2 (country specific labour elasticity)					
All	0.83 (8.12)	-0.18 (-5.80)	0.90 (2.94)	-0.22 (-7.05)	81.0%
Small	1.13 (6.30)	-0.17 (-4.61)	1.28 (3.20)	-0.24 (-5.68)	66.8%
Large	0.64 (4.73)	-0.19 (-2.56)	0.59 (1.35)	-0.18 (-2.27)	91.5%
Europe	0.81 (6.97)	-0.22 (-6.59)	1.16 (3.56)	-0.25 (-6.80)	24.8%
Subset	0.60 (5.11)	-0.24 (-4.97)	0.82 (4.77)	-0.24 (-4.52)	8.3%
Specification 3 (labour elasticity equal to cost share)					
All	0.08 (2.71)	-0.18 (-7.63)	-0.01 (-0.07)	-0.22 (-8.42)	46.2%
Small	0.15 (3.83)	-0.20 (-5.40)	-0.03 (-0.21)	-0.21 (-5.62)	23.7%
Large	0.04 (1.03)	-0.20 (-3.05)	-0.09 (-0.45)	-0.22 (-3.03)	47.7%
Europe	0.11 (3.32)	-0.20 (-6.35)	0.00 (0.03)	-0.23 (-6.70)	42.6%
Subset	0.07 (1.66)	-0.23 (-4.25)	0.04 (0.26)	-0.27 (-4.34)	84.0%

Notes: t-values are in parentheses. 'Hausman p-val' is the probability from testing whether the pooled mean group and mean group coefficients are equal across countries. Specification 1 corresponds to Equation (6.3) with a constant labour elasticity across industries (omitted from the table). Specification 2 corresponds to Equation (6.3) with varying labour elasticity across industries (omitted from the table). Specification 3 corresponds to Equation (6.4)

Figure 6.1. Estimated output elasticity of public capital in a panel of 22 countries, with 95% confidence intervals



Notes: 'Small' excludes US, UK, France, Germany and Japan, 'Large' only covers those countries. 'Europe' includes the 15 (old) EU member countries and 'Subset' includes Australia, France, Germany, Netherlands, UK and US. Spec. 1 corresponds to Equation (6.3) with a constant labour elasticity estimate across countries, Spec. 2 corresponds to Equation (6.3) with varying labour elasticity estimate across countries (omitted from the table). Spec. 3 corresponds to Equation (6.4).

of unequal public capital elasticities, but because of unequal labour elasticities.[11]

The results for France are quite sensitive to the inclusion of the sectors 'food products' and 'furniture and misc. manufacturing'. Without these two sectors, the public capital output elasticity estimates for the first and third specification are 0.18 and 0.26, still relatively high, but not as high as 1.29 and 1.52. Also the results for the UK are sensitive to the exclusion of two outliers. Dropping 'food products' and 'machinery' for the UK changes the first and second results to -0.28 and -0.02. Despite the large impact on the estimates of certain industries in certain countries, there is little reason to exclude these industries systematically. As in the case of the country-level estimates, the large variation of estimates across countries and specifications raises doubts about the usefulness of these estimates.

[11] The Hausman test p-values for equal public capital coefficients are 0.35 and 0.63 for all countries and the UK in the first specification. The Hausman test p-values for equal labour coefficients are 0.02 and 0.01.

Table 6.4. Industry-level estimates of the output elasticity of public capital

	Pooled Mean Group		Mean Group		Hausman
	Output elasticity	Error correction	Output elasticity	Error correction	p-val
Specification 1 (common labour elasticity)					
All countries	-0.23 (-2.87)	-0.55 (-16.46)	-3.73 (-1.00)	-0.80 (-16.64)	6.6%
Netherlands	-0.37 (-1.59)	-0.64 (-7.32)	0.28 (0.30)	-0.99 (-6.69)	76.7%
France	1.29 (5.23)	-0.39 (-4.62)	0.17 (0.10)	-0.85 (-5.14)	2.8%
Germany	0.12 (1.21)	-0.69 (-6.28)	0.42 (0.52)	-0.89 (-8.42)	86.4%
UK	-0.52 (-3.01)	-0.45 (-5.16)	-0.15 (-0.18)	-0.65 (-6.28)	2.4%
Australia	-2.38 (-4.31)	-0.48 (-7.44)	-22.55 (-1.01)	-0.56 (-9.38)	14.5%
USA	-0.66 (-4.62)	-0.79 (-8.56)	-0.55 (-0.58)	-0.89 (-10.71)	16.0%
Specification 2 (country specific labour elasticity)					
All countries	-0.04 (-0.84)	-0.71 (-16.93)	-3.73 (-1.00)	-0.80 (-16.64)	32.5%
Netherlands	-2.18 (-13.51)	-0.80 (-5.70)	0.28 (0.30)	-0.99 (-6.69)	0.9%
France	-1.10 (-5.78)	-0.75 (-4.90)	0.17 (0.10)	-0.85 (-5.14)	47.2%
Germany	0.25 (2.63)	-0.79 (-7.15)	0.42 (0.52)	-0.89 (-8.42)	83.1%
UK	0.06 (0.98)	-0.57 (-5.28)	-0.15 (-0.18)	-0.65 (-6.28)	79.4%
Australia	-1.75 (-3.69)	-0.58 (-9.01)	-22.55 (-1.01)	-0.56 (-9.38)	35.1%
USA	-0.70 (-5.04)	-0.82 (-9.58)	-0.55 (-0.58)	-0.89 (-10.71)	87.9%
Specification 3 (labour elasticity equal to cost share)					
All countries	-0.36 (-4.05)	-0.50 (-18.29)	-0.72 (-1.01)	-0.57 (-19.24)	61.0%
Netherlands	0.21 (0.83)	-0.52 (-8.83)	1.58 (1.94)	-0.65 (-8.57)	7.7%
France	1.53 (8.07)	-0.34 (-4.37)	1.22 (0.44)	-0.50 (-5.96)	91.0%
Germany	0.34 (2.41)	-0.57 (-7.53)	0.59 (0.98)	-0.62 (-8.48)	67.1%
UK	-0.84 (-4.83)	-0.42 (-7.04)	-3.14 (-1.70)	-0.49 (-7.81)	21.2%
Australia	-2.14 (-4.37)	-0.51 (-8.40)	-1.64 (-1.12)	-0.57 (-10.38)	71.9%
USA	-1.65 (-7.57)	-0.54 (-5.95)	-2.96 (-1.50)	-0.62 (-7.25)	50.4%

See notes table 6.3 for definitions.

Extended estimation results

The previous subsection showed that allowing for heterogeneity between industries within a country increases the variation in estimates across countries and specifications. Although the Hausman tests do not reject equal public capital coefficients in general, it might help if we allow for heterogeneity between sectors regarding the effect of public capital. Table 6.5 shows the estimation results for equations (6.11) and (6.13). The estimated coefficient of interest in these equations is ϕ, which can be multiplied by the average transport equipment share to find the corresponding output elasticity of public capital as in the previous tables. Specification 4 in Table 6.5 shows the results of Equation (6.11), the relation in levels. The last column shows the implied output elasticities of public capital and they are within the range of estimates in Table 6.4. Specification 5 shows the results of the estimating Equation (6.13), Fernald's equation in growth rates, de-meaned to account for possible reverse causality. The point estimate for ϕ for the USA, 9.17, lies well within the range reported by Fernald (1999), although we used a slightly different specification and a different estimation procedure. Also the results for Australia, the UK, and Germany are within the expected interval, although the implied output elasticities are relatively high, especially for the UK, Australia, and the USA. The estimate for the Netherlands is very low, -21.9, implying an output elasticity of -0.53. Moreover, this result does not seem to be driven by outliers. While in Table 6.4, the Hausman tests did not reject equality of output elasticities across industries, the Hausman tests in Table 6.5 also do not reject equality of ϕ across industries. So while Fernald's (1999) identification scheme is attractive on conceptual grounds, there are no statistical reasons to prefer either model.

The results in Tables 6.3–6.5 do not support even the most general conclusions about the impact of public capital on private productivity, since the output elasticity estimates range from negative to positive across specifications. As a further robustness check, we also collected data on infrastructure investment, rather than the broader category of public capital, and road length. However, these alternative capital measures only reinforce the finding that very few robust conclusions can be drawn from this type of regression analysis.

