

# Optimal education in times of ageing: The dependency ratio in the Uzawa–Lucas growth model

Citation for published version (APA):

von Gaessler, A. E., & Ziesemer, T. (2016). Optimal education in times of ageing: The dependency ratio in the Uzawa–Lucas growth model. *The Journal of the Economics of Ageing*, 7, 125-142.  
<https://doi.org/10.1016/j.jeoa.2016.03.001>

## Document status and date:

Published: 01/04/2016

## DOI:

[10.1016/j.jeoa.2016.03.001](https://doi.org/10.1016/j.jeoa.2016.03.001)

## Document Version:

Accepted author manuscript (Peer reviewed / editorial board version)

## Document license:

CC BY-NC-ND

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# Optimal Education in Times of Ageing: The Dependency Ratio in the Uzawa-Lucas growth model

---

Anne Edle von Gaessler<sup>a</sup>

Thomas Ziesemer<sup>b</sup>

## Abstract

The increasing share of retirees puts pressure on the shrinking working generation which will need to produce more output per worker to ensure a constant standard of living. We investigate the influence of a changing dependency ratio has on the time individuals spend in education and production. Longer education will increase productivity in the future, but will lower production in the short run, whereas an increase in labor input at the cost of education will provide more production immediately. We introduce a dependency ratio into a discrete-time Uzawa-Lucas model with international capital movements, human capital externalities and decreasing returns to labor in human capital formation. The dependency ratio is defined as the fraction between inactive and active individuals in regard to work or education. By calibration of the model, we find multiple steady states indicated by a u-shaped relation between education time-shares and the growth rate of the dependency ratio. Near the stable, high-level steady state, the optimal response to higher growth of the dependency ratio is more education to enhance productivity. We find evidence for this relation for 16 OECD countries. As a model extension, a debt-dependent interest rate has been introduced and estimated.

**Key Words:** Demographic Change, Education, Endogenous Growth,  
Human Capital Development

**JEL codes:** O15, J11

<sup>a</sup> Maastricht University, Department of Economics, P.O. Box 616, NL 6200 MD Maastricht  
Email: A.Edlevongaessler@maastrichtuniversity.nl

<sup>b</sup> Maastricht University, Department of Economics, and UNU-MERIT P.O. Box 616, NL 6200 MD Maastricht  
Email: T.Ziesemer@maastrichtuniversity.nl

## 1 Introduction

During the past decades concern has been raised towards the question whether and how the changing age structure in the USA, Europe, China and other countries will change economic behavior. Demographic change is caused by increased longevity on the one hand, and lower birth rates on the other (R. Lee, 2003). With increased longevity, a higher share of one's lifetime is spent in retirement as long as the retirement age is constant. Lower birth rates cause a decreasing share of population who will later be able to generate output to support the rest of the population. Also, higher life expectancy drives up the dependency ratio independently. This may result in shortcomings in supporting the non-working members of society. Instead of analyzing the cause of demographic change, this paper looks at its effects on economic growth and education. Consequently, demographic change is conceptualized by analyzing the two major consequences of the transition; the increasing number of retirees and the decreasing workforce. One way to measure these dynamics is to look at the dependency ratio and its development. It indicates how many people need to be supported relative to the number of people who are working. This concept will be introduced into an Uzawa-Lucas growth model.

Several studies have supported the pessimistic view of diminishing real output per capita and national savings rates due to population ageing within the next years if there is no impact on technical change (Bloom et al., 2010; Fayissa and Gutema, 2010; Hviding and Mérette, 1998; Muysken and Ziesemer, 2013, 2014). Wright et al. (2014) calculate the loss of per capita output caused by demographic change in absence of technological shifts to be more than 15% within the next 100 years. If the goal is to keep consumption per capita constant or even growing, production per worker needs to increase in order to keep up with the rising number of non-working members of the population. One way to increase production is to keep productivity per worker constant, but increase the time spent in production. The longer the production time, the more output can be generated with a constant productivity, implying a higher retirement age or more working hours per year. Another way is to increase the productivity of each worker and, hence, increase output per working-hour. This way, no extra time in production is needed, but this productivity increase comes at a cost. Workers need more education to learn how to produce more productively, indicating a higher time share devoted to education rather than production. The educational optimum is likely to change in times of ageing. Galasso (2008) analyzes the effect of a higher retirement age in several OECD countries, implying a constant educational share. His results are in favor of a higher retirement age, though, he neglects the possibility of a higher output through higher productivity. This paper focuses on the impact of the demographic change on education. The literature on the interaction of the demographic change and schooling is diverse. Through their vintage human capital model with a realistic survival law, Boucekkine et al. (2002) find that an increase in longevity results in longer schooling and a later retirement. The question remains, whether longer schooling is proportionate to the longer life span, or if there is a shift towards a higher or lower share of education. De la Croix and

Licandro (1999) show that an increasing life expectancy has a positive effect on the individual time devoted to schooling, but may have a negative effect on participation rates.

We do not deal explicitly with life expectancy, retirement age, fertility, and working hours before retirement. They are included implicitly in the exogenous growth rates of labour and population which we use. Other things constant, an earlier retirement leads to a smaller labour force and a higher dependency ratio. Similarly, living longer without working longer has the same effect on the dependency ratio because population growth is higher because of a lower mortality rate. More children at pre-school age also enhance the dependency ratio. Our formulation of the ratio does not require making all these details explicit as microeconomics studies, demographers and pension funds usually do. Countries differ in all these details because of differences in institutions. We do not go into all the institutional details in the following literature discussion, but this literature points to the importance of institutional aspects regarding health, (public) education, labour (see retirements discussion in the introduction and section 4 below) and capital markets. (Hansen and Lønstrup, 2012) show for a utility maximizing household that without student loans and correspondingly less education, an increase in the probability to go to the third period, during which a retirement decision is taken, leads to less of an increase of savings than with student loans and therefore the final period's utility is maximized by earlier rather than later retirement. Cervellati and Sunde (2013) extend this household decision model with a positive death probability in every period and perfect capital markets. In their model, decreasing mortality and reduction of labour time during the working period increase the benefits and reduce the opportunity costs of education. The resulting positive income effect yields more educated people and less labour supply. The latter is inessential for the impact of life expectancy on the share of people that receive schooling at given equilibrium wages represented by a Mincer equation. Cervellati and Sunde (2015a) add a decision on the hours spent on education to the schooling decision, but hours not spent on education cannot be invested in technical change because the demand side is not needed for the purpose of their paper. In contrast, the problem of ageing is to cause a general equilibrium effect of an increased dependency ratio driving up wages for human capital, and reducing labour demand of producers and providing a stronger incentive, but also leading to larger opportunity costs for a shifting time input into education. In a general equilibrium model by Cervellati and Sunde (2015b) fertility does not feedback into education, and because technical change is a byproduct of education it is not possible to move a factor out of production of output into that of technical change as both are increasing with the share of educated people when the economy is in a developed stage; once all people have education technical change is at a fixed rate. In the models of Cervellati and Sunde (2015a,b) and Bonneuil and Boucekkiné (2014) life expectancy has an impact on public education under high life expectancy, low mortality and fertility, but in both no other strategy like capital accumulation under international capital movements, or investing in technical change exists. In the Uzawa-Lucas model, which we use, shifting labour from production of output to that of human capital, identical to technical change, is possible. Decreasing returns to education time in human

capital formation and human capital externalities in final output production become crucial. Empirically, the positive effects dominate and education is enhanced. This generates technical change under an infinite horizon including all future generations. Limited life expectancy and mortality appear in our model only implicitly in the form of a positive rate of depreciation of human capital. What matters in the end for all countries' ageing, is the dependency ratio and its growth, unless one is interested in the details of the distribution across a continuum of generations, which is beyond the scope of this paper. As a byproduct of not using life expectancy and fertility, in our empirical parts we do not have to choose between different measures of life expectancy (age 5 or at birth) and mortality (adult, maternal, at working age, under 5, or infant), which have theoretically ambiguous and empirically controversial effects (Hazan, 2012).<sup>1</sup>

Therefore, in this paper a dependency ratio, defined as the population/labour ratio without the explicit modeling of the underlying age distribution, is introduced into a discrete-time Uzawa-Lucas model (Frenkel et al., 1996; Lucas, 1988; Uzawa, 1965) with capital movements, decreasing returns to labor in human capital formation, and human capital externalities in final output production to find out how the economy reacts to the new challenges. We find two optimal shares of education. In the lower steady state, the economy faces high participation rates in production with relatively little time in education, whereas the other steady state is characterized by high schooling. The latter one is stable. We are specifically interested in the optimal share of education and its implied effects on the economy, in particular its level and growth of productivity. A virtuous circle of education and productivity in connection with favorable demographic trends is most popular in the literature explaining the escape from per-industrial low income and growth. For example in the model of Becker et al. (1990) a human capital shock leads to lower fertility and a lower discount rate per child but a higher discount rate for all children together, which induces a reduction of a high number of children in favour of more education for each child, leading to more growth and less fertility again. Similarly, in the model by Cervellati and Sunde (2005) an improvement in life expectancy encourages human capital formation, which increases technical progress, which in turn encourages more human capital formation. In contrast, in our model for rich countries an unfavourable development of ageing triggers an increase in education and productivity.

The paper is structured as follows: In Section 2 the model will be set up. Section 3 analyzes the existence and stability of multiple steady states. Section 4 will introduce the data for the dependency ratio and the share in education, and shows the relation between the model and the data. Section 5 addresses the dynamics of foreign debt and Section 6 concludes.

---

<sup>1</sup> The complications stem from the diverse causes of the change in life expectancy. Hansen (2013) suggests that in regression analysis of the effect of life expectancy on schooling life expectancy should be instrumented by its causes, medical innovations related to childhood diseases (Hansen 2013); Cervellati and Sunde (2013) emphasize the shifting peak towards higher age of the survival distribution during working age.

## 2 The Model

We use a model by Frenkel et al. (1996), which is a discrete-time version of the Uzawa-Lucas with perfect international capital movements and non-increasing returns to labor<sup>2</sup> in human capital formation. Besides allowing for imperfect capital movements, we introduce exogenous growth rates of the active population and the total population. This will show how the long run outcome evolves if there is a discrepancy between the two rates, indicating demographic changes. They will be varied numerically to see the effect on the allocation and growth of the economy. The economy consists of representative households and firms maximizing their respective utility and profits.

The output of the economy is determined by capital,  $K_t$ , and human capital,  $H_t$ , and formed by a Cobb-Douglas production function, where human capital,  $H_t = h_t L_t$ , is the number of the members in the active population,  $L_t$ , times their respective skill level,  $h_t$ . With the total population denoted as  $N_t$ , the dependency ratio is defined as the inactive population over the active population,  $D_t = \frac{N_t - L_t}{L_t}$ . The active population is defined as the part of the population that is actively engaged in either the production of output, or education. The inactive population contains all others. This definition can be rearranged in terms of  $L_t$  to see how the active population interacts with the dependency ratio,  $L_t = \frac{N_t}{1 + D_t}$ . For a given population size, the active population decreases if the dependency ratio increases. Replacing this in the previous definition for total human capital gives:  $H_t = h_t \frac{N_t}{1 + D_t}$ . The production function then is:

$$Y_t = A(K_t)^{1-\alpha} \left( (1 - e_t) h_t \frac{N_t}{1 + D_t} \right)^\alpha \bar{h}_t^\epsilon \quad (1)$$

The productivity level,  $A$ , is assumed to be constant and  $h_t$  may grow. The  $L_t$  agents in this economy decide to spend their time either in education  $e_t$ , or production  $(1 - e_t)$ . Equivalently, we can think of  $e_t$  as the share of the active population in education and  $(1 - e_t)$  as the share in production in a given time frame (i.e. one year). By implication, our definition of the variable  $L_t$  deviates from the standard labour market definition, where the active population does not include those in education. We do this because classifying them as inactive would ignore the fact that this part of the labour time contributes a period later via human capital to output production, whereas the difference between  $N_t$  and  $L_t$  remains inactive.  $N_t - L_t$  can be interpreted as the number of old people who will not return to schooling because they are too old. Average human capital  $\bar{h}_t^\epsilon$  contributes to the productivity of all factors and is modeled after Lucas (1988). As no single person can influence average human capital, the representative optimizing agent takes the externality  $\bar{h}_t^\epsilon$  as given when deciding on their optimal time spent in education, which leads to a

---

<sup>2</sup> Uzawa (1965) also used decreasing returns; Lucas (1988) simplified to assuming constant returns in order to allow for an explicit solution for the long-term growth rates in equilibrium and optimum.

second best solution. Hence, what we call an optimal solution is in fact a second best solution. Through including externalities of human capital formation into the production function for final output there may be two steady states for each growth rate of the dependency ratio. Xie (1994) establishes the possibility of multiple steady states for a large enough external effect of human capital. By implication, it is an empirical question, whether education should be increased or decreased in response to ageing. In general, preferences, technologies, endowments and institutions play a role in this decision. In the Uzawa-Lucas model, there is a special role of a positive human capital externality in final output production, which encourages human capital production from the demand side, but there is also a decreasing marginal effect of more education time in human capital formation, which discourages human capital formation. The question of which one is stronger, is not obvious without theoretical and empirical investigation in a multiple steady state model connecting them to other assumptions on preferences, technologies, endowments of active and inactive population parts, and institutional assumptions. In our model, we do not consider health aspects, education is a private decision, capital markets are perfect up to the interest rate mark up when the debt/GDP ratio increases; the labour market is Walrasian up to the human capital externality on the demand side.

The consumer in this economy owns physical and human capital which (s)he supplies to the production sector. The demand for these is determined by the firms which solve a static maximization program:

$$\max_{(1-e_t), K_t} \pi = A(K_t)^{1-\alpha} \left( (1-e_t)h_t \frac{N_t}{1+D_t} \right)^\alpha \bar{h}_t^\epsilon - \omega_t (1-e_t)h_t \frac{N_t}{1+D_t} - r_{kt} K_t \quad (2)$$

The first term on the RHS is the output of the firms, the second term is the cost of wages and the third term is the cost of capital.

The first-order conditions for  $(1-e_t)$  and  $K_t$  are

$$\omega_t = \frac{\alpha Y_t}{(1-e_t)h_t \frac{N_t}{1+D_t}} \quad (3)$$

$$r_{kt} = (1-\alpha) \frac{Y_t}{K_t} \quad (4)$$

In (3) the equilibrium consequences of an increase in the dependency ratio are observable. If the dependency ratio increases, *ceteris paribus*, equilibrium wages will also increase. Alternatively, for given wages and output, either time spent in production goes up (increase in  $(1-e_t)$ ), or individual human capital increases over time. An increase in human capital can be obtained by spending more time in education (see equation (6) below). Here the ambivalence of how time should be optimally spent becomes clear.

The consumers' utility is given by an isoelastic utility function,  $U_t = \sum_{t=0}^{\infty} \beta^t N_t \frac{c_t^{1-\sigma}}{(1-\sigma)}$ , with  $0 < \beta < 1$  as the subjective discount factor and  $\sigma > 0$  as the intertemporal elasticity of substitution in consumption.

The consumers' budget constraint is

$$N_t c_t + K_{t+1} - (1 - \delta_k)K_t = \omega_t(1 - e_t)h_t \frac{N_t}{1+D_t} + r_{kt}K_t + B_{t+1} - \left(1 + r\left(\frac{B_t}{Y_t}\right)\right)B_t \quad (5)$$

Where expenses for consumption in period  $t$ ,  $N_t c_t$ , and savings,  $K_{t+1} - (1 - \delta_k)K_t$ , must equal the income from wages,  $\omega_t(1 - e_t)h_t \frac{N_t}{1+D_t}$ , and capital,  $r_{kt}K_t$ , plus the borrowings,  $B_{t+1}$ , minus the debt service,  $(1 + r\left(\frac{B_t}{Y_t}\right))B_t$ . The interest rate  $r\left(\frac{B_t}{Y_t}\right)$  is an increasing function of the debt to GDP ratio. For now,  $r\left(\frac{B_t}{Y_t}\right)$  is assumed to be constant. This assumption will be relaxed in Section 5 in which a realistic function of  $r\left(\frac{B_t}{Y_t}\right)$  will be estimated. Taking care of the impact of the debt/GDP ratio on the interest rate can be interpreted as a household or government activity. But as there is no other separate government activity we integrate this task into the household decision assuming households take care of their impact on the interest rate. As there is only one household who owns the firm it can also take over this small government task. Moreover, the household cares for the old. As we do not have many heterogeneous households we do not have to decide on which pension model to use. All real-world pension models have been mis-designed and all ran into problems whenever ageing came about. Designing pension fund models is beyond the scope of this paper. The only assumption we make in this regard is that the utility function below has the same per capita consumption for the inactive  $N_t - L_t$  (old and pre-school children) and the active  $L_t$  (young, working and in school).<sup>3</sup>

Human capital formation is described as

$$h_{t+1} = F e_t^\gamma h_t + (1 - \delta_h)h_t \quad (6)$$

$F$  is the knowledge efficiency coefficient,  $\delta_h$  is the depreciation rate of human capital and  $\gamma$  is the productivity parameter, with  $\gamma \leq 1$ , to assure diminishing or constant returns to education time.

The consumers maximize their utility subject to the budget constraint (5) and the development of human capital (6) for given initial values  $K_0, h_0, B_0$ .

---

<sup>3</sup> Muysken and Ziesemer (2014) assume that the proportion of consumption of old and young is fixed, but not necessarily unity.

$$\begin{aligned} \max_{c_t, e_t, B_{t+1}, K_{t+1}, h_{t+1}} \sum_{t=0}^{\infty} \beta^t \left( N_t \frac{c_t^{1-\sigma}}{(1-\sigma)} \right. \\ \left. - \mu_t \left[ N_t c_t + K_{t+1} - (1-\delta_k)K_t - \omega_t(1-e_t) \frac{N_t}{1+D_t} h_t - r_{kt}K_t - B_{t+1} \right. \right. \\ \left. \left. + \left( 1 + r \left( \frac{B_t}{Y_t} \right) \right) B_t \right] - \mu_{ht} [h_{t+1} - F e_t^\gamma h_t - (1-\delta_h)h_t] \right) \end{aligned}$$

The first order conditions are:

$$c_t: c_t^{-\sigma} = \mu_t \quad (7)$$

$$e_t: \mu_t \omega_t \frac{N_t}{1+D_t} = \mu_{ht} F \gamma e_t^{\gamma-1} \quad (8)$$

$$B_{t+1}: \mu_t = \beta \mu_{t+1} \left( 1 + r \left( \frac{B_{t+1}}{Y_{t+1}} \right) \right) + \beta \mu_{t+1} \frac{B_{t+1}}{Y_{t+1}} r' \left( \frac{B_{t+1}}{Y_{t+1}} \right) \quad (9)$$

$$K_{t+1}: \mu_t = \beta \mu_{t+1} (1 - \delta_k + r_{kt+1}) \quad (10)$$

$$h_{t+1}: \mu_{ht} = \beta \left[ \mu_{t+1} \omega_{t+1} (1 - e_{t+1}) \frac{N_{t+1}}{1+D_{t+1}} + \mu_{ht+1} F e_{t+1}^\gamma + (1 - \delta_h) \mu_{ht+1} \right] \quad (11)$$

The following transversality conditions must hold<sup>4</sup>:

- I.  $\lim_{t \rightarrow \infty} \beta^t \mu_t K_t = 0$
- II.  $\lim_{t \rightarrow \infty} \beta^t \mu_{ht} h_t = 0$

Equations (1)-(11) define a system of 11 equations for 10 endogenous variables,  $Y_t, K_t, h_t, e_t, \omega_t, r_{kt}, c_t, B_t, \mu_t$ , and  $\mu_{ht}$ . Due to Euler's theorem, equations (1)-(4) are linearly dependent, if (2) equals zero. Hence, one of them ((2) in this case) can be dropped from the system. This leaves 10 variables and 10 equations. From the 10x10 equation system we can derive the rates of return to physical capital, bonds and human capital.

$$\frac{1}{\beta} \left( \frac{c_{t+1}}{c_t} \right)^\sigma = \frac{\mu_t}{\beta \mu_{t+1}} = R_{B_{t+1}} = R_{K_{t+1}} = R_{H_{t+1}} \quad (12a)$$

$$R_{B_{t+1}} = 1 + r \left( \frac{B_{t+1}}{Y_{t+1}} \right) (1 + \eta_{rb}) \quad (12b)$$

$$R_{K_{t+1}} = 1 - \delta_k + (1 - \alpha) \frac{Y_{t+1}}{K_{t+1}} \quad (12c)$$

$$R_{H_{t+1}} = (1 + g_\omega) \frac{1+g_N}{1+g_{1+D}} F \gamma e_t^{\gamma-1} \left[ 1 - e_{t+1} + \frac{1}{\gamma} e_{t+1} + \frac{1-\delta_h}{F \gamma e_{t+1}^{\gamma-1}} \right] \quad (12d)$$

The first equation in (12a) is derived from (7) and the other expressions follow from (12b-d), as equations (9), (10) and (11) are rearranged to equal  $\frac{\mu_t}{\beta \mu_{t+1}}$ . Equation (12b) is derived from

<sup>4</sup> For proof that the utility function has a finite integral, and hence has an interior maximum see Appendix A.

(9) with  $\eta_{rb} = \frac{B_{t+1}}{Y_{t+1}} \frac{r' \left( \frac{B_{t+1}}{Y_{t+1}} \right)}{r \left( \frac{B_{t+1}}{Y_{t+1}} \right)}$ , (12c) is derived from (10) where  $r_{kt+1}$  is replaced by the expression in (4) and equation (12d) is derived by (11) where  $\mu_{ht}$  is replaced by the relation in (8).

Because  $r \left( \frac{B_{t+1}}{Y_{t+1}} \right)$  is assumed to be constant for now  $r' \left( \frac{B_{t+1}}{Y_{t+1}} \right) = 0$ . This leads to  $\eta_{rb} = 0$ . For completeness of the model and because it will become crucial in later sections, the  $\eta_{rb}$  term is still carried along and will be set to zero whenever necessary. (12b) implies a constant growth rate of  $c_t$  for constant interest rate  $r$ , which implies constant  $R_{Kt+1}$  and  $R_{Ht+1}$  in (12c) and (12d).

This implies constancy of  $\frac{Y_{t+1}}{K_{t+1}}$  in (12c), from which follows that output and capital grow at

the same rate. It follows that  $\frac{Y_{t+1}}{K_{t+1}} = A \left( \frac{K_{t+1}}{(1-e_{t+1})h_{t+1} \frac{N_{t+1}}{1+D_{t+1}}} \right)^{-\alpha} \bar{h}_{t+1}^\epsilon$  must be constant.

Hence, we can derive how the growth rates of  $K_t$ ,  $e_t$  and  $h_t$  are related. Note that  $\bar{h}_{t+1} = h_{t+1}$  since  $\bar{h}_{t+1}$  is the average human capital of identical households. This implies

$$\frac{Y_{t+1}}{K_{t+1}} = A \left( \frac{K_{t+1}}{(1-e_{t+1})h_{t+1} \frac{N_{t+1}}{1+D_{t+1}}} \right)^{-\alpha}$$

Constancy implies equality of the growth rates of the numerator and the denominator:

$$1 + g_Y = 1 + g_K = (1 + g_{1-e})(1 + g_h)^{1+\frac{\epsilon}{\alpha}} \frac{1+g_N}{1+g_{1+D}} \quad (12c)^I$$

(12c)<sup>I</sup> shows the growth rate of capital in relation to the growth rates of time spent in production,  $(1 - e_t)$ , individual human capital,  $h_t$ , from (6), population,  $N_t$  and the dependency ratio,  $(1 + D_t)$ . An increase in the growth rate of the dependency ratio decreases the growth rate of capital, ceteris paribus. This can, once again, either be offset by increasing the growth rate of time spent in production (which is clearly only a short term measure, as time spent in production cannot exceed 100%), or by increasing the growth rate of human capital. This can be increased by increasing  $e_t$ , the share of education (see (6)). (12c)<sup>I</sup> shows the trade-off of time spent in the different sectors when keeping  $g_K$  constant.

Equation (12d) is the central equation to find a steady state expression for  $e_t$ . (3) implies:

$$1 + g_\omega = \frac{1+g_Y}{(1+g_{1-e})(1+g_h) \frac{1+g_N}{1+g_{1+D}}} \quad (3)^I$$

It has been established above that output and capital grow at the same rate, inserting (12c)<sup>I</sup> into (3)<sup>I</sup> leads to:

$$1 + g_\omega = (1 + g_h)^{\frac{\epsilon}{\alpha}} \quad (3)^{II}$$

Together with (6) this shows that  $g_\omega$  is constant if  $e_t$  is constant (i.e. if  $g_e = 0$ ).  $(1 + g_\omega)$  can be replaced in (12d) to relate  $e_t$  and  $e_{t+1}$  to only exogenous variables:

$$R_{Ht+1} = (F e_t^\gamma + (1 - \delta_h))^\frac{\epsilon}{\alpha} \frac{1+g_N}{1+g_{1+D}} F\gamma e_t^{\gamma-1} \left[ (1 - e_{t+1}) + \frac{1}{\gamma} e_{t+1} + \frac{(1-\delta_h)}{F\gamma e_{t+1}^{\gamma-1}} \right] \quad (12d)'$$

As this expression includes time spent in education of the current and future period, no analysis can be done yet about their behavior in and around the steady state. To do this  $e_t$  will be linked to its growth rate. Multiplying  $e_t^{\gamma-1}$  into the brackets yields:

$$R_{Ht+1} = (F e_t^\gamma + (1 - \delta_h))^\frac{\epsilon}{\alpha} \frac{1+g_N}{1+g_{1+D}} F\gamma \left[ e_t^{\gamma-1} - e_t^\gamma \frac{e_{t+1}}{e_t} + e_t^\gamma \frac{1}{\gamma} \frac{e_{t+1}}{e_t} + \left( \frac{e_t}{e_{t+1}} \right)^{\gamma-1} \frac{(1-\delta_h)}{F\gamma} \right]$$

Using  $\frac{e_{t+1}}{e_t} = 1 + g_e$  we get:

$$R_{Ht+1} = (F e_t^\gamma + (1 - \delta_h))^\frac{\epsilon}{\alpha} \frac{1+g_N}{1+g_{1+D}} F\gamma \left[ e_t^{\gamma-1} - e_t^\gamma (1 + g_e) + e_t^\gamma \frac{1}{\gamma} (1 + g_e) + \left( \frac{1}{1+g_e} \right)^{\gamma-1} \frac{(1-\delta_h)}{F\gamma} \right] \quad (12d)''$$

With  $R_{Ht+1}$  constant, the RHS of (12d)'' is constant as well. This is the dynamic equation that shows how  $e_t$  develops over time, depending on several parameters and exogenous variables. Note that  $g_N$ ,  $g_{1+D}$  and all parameters are exogenous and constant. With help of this equation, the stability of  $e_t$  in the steady state can be analyzed. Unfortunately, the above expression cannot be solved for  $e_t$ , or  $g_e$  analytically. In the next section the model will be calibrated to find and analyze the steady state conditions.

### 3 Existence and Partial Stability of Multiple Steady States

In this section the relation between  $g_e$  and  $e_t$  will be analyzed. By calibration of the model two steady states are found. The RHS of the steady state condition (12d)'' is constant if either  $g_e = 0$  and hence  $e_t$  is constant.

Understanding the interactions between  $g_e$  and  $e_t$  is crucial for finding possible steady states. Since (12d)'' can neither be solved for  $g_e$  nor for  $e_t$ , it needs to be differentiated implicitly to find how the relationship behaves. (13) shows the derivative of  $g_e$  with respect to  $e_t$ .<sup>5</sup>

$$g_e' = \frac{\frac{\epsilon}{\alpha} (F e_t^\gamma + (1 - \delta_h))^{-1} F\gamma e_t^{\gamma-1} \left[ e_t^{\gamma-1} - e_t^\gamma (1+g_e) + e_t^\gamma \frac{1}{\gamma} (1+g_e) + \left( \frac{1}{1+g_e} \right)^{\gamma-1} \frac{(1-\delta_h)}{F\gamma} \right] - [(\gamma-1)e_t^{\gamma-2} - \gamma e_t^{\gamma-1} (1+g_e) + e_t^{\gamma-1} (1+g_e)]}{\left[ -e_t^\gamma + e_t^\gamma \frac{1}{\gamma} + (1-\gamma)(1+g_e) - \gamma \frac{(1-\delta_h)}{F\gamma} \right]} \quad (13)$$

<sup>5</sup> The full derivation is in Appendix B.

For our purposes it proves helpful to do a full analysis of this relation. To find a turning point (maximum, or minimum),  $g'_e = 0$  is required:

$$-\frac{\epsilon}{\alpha} \left( F e_t^\gamma + (1 - \delta_h) \right)^{-1} F \gamma \left[ e_t^{\gamma-1} + e_t^\gamma (1 + g_e) \left( \frac{1}{\gamma} - 1 \right) + \left( \frac{1}{1+g_e} \right)^{\gamma-1} \frac{(1-\delta_h)}{F\gamma} \right] = (1 - \gamma)(1 + g_e - e_t^{-1}) \quad (13)^1$$

This is the most this equation can be simplified without making assumptions about the magnitude of the parameters. Some parameters are set in the literature. The following common values have been applied:  $\alpha = 0.6$  and  $\delta_h = 0.03$ . Lucas (1988) calibrates  $\epsilon$  to US data for the case  $\gamma = 1$ . As we use  $\gamma < 1$ , in order to get empirically realistic results we will use twice Lucas' assumption and set  $\epsilon = 0.834$ , which is necessary when the rate of depreciation for human capital is not zero and  $\gamma < 1$ .<sup>6</sup> Alternatively, we could have set a smaller rate of depreciation and/or a higher  $\gamma$  explained next.  $F$  and  $\gamma$  are interdependent through equation (6) for given other parameters and data. In order to find reasonable values for  $F$  and  $\gamma$  to fulfill condition (6) for given data,  $g_h$  is set to 0.011, close to Denison's estimate for the United States (Denison, 1962), also used by Lucas. Finding combinations of values for  $F$  and  $\gamma$ , which satisfy equation (6) is done for given  $e$ . We set the current value to  $= 0.334$ , the panel mean of 16 OECD countries for the period 1985 to 2010 (see next section for details on data). For our purposes, from a set of combinations of values of data for  $F$  and  $\gamma$  we choose  $F = 0.055$  and  $\gamma = 0.268$  (see Appendix C for justification) to ensure greatest possible similarity between the simulation and data analysis in Section 4 below. Whereas Mankiw et al. (1992) assume the same rate of depreciation for physical and human capital, Lucas (1988) clearly prefers a lower one for human capital because of the intergenerational transfer of knowledge and sets it equal to zero. In the remainder of this paper, we follow Mankiw et al. (1992) and set the depreciation rate of human capital equal to the depreciation rate of physical capital,  $\delta_h = \delta_k = 0.03$ . In Appendix D however, we take an intermediate position to show that the choice of the rate of depreciation affects the adequate discount rate and the values of the solution of the model only marginally. We also set  $r \left( \frac{B_{t+1}}{Y_{t+1}} \right) = 0.05$  and  $g_N = 0.002$ , in line with the literature for OECD countries. For the steady state, and in line with the panel average of 16 OECD countries 1985 to 2010, it is reasonable to assume that the growth rate of the dependency ratio will be stable, while ageing populations are possibly a transitional phenomenon. Hence, we set  $g_{1+D} = 0$ . This assumption will be relaxed later on.