Although the detailed results are available upon request, we will use the estimated coefficients from these regressions, as well as the earlier ones to illustrate the more general conclusion about the robustness of output elasticity estimates of public and infrastructure capital. The previous tables have shown a large range of estimates, both across countries and across different specifications. We have no in-

Table 6.5. Extended industry level output elasticity estimates

	Pooled Mean Group				Hausman	Output
	$\hat{\phi}$		Error correction		p-val	elasticity
Specification 4, Fernald (1999) in log levels						
All countries	-10.90	(-3.39)	-0.49	(-18.03)	63.2%	-0.24
Netherlands	-9.31	(-1.38)	-0.52	(-8.66)	14.1%	-0.23
France	0.65	(0.07)	-0.38	(-5.97)	29.9%	0.01
Germany	26.27	(2.96)	-0.57	(-7.82)	32.8%	0.55
UK	-30.54	(-4.05)	-0.42	(-6.56)	14.7%	-0.90
Australia	-35.52	(-3.40)	-0.51	(-8.46)	79.2%	-1.49
USA	-64.60	(-4.42)	-0.52	(-6.30)	27.0%	-1.74
Specification 5, Fernald (1999) in growth rates						
All countries	1.93	(0.69)	-1.03	(-39.35)	7.3%	0.04
Netherlands	-21.85	(-2.95)	-0.99	(-18.12)	54.2%	-0.53
France	-2.93	(-0.30)	-0.94	(-19.41)	31.7%	-0.06
Germany	3.06	(0.56)	-1.04	(-20.34)	14.0%	0.06
UK	7.49	(0.99)	-0.94	(-13.18)	17.8%	0.22
Australia	4.61	(0.72)	-1.13	(-20.18)	96.0%	0.19
USA	9.17	(1.50)	-1.14	(-13.15)	3.3%	0.25

Notes: t-values are in parentheses. 'Hausman p-val' is the probability from testing whether the pooled mean group and mean group coefficients are equal across countries. Specification 4 corresponds to Equation (6.11) and Specification 5 corresponds to Equation (6.13).

dependent information to verify whether the cross-country variation in estimates is reasonable or not, but the cross-specification variation can be used, since each specification attempts to uncover the same underlying parameter, namely the output elasticity of public capital.[12]

Table 6.6 illustrates this cross-specification variability. The first row, 'country-level', is based on the elasticity estimates from Table 6.3. For each country group, the average across specifications is first subtracted to focus on the variation across specifications. From this set of 15 parameters, the 25th and 75th percentile is determined to get a relatively robust measure of the spread of the distribution and the inter-quartile range is determined.[13] So for a given country group, a change in

[12] In specifications 4 and 5 in Table 6.5, the output elasticity is not estimated directly, but is derived given the average share of transport capital in value added.

[13] There is overlap in country coverage between the different estimates, as for example, the 'all countries' as well as the 'small' and 'large' countries parameters are included. If anything, this is likely to reduce the inter-quartile range, since the 'all countries' parameter is a (weighted) average of the 'small' and 'large' parameter.

Table 6.6. Cross-specification variation in estimates of the output elasticity of public capital

	Inter-quartile range
Country-level	0.56
Industry-level, specifications 1–3	0.49
Industry-level, specifications 1–5	
Only public capital	0.62
Public capital, infrastructure and roads	0.75

Notes: The inter-quartile range compares the 75th to the 25th percentile of the parameters. The average across specifications for each country (group) is subtracted to put the estimates on a comparable basis. For the final row, this average is different for the definitions of capital. The first figure is based on the parameters in Table 6.3, the second on Table 6.4 and the third on Tables 6.4 and 6.5.

specification could change the elasticity estimate substantially, by 0.56 when moving from the 25th to the 75th percentile. While Figure 6.1 showed that specifications 1 and 2 gave systematically higher elasticity estimates than specification 3, the industry-level parameter estimates are more randomly distributed across specifications. In all cases, relatively minor changes in specification can lead to radically different elasticity estimates. Different specifications can be used to defend the thesis that public capital has no impact on output at all or the same impact as labour.

6.5 Concluding remarks

The studies surveyed in the previous chapter reached very different conclusions about the effect of public capital on productivity. Given the large differences across studies in methodology, data, country and time coverage, this chapter has examined the effect of public capital in a production function framework, using state-of-the-art econometric methodology and two data sets covering a broad range of countries, industries and years. In particular, we use the pooled mean group estimator (PMG) to find the long-run effect of public capital on productivity. Even within this framework, reliable estimates remain elusive, The question arises whether this type of analysis is ever likely to yield estimates that are useful for informing infrastructure investment decisions.

A number of different specifications are tested using country-level and industry-level data. Estimates vary wildly across specifications and country groups and individual countries (using industry data). The cross-country variation is hard to discount, as no independent information is available to confirm whether this variation is reasonable or not. However, the cross-specification variation raises more fundamental questions about using production function estimates to gauge the impact of infrastructure on productivity. In all cases, the regressions aim to find the same underlying output elasticity of public capital and there are no a priori reasons to discount certain specifications entirely. While some would prefer some specifications over others, such as allowing for country-specific output elasticities of labour, this does not seem like a solid basis for drawing radically different conclusions based on the same data and econometric methodology.

And the conclusions are indeed radically different: simply moving from one specification to another for a given country or country group may increase the elasticity estimate by as much as 0.6.[14] In comparison, in most (value added) production function estimates, the output elasticity of labour is about 0.6. Furthermore, it seems unlikely that advances in data or econometrics will improve this situation. Already, the industry data cover about a quarter of century and 24 industries. Furthermore, the longer country-level dataset (more than 40 years) does not fare much better than the industry estimates. The conclusion then has to be that public capital is not important enough for productivity to yield more than the occasional spurious correlation.

[14] Moreover, this range is not based on the extreme estimates but on the 25th and 75th percentiles of the estimates.

Chapter 7

Conclusion and Discussion

In this dissertation we have focused on two aspects of dynamic macroeconomics: the role of demographics and demographic change as one of the main determinants of intergenerational redistribution, and the impact of public capital on economic growth.

7.1 Realistic demographics in overlapping generations models

In Part I of this dissertation, we developed an extended version of the Blanchard-Yaari-type overlapping generations (OLG) model. We incorporated a general description of the mortality process, overcoming one of the main drawbacks of the standard model: the perpetual youth assumption. In Chapter 2 we showed that incorporating a realistic demographic structure is quite feasible as long as we restrict our attention to a small open economy facing a constant world interest rate. One of the most attractive features of our extended model is that at the level of individual households, a realistic description of the mortality process reinstates the classic life-cycle saving insights of Modigliani and co-workers. The added complexity does not destroy the main strength of simple theoretical models, we can still analytically track the effect of various macroeconomic shocks on the main variables in the model, both on the individual as on the aggregate level.

The recently developed tractable OLG models with realistic demographics (like Kalemni-Ozcan et al. (2000) and Boucekkine et al. (2002)) have not yet been applied to welfare analysis of policy shocks and exogenous shocks, this in contrast to the perpetual youth models (see e.g. Heijdra and Meijdam (2002) and Bovenberg

(1993)). Using our standard model in Chapter 2, we find that there are significant differences in welfare effects of different shocks between the perpetual youth models and models with a realistic mortality process. First of all, transition is much faster in models with realistic demographics than in models with the perpetual youth assumption, because expected remaining lifetimes are much lower. Second, the perpetual youth model neglects the fact that old generations do not value future income gains or losses as high as new or unborn generations, because the conditional survival function is equal for all cohorts. Finally, we have demonstrated that the demographic details do not 'wash out' at the aggregate level. The impulse-response functions for the different shocks are quite different for the Blanchard and the Gompertz-Makeham models, especially the ones for per capita consumption and financial assets.

In Chapter 3 we used the extended OLG model to analyse the effects on the economic growth performance of a small open economy of demographic shocks of the type and magnitude that hit the Western world over the last decades. Following Lucas (1988), we assumed that schooling is the main mechanism that causes growth. Individuals spend their first years at school, which increases their productivity and earnings potential later in life. Our analysis shows that only for a unrealistically strong intergenerational knowledge spillover, policy changes and demographic shocks lead to a permanent higher (or lower) growth rate. Moreover, if the intergenerational spillover is unrealistically large, the link between longevity and economic growth is non-monotonic. A higher life expectancy at birth causes a lower long-run growth rate in most developed countries.

As a second application we extended the basic framework of Chapter 2 with a retirement decision as proposed by Sheshinski (1978) and a pension system that fits the stylised facts of Gruber and Wise (1999, 2004, 2005). A consumption and leisure loving individual chooses a retirement age that maximises his lifetime utility. With our simple model we analysed the effects of a baby bust and a longevity shock on a hypothetical economy. Both shocks lead to an ageing society and renders the pension system unsustainable. We showed that at a microeconomic level it is under most pension systems in the Western world optimal for people to retire at the age where retirement benefits are first available (the early eligibility age, EEA). Large policy reforms are necessary to make people work longer, without forcing them by increasing the EEA.

Although our models are far too simple for real world policy evaluation, they do provide useful insights. The main advantage of our models is that at least the steady state effects are analytically tractable. Moreover, it is usually even possible to analytically distinguish between various phases in the transition process, making it easier to understand what is happening when and more importantly why.

Various models already exist for policy evaluation. These large computable general equilibrium models like the well-known Auerbach-Kotlikoff model (Auerbach and Kotlikoff (1987), see Altig et al. (2001) for an enhanced version of this model), IMF's MULTIMOD (Laxton et al., 1998) and specifically for the Netherlands, the IMAGE model (Broer, 1999) often include highly detailed institutions and are calibrated to fit the real world as closely as possible. Their high level of detail and the corresponding complexity is both a strength and a weakness of these models. The advantage is that these models can be used for real world scenarios evaluations and to determine the effect of various policy shocks on different agents. The disadvantage is that the high level of realism is bought at the cost of tractability; due to their inherent complexity, the interpretation of the observed effect is very difficult. A second drawback of these models is that they take a long time to solve and this makes them ill-equipped for sensitivity analysis and quick calculations. Analytically tractable overlapping generation models do not suffer from these drawbacks and if a realistic mortality process is incorporated, these models can take real intergenerational links into account that will provide valuable insights in the ongoing ageing process of the developed world. Furthermore, with relaxed assumptions, overlapping generation models will be better suited to analyse the historical long-run relationship between demographic change and economic growth.

The drawbacks of these large scale CGE models are exactly the strengths of our simple models. It is our opinion that the framework we develop in Chapter 2 and the extensions in Chapters 3 and 4 are a useful addition to the large scale CGE-models. Our models can be used to identify the main effects observed in the large models and to better understand how these models work.