---

<sup>6</sup> An implication of these assumptions is that we do not get an indeterminacy result for  $\epsilon > \alpha$  as the literature does under the assumptions made by Lucas (Benhabib and Perli, 1994; Xie, 1994).

## Optimal Education in Times of Ageing

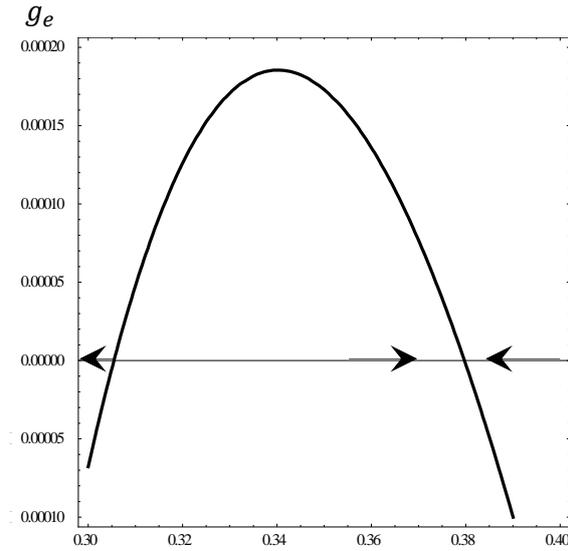


Figure 1 – Dynamics in  $e_t$

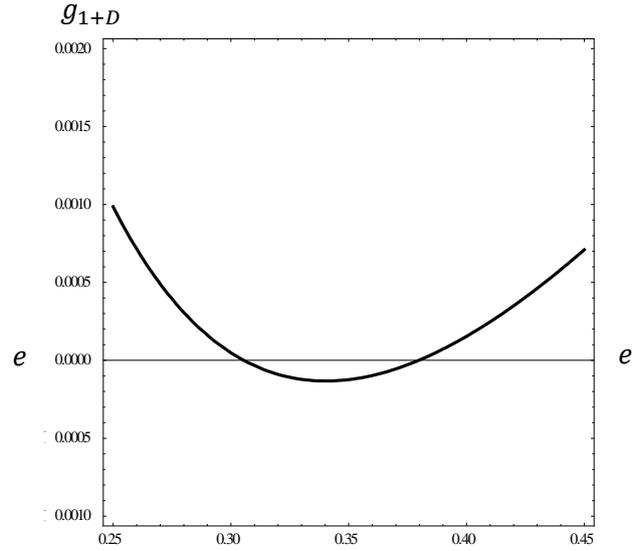


Figure 2 – Growth rate of the activity ratio

These values are inserted into (13)<sup>I</sup>, which then can be solved for  $e_t = 0.354$ . This shows that there is a single value at which  $g_e' = 0$ , which must then be the only maximum, or minimum of the function (12d)<sup>II</sup>. With the given parameters and the assumption  $g_e = g_{1+D} = 0$ , the roots of function (12d)<sup>II</sup> are found to be  $e_1 = 0.305$  or  $e_2 = 0.380$ . As  $g_e'$  is negative for  $e_t > e_2$  and positive for  $e_t < e_2$ ,  $g_e$  is positive according to (12d)<sup>II</sup> for values of the parameters presented above and of  $e_t$  below  $e_2$  and above  $e_1$ , and negative above  $e_2$  and below  $e_1$ , the possibility of a minimum can be excluded, as it would require the opposite values of  $g_e$  and  $g_e'$  for the respective regions. This establishes that the function (12d)<sup>II</sup> must have a maximum at  $e_t = 0.354$  in the  $g_e - e$  plane. Figure 1 can then be drawn in the  $g_e - e$  plane with a maximum at  $e_{max} = 0.354$ , and roots at  $e_1 = 0.305$  or  $e_2 = 0.380$ . Because  $g_e$  is the growth rate of  $e_t$ , the stability of both steady states can be evaluated.  $e_2 = 0.380$  is stable and  $e_1 = 0.305$  is not, again because  $g_e$  is positive for values of  $e_t$  below  $e_2$  and above  $e_1$ , and negative above  $e_2$  and below  $e_1$ . Once out of the steady state, the economy will always return to  $e_2$  if the starting point is to the right of  $e_1$ . This shows the existence of multiple steady states with a stable one at  $e_2 = 0.380$  for the parameters indicated above.

The steady state value of  $e$  can be interpreted in two different ways as the labour endowment can be seen as total hours available for work and education or as total number of persons on a job or in education. It is either the time share each individual spends in education per year, or the share of the active persons in the population that is engaged in education in a given year. As we have to link it to the available data, we prefer the second interpretation, as it gives a more exact indication in times of changing demographics and it is closer to the spirit of the model.  $e_2 = 0.380$  can then be interpreted as 38% of the active population are engaged in education and, hence, 62% are in production in the steady state. The data for education time that we use below includes continuous vocational training and on the job training.

In the model there are 10 endogenous variables:  $Y_t, K_t, h_t, e_t, \omega_t, r_{kt}, c_t, B_t, \mu_t$  and  $\mu_{ht}$ . Their growth rates need to be determined within the model. The numerical values of the stable steady state will be derived. In addition to values chosen above, the value for  $\sigma$  is set to  $1.06^7$  and  $\beta = 0.982$ .

The steady state values in Table 1 indicate that an increase in the growth rate of the dependency ratio  $g_{1+D}$  starting from zero, leads to a decrease in the growth rates of output, capital and the shadow price of human capital in the long run, ceteris paribus.

<b>Steady State Relations</b>	<b>From Equation</b>	<b>Numerical</b>
$1 + g_Y = \left( F e_2^\gamma + (1 - \delta_h) \right)^{1+\frac{\epsilon}{\alpha}} \left( \frac{1+g_N}{1+g_{1+D}} \right)$	(1), (12c)', (6)	$g_Y = \mathbf{0.031}$
$1 + g_K = \left( F e_2^\gamma + (1 - \delta_h) \right)^{1+\frac{\epsilon}{\alpha}} \left( \frac{1+g_N}{1+g_{1+D}} \right)$	(12c)', (1), (6)	$g_K = \mathbf{0.031}$
$1 + g_h = F e_2^\gamma + (1 - \delta_h)$	(6)	$g_h = \mathbf{0.011}$
$1 + g_\omega = \left( F e_2^\gamma + (1 - \delta_h) \right)^\frac{\epsilon}{\alpha}$	(3)'', (6)	$g_\omega = \mathbf{0.017}$
$1 + g_\mu = \left( \beta \left( 1 + r \left( \frac{B_{t+1}}{Y_{t+1}} \right) (1 + \eta_{rb}) \right) \right)^{-1}$	(9)	$g_\mu = \mathbf{-0.030}$
$1 + g_{\mu h} = \frac{1}{\beta \left( 1 + r \left( \frac{B_{t+1}}{Y_{t+1}} \right) (1 + \eta_{rb}) \right)} \left( F e_2^\gamma + (1 - \delta_h) \right)^\frac{\epsilon}{\alpha} \frac{1+g_N}{1+g_{1+D}}$	(8)	$g_{\mu h} = \mathbf{-0.012}$
$1 + g_c = \left[ \beta \left( 1 + r \left( \frac{B_{t+1}}{Y_{t+1}} \right) (1 + \eta_{rb}) \right) \right]^\frac{1}{\sigma}$	(7) and (9)	$g_c = \mathbf{0.029}$

**Table 1 - Steady states**

Note to Table 1:  $r \left( \frac{B_{t+1}}{Y_{t+1}} \right) (1 + \eta_{rb}) = 0.05$ ,  $\alpha = 0.6$ ,  $\delta_h = 0.03$ ,  $g_N = 0.002$ ,  $g_{1+D} = 0$ ,  $F = 0.055$ ,  $\gamma = 0.268$ ,  $\epsilon = 0.834$ ,  $\sigma = 1.06$ ,  $\beta = 0.982$  and  $e_2 = 0.380$ . For the transversality conditions to hold, the growth rate of  $\beta^t \mu_t K_t$  and  $\beta^t \mu_t h_t$  must be negative. With  $\beta = 0.982$ , the growth rate of  $\beta^t$  is  $-0.018$ . With  $g_\mu = -0.030$  and  $g_K = 0.031$ , the growth rate of the first expression is negative. The growth rate of the second product is also negative because  $g_{\mu h} = -0.012$  and  $g_h = 0.011$ . The transversality conditions are, hence, fulfilled.

The detailed calculations of Table 1 can be understood as follows in connection with the parameter values below Table 1. The growth rate for human capital  $h$  follows from (6), using the steady state value  $e_2 = 0.38$ . Equation (12c)' together with the assumption of a constant interest rate and the underlying debt/GDP ratio, leads to a constant capital-output ratio; using the production function and again the previous result from (6) yields the growth rates for  $Y$  and  $K$ . The result from (6) in (3)'' yields the growth rate of wages. The growth rate for the marginal utility of income,  $1 + g_\mu$ , follows directly from (9). The latter result together with (8) and the growth rates for wage and human capital yield the growth rate of the value of the shadow price of human capital. Finally, (7) and (9) yield the growth rate of consumption. The constancy of all these growth rates depends on those of the labour share in education  $e$ , the debt/GDP ratio,  $g_N$  and  $g_{1+D}$ . Existence, uniqueness and stability of the consumption and debt part of the steady state of the model will be discussed in Section 5.

<sup>7</sup> Climate change papers use values between one and two. Denk and Weber (2011) use 2,5 as value for  $\sigma$ . We use 1.06 to ensure stability in the debt dynamics in Section 5.

In Figure 1, the steady state was analyzed for  $g_{1+D} = 0$ . If this assumption is relaxed, the relation between steady state values for  $e$  and  $g_{1+D}$  can be plotted using (12d)<sup>11</sup> as in Figure 2 above for  $g_e = 0$  and given values of  $r \left( \frac{B_{t+1}}{Y_{t+1}} \right) (1 + \eta_{rb})$ ,  $g_N$ ,  $\alpha$ ,  $\delta_h$ ,  $F$ ,  $\gamma$  and  $\epsilon$ .

The u-shaped curve of Figure 2 reflects the two steady-state time shares of education for each growth rate of the activity ratio. The active population can either be less educated and spend more time in production to increase output or invest more time in education and be more productive in the future. As has been analyzed in this section, the stable steady states are on the upward sloping part, indicating a higher optimal share of education for a higher growth rate of the activity ratio if only taking stable equilibria into account.

Xie (1994) sees the reason for multiple steady states in the effects of the externality ( $\epsilon$ ). By disentangling the effects of the externality and the diminishing returns to time spent in human capital formation ( $\gamma$ ), it becomes clear that the interaction of these two parameters is vital to the existence of the two steady states. The contribution of both will be shown by setting first only  $\epsilon$  equal to zero and second only  $\gamma$  equal to one.

Figure 3a shows the steady-state relationship (for constant  $g_e$ ) between  $g_{1+D}$  and  $e_t$  with constant returns to time spent in human capital formation ( $\gamma = 1$ ) but positive externality and all other parameters as in Table 1. By setting  $\gamma$  equal to one, the relationship becomes almost linear and upward sloping,<sup>8</sup> indicating an unambiguous increase in the share of education if the growth rate of the activity ratio increases.<sup>9</sup> For a stable activity ratio the steady state share of education is 0.023, suggesting an unrealistically low share in education when compared with the data considered in detail in the next section. Gruescu (2006) also finds a linear relationship between the growth rate of the dependency ratio and the time shares. In her model she considers neither the externality effect, nor diminishing returns to education time in human capital formation. However, when the more realistic case of decreasing returns is allowed for as we do here, the education policy decision is complicated by a non-linearity implying multiple steady states. This requires ruling out the policy alternative of increasing labor input in production and decreasing it in education as shown in Figure 3b and above in Figures 1 and 2, where an increase in  $g_{1+D}$  has two possible reactions. Figure 3b displays the steady-state relationship (for constant  $g_e$ ) between  $g_{1+D}$  and  $e_t$  if the externality effect is set to zero, but diminishing returns to time spent in human capital formation are present. Without the externality effect of human capital, the relationship between the share of education and the growth rate of the activity ratio is negative. With diminishing returns to time spent in human capital formation and no positive externalities it is not profitable anymore to invest more time in education in times of ageing.

<sup>8</sup> The exact formula is  $g_{1+D} = 1.0392(0.119e+0.97)^{1.39} - 1$ .

<sup>9</sup> In the calibration, in order for (6) to hold, the value of  $F$  needs to be adjusted to 0.119 in the same manner as done before.

The two effects together create the two steady states that we observe in the model.<sup>10</sup> This is in line with the findings of Xie (1994) and Zhang (2013), who prove the existence and (in-) stability of multiple equilibria, if “the external effect of human capital in goods production is sufficiently large” (Xie, 1994). Finding a realistic combination of the most important parameters is of utmost importance for a good policy decision in the presence of multiple steady states.

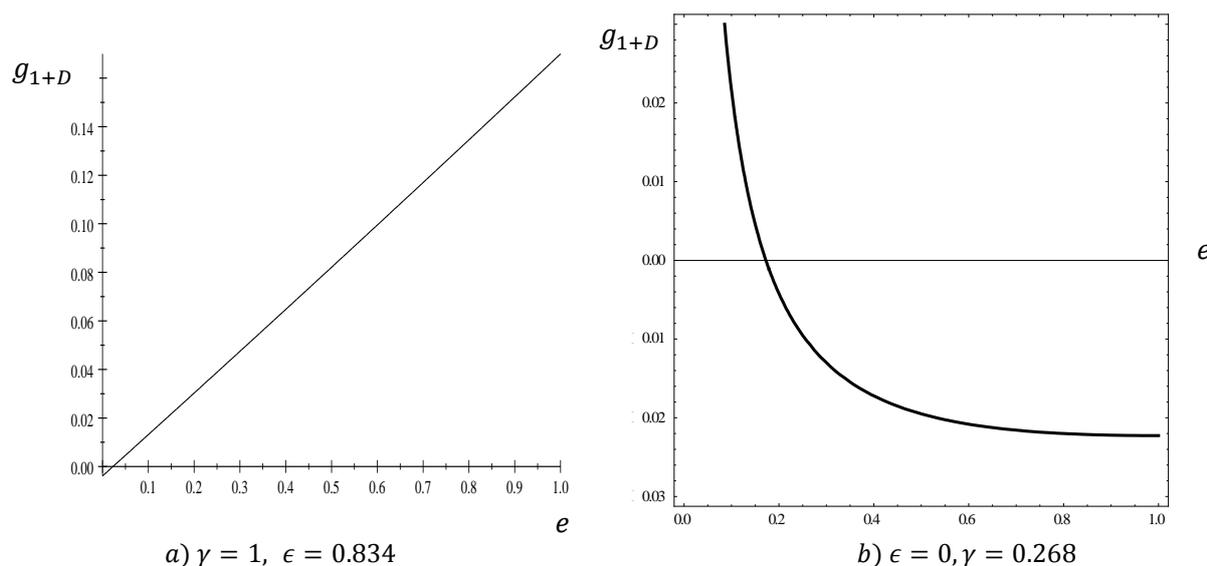


Figure 3 – Case Study for  $\gamma$  and  $\epsilon$

## 4 Empirical Analysis

Because of the importance of the empirical details for policy under multiple steady states this section provides some empirical insights into the relationship between the growth rate of the dependency ratio and the time spent in education as defined in the model. This section will show that the optimal response to a changing dependency ratio, calculated above and captured in Figures 1 and 2, also holds empirically. If this were not the case, we should have used different parameter values. So far, there is no readily constructed data to display the variable time spent in education,  $e_t$ , nor for the growth rate of the dependency ratio,  $g_{1+D}$ , as displayed in the model. Consequently, the steady state relation has not yet been estimated to this extent. This section will first provide insights into the construction of the data before the steady state relation will be estimated.