7.1.1 Limitations and future extensions

All three models share two common limitations: they only apply to an open economy that faces an exogenous world interest rate and we need the existence of an actuarially fair insurance system. Unfortunately, it is far from trivial to drop these assumptions.

As we showed in the Chapter 2, a mortality process with a realistic probability of death results in a life-cycle profile of savings. If we combine this life-cycle savings profile with ageing, it is expected that ageing will result in a lower capital supply (IMF, 2004; Poterba, 2001). If capital is not perfectly mobile or if the ageing problem is of such a scale that it becomes a global phenomenon, the assumption of an exogenous interest rate might be too strong. The introduction of a realistic mortality process in a *closed* economy is complicated by the fact that exact aggregation of the consumption function is impossible. Of course, the steady state can still be characterized analytically, it is the same as the steady state of the open economy with one extra restriction, namely that foreign assets should be zero. The transitional effects of various shocks are, however, much more difficult to compute due to the fact that equilibrium factor prices will generally be time-varying. In the near future we wish to investigate whether approximate aggregation of the key behavioural relationships is feasible for particular shock parametrizations. If that fails, numerical methods will be employed to characterize transitional dynamics.

The other limitation concerns the assumed availability of actuarially fair annuities that agents can use to insure themselves against dying indebted. Mitchell et al. (1999) document how unattractive private annuity contracts are in the United States, making the assumption of actuarial fair notes rather far-fetched, even in modern developed financial markets (see also Davidoff et al., 2005). Things become even worse if one realises that one objective of introducing realistic demographics is to explain the historical non-monotonic relation between demographic change, human capital accumulation and economic growth (Kalemni-Ozcan et al., 2000; Boucekkine et al., 2002; Fuster et al., 2005). Especially in a historical context, the assumption of actuarially fair annuities is unjustified. Kalemni-Ozcan and Weil (2002) present an overlapping generation model that they use to analyse the effect of mortality change on retirement without insurance possibilities. Unfortunately, to keep the model analytically tractable they have to assume that the mortality rate is constant (Blanchard's perpetual youth model) and that income and the interest rate is high enough to prevent the liquidity constraint to be binding. They fall back on simulation techniques to study a more complex version of their model, with a realistic mortality rate. The whole problem boils down to solving a dynamic optimization problem with constraints on a state variable; individuals are not allowed to die in debt. This class of models is inherently difficult to solve, but since there is no aggregate risk, only lifetime is stochastic, we expect that it is possible to solve the macroeconomic model.

In conclusion, we hope that the extended Blanchard-Yaari model we constructed will prove to be a useful addition to the toolbox of both theoretical economists and policy practitioners alike. At least in the context of a small open economy, there is no justification whatsoever to use models based on a blatantly unrealistic description of demography. Had mortality not caught up with him, Benjamin Gompertz would probably support that conclusion!

7.2 Public capital and economic growth: an empirical analysis

In Part II of this dissertation we ignored intergenerational issues and focused entirely on public capital, one of the determinants of economic growth. The central issue in Chapter 5 and 6 was whether a robust long-run empirical relationship exists between public capital and economic performance.

Chapter 5 provides a review of the recent studies that examine the relationship between public capital and economic growth. Although not all studies find a growth-enhancing effect of public capital there is more consensus in the recent literature than in the older literature that public capital does spur economic growth. The impact is also much lower than found by Aschauer (1989), which is generally considered to be the starting point of this line of research.

Many studies report that there is heterogeneity: the effect of public investment differs across countries, regions, and sectors. This is perhaps not a surprising result. After all, the effects of new investment spending will depend on the quantity and quality of the capital stock in place. In general, the larger the stock and the better its quality, the lower will be the impact of additions to this stock. Some studies also suggest that the effect of public investment spending may also depend on institutional and policy factors. The network character of public capital, notably infrastructure, causes non-linearities. The effect of new capital will crucially depend on the extent to which investment spending aims at alleviating bottlenecks in the existing network.

Chapter 6 investigated whether it is possible to arrive at robust estimates of the long-run output elasticity of public capital using internationally comparable aggregate and industry data for a considerable number of developed countries. State of the art econometric techniques and economic modelling insights were used to correct for heterogeneity regarding time, country and industry. Unfortunately, we had to conclude that stable output elasticity estimates based are elusive. The estimated parameters vary wildly between equally plausible econometric specification and range between -2 and 2. The aggregate estimates tend to be more stable, but even those range between 0.04 and 1.13. While it is hard to discount cross-country variation, the cross-specification variation we found suggests extreme sensitivity to conceptually innocuous specification choices. Overall, this suggests that production function estimates of the impact of infrastructure are not well-suited to be used for infrastructure policy recommendations.

The question arises why Chapter 5 concluded that 'there is more consensus' and Chapter 6 concluded that the production function estimates vary wildly? There are several possible reasons. First, most studies that use a production function approach in a panel setting estimate a single equation for each country and take the average for all countries. These studies tend to present only one specification and as we showed in Chapter 6, it is possible to arrive at any outcome. Those studies that use the cross-section variation in a panel setting usually do not present elasticity estimates, but focus on the question whether the current public capital stock is optimal.

7.2.1 Limitations and future research

There are a few problems that have not received much attention in the literature on public capital and economic growth. Three of these problems are institutional factors that play a role in the creation of public capital, the role of maintenance on public capital, and the lack of sound theoretical foundations in most empirical studies.

Part of the heterogeneity between countries and regions can be explained by the large differences in the quantity and quality of the public capital stock. Attempts at explaining existing differences in capital stocks are only in their infancy. A possible complicating factor is that certain questionable political practices may determine where and what is invested in public capital. According to Estache (2006, p. 5), 'there is strong anecdotal evidence now that politics matter. Experiences in Asia, Eastern Europe or Latin America show that politicians will never give up the control of a sector that buys votes in democratic societies. Moreover, in societies in which corruption is rampant, they will not give up control of a sector involving large amounts of money and in which contract award processes often provide opportunities for unchecked transactions.' These practices are not confined to developing countries, others (e.g. Cadot et al., 2002 and Kemmerling and Stephan, 2002) find also in industrialised countries evidence of pork barrel politics. Attempts that try to explain differences between countries must take these political factors into account.

As pointed out by Kalaitzidakis and Kalavitis (2005), in most theoretical studies public capital deterioration is considered as an exogenously given technical relationship, thereby neglecting a crucial choice concerning the implementation of public investment decisions, namely the choice between investing in 'new' public capital and extending the durability of the existing public capital stock via maintenance. There is a small, but very interesting line of literature on maintenance. In contrast to standard results derived by endogenous growth models with public infrastructure, the optimal tax burden that maximizes the long-run growth rate of the economy is now larger than the elasticity of infrastructure in the production function. The lack of studies that take maintenance into account can be explained by the fact that published data on maintenance are very scarce due to inherent problems in the measurement of the maintenance expenditures. Kalaitzidakis and Kalavitis (2005) use data from a Canadian survey which contains evidence on maintenance expenditures of both private firms and government organizations, to test the impact of total public capital expenditures and their components on growth. Their results indicate that the Canadian economy would benefit from a fall in total expenditures on both 'new' capital and maintenance, and that the aggregate share of maintenance in total expenditures should be lower over the period under consideration.

A striking result is that only a few of the enormous bulk of studies on the output effects of infrastructure base their estimates on solid theoretical models. A problem is that the theoretical papers that link public capital to economic growth, starting with the work of Uzawa (1974) usually neglect the channels through which infrastructure affects economic growth. They simply postulate an aggregate production function with public capital as an extra input. This neglects the usually complex links, after all, government roads as such do not produce anything. A major exception are the spatial economics models. In these models extra public capital, mostly infrastructure, reduces transportation costs, which boost productivity of the other production factors.

Finally, we must conclude that although there is more consensus on public capital having a positive impact on economic growth, the size of the effect is still not clear. Still, not all possible methods of research have been explored. What is certain is that aggregate level methods are probably too crude to provide policy makers with useful information.

Bibliography

Aaron, H. J. (1990). Discussion. In A. H. Munnell (Ed.), *Is there a shortfall in public capital?* Boston: Federal Reserve Bank of Boston.

Agénor, P., Nabli, M. K., and Yousoef, T. M. (2005). *Public infrastructure and private investment in the middle east and north africa* (Policy Research Paper No. 3661). Washington: World Bank.

Ai, C., and Cassou, S. P. (1995). A normative analysis of public capital. *Applied Economics*, *27*, 1201–1209.

Albala-Bertrand, J. M. (2004). *Can the composition of capital constrain potential output? A gap approach* (Working Paper No. 510). London: Queen Mary University of London, Department of Economics.

Albala-Bertrand, J. M., and Mamatzakis, E. C. (2004). The impcat of public infrastructure on the productivity of the Chilean economy. *Review of Development Economics*, *8*(2), 266–278.

Altig, D., Auerbach, A. J., Kotlikoff, L. J., Smetters, K. A., and Walliser, J. (2001). Simulating fundamental tax reform in the United States. *American Economic Review*, *91*(3), 574–595.

Aschauer, D. A. (1989). Is public expenditure productive? *Journal of Monetary Economics*, *23*, 177–200.

Aschauer, D. A. (2000). Do states optimise? Public capital and economic growth. *The Annals of Regional Science*, *34*, 343–363.

Auerbach, A. J., and Kotlikoff, L. J. (1987). *Dynamic Fiscal Policy*. Cambridge: Cambridge University Press.

Auerbach, A. J., Gokhale, J., and Kotlikoff, L. J. (1994). Generational accounting: A meaningful way to evaluate fiscal policy. *Journal of Economic Perspectives*, *8*(1), 73–94.

Azariadis, C., and Drazen, A. (1990). Threshold externalities in economic development. *Quarterly Journal of Economics*, *105*, 501–526.

Barr, N., and Diamond, P. (2006). The economics of pensions. *Oxford Review of Economic Policy*, *22*(1), 15–39.

Barro, R. J. (1991). Economic growth in a cross section of countries. *The Quarterly Journal of Economics*, *106*(2), 407–443.

Bassanini, A., and Scarpetta, S. (2002). Does human capital matter for growth in OECD countries? A pooled mean-group approach. *Economic Letters*, *74*(3), 399–405.

Batina, R. G. (1998). On the long run effects of public capital and disaggregated public capital on aggregate output. *International Tax and Public Finance*, *5*(3), 263–281.

Batina, R. G. (1999). On the long run effects of public capital on aggregate output: Estimation and sensitivity analysis. *Empirical Economics*, *24*, 711–717.

Bellman, R., and Cooke, K. L. (1963). *Differential-Difference Equations*. New York: Academic Press.

Belloc, M., and Vertova, P. (2006). Public investment and economic performance in highly indebted poor countries: An empirical assessment. *International Review of Applied Economics*, *20*(2), 151–170.

Ben-Porath, Y. (1967). The production of human capital and the life cycle of earnings. *Journal of Political Economy*, *75*(4, Part 1), 352–365.

Berndt, E. R., and Hansson, B. (2004). *Measuring the contribution of public infrastructure capital in Sweden* (Working Paper No. 3842). Cambridge, MA: NBER.

Bettendorf, L. J. H., and Heijdra, B. J. (2001a). Intergenerational and international welfare leakages of a product subsidy in a small open economy. *International Tax and Public Finance*, *8*, 705–729.

Bettendorf, L. J. H., and Heijdra, B. J. (2001b). Intergenerational welfare effects of a tariff under monopolistic competition. *Journal of Economics*, *73*, 313–346.

Bettendorf, L. J. H., and Heijdra, B. J. (2006). Population ageing and pension reform in a small open economy with non-traded goods. *Journal of Economic Dynamics and Control*, *30*, 2389–2424.

Bils, M., and Klenow, P. J. (2000). Does schooling cause growth? *American Economic Review*, *90*, 1160–1183.

Blanchard, O. J. (1985). Debts, deficits, and finite horizons. *Journal of Political Economy*, *93*(2), 223–247.

Blanchard, O. J., and Fischer, S. (1989). *Lectures on Macroeconomics*. Cambridge, MA: MIT Press.

Bonaglia, F., La Ferrara, E., and Marcellino, M. (2000). *Public capital and economic* *performance: Evidence from Italy* (Working Paper No. 163). Milano: IGIER.

Boscá, J. E., Escribá, F. J., and Murgui, M. J. (2000). The effect of public infrastructure on the private productive sector of spanish regions. *Journal of Regional Science*, *42*, 301–326.

Boucekkine, R., de la Croix, D., and Licandro, O. (2002). Vintage human capital, demographic trends, and endogenous growth. *Journal of Economic Theory*, *104*, 340–375.

Bougheas, S., Demetriades, P. O., and Morgenroth, E. L. (1999). Infrastructure, transport costs and trade. *Journal of International Economics*, *47*, 169–189.

Bovenberg, A. L. (1993). Investment promoting policies in open economies: The importance of intergenerational and international distributional effects. *Journal of Public Economics*, *51*(1), 3–54.

Bovenberg, A. L. (1994). Capital taxation in the world economy. In F. van der Ploeg (Ed.), *Handbook of International Macroeconomics.* Oxford: Basil Blackwell.

Broer, P. (1999). *Growth and welfare distribution in an ageing society: An applied general equilibrium analysis for the Netherlands* (Research Memorandum No. 9908). Rotterdam: OCFEB.

Buiter, W. H. (1988). Death, birth, productivity growth and debt neutrality. *Economic Journal*, *98*, 279–293.

Burbidge, J. B., and Robb, A. L. (1980). Pensions and retirement behaviour. *Canadian Journal of Economics*, *13*(3), 421–437.

Cadot, O., Röller, L. H., and Stephan, A. (2002). *Contribution to productivity or pork barrel? The two faces of infrastructure investment* (Discussion Paper No. FS IV 02-09). Berlin: WZB.

Calderón, C., and Servén, L. (2002). *The output cost of Latin America's infrastructure gap* (Working Paper No. 186). Santiago: Central Bank of Chile.

Calvo, G. A., and Obstfeld, M. (1988). Optimal time-consistent fiscal policy with finite lifetimes. *Econometrica*, *56*, 411–432.

Canaleta, C. G., Arzoz, P. P., and Gurate, M. R. (2002). *Public capital, regional productivity and spatial spillovers* (Working Paper No. 9811). Pamplona: Universidad Pública de Navarra, Lan Gaiak Departamento de Economía.

Canning, D. (1998). A database of world infrastructure stocks, 1950-1995. *World Bank Economic Review*, *12*, 529–547.

Canning, D. (1999). *The contribution of infrastructure to aggregate output* (Working Paper No. 2246). Washington, DC: World Bank.

Canning, D., and Bennathan, E. (2000). *The social rate of return on infrastructure* *investments* (Working Paper No. 2390). Washington, DC: World Bank.

Canning, D., and Pedroni, P. (1999). *Infrastructure and long run economic growth.* (Mimeo, World Bank, Washington DC.)

Charlot, S., and Schmitt, B. (1999). *Public infrastructure and economic growth in France's regions.* (Paper (#129) for ERSA 39th Congress, Dublin, Ireland)

Cohen, J., and Morrison Paul, C. (2004). Public infrastructure investment, interstate spatial spillovers, and manufacturing costs. *The Review of Economics and Statistics*, *86*(2), 551–560.

Conrad, K., and Seitz, H. (1994). The economic benefits of public infrastructure. *Applied Economics*, *26*, 303–311.

Crowder, W., and Himarios, D. (1997). Balanched growth and public capital: An empirical analysis. *Applied Economics*, *29*(8), 1045–1053.

d'Albis, H., and Augeraud-Véron, E. (2005). *Endogenous retirement and monetary cycles.* (Mimeo, GREMAQ, University of Toulouse, I (October))

Davidoff, T., Brown, J. R., and Diamond, P. A. (2005). Annuities and individual welfare. *American Economic Review*, *95*(5), 1573-1590.

de Haan, J., Sturm, J. E., and Sikken, B. (1996). Government capital formation: Explaining the decline. *Weltwirtschaftliches Archiv*, *132*, 55–74.

de la Croix, D., and Licandro, O. (1999). Life expectancy and endogenous growth. *Economics Letters*, *65*, 255–263.

de la Fuente, A., and Doménech, R. (2006). Human capital in growth regressions: How much difference does data quality make? *Journal of the European Economic Association*, *4*, 1–36.

de la Fuente, A., and Vives, X. (1995). Infrastructure and education as instruments of regional policy: Evidence from Spain. *Economic Policy*, *10*(20), 13–51.

Delgado Rodriguez, M., and Álvarez Ayuso, I. (2000). *Public productive infrastructure and economic growth.* (Paper for ERSA 40th Congress)

Demetriades, P., and Mamuneas, T. (2000). Intertemporal output and employment effects of public infrastructure capital: Evidence from 12 OECD economies. *Economic Journal*, *110*, 687–712.

Devarajan, S., Swaroop, V., and Zou, H. (2000). The composition of public expenditure and economic growth. *Journal of Monetary Economics*, *37*, 313–344.

Diamond, P. A. (1965). National debt in a neoclassical growth model. *American Economic Review*, *55*, 1126–1150.

Driffill, E. J. (1980). Life-cycles with terminal retirement. *International Economic Review*, *21*, 45–62.

Duggal, V., Saltzman, C., and Klein, L. (1999). Infrastructure and productivity: A non-linear approach. *Journal of Econometrics*, *92*, 47–74.

Duval, R. (2003). *The retirement effects of old-age pension and early retirement schemes in OECD countries* (Economics Department Working Paper Series No. 370). Paris: OECD.

Easterly, W., and Rebelo, S. (1993). Fiscal policy and economic growth: An empirical investigation. *Journal of Monetary Economics*, *32*, 417–458.

Eberts, R. (1986). *Estimating the contribution of urban public infrastructure to regional growth* (Working Paper No. 8610). Cleveland, OH: Federal Reserve Bank of Cleveland.

Egger, H., and Falkinger, J. (2003). *The role of public infrastructure for firm location and international outsourcing* (Working Paper No. 970). Munchen: CESIfo.

Esfahani, H., and Ramíres, M. (2003). Institutions, infrastructure and economic growth. *Journal of Development Economics*, *70*, 443–477.

Estache, A. (2006). *Infrastructure. a survey of recent and upcoming issues.* (Mimeo, World Bank, Washington DC.)

Evans, P., and Karras, G. (1994). Government capital productive? Evidence from a panel of seven countries. *Journal of Macroeconomics*, *16*, 271–279.

Everaert, G. (2003). Balanced growth and public capital: An empirical analysis with I(2) trends in capital stock data. *Economic Modelling*, *20*, 741–763.