### 4.1 International Panel Data

The dependency ratio displays the relation of the non-working to the working population. It gives an indication of how many people in the economy need to be supported by the workforce. The higher the dependency ratio, the more people need to be supported. In many data sources the dependency ratio is typically referred to as the dependence between

<sup>10</sup> Combining the two cases  $\epsilon = 0$  and  $\gamma = 1$  leads to solution-problems in (12d)<sup>l</sup>. There is no variable to ensure equality because only exogenous variables are left.

age groups, indirectly making very specific assumptions about the relation between age and employment. The age-dependency ratio of e.g. the OECD and the World Development Indicators is defined as the number of people of age 65 and older over the number of people between 15 and 64 years of age. This reveals several assumptions: (i) children under the age of 15 are not working (typically engaged in education), (ii) people between the ages 15 and 64 are engaged in production and (iii) the retirement age is 65 years. A look at the official retirement ages in OECD countries shows that assuming a universal retirement age at 65 is not accurate enough.

<b>Australia</b>	65,0	<b>Hungary</b>	60,0	<b>Norway</b>	67,0
<b>Austria</b>	65,0	<b>Iceland</b>	67,0	<b>Poland</b>	65,0
<b>Belgium</b>	60,0	<b>Ireland</b>	65,0	<b>Portugal</b>	65,0
<b>Canada</b>	65,0	<b>Italy</b>	59,0	<b>Slovak Republic</b>	62,0
<b>Czech Republic</b>	61,0	<b>Japan</b>	65,0	<b>Spain</b>	65,0
<b>Denmark</b>	65,0	<b>Korea</b>	60,0	<b>Sweden</b>	65,0
<b>Finland</b>	65,0	<b>Luxembourg</b>	60,0	<b>Switzerland</b>	65,0
<b>France</b>	60,5	<b>Mexico</b>	65,0	<b>Turkey</b>	44,9
<b>Germany</b>	65,0	<b>Netherlands</b>	65,0	<b>United Kingdom</b>	65,0
<b>Greece</b>	57,0	<b>New Zealand</b>	65,0	<b>United States</b>	66,0
<b>OECD average</b>	<b>63,0</b>				

**Table 2 – Official Retirement Ages in the OECD**

Source: National officials, OECD calculations and Turner (2007)

Table 2 shows that the official retirement ages range between 44.9 years in Turkey and 67 years in Iceland and Norway. The difference of 22.1 years is too big to be neglected. There are many ways in which the ratio of dependency can be redefined. See Quesada and García-Montalvo (1996), Crespo Cuaresma et al. (2014) and Hammer et al. (2015) for an overview of different definitions and their importance. To calculate the dependency ratio as defined in the model, the active and inactive parts of the population need to be identified. The active population, in line with the model, is divided into two parts; the active workforce ( $(1 - e_t)L_t$ ) and people in education ( $e_tL_t$ ). We include both into the active population because we want to focus on the allocation of time between working and schooling as in the model. The rest is the non-active population, i.e. children before attending school, retirees, unemployed and other people. We have data for 16 countries<sup>11</sup>. For the entire set, data from 1985 to 2011 is used.

The variable “education” is constructed as follows: It includes any kind of participation in educational programs. Next to the regular school career, this includes vocational training and on-the-job training. To identify the regular students, the OECD iLibrary (Dataset: Students enrolled by age) is used. This includes enrollments in ISCED levels 0-6. Some

<sup>11</sup> The countries are: Australia, Canada, Denmark, Spain, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Sweden, the UK and the US. We refrain from using compounded data (like EU, or OECD data) because it is impossible to track missing data. Since we are using absolute values, missing data by just one country would make a big difference.

countries have a dual system of apprenticeships in which students work parts of the week and go to school for the rest. In the available data sets, apprentices are generally counted as full-time students, regardless of their contribution to output. To include their contribution to output in our data, we account 3 days per week in education (60%) and 2 days in production (40%). The 40% of the production side need to be added to the data on the working population and subtracted from educational data. In line with the concept of lifelong learning, many employers provide continuous vocational training (CVT) to their employees. This is usually done in courses that take up a varying amount of time (a couple of hours to several weeks or months per year). Eurostat has conducted a survey in which they track the percentage of time the average employee spends in training during the year. The scope of continuous vocational training has been measured in two years (2005 and 2010). We use the 2005 data as a markup for prior years. The time spent in CVT will be added to education and subtracted from employment. Unfortunately this data is only available in Europe. We use UK data as estimate for the time spent in CVT in three non-European countries, Australia, USA and Canada.

To calculate the growth rate of the dependency ratio, the variable  $L_t$  (the active population) is constructed by adding people in education and in production. Together with the total population ( $N_t$ ) the dependency ratio can be calculated with the relation used in the model,  $D_t = \frac{N_t - L_t}{L_t}$ . To stay in line with the model, we calculate the growth rate of the activity ratio ( $1 + D_t$ ) in the same way as in the discrete-time model, to ensure comparability:  $\frac{x_{t+1}}{x_t} = 1 + g_x$ .

## 4.2 Descriptive Statistics

The average growth rate of the activity ratio,  $g_{1+D}$ , over the whole sample starting 1985 to 2010 for 16 OECD countries is  $-0.001$ , with data ranging from  $-0.062$  to  $0.059$ . Temporarily the growth rate of the activity ratio may be quite low or high, but the panel mean over the last 25 years is close to zero. A temporarily diverging growth rate of the dependency ratio from zero is important to analyze transitions from a younger to an older population, but in the long run, the dependency ratio may be stable. The variable education has a panel mean of  $0.344$ , ranging between  $0.284$  and  $0.518$  for the countries' cross-section mean. The real interest rate,  $r_t$ , is taken from the World Development Indicators (WDI) and ranges between  $-0.025$  and  $0.146$  with a mean of  $0.056$ . In most European countries this data series is only available until the early 2000s. Only for the non-European countries and Italy, the Netherlands and the UK the time range until 2010 is available. The growth rate of the population,  $g_N$ , is also taken from the WDI. It ranges between  $-0.004$  and  $0.029$ .

## 4.3 Empirical Model and Results

As in the theoretical model above, our focus lies on the steady state relation (12d)<sup>||</sup>, which is the equation we would like to estimate. As mentioned above, it is not possible to solve for  $e_t$  or its growth rate. Hence, we estimate an approximation in logs. If (12d)<sup>||</sup> could be solved, the resulting equation would most probably not be linear in the expressions for interest

rates, population growth and dependence. Therefore, linear, quadratic and cubic terms are used in the estimation, because polynomials of the third degree have enough flexibility to capture many forms of non-linearity. To allow for some flexibility, 4 models have been estimated with stepwise elimination of the most insignificant variables after the first estimation. Model 1 uses the polynomials of the third degree for all variables and Model 2 does so only for the education variable in natural logs and is linear in the logs of the other variables of equation (12d)<sup>11</sup>. Models 3 and 4 are obtained from stepwise elimination of the insignificant variables:

<b>Model 1</b>	$\log(e_{it}) = c_0 + c_1 \log(e_{i(t-1)}) + c_2 \left( \log(e_{i(t-1)}) \right)^2 + c_3 \left( \log(e_{i(t-1)}) \right)^3 + c_4 \log(1 + g_{(1+D)i(t-1)}) + c_5 \log(1 + g_{(1+D)i(t-1)})^2 + c_6 \log(1 + g_{(1+D)i(t-1)})^3 + c_7 \log(1 + g_{Ni(t)})^2 + c_8 \log(1 + g_{Ni(t)})^3 + c_9 \log(1 + r_{it}) + c_{10} \log(1 + r_{it})^2 + c_{11} \log(1 + r_{it})^3 + \eta_i + \varphi_t + \epsilon_{it}$
<b>Model 2</b>	$\log(e_{it}) = c_0 + c_1 \log(e_{i(t-1)}) + c_2 \left( \log(e_{i(t-1)}) \right)^2 + c_3 \left( \log(e_{i(t-1)}) \right)^3 + c_4 \log(1 + g_{(1+D)i(t-1)}) + c_5 \log(1 + g_{Ni(t-1)}) + c_6 \log(1 + r_{it}) + \eta_i + \varphi_t + \epsilon_{it}$
<b>Model 3</b>	$\log(e_{it}) = c_0 + c_1 \log(e_{i(t-1)}) + c_2 \left( \log(e_{i(t-1)}) \right)^2 + c_3 \left( \log(e_{i(t-1)}) \right)^3 + c_4 \log(1 + g_{(1+D)i(t-1)}) + c_5 \log(1 + g_{Ni(t-1)}) + \eta_i + \varphi_t + \epsilon_{it}$
<b>Model 4</b>	$\log(e_{it}) = c_0 + c_1 \log(e_{i(t-1)}) + c_2 \left( \log(e_{i(t-1)}) \right)^2 + c_3 \left( \log(e_{i(t-1)}) \right)^3 + c_4 \log(1 + g_{(1+D)i(t-1)}) + \eta_i + \varphi_t + \epsilon_{it}$

Table 3 – Models to be estimated

Where  $\eta_i$  is the unobservable individual effect and  $\epsilon_{it}$  is a disturbance term and  $\varphi_t$  are period fixed effects. As our data fulfills  $T > N$ , fully modified OLS (FMOLS) might be the most adequate estimation procedure, assuming cointegration. It would deal with endogeneity, contemporaneous correlation and serial correlation through a data transformation (see Baltagi (2008) chap.12). However, its usage leads to a loss of data, leaving only 9 or even less countries in the sample. Instead, we use the System GMM method. It estimates two equations simultaneously, namely the given equation estimated in levels, but instrumented with their first-differences and the given equation in first-differences, instrumented with their levels, or, alternatively in orthogonal deviations. It is applied because the lagged dependent variable ( $\log(e_{t-1})$ ) is correlated with the error term by definition which would lead to a downward bias of  $1/T$  for a fixed effects estimator (Nickell, 1981). With the 24 periods in the given sample, this leads to a bias of  $\frac{1}{24} = 0.042$ . All 4 models are estimated using the orthogonal deviations<sup>12</sup> version of System GMM, which replaces the difference equation by one subtracting a weighted sum of future residuals from the current residual (see Arellano and Bover (1995)). We use one instrumental variable per

<sup>12</sup> Using orthogonal deviations reduces the loss of data, when faced with sporadically missing data (Roodman, 2009a)

regressor, the second lag for the lagged dependent variable and the first lag for the other regressors, in line with Okui (2009), who suggest this for very short panels.

	Model 1	Model 2	Model 3	Model 4
$\log(e_{i(t-1)})$	0.607 (0.0855)***	14.070 (5.883)**	14.869 (5.730)**	15.830 (5.773)***
$(\log(e_{i(t-1)}))^2$		13.402 (5.992)**	14.741 (5.979)**	15.683 (6.044)**
$(\log(e_{i(t-1)}))^3$		4.467 (2.017)**	5.078 (2.049)**	5.386 (2.076)**
$\log(1 + g_{(1+D)(t-1)})$		0.167 (0.093)*	0.414 (0.121)***	0.429 (0.124)***
$(\log(1 + g_{(1+D)(t-1)}))^2$	5.150 (3.181)			
$\log(1 + g_{N(t-1)})$		-0.876 (0.606)	-0.634 (0.748)	
$(\log(1 + g_{Ni(t)}))^2$	-414.339 (156.536)***			
$(\log(1 + g_{Ni(t)}))^3$	13230.62 (7325.986)*			
$\log(1 + r_t)$		0.145 (0.1363)		
$(\log(1 + r_t))^2$	2.208 (1.248)*			
S.E. of regression	0.016	0.015	0.022	0.023
Instrument Rank	50	51	52	51
J-statistic	22.599	34.923	22.312	21.580
Prob(J-statistic)	0.425	0.039	0.501	0.580
Cross sections	14	14	16	16
Periods	1988-2010	1988-2011	1988-2011	1988-2011
Obs.	159	158	235	235

**Table 4 – Estimations with Panel Generalized Methods of Moments, Transformations: Orthogonal Deviations**

**Note to table:** In Models 1 and 2 the countries Luxembourg and Portugal are lost because data is not matching. The few observations, compared to the time frame and cross sections available, are due to many asymmetrically missing values.

**Instruments:** Model 1:  $\log(e_{i(t-2)}), (\log(1 + g_{(1+D)(t-1)}))^2, (\log(1 + g_{Ni(t-1)}))^2, (\log(1 + g_{Ni(t-1)}))^3, (\log(1 + r_{t-1}))^2$   
 Model 2:  $\log(e_{i(t-1)}), (\log(e_{i(t-2)}))^2, (\log(e_{i(t-2)}))^3, \log(1 + g_{(1+D)(t-1)}), \log(1 + g_{N(t-1)}), \log(1 + r_{t-1})$   
 Model 3:  $\log(e_{i(t-1)}), (\log(e_{i(t-2)}))^2, (\log(e_{i(t-2)}))^3, \log(1 + g_{(1+D)(t-1)}), \log(1 + g_{N(t-1)})$   
 Model 4:  $\log(e_{i(t-1)}), (\log(e_{i(t-2)}))^2, (\log(e_{i(t-2)}))^3, \log(1 + g_{(1+D)(t-1)})$

After having dropped the other insignificant regressors in Model 1, we find that the squared growth rate of the dependency ratio is just insignificant. If we also drop it, the other variables become insignificant as well and this version of stepwise regression collapses.<sup>13</sup> For the linear Model 2, we find that interest rates are most insignificant. Dropping them leads to

<sup>13</sup> Using current or lagged values of the regressors does not improve this result.

Model 3 where the population growth rate is insignificant. Dropping that one also leads to Model 4, in which all variables are significant. The p-values of the J-statistic indicate that it is not too high (or p-values too low) to have a chi-square distribution (Davidson and MacKinnon, 2004), except for the first regression where p is too low, because of a low number of observations. Otherwise we would cast doubt on the instruments or the specification. They also indicate that the p-values are not too high (or the J-statistic too low), which would indicate that the over-identifying instruments are ineffective in correcting the bias (Roodman, 2009b). As it is significant in all variables we prefer Model 4 to compare Figure 1 and 2 in Section 3 with the empirical outcomes.

Model 4 gives a specification of<sup>14</sup>

$$\log(e_t) = 4.532 + 15.830 * \log(e_{i(t-1)}) + 15.683 * (\log(e_{i(t-1)}))^2 + 5.386 * (\log(e_{i(t-1)}))^3 + 0.429 * \log(1 + g_{1+Dit})$$

For a comparison with Figure 1 the model needs to be solved for the growth rate of  $e_t$ . With  $1 + g_e = \frac{e_{t+1}}{e_t} \leftrightarrow \log(1 + g_e) = \log(e_t) - \log(e_{t-1}) \approx g_e$  this becomes:

$$\log(1 + g_e) = 4.532 + (15.830 - 1) * \log(e_{i(t-1)}) + 15.683 * (\log(e_{i(t-1)}))^2 + 5.386 * (\log(e_{i(t-1)}))^3 + 0.429 * \log(1 + g_{1+Dit}) \quad (14)$$

This compares to Figure 1 with  $\log(1 + g_e) \approx g_e$  on the vertical axis. Up until  $e_t \approx 0.45$ , it has a similar shape as Figure 1. It has a (local) maximum at  $e_t = 0.324$ . The calibrated model of Figure 1 above shows a similar maximum at 0.354. The steady state values are  $e_1 = 0.294$  and  $e_2 = 0.367$  comparing to the very similar ones of the calibration in Figure 1,  $e_1 = 0.305$  or  $e_2 = 0.380$  in Figure 1. This approximately confirms the calibrated values by estimation, with impacts of interest rates and population growth rates close to zero though. If both of these were significant instead, interest rates would shift the curve up and population growth rates would shift it down. If we use Model 2 instead, the steady-state values are  $e_1 = 0.263$  and  $e_2 = 0.359$ , which are very similar again. Only few data points have higher values of  $e$  than 0.45, they mostly belong to Ireland and to a lesser extent to Spain. This is an indication that the upward trend above educational shares of 45% per cent may not be relevant.