Everaert, G., and Heylen, F. (2004). Public capital and long-term labour market performance in Belgium. *Journal of Policy Modelling*, *26*, 95–112.

Faruqee, H. (2003). Debt, deficits, and age-specific mortality. *Review of Economic Dynamics*, *6*, 300–312.

Faruqee, H., and Laxton, D. (2000). *Life cycles, dynasties, saving: Implications for closed and small, open economies* (Working Paper No. WP/00/126). Washington, DC: International Monetary Fund.

Fernald, J. (1999). Assessing the link between public capital and productivity. *American Economic Review*, *89*(3), 619–638.

Finn, M. (1993). Is all government capital productive? *Federal Reserve Bank of Richmond Economic Quarterly*, *79*, 53–80.

Fisher, I. (1930). *The Theory of Interest*. New York: Macmillan. (Reprinted by Augustus M. Kelley, Fairfield, NJ, 1986.)

Flores de Frutos, R., Garcia-Diez, M., and Perez-Amaral, T. (1998). Public capital and economic growth: An analysis of the Spanish economy. *Applied Economics*, *30*(8), 985–994.

Ford, R., and Poret, P. (1991). Infrastructure and private-sector productivity. *OECD Economic Studies*, *17*, 63–89.

Fujita, M., Krugman, P. R., and Venables, A. (1999). *The Spatial Economy*. Cambridge, MA: MIT Press.

Fuster, L., Imrohoroglu, A., and Imrohoroglu, S. (2005). *Personal security accounts and mandatory annuitization in a dynastic framework* (Working Paper No. 1405). Munchen: CESifo.

Gardia, E. (1991). The dynamics of a small open economy in response to monetary, fiscal, and productivity shocks. *Journal of Monetary Economics*, *28*, 411–434.

Gavrilov, L., and Gavrilova, N. (1991). *The Biology of Life Span: A Quantitative Approach*. Chur, Switserland: Harwood Academic.

Ghali, K. (1998). Public investment and private capital formation in a vector error-correction model of growth. *Applied Economics*, 837–844.

Gomme, P., Rogerson, R., Rupert, P., and Wright, R. (2005). The business cycle and the life cycle. *NBER Macroeconomics Annual 2004*, *19*, 415–461.

Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality. *Philosophical Transactions of the Royal Society of London, Series A*, *115*, 513–580.

Gramlich, E. (1994). Infrastructure investment: A review essay. *Journal of Economic Literature*, 1176–1196.

Gruber, J., and Wise, D. A. (1999). Introduction and summary. In J. Gruber and D. A. Wise (Eds.), *Social Security and Retirement Around the World* (pp. 1–35). Chicago, IL: University of Chicago Press.

Gruber, J., and Wise, D. A. (2004). Introduction and summary. In J. Gruber and D. A. Wise (Eds.), *Social Security and Retirement Around the World, Micro-Estimation* (pp. 1–35). Chicago, IL: University of Chicago Press.

Gruber, J., and Wise, D. A. (2005). (Working Paper No. 11290). Cambridge, MA: NBER.

Gustman, A. L., and Steinmeier, T. L. (1986). A structural retirement model. *Econometrica*, *54*, 555–584.

Gwartney, J. D., Holcombe, R. G., and Lawson, R. A. (2006). Institutions and the impact of investment on growth. *Kyklos*, *59*, 255–273.

Haughwout, A. F. (2002). Public infrastructure investments, productivity and welfare in fixed geographic areas. *Journal of Public Economics*, *83*, 405–428.

Hausman, J. (1978). Specification tests in econometrics. *Econometrica*, *46*, 1251–1271.

Heckman, J. J. (1976). A life-cycle model of earnings, learning, and consumption.

Journal of Political Economy, *84*, S11–S44.

Heckman, J. J., Lochner, L., and Taber, C. (1998). Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents. *Review of Economic Dynamics*, *1*, 1–58.

Heijdra, B. J., and Meijdam, L. (2002). Public investment and intergenerational distribution. *Journal of Economic Dynamics and Control*, *26*, 707–735.

Heijdra, B. J., and van der Ploeg, F. (2002). *Foundations of Modern Macroeconomics*. Oxford: Oxford University Press.

Holtz-Eakin, D. (1994). Public-sector capital and the productivity puzzle. *Review of Economics and Statistics*, *76*, 12–21.

Holtz-Eakin, D., and Lovely, M. (1996). Scale economics, returns to variety, and the productivity of public infrastructure. *Regional Science and Urban Economics*, *26*, 105–123.

Holtz-Eakin, D., and Schwartz, A. (1995). Infrastructure in a structural model of economic growth. *Regional Science and Urban Economics*, *25*, 131–151.

Hooker, P. (1965). Benjamin Gompertz: 5 March 1779–14 July 1865. *Journal of the Institute of Actuaries*, *91*, 203–212.

Hulten, C. R. (1996). *Infrastructure capital and economic growth: How well you use it may be more important than how much you have* (Working Paper No. 5847). Cambridge, MA: NBER.

Hulten, C. R., and Schwab, R. M. (1991). *Is there too little public capital? Infrastructure and economic growth.* (American Enterprise Institute Discussion Paper, February)

Human Mortality Database. (2006). *Data by country: The Netherlands (life tables, total).* University of California, Berkeley, CA. (Available from: `www.mortality.org`. (Data downloaded February 27, 2006))

Hurd, M. D. (1990). Research on the elderly: Economic status, retirement, and consumption and saving. *Journal of Economic Literature*, *28*(2), 565–637.

Hurd, M. D. (1997). The economics of individual aging. In M. R. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economics* (pp. 891–966). Amsterdam: North-Holland.

IMF. (2004). *The global demographic transition* (World Economic Outlook). Washington, DC: IMF.

Inklaar, R. C., Timmer, M. P., and van Ark, B. (2006). *Mind the gap! International comparisons of productivity in services and goods production.* (paper presented at NBER/CRIW Summer Institute 2006, downloadable at http://www.nber.org/ confer/2006/si2006/prcr/vanark.pdf)

Kalaitzidakis, P., and Kalavitis, S. (2005). Financing 'new' public investment and/or maintenance in public capital for long-run growth? The Canadian experience. *Economic Inquiry*, *43*(3).

Kalemni-Ozcan, S., Ryder, H. E., and Weil, D. N. (2000). Mortality decline, human capital investment and economic growth. *Journal of Development Economics*, *62*(1), 1–23.

Kalemni-Ozcan, S., and Weil, D. N. (2002). *Mortality change, the uncertainty effect, and retirement* (Working Paper No. 8742). Cambridge, MA: NBER. (latest version August 2004)

Kamps, C. (2004). *The dynamic effects of public capital: VAR evidence for 22 OECD countries* (Working Paper No. 1224). Kiel: Kiel Institute of World Economics.

Kamps, C. (2005). Is there a lack of public capital in the european union? *EIB Papers Volume*, *10*(1), 72–93.

Kamps, C. (2006). New estimates of government net capital stocks for 22 OECD countries 1960–2001. *IMF Staff Papers, forthcoming*. (An earlier version is available as IMF Working Paper 04/67)

Kelley, A. C., and Schmidt, R. M. (1995). Aggregate population and economic growth correlations: The role of the components of demographic change. *Demography*, *32*(4), 543–555.

Kemmerling, A., and Stephan, A. (2002). The contribution of local public infrastructure to private productivity and its political-economy: Evidence from a panel of large German cities. *Public Choice*, *113*, 403–422.

Kingston, G. H. (2000). Efficient timing of retirement. *Review of Economic Dynamics*, *3*, 831–840.

Kirkwood, T. B. (2001). Where will it all end? *The Lancet*, *357*(9256), 576.

Kreyszig, E. (1999). *Advanced Engineering Mathematics* (eight ed.). New York: John Wiley.

Krugman, P. R. (1991). Increasing returns and economic geography. *Journal of Political Economy*, *99*, 403–422.

La Ferrara, E., and Marcellino, M. (2000). *TFP, costs, and public infrastructure: An equivocal relationship* (Working Paper No. 176). Milano: IGIER.

Laxton, D., Isard, P., Faruqee, H., Prasad, E., and Turtelboom, B. (1998). *MULTIMOD Mark III, the core dynamic and steady-state models* (Occasional Paper No. 164). Washington, DC: IMF.

Lazear, E. P. (1979). Why is there mandatory retirement? *Journal of Political Economy*, *87*(6), 1261–1284.

Lazear, E. P. (1999). Retirement from the labor force. In O. Ashenfelter and R. Layard (Eds.), *Handbook of Labor Economics* (Vol. I, pp. 305–355). Amsterdam: North-Holland.

Leamer, E. (1983). Let's take the con out of econometrics. *American Economic Review*, *73*(3), 31–43.

Lee, R. (2003). The demographic transition: Three centuries of fundamental change. *Journal of Economic Perspectives*, *17*(Fall), 167–190.

Levine, R., and Renelt, D. (1992). A sensitivity analysis of cross-country growth regressions. *American Economic Review*, *82*, 942–963.

Ligthart, J. E. (2002). Public capital and output growth in Portugal: An empirical analysis. *European Review of Economics and Finance*, *1*(2), 3–30.

Lindbeck, A., and Persson, M. (2003). The gains from pension reforms. *Journal of Economic Literature*, *49*, 74–112.

Linz, P. (1985). *Aalytical and Numerical Methods for Volterra Equations*. Philadelphia: SIAM.