---

<sup>14</sup> The constant is the average of country-specific intercepts, not reported in Table 4 (see Greene (2008) for the procedure of calculation). Bun and Windmeijer (2010) state that the variance of the fixed effect should be similar to the variance of the regression. A simple check shows that the ratio in our case is  $\frac{(0.0226)^2}{(0.0223)^2} = 1.0271$  for Model 4, which is close to unity, which was used in the Monte Carlo simulation studies underlying the validity of System GMM.

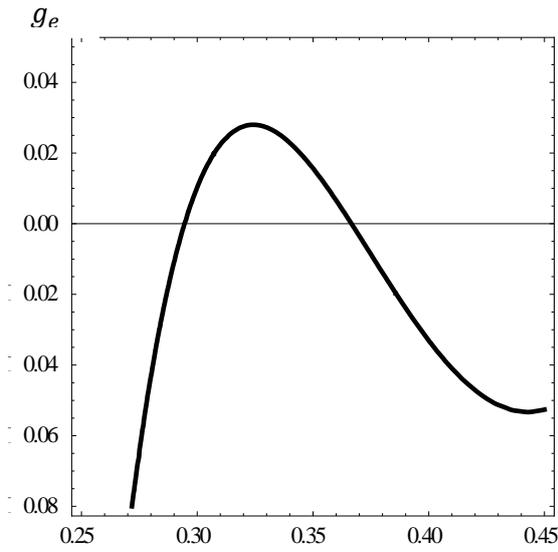


Figure 4

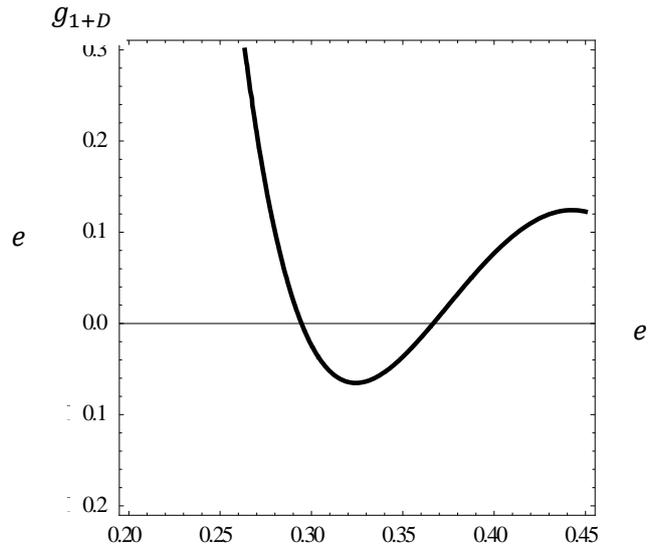


Figure 5

To show the relation to Figure 2,  $g_e$  is set to zero and (14) is solved for  $g_{(1+D)t}$ .

$$g_{(1+D)t} = \frac{4.532 + (15.830 - 1) * \log(e_{i(t-1)}) + 15.683 * (\log(e_{i(t-1)}))^2 + 5.386 * (\log(e_{i(t-1)}))^3}{-0.429}$$

This is plotted in the  $g_{(1+D)} - e_t$  plane.

Figure 5 relates to Figure 2, with  $g_{(1+D)t}$  on the vertical axis. Up until  $e_t \approx 0.45$ , this has a similar shape as Figure 2, with a minimum at  $e = 0.324$  and the same roots as Figure 4,  $e_1 = 0.294$  and  $e_2 = 0.367$ . The panel of the 16 countries shows a similar structure as the simulation of the model in Figure 1 and 2. This shows, as the first major result of this section, that the choice of parameters is in line with Lucas (1988). Specifically  $F$ ,  $\gamma$  and  $\epsilon$  of the model were close to the mark, whereas other models, without externalities are not.<sup>15</sup> If our choices for parameters values leading to Figures 1 and 2 had not been realistic Figures 1 and 4 as well as 2 and 5 would differ much more from each other. Before getting this empirical support we could not have excluded the possibility that the empirical figures might have looked like those of Figures 3 a or b.

In the theoretical framework, we were forced to do two case studies of either setting  $g_e = 0$ , or  $g_{1+D} = 0$  for analytical reasons. The empirical model allows us to include both cases of the theoretical framework. The second major finding of this section is that whereas an increase in the growth rate of population shifts the  $g_e - e$  curve in Figure 4 downward, an increase in the growth of the dependency ratio and the interest rate shift the curve upward leading to a faster second best growth rate of  $e_t$  in the transition to the steady state and a higher steady state value. By implication, the second-best policy response to ageing is more

<sup>15</sup> See Figure 3b for graphical representation of the case without externality.

education for more human capital and productivity growth.<sup>16</sup> This partial result holds for a constant interest rate, which will be treated as endogenous next.

## 5 Debt Dynamics and general equilibrium

In sections 3 and 4 one of the assumptions has been a zero elasticity of the interest rate with respect to the debt ratio,  $\eta_{rb} = 0$ , and thereby a fixed interest rate given by the world market because a small and open economy has been considered. This assumption was made in order to postpone the discussion of feedback effects of changes in  $e$  on those of the debt/GDP ratio and the interest rate  $r$ , which would have an effect on the  $g_e - e$  relation according to equation (12d)<sup>11</sup> and Figure 1, although the estimated model shows that this effect is insignificant and therefore at best small. If the elasticity is assumed to be zero in some sub-sections, this feedback effect on the interest rate is interrupted and partial equilibrium analysis gives adequate results. In this section this assumption will be loosened to a flexible interest rate for several reasons. First, price-takership taken literally requires countries to be atomistically small, which they never are; even if there is no attempt to exploit an impact on the interest rate, a country's impact on the world market interest rate can never be exactly zero. Second, EU countries have understood that each country's cumulated government deficit will increase the interest rate for all of them because of the size effect just explained and, therefore, it is rational to consciously limit government and foreign debt as modeled by a higher interest rate at higher debt/GDP ratios in Bardhan (1967). Third, building on the first two points raised here, if a country's debt is large, the risk of debt service rescheduling, moratoria or repudiation is also large, as was shown by Eaton and Gersovitz (1981); more empirical support was obtained by Edwards (1984) and S. H. Lee (1991); (1993). We will follow this line of reasoning in our modelling strategy as others, cited below, did. It will first be shown how debt behaves in an unstable manner under a fixed and given interest in order to illustrate the stability problem. In a second step, the model will be extended to include an endogenous interest rate depending on the magnitude of debt, which has not been done in an Uzawa-Lucas model so far. Third, the interest-debt relation will be estimated. Fourth, we consider the stability of the consumption-debt dynamics with endogenous interest rate. Fifth, the general equilibrium effect of ageing on all steady-state variables of the model will be derived.

### 5.1 Fixed Interest Rate

This sub-section illustrates the stability problem in the presence of foreign debt. The special case of no effect of debt on the interest rate,  $\eta_{rb} = 0$ , is instructive here. In the model economy the current account equation must hold. It states that next period's debt/lending

---

<sup>16</sup> Seeing the threat of an increasing  $g_{1+D}$ , pension funds in countries like the Sweden, Netherlands and Germany have announced to pay lower pensions per working year. In response to that announcement, people have started working more years (but not so in Austria). With more years, longer education time becomes profitable. In countries with less ageing or more immigration this may be different.

must balance the deficit/surplus in production after consumption, saving and previous interest payments have been accounted for.

$$B_{t+1} = N_t c_t + K_{t+1} - (1 - \delta_k)K_t - Y_t + \left(1 + r \left(\frac{B_t}{Y_t}\right)\right) B_t \quad (15)$$

(15) is the equivalent to equation (5) with  $\omega_t$  and  $r_{Kt}$  at their equilibrium values of equations (3) and (4) and, hence, with  $\omega_t(1 - e_t)h_t \frac{N_t}{1+D_t} + r_{kt}K_t = Y_t$ .

Because the previously determined relations hold,  $K_{t+1}$  and  $K_t$  can be replaced by  $K_{t+1} = \frac{1-\alpha}{r\left(\frac{B_{t+1}}{Y_{t+1}}\right)(1+\eta_{rb})+\delta_k} Y_{t+1}$  from equation (12c) and (12b) with  $R_{Ht+1} = R_{Bt+1} = 1 + r\left(\frac{B_{t+1}}{Y_{t+1}}\right)(1 + \eta_{rb})$  for their respective periods. The interest rate is given and constant in this subsection, requiring the assumption  $\eta_{rb} = 0$ . If the debt to GDP ratio is defined as  $b_t = \frac{B_t}{Y_t}$ , then

$$b_{t+1} = \frac{N_t c_t}{Y_{t+1}} - \frac{\theta}{1+g_Y} + \frac{1+r}{1+g_Y} b_t \quad (15)^1$$

Where  $\theta = (1 - \delta_k) \frac{1-\alpha}{r(1+\eta_{rb})+\delta_k} + 1 - (1 + g_Y) \frac{1-\alpha}{r(1+\eta_{rb})+\delta_k} < 1$  and is constant in steady-state.

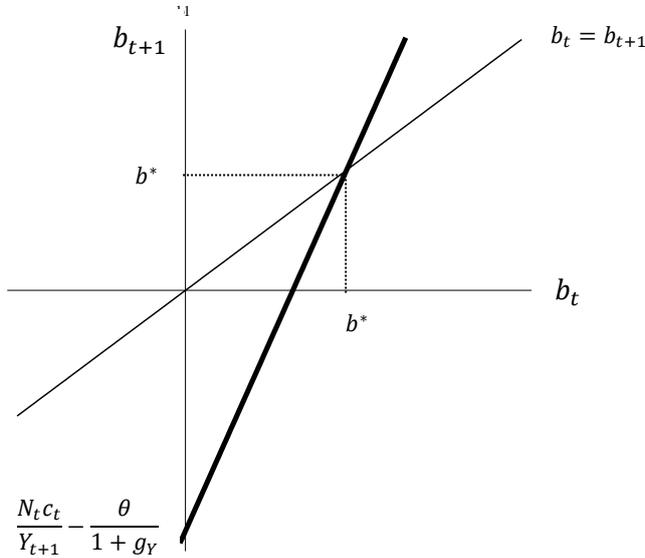


Figure 6

Equation (15)<sup>1</sup> describes the development of the debt-to-GDP ratio over time for the previously discussed special values of the interest rate and the constant steady state growth rates of output and consumption, for which the slope of (15)<sup>1</sup> is constant and equal to 1.018.

The intercept is  $\frac{N_t c_t}{Y_{t+1}} - \frac{\theta}{1+g_Y}$ . The marginal propensity to consume is constant as  $N_t c_t$  grows at the rate  $(1 + g_N)(1 + g_c) = 1.031$  which is the same rate as that of the denominator,  $1 + g_Y = 1.031$ . The last term of the intercept only contains exogenous variables and

parameters and the growth rate of output which is constant in steady state. These special assumptions for only this sub-section boil down to a partial equilibrium analysis and imply a stable average propensity to consume and a stable intercept. Because the slope is larger than unity,  $b^*$  is an unstable steady state value, indicating that any deviation from the steady state value cannot lead back to  $b^*$  if all other parameters are unchanged. It leads to ever increasing debt if the process starts to the right of  $b^*$  as described by Blanchard (1983) using the Cass-Koopmans model. Without extra assumptions about jumping to  $b^*$  or adjustment of the magnitude of the marginal propensity to consume, no conclusions about the value of  $b_*$  can be drawn. For higher consumption rates this would imply lower debt-GDP ratios, because with higher consumption there is less money for debt services; the current account surplus would be lower. However, in general the dynamics in  $b$  show an unstable steady state because of the parameter choices. For different parameter choices, especially  $\sigma$  and  $\beta$ , the average propensity to consume is not constant anymore. If  $\sigma = 2$  and  $\beta = 0.97$ , the growth rate of consumption is 0.009 instead of 0.029, which leads to a decreasing average propensity to consume as the numerator grows with  $(1 + g_N)(1 + g_c) = 1.011$  and the denominator with  $1 + g_Y = 1.031$ . The intercept of the debt to GDP function (15)' thus becomes more and more negative causing an outward shift of the debt to GDP line with no steady state. By implication, a jump onto a certain value of the debt/GDP ratio cannot be a general solution to the stability problem and more adjustment is needed. If the interest rate increases in  $r$ , removing the zero-elasticity assumption, the slope increases with increasing debt/GDP ratio, which would aggravate the stability problem.  $\theta$  would increase, moving the line to the left, which would also destabilize the process. Adjustment then must mainly come from decreasing the consumption share, because changes of education shares  $e$  and effects from  $e$  on  $g_Y$  are small.

## 5.2 Endogenized interest rate

Section 5.1 shows the solution to the model as an unstable steady state for one specific parameter set. With a constant interest rate it would have been possible to keep borrowing to finance infinite consumption, implying that no utility maximum exists. However, an interest rate depending on the debt to GDP ratio avoids this problem of infinite consumption and stabilizes the unstable debt process and acts as a borrowing restriction as assumed in the literature (Bhandari et al., 1990; Hamada, 1969; Philippopoulos, 1991). Therefore, to analyze the dynamics around the steady state it is useful to assume a non-fixed interest rate which depends on the debt of the country. In this case of a model extension, the interest rate depends on the world interest rate  $\bar{r}$  and the spread of the interest rate which in turn depends on debt over GDP,  $r(b_t) = \bar{r} + r_s(b_t)$ . To our knowledge this is the first time that an empirically founded endogenous interest rate has been introduced to an Uzawa-Lucas model. In this subsection we will first understand how the creditor chooses the lending rate for a specific country and then find an expression for  $r(b_t)$  that holds empirically which will then be introduced into the model at hand.

### 5.2.1 The creditor chooses the lending rate

So far, the exact function for  $r(b_t)$  has not been specified. To see how the schedule for the creditor's specific lending rate ( $r$ ) behaves, we need to go more into depth on how the creditor chooses the lending rate. The creditor is assumed to maximize his profit in a competitive market. His profit is the revenue he gets,  $r_c B$ , times the repayment probability  $p(b)$  minus the cost. It is assumed that  $p(b)$  is decreasing in  $b$  since it is more difficult to repay if the debt/GDP ratio rises. The creditors costs are the world market interest rate times the debt issued,  $\bar{r}B$ . Because the payback probability of the country may be less than one under a temptation to repudiate, the interest charged,  $r_c$ , may be higher than the world interest rate. The creditor then maximizes:

$$\max_B p(b)r_c B - \bar{r}B$$

Where  $p(b)$  is the probability of debt repayment,  $\bar{r}$  is the world interest rate and  $r_c$  is the lending rate of the creditor. The first-order condition is  $p(b)r_c + p'(b)r_c \frac{B}{Y} - \bar{r} = 0$  which can be solved for the lending rate of the creditor:

$$r_c = \frac{\bar{r}}{p(b)[1+\eta_{pb}]}$$

with  $b \frac{p'(b)}{p(b)} = \eta_{pb} < 0$  as the elasticity of the repayment probability to debt per GDP. The lending rate formula shows the supply schedule for debt if the creditor maximizes profit in a competitive market. If debt per GDP increases, the probability of payment decreases and hence the optimal lending rate increases, given a constant elasticity  $\eta_{pb}$ . Subtracting the world market interest rate from the lending rate yields

$$r_c - \bar{r} = \frac{\bar{r}}{p(b)[1+\eta_{pb}]} - \bar{r}$$

The spread is equal to the right-hand side. In case of a competitive creditor's market, the profits covering some fixed costs are expected to be zero. Then it follows from the definition of non-negative profits, that

$$r_c - \frac{\bar{r}}{p(b)} > 0$$

The left-hand side of the spread formula is even larger under positive profits.