Lucas, R. E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, *22*(1), 3–42.

Lumsdaine, R. L., and Mitchell, O. S. (1999). New developments in the economic analysis of retirement. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics* (Vol. III, pp. 3261–3307). Amsterdam: North-Holland.

Lumsdaine, R. L., Stock, J. H., and Wise, D. A. (1992). Three models of retirement: Computational complexity versus predictive validity. In D. A. Wise (Ed.), *Topics in the Economics of Aging* (pp. 21–60). Chicago, IL: University of Chicago Press.

Makeham, W. M. (1860). On the law of mortality and the construction of annuity tables. *Assurance Magazine*, *8*, 301–310.

Mamatzakis, E. C. (1999a). *Public infrastructure, private input demand, and economic performance of the Greek industry* (Working Paper No. 406). London: Queen Mary & Westfield College.

Mamatzakis, E. C. (1999b). Testing for long run relationship between infrastructure and private capital productivity: A time series analysis for the Greek industry. *Applied Economics Letters*, *6*(4), 243–246.

Matsuyama, K. (1987). Current account dynamics in a finite horizon model. *Journal of International Economics*, *23*, 299–313.

Milbourne, R., Otto, G., and Voss, G. (2003). Public investment and economic growth. *Applied Economics*, *35*(5), 527–540.

Miller, J., and Tsoukis, C. (2001). On the optimality of public capital for long-run economic growth. *Applied Economics*, *33*, 1117–1129.

Mitchell, O. S., and Fields, G. S. (1984). The economics of retirement behavior. *Journal of Labor Economics*, *2*(1), 84–105.

Mitchell, O. S., Poterba, J. M., Warshawsky, M. J., and Brown, J. R. (1999). New evidence on the moneys worth of individual annuities. *American Economic Review*, *89*(5), 1299–1318.

Mittnik, S., and Neumann, T. (2001). Dynamic effects of public investment: Vector autoregressive evidence from six industrialized countries. *Empirical Economics*, *26*, 429–446.

Moreno, R., López-Bazo, E., and Artís, M. (2003). On the effectiveness of private and public capital. *Applied Economics*, *35*(5), 727–740.

Mulligan, C. B. (1999). Substitution over time: Another look at life cycle labor supply. *NBER Macroeconomics Annual 1998*, *13*, 75–134.

Munnell, A. H. (1990a). How does public infrastructure affect regional economic performance? *New England Economic Review*, 11–32.

Munnell, A. H. (1990b). Why has productivity growth declined? Productivity and public investment. *New England Economic Review*, 2–22.

Munnell, A. H. (1992). Policy watch. infrastructure investment and economic growth. *Journal of Economic Perspectives*, *6*, 189–198.

OECD. (1998). Work-force ageing in OECD countries. *OECD Employment Outlook*, 123–150.

OECD. (2005). *Pensions at a Glance: Public Pensions across OECD Countries*. Paris: OECD.

O'Mahony, M., and Vecchi, M. (2005). Is there an ICT impact on TFP? A heterogeneous dynamic panel approach. *Economica*, *72*, 615–633.

Oxley, H., and Martin, J. (1991). Controlling government spending and deficits: Trends in the 1980s and prospects for the 1990s. *OECD Economic Studies*, *17*, 145–189.

Pereira, A. M. (2000). Is all public capital created equal? *Review of Economics and Statistics*, *82*(3), 513–518.

Pereira, A. M. (2001). On the effects of public investment on private investment: What crowds in what? *Public Finance Review*, *29*(1), 3–25.

Pereira, A. M., and Andraz, J. M. (2003). On the impact of public investment on the performance of U.S. industries. *Public Finance Review*, *31*(1), 66–90.

Pereira, A. M., and Flores de Frutos, R. (1999). Public capital accumulation and private sector performance. *Journal of Urban Economics*, *46*(2), 300–322.

Pereira, A. M., and Roca-Sagales, O. (1999). Public capital formation and regional development in Spain. *Review of Development Economics*, *3*(3), 281–294.

Pereira, A. M., and Roca-Sagales, O. (2001). Infrastructures and private sector performance in Spain. *Journal of Policy Modelling*, *23*(4), 371–384.

Pereira, A. M., and Roca-Sagales, O. (2003). Spillover effects of public capital formation: Evidence from the Spanish regions. *Journal of Urban Economics*, *53*(2), 238–256.

Pesaran, M., Shin, Y., and Smith, R. P. (1999). Pooled mean group estimation of dynamic heterogeneous panels. *Journal of the American Statistical Association*, *94*, 621–634.

Phelps, E. S. (1961). The golden rule of accumulation: A fable for growthmen. *American Economic Review*, *51*, 638–643.

Phillips, P. (1998). Impulse response and forecast error variance asymptotics in nonstationary VARs. *Journal of Econometrics*, *83*(1-2), 21–56.

Pina, A. M., and St. Aubyn, M. (2005). Comparing macroeconomic returns on human and public capital: An empirical analysis of the Portuguese case (1960–2001). *Journal of Policy Modelling*, *27*, 585–598.

Pina, A. M., and St. Aubyn, M. (2006). How should we measure the return on public investment in a VAR? *Economics Bulletin*, *8*(5), 1–4.

Poterba, J. (2001). Demographic structure and asset returns. *Review of Economics and Statistics*, *83*(4), 565–584.

Preston, S. H., Heuveline, P., and Guillot, M. (2001). *Demography: Measuring and Modeling Population Processes*. Oxford: Blackwell Publishers.

Pritchett, L. (1996). *Mind your P's and Q's* (Policy Research Paper No. 1660). Washington, DC: World Bank.

Razin, A. (1972). Optimum investment in human capital. *Review of Economic Studies*, *39*, 455–460.

Royal Swedish Academy of Sciences. (2006). *Advanced information on Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel: Edmund Phelps's contribution to macroeconomics.* (Kungl. Vetenskapsakademien, Stockholm)

Ruschinski, M., and Sturm, J. E. (2004). Foreign direct investment and growth: Panel data evidence for 22 OECD countries. In G. H. C. Dreger (Ed.), *Advances in Macroeconometric Modeling* (pp. 13–24). Baden-Baden: Nomos.

Sala-i-Martin, X. (1997). I just ran two millions regressions. *American Economic*

Review, *87*(2), 178–183.

Sanchez-Robles, B. (1998). Infrastructure investment and growth: Some empirical evidence. *Contemporary Economic Policy*, *16*, 98–108.

Seung, C., and Kraybill, D. (2001). The effects of infrastructure investment: A two sector dynamic computable general equilibrium analysis for ohio. *International Regional Science Review*, *24*(2), 261–281.

Sheshinski, E. (1978). A model of social security and retirement decision. *Journal of Public Economics*, *10*, 337–360.

Shioji, E. (2001). Public capital and economic growth: A convergence approach. *Journal of Economic Growth*, *6*, 205–227.

Sims, C., Stock, J., and Watson, M. (1990). Inference in linear time series models with some unit roots. *Econometrica*, *58*, 113–144.

Stephan, A. (2000). Regional infrastructure policy and its impact on productivity: A comparison of Germany and France. *Applied Economics Quarterly*, *46*, 327–356.

Stephan, A. (2003). Assessing the contribution of public capital to private production: Evidence from the German manufacturing sector. *International Review of Applied Economics*, *17*, 399–418.

Stock, J. H., and Wise, D. A. (1990a). The pension inducement to retire: An option value analysis. In D. A. Wise (Ed.), *Issues in the Economics of Aging* (pp. 205–224). Chicago: University of Chicago Press.

Stock, J. H., and Wise, D. A. (1990b). Pensions, the option value of work, and retirement. *Econometrica*, *58*(5), 1151–1180.

Sturm, J. E., and de Haan, J. (1995). Is public expenditure really productive? New evidence for the US and the Netherlands. *Economic Modelling*, *12*, 60–72.

Sturm, J. E., and de Haan, J. (2005). Determinants of long-term growth: New results applying robust estimation and extreme bounds analysis. *Empirical Economics*, *30*(3), 597–617.

Sturm, J. E., Jacobs, J. P., and Groote, P. G. (1999). Output effects of infrastructure investment in the Netherlands 1853–1913. *Journal of Macroeconomics*, *21*(2), 355–380.

Sturm, J. E., Kuper, G. H., and de Haan, J. (1998). Modelling government investment and economic growth on a macro level: A review. In S. Brakman, H. van Ees, and S. K. Kuipers (Eds.), *Market Behaviour and Macroeconomic Modelling.* London: MacMillan Press Ltd.

Takayama, A. (1985). *Mathematical Economics* (2nd ed.). Cambridge: Cambridge University Press.

Tamura, R. (1991). Income convergence in an endogenous growth model. *Journal of Political Economy*, *99*, 522–540.

Tatom, J. (1991). Public capital and private sector performance. *Federal Reserve Bank of St. Louis Review*, *73*, 3–15.

Temple, J. (2000). Growth regressions and what the textbooks don't tell you. *Bulletin of Economic Research*, *52*(3), 181–205.

Tinbergen, J. (1956). The optimum rate of saving. *Economic Journal*, *66*(264), 603–609.

Uzawa, H. (1974). Sur la théorie économique du capital collectif social. *Cahiers du Séminaire d'Économêtrie*.

Vijverberg, W., Vijverberg, C., and Gamble, J. (1997). Public capital and private production. *Review of Economics and Statistics*, *79*(2), 267–278.

Voss, G. (2002). Public and private investment in the United States and Canada. *Economic Modelling*, *19*, 641–664.

Weil, D. N. (1997). The economics of population aging. In M. R. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economics* (pp. 967–1014). Amsterdam: North-Holland.

Weil, P. (1989). Overlapping families of infinite-lived agents. *Journal of Public Economics*, *38*, 183–198.

Weiss, Y. (1972). On the optimal lifetime pattern of labour supply. *Economic Journal*, *82*(328), 1293–1315.

World Bank. (1994). *Annual Report*. Washington DC: World Bank.

Yaari, M. E. (1965). Uncertain lifetime, life insurance, and the theory of the consumer. *Review of Economic Studies*, *32*, 137–150.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regressions and test for aggregation bias. *Journal of the American Statistical Association*, *57*, 348–368.

Samenvatting

Hoe moeten de belangen van de verschillende generaties tegen elkaar worden afgewogen? Wat zijn de korte- en lange-termijn effecten van overheidsinvesteringen? En wie betaalt de kosten van de vergrijzing, de huidige ouderen, de jongeren, of zelfs toekomstige generaties? De intertemporele macro-economische wetenschap probeert deze vragen te beantwoorden en beleidsmakers te adviseren.

Dit proefschrift beperkt zich tot twee thema's binnen de intertemporele macroeconomie. Deel I onderzoekt de rol van de demografische opbouw en verandering van de bevolking in de verdeling van baten en lasten over de bestaande en toekomstige generaties. Deel II richt zich op de effecten van overheidskapitaal op de economische groei van een land.

Demografische verandering

Onzekerheid over het moment van het onvermijdelijke einde, de dood, heeft niet alleen emotionele maar ook economische gevolgen. In hoeverre houdt men rekening met gebeurtenissen in de toekomst? Wat gebeurt er met de erfenis? En hoeveel risicopremie betalen mensen extra op hun lening doordat vermogensverschaffers het risico lopen hun lening niet terug te krijgen als de lener overlijdt?

Yaari (1965) heeft de gevolgen van de onzekerheid over het moment van sterven geanalyseerd op micro-economisch niveau. Zijn boodschap was tweeledig. Allereerst concludeerde Yaari dat voor een rationeel individu toekomstige gebeurtenissen steeds minder zwaar tellen, niet alleen vanwege een tijdsvoorkeur, maar ook door de kans op overlijden in de tussentijd. Zijn tweede conclusie was dat het optimaal is om het gehele vermogen te investeren in zogenaamde annuïteiten of in het geval van schuld, een levensverzekering af te sluiten waarmee de schulden worden afgelost na overlijden. Normaal gesproken zouden financiële mark-ten schulden niet toestaan vanwege het risico dat de lener niet terugbetaalt door overlijden. Levensverzekeringen lossen dit probleem op. Annuïteiten zorgen voor het tegenovergestelde, deze maken sparen aantrekkelijk door een hoger bedrag uit te keren dan een normale investering. De extra kosten worden gedekt doordat bij overlijden het gehele vermogen vervalt aan de annuïteitenverschaffer, er is geen erfenis. Concurrentie zorgt er voor dat het rendement op de annuïteiten en de premie op verzekeringen exact gelijk is aan de risicovrije rente plus de kans op overlijden; ze zijn actuarieel eerlijk.

Yaari's inzichten zijn door Blanchard (1985) toegepast in een algemeen evenwichtsmodel waarin op elk tijdstip meerdere generaties leven en er expliciet rekening mee wordt gehouden houden dat mensen een tijdelijk leven hebben. Dit zijn de zogenaamde continue tijd overlappende generatie modellen (OLG). Dit Blanchard-Yaari raamwerk is de afgelopen 20 jaar uitgegroeid tot één van de meest toegepaste modellen in de macro-economie. Helaas was het niet mogelijk het micro-economische Yaari-model probleemloos te implementeren op een macro-economisch niveau. Exacte aggregatie van de micro-economische gedragsrelaties was slechts mogelijk als individuen een constante, niet leeftijd afhankelijke, kans hebben op overlijden. Dit betekent dat iedereen toekomstig nut en inkomen op dezelfde wijze disconteert en dus dat mensen – althans in economisch opzicht – de eeuwige jeugd hebben. Exacte aggregatie is noodzakelijk in een gesloten economie; het is immers de combinatie van geaggregeerde vraag en aanbod die factorprijzen bepaalt. Echter, in een open economie worden factorprijzen op de wereldmarkt bepaald en is aggregatie niet noodzakelijk of slechts op het laatste moment op het niveau van de overheidsbudgetbeperking.

In Deel I van dit proefschrift maken we gebruik van dit laatste inzicht om in een OLG model Blanchards onrealistische demografische proces te vervangen door een algemene demografie. In hoofdstuk 2 ontwikkelen we ons basismodel. Dit basismodel wordt in hoofdstuk 3 uitgebreid met een scholingsbeslissing en gebruikt om de effecten van vergrijzing op scholing en economische groei te analyseren. In hoofdstuk 4 breiden we ons basismodel van hoofdstuk 2 uit met een pensioneringsbeslissing. Dit model gebruiken we om de gevolgen van vergrijzing voor de houdbaarheid van pensioenen te onderzoeken.

In hoofdstuk 2 tonen we aan dat het goed mogelijk is om een overlappend generaties model te ontwikkelen met een algemene beschrijving van het demografisch proces, mits we ons beperken tot een open economie. Blanchards eeuwige jeugd is een speciaal geval van ons model. Met een exogene rente volgt uit de aanbodszijde van de economie ook een exogene loonvoet. Met rente en loon gegeven weet elk individu ook wat de verdisconteerde waarde is van het toekomstige inkomen. Een verschil met Blanchard is dat de verdisconteerde waarde van het toekomstige inkomen daalt naarmate men ouder wordt, doordat de verwachte resterende levensduur daalt.

Het belangrijkste verschil is echter dat de marginale consumptie quote, de verhouding tussen consumptie en de som van de financiële activa en menselijk vermogen (het totale vermogen), niet leeftijdsonafhankelijk is. Ouderen consumeren een groter deel van hun totale vermogen dan jongeren, omdat zij een kortere tijdshorizon hebben en het zinloos is vermogen mee te nemen in het graf. Dit leidt tot een klokvormig profiel voor besparingen, zoals in Modigliani's klassieke levensloop model.

Net als in Blanchards eeuwige jeugd model is het in ons basismodel mogelijk om de lange-termijn effecten van macro-economische schokken analytisch te bepalen. Onze comparatieve statica analyse toont aan dat de lange-termijn effecten van ons veralgemeniseerde model kwalitatief hetzelfde zijn als in het standaard model. Er zijn echter grote verschillen tijdens de transitieperiodes. De individuele impuls-respons functies na macro-economische schokken verschillen sterk. De transitieperiodes bij realistische demografische processen zijn veel korter, doordat de levensverwachting van elk individu korter is. De standaard Blanchard-Yaari modellen voorspellen gewoonlijk transitieperiodes van een eeuw en soms nog langer. In onze model is dit gereduceerd tot 50 jaar. Tijdens de transitieperiodes zijn de fluctuaties van de macro-economische variabele echter wel groot, soms tot twee keer zo groot.

In hoofdstuk 3 breiden we het basismodel – met realistische demografie – van hoofdstuk 2 uit met een scholingsbeslissing. Iedereen gaat tijdens zijn jeugd naar school en hoe langer de opleiding, hoe hoger de productiviteit – en dus het loon – tijdens het werkzame leven. Mensen hebben geen voorkeur voor school of werk, dus wordt de scholingsperiode dusdanig gekozen dat het verwachte nut uit consumptie wordt gemaximaliseerd.

Een hogere opleiding heeft externe effecten zoals in Azariadis en Drazen (1990). Hoe hoger de opleiding van de mentor, hoe meer effect een extra jaar opleiding heeft op het niveau van de leerling – het 'standing on the shoulders of giants' effect. Indien dit externe effect groot genoeg is, dan zou een eenmalige schok aanleiding geven tot permanente verhoging van de economische groei, een zogenaamd endogene groei proces. Empirische schattingen geven echter aan dat dit zeer onwaarschijnlijk is. Waarschijnlijker is dat het externe effect wel aanwezig is, maar niet voldoende voor endogene groei. Na verloop van tijd keert convergeert de groeivoet van de economie naar zijn exogeen bepaalde niveau.

We gebruiken ons groeimodel om de effecten van demografische en beleidsmatige schokken te bepalen tijdens de transitieperiode en op de lange termijn. Ondanks dat dit model gecompliceerder is dan het basismodel uit hoofdstuk 2 is het toch mogelijk om de interessante lange-termijn effecten analytisch te bepalen. Eén van de meer verrassende resultaten is een zeer niet-linear – zelfs niet-monotoon – verband tussen de verwachte levensduur en economische groei. Een lagere sterftekans zou kunnen leiden tot een lagere economische groei. Sterker nog, dit is het geval bij modelparameters zoals die typisch worden gebruikt in simulatiemodellen voor ontwikkelde landen. Verder is het transitieproces naar het nieuwe lange-termijn evenwicht zeer volatiel en neemt dit proces vaak meerdere decennia in beslag. Dit niet-monotone aanpassingsproces gecombineerd met de zeer lange transitieperiode kan verklaren waarom het erg moeilijk is om een robuust empirisch verband te vinden tussen demografische variabelen en economische groei (Kelley en Schmidt, 1995). De bestaande tijdreeksen zijn simpelweg te kort om de lange-termijn effecten te schatten. Door het niet-monotone aanpassingspad bepaalt de lengte van de tijdsreeks het gevonden effect.