Under costless perfect information the creditors may also take the debt dynamics (15)<sup>1</sup> into account. However, in practice country studies are expensive and it is cheaper to know only the value of the debt/GDP ratio as sufficient information in our model, which is sufficiently simple to allow linking it to the Uzawa-Lucas growth model. If  $r_t$  is a suitable function of  $b_t$ , it is constant in the steady state with constant  $b$ .

### 5.2.2 Estimation of the interest-debt relation

The basic idea in this section and the related literature is that interest rates increase with the debt/GDP ratio either because the country has a large market share in the capital market

(Bardhan, 1967) or because higher debt relative to the potential to pay the debt service increases the risk of financial problems. Whereas this is widely accepted, the second and third derivatives of the relation are discussed much less and often imposed by mere assumption. Theoretically, Stiglitz and Weiss (1981) balance the risk compensation against the increase of the probability being unable to pay, leading to credit rationing. In the model of the previous section, the interest rate is linear in that of the world market, but the probability to pay debt service has a negative relation with the debt/GDP ratio. Second order conditions are more likely to be fulfilled if the second derivative is also negative in particular for low values of the interest rate. But this is less necessary under higher interest rates, and the Stiglitz-Weiss argument would say that the second-derivative should be decreasing when debt/GDP is high. Thus, second derivatives may be changing and third derivatives are likely to increase them from negative to positive. In order to solve our endogenous growth model we need an estimate for the same countries for which we have considered education data. We chose a cubic specification because it is very flexible and allows for changing signs of derivatives. Assume a positive transformation of  $\log[1 + r_t] = c_1 + c_2 \log[1 + \bar{r}] + c_3 \log[2 + b_t] + c_4(\log[2 + b_t])^2 + c_5(\log[2 + b_t])^3$ , which may balance the consumption reduction effect and the partial instability of  $b(b(-1))$  dynamics. Adding the quadratic and cubic term is an empirical decision for a non-linear effect. In order to not lose any of the already scarce data, the measurements have been adapted. We chose to regress on  $\log(1 + r_t)$  instead of  $\log(r_t)$ , because  $r_t$  ranges between -0.025 and 0.146 and some of the data would be lost if  $\log(r_t)$  was employed. For the same reason  $\log(2 + b_t)$  has been used in the model as  $b_t$  ranges between -1.6 and 1.7.

For the variable  $b_t$  we choose the measure “Net foreign debt” from the updated and extended version of the dataset constructed by Lane and Milesi-Ferretti (2007). It accounts for a country’s assets held within and outside of its borders, where a positive value indicates a borrowing position of the country and a negative value a lending position. The interest rate,  $r$ , is taken from the World Development Indicators (as above).  $\bar{r}$  is approximated by the interest rate of the United States,  $r_{USA}$ , which is 0.05 on average. This way, the part of the movement of the world interest rate which is not caused by debt movements of the specific country is accounted for by the movements in the US interest rate.

We apply Fully Modified Ordinary Least Square (FMOLS) estimation method to the data. Phillips and Hansen (1990) first propose the FMOLS estimator as a consistent estimator for time-series samples with high endogeneity and serial correlation. The cointegrated panel estimator is also consistent for small samples with contemporaneous correlation of the residuals (Pedroni, 2004). It is implemented taking into account fixed effects by de-meaning regressors by country averages as in the within estimator.

Testing for cointegration between the interest rate and debt yields mixed results. We conduct the Pedroni (1999) test for cointegration in panel data, which is Engle-Granger (1987) based. Eight of the 11 tests are significant, leading to a rejection of the null

hypothesis of “no cointegration” in all but three cases. Because the other tests with the same hypothesis show very low p-values and the number of observations is low, cointegration between the two variables can be assumed.<sup>17</sup>

The estimated equation is reported below. With the truncated-uniform kernel<sup>18</sup>, the regression shows the highest adjusted R<sup>2</sup> and the most significant estimates. The estimation has an adjusted R<sup>2</sup> of 0.1196 and a DW statistics of 0.6035. Both values are not satisfactory, but given the scarcity of the available data, this is the best possible. Standard errors are reported in parentheses below the estimate.

$$\begin{aligned} \log[1 + r_t] = & -0.204 + 0.566 \log[1 + \bar{r}] + 0.605 \log[2 + b_t] - 0.564(\log[2 + b_t])^2 \\ & (0.002)^{***} \quad (0.041)^{***} \quad (0.038)^{***} \quad (0.049)^{***} \\ & + 0.214(\log[2 + b_t])^3 \\ & (0.059)^{***} \end{aligned} \quad (16)$$

With this specification the development of the interest rate with respect to the debt to GDP ratio is

$$r(b_t) = 0.815 \exp^{-0.564 \log[2+b_t]^2 + 0.214 \log[2+b_t]^3} (2 + b_t)^{0.605} (1 + \bar{r})^{0.566} - 1 \quad (16)^l$$

which is displayed in Figure 7.

Implementing  $\frac{N_t c_t}{Y_t} = X_t$  in (15)<sup>l</sup> together with (12b) it must hold that

$$(1 + g_b) = \frac{1}{1 + g_y} \left( \frac{1}{b_t} X_t - \frac{1 + \frac{1-\alpha}{r(b_t)(1+\eta_{rb}) + \delta_k} ((1-\delta_k) - (1+g_y))}{b_t} + 1 + r(b_t) \right) \quad (17)$$

<sup>17</sup> The variable  $\log(2 + b_t)$  shows the existence of a unit root with and without a trend added. The variable  $\log(1 + r_t)$  shows a unit root if no trend is added, but with an added trend there is mixed evidence.

<sup>18</sup> Estimating the long-run covariances using the Bartlett kernel with automatic bandwidth choice of Newey-West or Andrews yields very similar results.

## Optimal Education in Times of Ageing

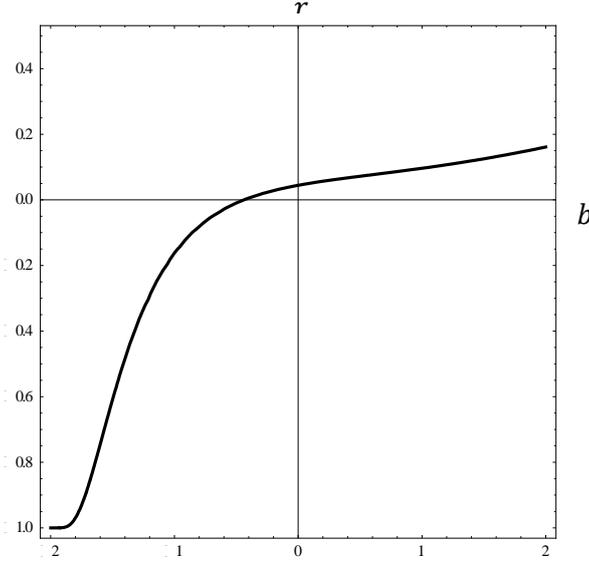


Figure 7

In steady state,  $g_b = 0$  and (17) solved for  $X_t$  becomes

$$X_t = \left(1 + g_Y - (1 + r(b_t))\right) b_t + \frac{(1-\alpha)(1+g_Y)}{r(b_t)(1+\eta_{rb})+\delta_k} \left(\frac{(1-\delta_k)}{1+g_Y} - 1\right) + 1 \quad (17)^1$$

With  $1 + g_Y = \left(F e_t^Y + (1 - \delta_h)\right)^{1+\frac{\epsilon}{\alpha}} \left(\frac{1+g_N}{1+g_{1+D}}\right)$  and  $\eta_{rb}$ , the elasticity of the interest rate

with respect to the debt to GDP ratio that is derived from (16)<sup>1</sup> with  $\eta_{rb} = \frac{B_{t+1}}{Y_{t+1}} \frac{r' \left(\frac{B_{t+1}}{Y_{t+1}}\right)}{r \left(\frac{B_{t+1}}{Y_{t+1}}\right)}$ .

$$\begin{aligned} \eta_{rb} = & b_t \left( \frac{0.494 \exp^{-0.564 \text{Log}[2+b_t]^2 + 0.214 \text{Log}[2+b_t]^3} (1+\bar{r})^{0.566}}{(2+b_t)^{0.395}} + \right. \\ & 0.815(2+b_t)^{0.605} \exp^{-0.564 \text{Log}[2+b_t]^2 + 0.214 \text{Log}[2+b_t]^3} (1+\bar{r})^{0.566} \left( -\frac{1.127 \text{Log}[2+b_t]}{2+b_t} + \right. \\ & \left. \left. \frac{0.642 \text{Log}[2+b_t]^2}{2+b_t} \right) \right) * \\ & \left. \frac{1}{(-1 + 0.815(2+b_t)^{0.605} \exp^{-0.564 \text{Log}[2+b_t]^2 + 0.214 \text{Log}[2+b_t]^3} (1+\bar{r})^{0.566})} \right) \end{aligned}$$

(17)<sup>1</sup> is plotted in the in the  $b - X$  plane in the left-hand panel of Figure 8 for the steady state value  $e = 0.380$  and the previously stated parameter set.

For (15)<sup>1</sup> which is as in the previous section, the dynamic equation describing  $b_t$ , to hold, not only the dynamics in  $e_t$  are important, but also those of the average propensity to consume.

Defining  $\frac{N_t c_t}{Y_t} = X_t$ ,  $1 + g_X$  by definition is

$$1 + g_X = \frac{(1+g_N)(1+g_C)}{1+g_Y} = \frac{1+g_N}{\left(F e_t^Y + (1-\delta_h)\right)^{1+\frac{\epsilon}{\alpha}} \left(\frac{1+g_N}{1+g_{1+D}}\right)} \left[\beta(1 + r(b_t)(1 + \eta_{rb}))\right]^{\frac{1}{\sigma}} \quad (18)$$

Where  $1 + g_Y$  is its steady state relation and  $1 + g_C$  is derived from (7) and (9)<sup>l</sup>. In steady state  $g_X = 0$ , with the expression for  $r(b)$  of equation (16), (18) becomes:

$$\frac{1}{\beta} \left( (F e_t^\gamma + (1 - \delta_h))^{1+\frac{\epsilon}{\alpha}} \left( \frac{1}{1+g_{1+D}} \right) \right)^\sigma = (1 + r(b_t)(1 + \eta_{rb})) \quad (18)^l$$

This cannot be solved analytically for either  $b_t$  or  $e_t$  and hence is plotted as  $g_X = 0$  line in Figure 8 for the assumed values of the parameters.

The dynamics of  $g_e$  are given by equation (12d)<sup>ll</sup>, where  $R_{Ht+1} = 1 + r(b_t)(1 + \eta_{rb})$  from (12b<sup>l</sup>). In steady state  $g_e = 0$ , then together with the estimation for  $r(b_t)$ , (12d)<sup>ll</sup> becomes

$$1 + r(b_t)(1 + \eta_{rb}) = (F e_t^\gamma + (1 - \delta_h))^{\frac{\epsilon}{\alpha}} \frac{1+g_N}{1+g_{1+D}} F\gamma \left[ e_t^{\gamma-1} - e_t^\gamma + e_t^\gamma \frac{1}{\gamma} + \frac{(1-\delta_h)}{F\gamma} \right] \quad (19)$$

(19) shows all steady state combinations of  $b_t$  and  $e_t$  for the given parameter set with  $\bar{r}$  set to 0.05, the average over the sample. It cannot be solved analytically for either  $b$ , or  $e$ . To analyze the relationship further, it has been plotted below in the  $b - e$  plane for the parameters assumed in Table 1. The  $g_e = 0$  - line in the right panel of Figure 8 represents all steady state pairs of  $b$  and  $e$ .

Both lines intersect at  $e = 0.380$  and  $b = 0.042$ . This implies an  $X_t$  of 0.693 and an  $r$  of 0.050.

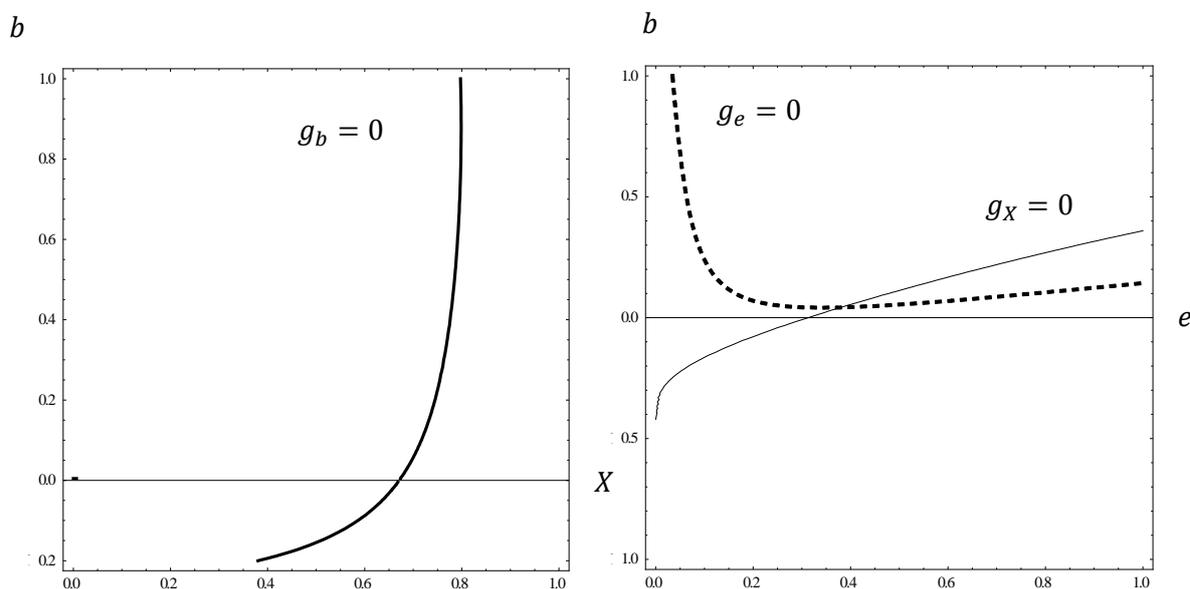


Figure 8

### 5.3 Partial Stability with Endogenous Interest Rate

In sections 5.1-5.3 the existence of a steady state in the case of an endogenized interest rate has been analyzed. This section will deal with the behavior of the economy out of the steady state and will analyze aspects of the stability of the system.

There are three mutually interacting variables that influence the dynamics:  $e_t$ ,  $b_t$  and  $X_t$ . The central equations are (12d)<sup>||</sup> for the dynamics in  $e_t$ , (15)<sup>||</sup> for the dynamics in  $b_t$  and (18) for the dynamics in  $X_t$ . Figure 1 in Section 3 shows the dynamics in  $e_t$  if  $r$  is a function of  $b_t$  and  $b_t$  is constant. For given  $b_t$  it has already been established, that the steady state value  $e^* = 0.380$  is stable.

Figure 9 shows the partial stability in  $X_t$  and  $b_t$ . The vertical line represents equation (18) for  $g_X = 0$  and hump shaped line is (15)<sup>||</sup> for  $g_b = 0$ . Where both lines intersect, for given  $e_t$ , the sub-system is stable in the saddle-point sense. If only  $b_t$  is off the  $g_X = 0$ -line in Figure 9, a larger (smaller) value of  $b_t$  implies a positive (negative) growth rate of  $X_t$  in (18). This is shown in the vertical movement-arrows. If  $X_t$  is larger (smaller) than its steady state value in (15)<sup>||</sup>,  $g_b$  becomes positive (negative). This is captured in the horizontal arrows above and below the  $g_b = 0$ -line. The arrows indicate that there is a saddle-point-stable trajectory. If the economy starts on this trajectory with given  $e_t$ ,  $X_t$  and  $b_t$ , it ends up in steady state. If the economy started to the right of this trajectory, debt would keep growing which violates the transversality condition. If the economy started on the left, it would move towards the origin, indicating that less is consumed than possible. This is also not optimal. Hence, the only optimal solution is to jump right on the saddle-point-stable trajectory.