In hoofdstuk 4 breiden we het basismodel van hoofdstuk 2 uit met een pensioneringsbeslissing zoals in Sheshinski (1978). We gebruiken dit model op een micro-economisch niveau om de effecten van vergrijzing en veranderingen in het pensioenstelsel op de pensioneringsleeftijd te analyseren. Onze toevoeging aan de micro-economische literatuur is dat we het standaard optimalisatieprobleem – iedereen kiest een pad van consumptie en een pensioneringsleeftijd – transformeren in een tweedimensionaal probleem met een eenvoudige grafische weergave. Zoals altijd vereenvoudigt een grafische weergave van het probleem de comparatieve statica. Het blijkt eenvoudig te verklaren waarom de overgrote meerderheid van de werknemers met pensioen gaat op de jongst mogelijk leeftijd – de leeftijd waarop voor het eerst aanspraak kan worden gemaakt op een (vervroegde) pensioenuitkering – en waarom mensen praktisch ongevoelig zijn voor financiële veranderingen in het pensioenstelsel (Gruber en Wise, 1999; Duval, 2003).

Op macro-economisch niveau gebruiken we ons model om de vereiste beleidsveranderingen te bepalen die noodzakelijk zijn om het pensioenstelsel solvabel te houden na demografische schokken zoals we die de afgelopen decennia hebben gezien in de Westerse wereld. Zoals goed gedocumenteerd is door o.a. Gruber en Wise (1999, 2004, 2005) en de OECD (2005) is de houdbaarheid van het huidige pensioenstelsel op de langere termijn twijfelachtig. Enerzijds leven mensen langer zonder dat hun pensioneringsleeftijd stijgt, wat zorgt voor hogere uitgaven. Anderzijds is het aantal geboortes gedaald, waardoor de hogere lasten door minder mensen worden gedragen.

Ons model geeft twee belangrijke inzichten. Allereerst heeft een daling van de geboortevoet direct een negatief effect op de financiële positie van het pensioenstelsel, terwijl een verlaging van de sterftekans pas na 50 tot 60 jaar leidt tot financiële problemen. Ten tweede voorspelt ons model dat een verhoging van de pensioengerechtigde leeftijd tot veel kleinere negatieve welvaartseffecten leidt dan welke andere oplossing ook. Doordat mensen langer moeten werken leidt een verhoging van de pensioengerechtigde leeftijd tot een welvaartsverlies. Dit wordt gedeeltelijk gecompenseerd doordat het totale inkomen stijgt. Hierdoor is een hoger consumptiepad mogelijk.

De modellen in dit eerste deel van het proefschrift zijn uiteraard te gestileerd om te worden gebruikt voor 'real world' voorspellingen. Het is echter niet ons doel om de zeer complexe simulatiemodellen van bijvoorbeeld Auerbach en Kotlikoff (1987) te vervangen. De gedetailleerdheid van deze modellen is zowel hun kracht als zwakte. Door hun details kunnen ze worden gebruikt om de effecten van beleidsmaatregelen te simuleren. Het nadeel van deze modellen is dat de voorspellingen zeer moeilijk zijn te interpreteren. Hier ligt de toegevoegde waarde van onze gestileerde modellen. Ze zijn simpel genoeg om de lange-termijn effecten analytisch te bepalen, hetgeen het begrip vergroot. De realistische demografie maakt het mogelijk om rekening te houden met de relatie tussen generaties. Zo kunnen onze simpele modellen worden gebruikt om meer inzicht te krijgen in de simulatie uitkomsten van de complexe numerieke modellen.

Overheidskapitaal

In deel II van dit proefschrift onderzoeken we de effecten van overheidskapitaal op economische groei. Overheidskapitaal, en dan met name infrastructuur, staan centraal in de activiteiten van zowel huishoudens als bedrijven (World Bank, 1994). Als infrastructuur niet de motor vormt van economische activiteit, dan toch zeker de wielen. Input-output tabellen tonen bijvoorbeeld aan dat telecommunicatie, elektriciteit en water in praktisch elk productieproces worden gebruikt en transportdiensten zijn cruciaal voor iedereen.

Toch is het verbazingwekkend hoe lang economen dit belang van overheidskapitaal hebben genegeerd in de empirische literatuur (Gramlich, 1994). Hier kwam een einde aan met het werk van Aschauer (1989). In zijn baanbrekende studie schatte hij dat een toename van 10 procent van het overheidskapitaal leidt tot een productiestijging van 4 procent, een elasticiteit van 40%! Aschauers resultaten spraken tot de verbeelding en momenteel bestaan er talloze studies die het effect van overheidskapitaal op de productiviteit van private productiefactoren en economische groei schatten.

Aschauers resultaten werden om verschillende redenen vrij snel in twijfel getrokken. Allereerst waren de door hem geschatte effecten van overheidskapitaal op productiviteit onwaarschijnlijk hoog. De vraag rees waarom niet meer private organisaties in infrastructuur investeerden als er echt zulke grote effecten zouden zijn. Ten tweede kleefden er een aantal methodologische en econometrische problemen aan zijn werk. De duidelijkste problemen waren toevallige correlatie door niet-stationariteit van de regressoren en het causaliteitsprobleem. Het bleek niet eenvoudig alle problemen op te lossen. Sturm et al. (1998) geven een overzicht van de eerste literatuur. Zij concluderen dat de schattingen van de marginale productiviteit van overheidskapitaal zeer uiteenlopen, variërend van negatief (Hulten en Schwab, 1991) tot onwaarschijnlijk hoog positief (Aschauer, 1989). De grote onzekerheid die hieruit volgt maakt deze schattingen vrijwel onbruikbaar voor beleidsbeslissingen.

Sinds het vorige literatuuroverzicht van Sturm et al. (1998) is een groot aantal studies verschenen die een aantal problemen uit de vroegere literatuur oplossen. Hoofdstuk 5 geeft een up-to-date overzicht van deze moderne literatuur. De centrale vragen in dit hoofdstuk zijn of er robuust empirische bewijs is gevonden dat overheidskapitaal economische groei bevorderd en of dit positieve verband overeind blijft als rekening wordt gehouden met het feit dat productiefactoren die naar overheidsinvesteringen gaan ook op andere manieren hadden kunnen worden gebruikt?

Uit ons overzicht blijkt dat de meeste recente studies goed rekening houden met de diverse econometrische en methodologische problemen. In de moderne literatuur bestaat redelijke consensus dat er inderdaad een positief verband is tussen overheidskapitaal en productiviteit, maar dat het positieve effect lang niet zo groot is als gerapporteerd door Aschauer (1989).

Een tweede conclusie in hoofdstuk 5 is dat het productiviteitseffect van overheidskapitaal lijkt te variëren tussen landen, sectoren en over tijd. Deze variantie kan meerdere redenen hebben. Het kan het gevolg zijn van verschillende schattingsmethodes, verschillende manieren van dataconstructie, of doordat de effecten daadwerkelijk verschillen. In hoofdstuk 6 gebruiken we internationaal vergelijkbare data en een state-of-the-art econometrische methode om op nationaal en sectoraal niveau de productiviteitseffecten van overheidskapitaal te schatten voor een groot aantal landen en over een lange tijdsperiode.

Veel studies lossen het probleem van niet-stationaire tijdreeksen op door het verband tussen overheidskapitaal en productiviteit in eerste verschillen te schatten. Zoals echter al is opgemerkt door Munnell (1992) wordt op deze manier het lange-termijn verband genegeerd. De methode die wij gebruiken, de 'Pooled Mean Group' schatter (Pesaran et al., 1999), is ontwikkeld om juist dit lange-termijn verband te schatten. Het gebruikt de cross-sectie dimensie in een panel data set om tot een efficiëntere schatting te komen dan alleen op basis van de enkele tijdreeks mogelijk is. De PMG-schatter veronderstelt dat slechts het lange-termijn verband hetzelfde is in elke afzonderlijke tijdsreeks en staat heterogeniteit toe in het aanpassingsproces naar deze lange-termijn relatie.

Helaas moeten we in hoofdstuk 6 concluderen dat stabiele uitkomsten, zowel op macroniveau als op sectoraal niveau, een illusie zijn. De productie elasticiteiten van overheidskapitaal variëren op sectoraal niveau voor verschillende specificaties en landen tussen -2 en 2. De resultaten op nationaal niveau lijken iets stabieler, maar variëren ook nog tussen 0.04 en 1.13. Hoewel het goed mogelijk is dat een deel van de variantie wordt veroorzaakt door verschillen tussen landen en sectoren, suggereren de resultaten dat het gevonden effect extreem gevoelig is voor ogenschijnlijk onschuldige veranderingen in specificatie.

Ondanks het zeer grote aantal studies over de effecten van overheidskapitaal op economische groei zijn er toch een aantal vragen onbeantwoord. Drie van deze vragen betreffen de rol van instituties in de creatie van nieuw overheidskapitaal, de rol van onderhoud en het gebrek aan een theoretische fundering in de meeste empirische studies. Concluderend kunnen we stellen dat hoewel duidelijk is dat de overheid met zijn investeringen productiviteit kan bevorderen, het nog steeds niet duidelijk is wat de grootte van dit effect bepaalt.

RuG

University of Groningen
P.O. Box 800
9700 AV Groningen
The Netherlands