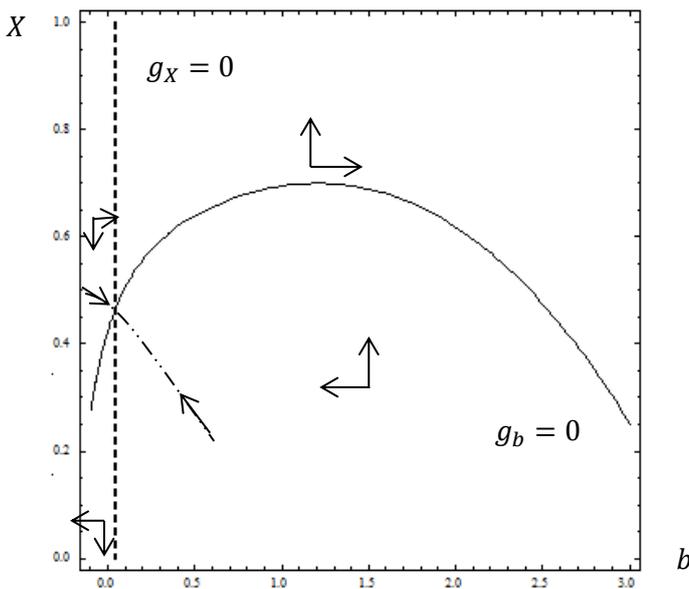


Figure 9

This shows the partial stability of the system for all three variables. To achieve stability  $X_t$  has to control  $b_t$  directly as in Figure 9 and via  $b_t$  also  $e_t$  indirectly as in Figure 1.

#### 5.4 The Influence of the Growth Rate of the Dependency Ratio

So far, we set  $g_{1+D} = 0$ . We are interested in what happens if the growth rate of the dependency ratio changes. For this we exemplarily plot 5 scenarios. Since the data of  $g_{1+D}$  ranges between -0.062 and 0.059 with an average of -0.001, the scenarios are chosen for the values -0.06, -0.03, 0, 0.03 and 0.06. From (18)<sup>||</sup> it follows that for given  $b$ , a higher  $g_{1+D}$  shifts the  $g_X=0$ -line to the right. From (19) it follows that for given  $e$ , the  $g_e = 0$  -line shifts down. The latter move dampens the increase in  $e^*$  from the first, and the former

dampens the fall in  $b^*$ . Whether these dampening effects even dominate can only be answered numerically. The same holds for  $X_t$  because the  $g_b = 0$ -line was drawn for steady-state values  $b^*$  and  $e^*$ . The answer is given in Table 5.

$g_{1+D}$	$e_t$	$b_t$	$X_t$	$r_t(1 + \eta)$	$1 + g_c$
-0.06	0.331	0.721	0.672	0.117	1.091
-0.03	0.354	0.351	0.674	0.082	1.059
0	0.380	0.042	0.693	0.050	1.029
0.03	0.407	-0.151	0.738	0.020	1.001
0.06	0.437	-0.276	0.887	-0.009	0.975

Table 5 – Summary of Influence of Changes in  $g_{1+D}$

With an increasing growth rate of the dependency ratio, the optimal time spent in education increases substantially. The optimal debt to GDP ratio decreases, indicating the need for different optimal practices when expecting ageing, and the interest rate decreases. The marginal propensity for current consumption increases and the growth rate of consumption decreases. Hence, the optimal reaction to ageing and the implied higher consumption is more education at the cost of consumption growth, leading to less credit demand and a lower interest rate, implying a lower burden of debt service.

In terms of figures, we can summarize the effect of the growth rate of the dependency ratio on education  $e$ , debt ratio  $b$  and consumption ratio  $X$  as follows, using the outcomes of Table 5 when several forces pull into opposite directions. First, in Figure 1, representing equation (12d)'', an increase in  $g_{1+D}$  shifts up the curve in figure 1 and  $e$  moves to the right.  $g_{1+D}$  also has the effect of moving the vertical line in Figure 9 for stationary  $X$ , equation (18), to the left, resulting in lower  $b$ . This is mitigated but not outweighed by the increase in  $e$ .  $g_{1+D}$  has no direct effect on the other inverted U-shape line in Figure 9. But the increase in  $e$  decreases the slope and increases  $X$  shifting the line upward, resulting in higher  $X$  at lower  $b$ . The lower values of the debt/GDP ratio  $b$  and interest rate  $r$  feed back into equation (12d)'' decreasing  $g_e$  for given  $e$  and therefore dampening the initial effect of  $g_{1+D}$  on  $e$ . However, as  $e$  is increasing according to Table 5, the feedback effect on  $e$  must be smaller than the initial effect, which is also in line with the insignificance of the interest term in model 2 of the regressions in Table 4. In short, ageing leads to higher education, more growth, and a lower debt/GDP ratio, a lower interest rate, a lower debt service and, if strong enough, turning into net interest income, and a lower growth rate of consumption and a higher level of consumption, where the latter happens to occur because of pulling forward the benefits of higher growth to early periods with more old people.

## 6 Conclusion

In this paper we have seen the consequences of an increasing age-independent dependency ratio introduced into a discrete-time Uzawa-Lucas model with international capital movements, human capital externalities and decreasing returns to schooling time in human capital formation. The model has been calibrated and compared to estimates; it shows patterns, which are similar to the education data of sixteen OECD countries since 1985. The economy turns out to have multiple steady states. Steady-state analysis has shown that only a high share in education is associated with a stable steady state. In the neighborhood of the

stable steady state, it is optimal to spend more time in education when the growth of the active part of the population lags behind that of the inactive part as it is the case in times of ageing. This increases the growth rates of human capital, GDP per capita, wages, and reduces the growth rate of consumption, interest rates and the debt/GDP ratio in order to reduce the burden of debt service. As a model extension and to ensure stability and non-exploding debt, the interest rate function has been estimated to cope endogenously with increasing debt.

In a broader perspective, when the process of ageing has started, it is important to care for the future. Higher education is one option that shifts the value of time via education as a store of value and technical change into the future. In principle, this could also be done through saving earlier or borrowing later. However, ageing in a utility maximizing framework requires keeping consumption high and therefore capital accumulation is not an attractive option, neither via saving early nor borrowing later. This holds at least under the specific assumptions of our model regarding decreasing returns to education time in human capital formation and education externalities in final output production. Therefore we have taken care of empirical properties of our model in the analysis.

Our infinite horizon approach avoids modeling decisions of generations separately. Therefore the recent emphasis of some literature on life-expectancy does not appear here. Once the life-expectancy discussion has included capital accumulation with international capital movements and fully endogenous technical change for ageing economies both approaches can perhaps be matched to further explore the effect of ageing on education. Similarly, decisions on retirement age do not appear in our model. Households can decide to let old or young people work longer. Explicit modeling of working time of young people is intimately related to union contracts. Working time of old people is related to those of pension fund rules, which differ per country. When unemployment is higher than the natural rate the suggestion for later retirement may be less convincing. Moreover, retirement decisions are 'undermined' by decisions to resume work after retirement (Kanabar, 2015) and by escape clauses for certain types of work leading to decisions of early retirement. These decisions are often softer rules than in the models. Therefore we leave this as a within-household decision. But, of course, if modeling approaches are able to capture heterogeneity of leisure preferences and labour market institutions this would lead to an enrichment of models and new possibilities to re-investigate our result suggesting enhancing education time. It would be in line with the life expectancy argument that later retirement makes education even more useful, provided retirement is not a perfect substitute for other labour market decisions, in case they capture all effects.<sup>19</sup>

As the literature discussed in the introduction has shown, also institutional aspects like those of health, education, labour markets, and capital markets are important. To the extent that family support and student loans are insufficiently available to achieve sufficiently much

---

<sup>19</sup> See Warren (2015) on some complements and substitutes for retirements.

education as assumed in Hansen and Lønstrup (2012) governments should improve the student loan system. Student loans do not only benefit the recipient families but also allow for the optimal response to ageing. Similarly, country specific market imperfections and institutions should be improved where ever it is beneficial.

A natural next step for future research would be to introduce more information on the growth rate of the dependency ratio. Furthermore, it would be interesting to use other endogenous growth models than that of Uzawa-Lucas. For the question whether or not there should be more education in reaction to ageing other models can only be those which do have endogenous human capital formation. Moreover, they need a clear empirical perspective regarding scale effects - defined as growing growth rates when factor endowments grow – and also regarding semi-endogenous growth - defined as zero growth when there is no population growth. We have avoided both of these extreme views because we think that the Lucas model is empirically more realistic. However, neither of these two alternative model classes has been entirely rejected by the literature, but the more well-known ones do not have endogenous human capital formation. Any such effort should combine calibration and estimation of models in order to ensure empirical robustness as far as possible.

## Acknowledgements

We are thankful for the helpful comments by the participants of the “Workshop on Human Capital and Aging” at the Harvard School of Public Health. Particularly, we would like to thank Jocelyn Finlay, David Canning, David Weil and Robert Lee. We also thank two anonymous referees for their valuable contributions. All remaining errors are our responsibility.

## References

- Arellano, Manuel, and Bover, Olympia. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, 68(1), 29-51.
- Baltagi, Badi. (2008). *Econometric analysis of panel data* (Vol. 1): John Wiley & Sons.
- Bardhan, Pranab. (1967). Optimum Foreign Borrowing, Essays on the Theory of Optimal Economic Growth (K. Shell ed.). *MIT Press, Cambridge, Massachusetts*, 274, 117-128.
- Benhabib, Jess, and Perli, Roberto. (1994). Uniqueness and indeterminacy: on the dynamics of endogenous growth. *Journal of Economic Theory*, 63(1), 113-142.
- Bhandari, Jagdeep S, Haque, Nadeem Ul, and Turnovsky, Stephen J. (1990). Growth, external debt, and sovereign risk in a small open economy. *IMF Staff Papers*, 37(2), 388-417.
- Blanchard, Olivier Jean. (1983). Debt and the current account deficit in Brazil *Financial policies and the world capital market: The problem of Latin American countries* (pp. 187-198): University of Chicago Press.
- Bloom, David E, Canning, David, and Fink, Günther. (2010). Implications of population ageing for economic growth. *Oxford Review of Economic Policy*, 26(4), 583-612.
- Bonneuil, Noël, and Boucekkine, Raouf. (2014). Longevity, Age-Structure, and Optimal Schooling.
- Boucekkine, Raouf, De la Croix, David, and Licandro, Omar. (2002). Vintage human capital, demographic trends, and endogenous growth. *Journal of Economic Theory*, 104(2), 340-375.

- Bun, Maurice JG, and Windmeijer, Frank. (2010). The weak instrument problem of the system GMM estimator in dynamic panel data models. *The Econometrics Journal*, 13(1), 95-126.
- Cervellati, Matteo, and Sunde, Uwe. (2005). Human capital formation, life expectancy, and the process of development. *American Economic Review*, 1653-1672.
- Cervellati, Matteo, and Sunde, Uwe. (2013). Life expectancy, schooling, and lifetime labor supply: theory and evidence revisited. *Econometrica*, 81(5), 2055-2086.
- Cervellati, Matteo, and Sunde, Uwe. (2015a). The effect of life expectancy on education and population dynamics. *Empirical Economics*, 48(4), 1445-1478.
- Cervellati, Matteo, and Sunde, Uwe (2015b). The Economic and Demographic Transition, Mortality, and Comparative Development *American Economic Journal: Macroeconomics*, 7(3), 189-225.
- Crespo Cuaresma, Jesus, Lábaj, Martin, and Pruzinský, Patrik. (2014). Prospective ageing and economic growth in Europe.
- Davidson, Russell, and MacKinnon, James G. (2004). *Econometric theory and methods* (Vol. 5): Oxford University Press New York.
- De la Croix, David, and Licandro, Omar. (1999). Life expectancy and endogenous growth. *Economics Letters*, 65(2), 255-263.
- Denison, Edward F. (1962). *The sources of economic growth in the United States and the alternatives before us* (Vol. 13): Committee for Economic Development New York.
- Denk, Michaela, and Weber, Michael. (2011). Avoid Filling Swiss Cheese with Whipped Cream: Imputation Techniques and Evaluation Procedures for Cross-Country Time Series. *IMF Working Papers*, 1-27.
- Eaton, Jonathan, and Gersovitz, Mark. (1981). Debt with potential repudiation: Theoretical and empirical analysis. *The Review of Economic Studies*, 289-309.
- Edwards, Sebastian. (1984). LDC's foreign borrowing and default risk: An empirical investigation (Vol. September, pp. 726-734): *American Economic Review*.
- Engle, Robert F, and Granger, Clive WJ. (1987). Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, 251-276.
- Fayissa, Bichaka, and Gutema, Paulos. (2010). Dependency Ratio and the Economic Growth Puzzle in Sub-Saharan Africa. *Department of Economics and Finance Working Paper series*. (Assessed on 22 December 2010 from <http://ideas.repec.org/p/mts/wpaper/201010.html>).
- Frenkel, Jacob A, Razin, Assaf, and Yuen, Chi-Wa. (1996). Fiscal policies and growth in the world economy. *The MIT Press, Cambridge*.
- Galasso, Vincenzo. (2008). Postponing retirement: the political effect of aging. *Journal of Public Economics*, 92(10), 2157-2169.
- Gong, Gang, Greiner, Alfred, and Semmler, Willi. (2004). The Uzawa–Lucas model without scale effects: theory and empirical evidence. *Structural change and economic dynamics*, 15(4), 401-420.
- Greene, William H. (2008). *Econometric analysis*: Granite Hill Publishers.
- Gruescu, Sandra. (2006). *Population ageing and economic growth: education policy and family policy in a model of endogenous growth* (Vol. 1431): Springer.
- Hamada, Koichi. (1969). Optimal capital accumulation by an economy facing an international capital market. *The Journal of Political Economy*, 77(4), 684-697.
- Hammer, Bernhard, Prskawetz, Alexia, and Freund, Inga. (2015). Production activities and economic dependency by age and gender in Europe: A cross-country comparison. *The journal of the economics of ageing*, 5, 86-97.
- Hansen, Casper Worm. (2013). Life expectancy and human capital: Evidence from the international epidemiological transition. *Journal of health economics*, 32(6), 1142-1152.
- Hansen, Casper Worm, and Lønstrup, Lars. (2012). Can higher life expectancy induce more schooling and earlier retirement? *Journal of Population Economics*, 25(4), 1249-1264.
- Hazan, Moshe. (2012). Life expectancy and schooling: new insights from cross-country data. *Journal of Population Economics*, 25(4), 1237-1248.

- Hviding, Ketil, and Mérette, Marcel. (1998). Macroeconomic Effects of Pension Reforms in the Context of Ageing Populations: Overlapping Generations Model Simulations for Seven OECD Countries Economics Department Working Papers No. 201. *WORKING PAPERS-ORGANISATION FOR ECONOMIC COOPERATION AND DEVELOPMENT ECONOMICS DEPARTMENT*.
- Kanabar, Ricky. (2015). Post-retirement labour supply in England. *The Journal of the Economics of Ageing*, 6, 123-132.
- Lane, Philip R, and Milesi-Ferretti, Gian Maria. (2007). The external wealth of nations mark II: Revised and extended estimates of foreign assets and liabilities, 1970–2004. *Journal of international Economics*, 73(2), 223-250.
- Lee, Ronald. (2003). The demographic transition: three centuries of fundamental change. *The Journal of Economic Perspectives*, 17(4), 167-190.
- Lee, Suk Hun. (1991). Ability and willingness to service debt as explanation for commercial and official rescheduling cases. *Journal of Banking & Finance*, 15(1), 5-27.
- Lee, Suk Hun. (1993). Are the credit ratings assigned by bankers based on the willingness of LDC borrowers to repay? *Journal of Development Economics*, 40(2), 349-359.
- Lucas, Robert E. (1988). On the mechanics of economic development. *Journal of monetary economics*, 22(1), 3-42.
- Mankiw, N Gregory, Romer, David, and Weil, David N. (1992). A Contribution to the Empirics of Economic Growth. *The Quarterly Journal of Economics*, 107(2), 407-437.
- Muysken, J, and Ziesemer, THW. (2013). A permanent effect of temporary immigration on economic growth. *Applied Economics*, 45(28), 4050-4059.
- Muysken, J, and Ziesemer, THW. (2014). The effect of immigration on economic growth in an ageing economy. *Bulletin of Applied Economics*, 1(1), 37-65.
- Nickell, Stephen. (1981). Biases in dynamic models with fixed effects. *Econometrica: Journal of the Econometric Society*, 1417-1426.
- Okui, Ryo. (2009). The optimal choice of moments in dynamic panel data models. *Journal of Econometrics*, 151(1), 1-16.
- Pedroni, Peter. (1999). Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors. *Oxford Bulletin of Economics and Statistics* 61, 653-670.
- Pedroni, Peter. (2004). Panel cointegration: asymptotic and finite sample properties of pooled time series tests with an application to the PPP hypothesis. *Econometric theory*, 20(03), 597-625.
- Philippopoulos, Apostolis. (1991). MACROECONOMIC DYNAMICS AND OPTIMAL EXTERNAL BORROWING IN AN IMPERFECT WORLD CAPITAL MARKET\*. *The Manchester School*, 59(1), 12-23.
- Phillips, Peter CB, and Hansen, Bruce E. (1990). Statistical inference in instrumental variables regression with I (1) processes. *The Review of Economic Studies*, 57(1), 99-125.
- Quesada, Javier, and García-Montalvo, José. (1996). Growth and population aging: The Spanish case: Instituto Valenciano de Investigaciones Económicas, SA (Ivie).
- Roodman, David. (2009a). How to do xtabond2: An introduction to difference and system GMM in Stata. *Stata Journal*, 9(1), 86.
- Roodman, David. (2009b). A note on the theme of too many instruments\*. *Oxford Bulletin of Economics and Statistics*, 71(1), 135-158.
- Stiglitz, Joseph E, and Weiss, Andrew. (1981). Credit rationing in markets with imperfect information. *The American economic review*, 393-410.
- Turner, John. (2007). Social security pensionable ages in OECD countries: 1949-2035. *International Social Security Review*, 60(1), 81-99.
- Uzawa, Hirofumi. (1965). Optimum technical change in an aggregative model of economic growth. *International economic review*, 6(1), 18-31.

- Warren, Diana A. (2015). Retirement decisions of couples in Australia: The impact of spousal characteristics and preferences. *The Journal of the Economics of Ageing*, 6, 149-162.
- Wright, Robert E, Kinnunen, Jouko, Lisenkova, Katerina, and Merette, Marcel. (2014). Population Ageing and Technological Change: EcoMod.
- Xie, Danyang. (1994). Divergence in economic performance: transitional dynamics with multiple equilibria. *Journal of Economic Theory*, 63(1), 97-112.
- Zhang, Pengfei. (2013). A Newfound Steady State in Standard Uzawa-Lucas Model with Externality. Available at SSRN 2262391.

## APPENDIX A

This Appendix shows the proof for the existence of maximum value of the utility integral for  $\sigma > 0$  in steady state. The utility function has a finite integral, and hence has an interior maximum, if the growth rate of discounted utility is smaller than zero ( $g_u < 0$ ). By definition, the growth rate of the individual utility function is  $1 + g_u = \frac{u_{t+1}}{u_t} = \beta \frac{c_{t+1}^{1-\sigma}}{c_t^{1-\sigma}} = \beta(1 + g_c)^{1-\sigma} \leftrightarrow g_u = \beta(1 + g_c)^{1-\sigma} - 1$ . For a positive growth rate of  $c$ ,  $\sigma > 1$  is sufficient.

## APPENDIX B

In this appendix, the derivation of equation (13), the derivative of (12d)<sup>II</sup> is shown.

Denote the RHS as  $J(e, g(e))$  and differentiate implicitly with  $g'(e) = -\frac{J'_1(e, g(e))}{J'_2(e, g(e))}$ :

Calculate partial derivatives w.r.t.  $g_e$  and  $e_t$ :

$$J'_1(e, g_e) = \frac{1+g_N}{1+g_{1+D}} F\gamma \left( \frac{\epsilon}{\alpha} (F e_t^\gamma + (1 - \delta_h)) \right)^{\frac{\epsilon}{\alpha}-1} F\gamma e_t^{\gamma-1} \left[ e_t^{\gamma-1} - e_t^\gamma (1 + g_e) + e_t^\gamma \frac{1}{\gamma} (1 + g_e) + \left( \frac{1}{1+g_e} \right)^{\gamma-1} \frac{(1-\delta_h)}{F\gamma} \right] + (F e_t^\gamma + (1 - \delta_h)) \left[ (\gamma - 1) e_t^{\gamma-2} - \gamma e_t^{\gamma-1} (1 + g_e) + e_t^{\gamma-1} (1 + g_e) \right]$$

$$J'_2(e, g_e) = \frac{1+g_N}{1+g_{1+D}} F\gamma (F e_t^\gamma + (1 - \delta_h)) \left[ -e_t^\gamma + e_t^\gamma \frac{1}{\gamma} + (1 - \gamma)(1 + g_e)^{-\gamma} \frac{(1-\delta_h)}{F\gamma} \right]$$

The ratio of the two partial derivatives is the derivative of  $g_e$  w.r.t.  $e_t$ .

$$g'_e = -\frac{J'_1(e,g(e))}{J'_2(e,g(e))} = -\frac{\frac{\epsilon}{\alpha}(F e_t^\gamma + (1-\delta_h))^{\frac{\epsilon}{\alpha}-1} F \gamma e_t^{\gamma-1} \left[ e_t^{\gamma-1} - e_t^\gamma (1+g_e) + e_t^{\frac{\gamma-1}{\gamma}} (1+g_e) + \left(\frac{1}{1+g_e}\right)^{\gamma-1} \frac{(1-\delta_h)}{F\gamma} \right]}{(F e_t^\gamma + (1-\delta_h))^{\frac{\epsilon}{\alpha}} \left[ -e_t^\gamma + e_t^{\frac{\gamma-1}{\gamma}} + (1-\gamma)(1+g_e)^{-\gamma} \frac{(1-\delta_h)}{F\gamma} \right]} - \frac{(F e_t^\gamma + (1-\delta_h))^{\frac{\epsilon}{\alpha}} \left[ (\gamma-1)e_t^{\gamma-2} - \gamma e_t^{\gamma-1} (1+g_e) + e_t^{\gamma-1} (1+g_e) \right]}{(F e_t^\gamma + (1-\delta_h))^{\frac{\epsilon}{\alpha}} \left[ -e_t^\gamma + e_t^{\frac{\gamma-1}{\gamma}} + (1-\gamma)(1+g_e)^{-\gamma} \frac{(1-\delta_h)}{F\gamma} \right]}$$

This can be simplified, by cancelling  $(F e_t^\gamma + (1 - \delta_h))^{\frac{\epsilon}{\alpha}}$ :

$$g'_e = -\frac{J'_1(e,g(e))}{J'_2(e,g(e))} = \frac{\frac{\epsilon}{\alpha}(F e_t^\gamma + (1-\delta_h))^{-1} F \gamma e_t^{\gamma-1} \left[ e_t^{\gamma-1} - e_t^\gamma (1+g_e) + e_t^{\frac{\gamma-1}{\gamma}} (1+g_e) + \left(\frac{1}{1+g_e}\right)^{\gamma-1} \frac{(1-\delta_h)}{F\gamma} \right] - [(\gamma-1)e_t^{\gamma-2} - \gamma e_t^{\gamma-1} (1+g_e) + e_t^{\gamma-1} (1+g_e)]}{\left[ -e_t^\gamma + e_t^{\frac{\gamma-1}{\gamma}} + (1-\gamma)(1+g_e)^{-\gamma} \frac{(1-\delta_h)}{F\gamma} \right]}$$

## APPENDIX C

In order to find reasonable values for  $F$  and  $\gamma$  to fulfill the condition (6),  $g_h$  is set to 0.011 with  $\delta_h$  being fixed at 0.03. This can only be done for given  $e_t$ . We choose  $e_t$  to reflect the panel average of the data in section 4:  $e_t = 0.344$ .

Gong et al. (2004) suggest estimated values for  $\gamma$  and  $\epsilon$  which differ substantially from ours. The reason may be that they include an exponent on individual human capital in (6), which is smaller than unity, as opposed to 1 in our model. Such a specification might change the estimated values for  $\gamma$  and  $\epsilon$  and leads actually to a semi-endogenous growth model as in Jones (1995). Lower values of  $\epsilon$ , as in Gong et al. (2004), only have one steady state. Moreover, they use different education data for their estimation than we do. As seen in Xie (1994), in the absence of depreciation though,  $\epsilon$  should be “large enough” to observe two steady states. Section 4 shows evidence for two steady states in the data and Figure 3a and 3b show sufficient conditions for uniqueness. We pick  $F = 0.055$  and  $\gamma = 0.268$ , because it ensures the closest fit to Figure 5 in Section 4.

## APPENDIX D

This appendix shows that a decrease in the depreciation rate of human capital only marginally affects the results. We set  $\delta_h = 0.01$ , then according to (6),  $F$  and  $\gamma$  need to be altered to ensure a close fit to the data,  $g_h = 0.011$  for  $e = 0.344$  as in the text. Below are Figures 1 and 2 for  $\delta_h = 0.01$ ,  $F = 0.035$  and  $\gamma = 0.436$ . All other parameters used in the derivation of Figures 1 and 2 are unchanged.

## Optimal Education in Times of Ageing

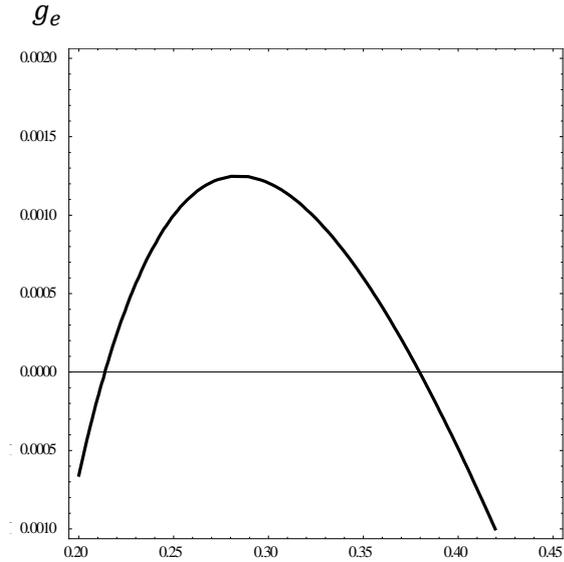


Figure D.9 – Dynamics in  $e_t$

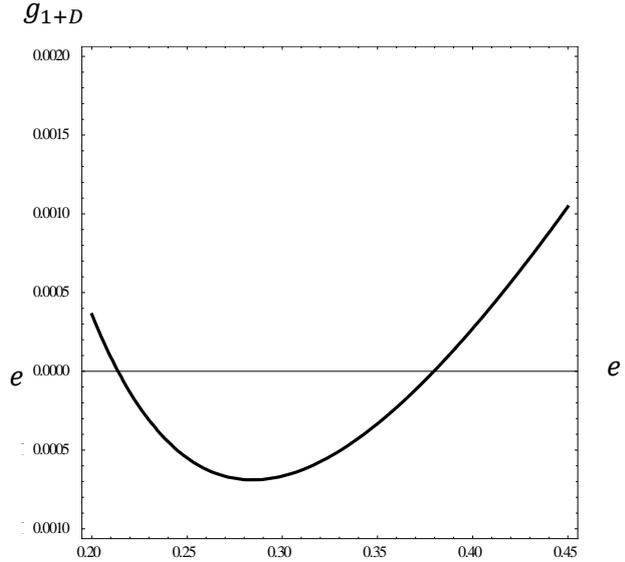


Figure D.10 – Growth rate of the activity ratio

While the upper (stable) steady state is close to the empirical value again ( $e = 0.380$  in this version of the calibrated and  $e = 0.367$  in in the estimated model), the lower steady state is a little further off ( $e = 0.214$  in this version of the model vs.  $e = 0.294$  in the estimated model). The calibrated model in the text with  $\delta_h = 0.03$  gave the same upper value for  $e$ , and the lower value  $e = 0.305$  is closer to that of the estimated model at  $0.294$ . In short, with  $\delta_h = 0.03$  instead of  $0.01$ , the calibrated model is closer to the estimated model at the lower steady state.

To make Figure D.3 compatible with the model (and comparable to Figure 8),  $\beta$  needs to be adjusted marginally from  $0.982$  to to  $0.984$ . Otherwise the  $g_e = 0$  and the  $g_X = 0$  lines do not intersect in the steady state.

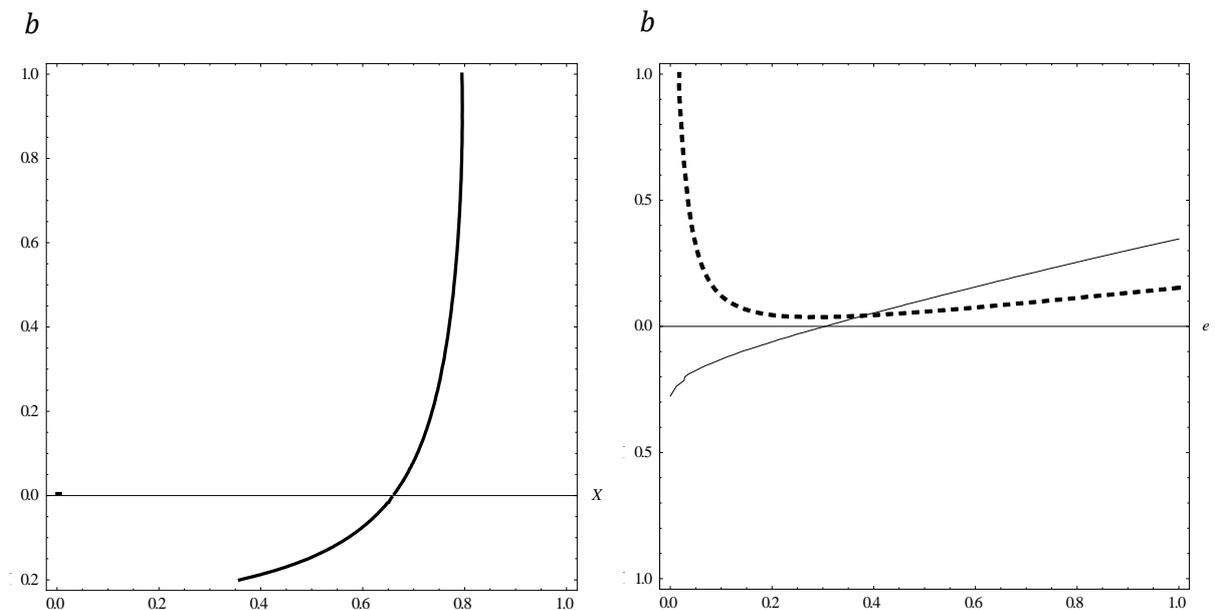


Figure D.3

## Optimal Education in Times of Ageing

The steady state values for the variables  $b$ , debt per GDP, and  $X$ , the marginal propensity to consume, change only marginally. The change in variable  $b$  is only visible if the precision is increased by one decimal. It decreases from 0.0418 in the case of  $\delta_h = 0.03$  to 0.0416 in the case of  $\delta_h = 0.01$ . The marginal propensity to consume decreases from  $X = 0.693$  in the case of  $\delta_h = 0.03$  to  $X = 0.683$  in the case of  $\delta_h = 0.01$ .