The Role of Endogenous Skill Choice in an Aging Society

Christian Jaag*

May 12, 2005

This paper analyzes the effects of an aging population on individual skill choices and the production structure by means of a dynamic general equilibrium model with overlapping generations and probabilistic aging. The model allows for capital-skill complementarity, which strongly affects the outcomes in a small open economy setting vs. a closed (or equivalently worldwide) economy. In an open economy with a fixed real interest rate, the necessary increase in the contribution rate discourages labor supply and depresses . With a variable real interest rate, however, capital usage increases and – by the capital-skill complementarity – also employment of high skilled labor. The mobilization of highly productive labor gives a boost to . Hence, the often cited adverse effects of aging are mitigated and can be overcome when taking into account a more realistic production structure.

*Instiitute of Public Finance and Fiscal Law, University of St.Gallen, Varnbüelstrasse 19, CH-9000 St.Gallen, christian.jaag@unisg.ch. Financial support from the Swiss National Science Foundation under project no. 1214-066928 is gratefully acknowledged. I thank Christian Keuschnigg, Mirela Keuschnigg, and seminar participants in St. Gallen for valuable comments.

1 Introduction

The purpose of this paper is to illuminate the macroeconomic relevance of education and human capital formation in an aging society. This topic has received much attention over the last years, mainly due to the expected sharp increase in the share of the retired over the working population (cf. figure 1).

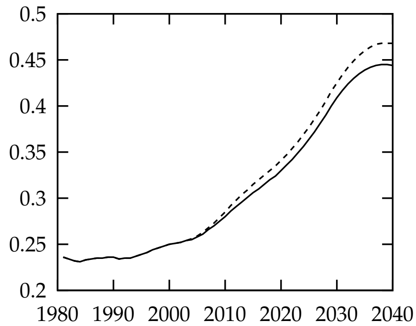

Figure 1

Dependency ratio projections for Switzerland. The solid line shows the trend extrapolation, the dashed line a more pessimistic scenario.
Source: (2001)

In order to compensate the adverse effects of the shift in the age structure on production, labor productivity of the working will need to increase.[1] We analyze the direct effect of a change in the age structure due to aging on the individuals' incentives to invest in education at the extensive margin. Indeed, aging itself strongly affects the skill structure of the workforce. However, depending on whether an economy is open or closed, additional policy measures towards increased human capital investment would have to be put in place in order to maintain total employment and to assure the financial viability of existing pension systems.

In this paper, individuals' extensive education decision is studied, along with their intertemporal consumption choice and labor supply. For simplicity, we abstract from the intensive training effort margin. To be able to simulate a concrete population scenario, the model realistically reproduces (1) individual life-cycle wage and consumption profiles; (2) the skill composition in the labor force; (3) skill premia in wages; (4) the change in relative prices and the production structure due to an aging population. We use an overlapping generations ( ) approach in general equilibrium in order to be able to analyze intergenerational distribution effects. In its complexity, the model relates to three strands of economic literature:

Extensive Education Decision Besides the literature on life-cycle human capital investment, which focuses on the intensive margin of education, there is a strand of literature concentrating on the extensive margin.[2] This mostly empirical literature is concerned with self-selection of students into skill types according to ability types which affects estimates of occupational choice and the distribution of earnings. The first source is Roy (1951), which has received subsequent elaboration, e.g. by Heckman and Honoré (1990). Willis and Rosen (1979) and Heckman, Lochner, and Taber (1998) derive a theoretical model of the demand for college attendance and empirically show that expected lifetime earnings indeed influence the college attendance choice decisively.

[1] Börsch-Supan (2003) argues that the decrease in the relative size of the economically active population cannot be balanced by higher capital intensity. Hence, the strengthening of human capital formation assumes high importance in the face of an aging population.

[2] Early references in the life-cycle human capital investment literature are Becker (1962) and Mincer (1974). Ben-Porath (1967) formulates a rigorous model which makes the analogy between human capital and investment in physical capital explicit. Weiss (1986) provides an extensive review of the theoretical literature while Mincer (1997) reviews the empirical literature.

Due to the specific nature of their sample, there is no ability bias in the estimation of the enrollment function, a problem which is often encountered in the empirical assessment of the rate of return to education.

Overlapping Generations Economic models with overlapping generations ( ) of households provide a good basis for the analysis of fiscal policy and intertemporal macroeconomics. There are basically two strands of literature: The one with a large number of generations and detailed life-cycle patterns pioneered by Auerbach and Kotlikoff (1987) with refinements by e.g. İmrohoroğlu, İmrohoroğlu, and Joines (1999) and Altig, Auerbach, Kotlikoff, Smetters, and Walliser (2001). The second strand bases upon the model by Ramsey (1928) with infinitely lived consumers. Blanchard (1985) introduces constant mortality hazard which leads to potentially finite lifetimes but eternal youth. Gertler (1999) models life-cycle behavior in this framework by allowing mortality only for the retired. We base our approach on its consequent enhancement by Grafenhofer, Jaag, Keuschnigg, and Keuschnigg (2005) who introduce probabilistic aging which allows for mortality among young ages and thereby exhibits a closer approximation of the demographic structure.

Demographic Transition The demographic transition in industrialized countries with lower fertility and higher life-expectancy poses challenges to pension systems and the economy as a whole. Börsch-Supan (2003) argues that in order to finance the pension system, contribution and tax rates will have to rise, which reduces the labor force participation of the young and destabilizes the pension system even more. In a model with imperfect substitutability between less and more experienced workers, Rojas (2005) shows that the effects of aging are less hazardous than with perfectly substitutable workers. Also, Conesa and Krueger (1999) consider heterogeneous agents and their impact on the political support for a funded pension system.

The remainder of this paper is organized as follows: Based on the probabilistic aging approach, we first present a framework for the analysis of the economic effects of aging. Then, we study the labor market and skill choice effects of an aging society. Initially, we abstract from an individual's extensive education decision and concentrate on the life-cycle decisions within a given skill-group before modeling the individuals' skill choice explicitly. Finally, we simulate the effects of aging on key macro variables in an open and closed economy setting.

2 Individual Optimization

2.1 Preferences

We assume a period utility over consumption and foregone leisure. Individuals decide how long to work per period, which is measured by the fraction of time δ spent at work. When not working, they receive a transfer payment e from the government. At young ages, this is an unemployment benefit, at old ages, it is pension income. While at work, they also choose their work intensity l. The arguments in the utility function are thus consumption C, intensive labor supply (work effort) l and extensive labor supply δ. The functional form of period utility is

$$Q \equiv C - \delta\varphi(l) - \phi(\delta), \tag{1}$$

with the budget constraint

$$C \leq (1-\tau)\,\delta\theta w l + (1-\delta)\,e, \tag{2}$$

where w is the wagerate per efficiency unit, θ is individual productivity and τ is the tax rate on labor income. The effort and retirement / unemployment cost functions are assumed to be of the form

$$\varphi(l) = (l_0)^{-1/v^W} \frac{(l)^{1+1/v^W}}{1+1/v^W}, \tag{3}$$

$$\phi(\delta) = \delta_0 v^R \exp\left(\frac{\delta}{v^R}\right). \tag{4}$$

In assuming intratemporally separable preferences, we eliminate all income effects on labor supply and are thus in a first step able to solve the individual intratemporal utility maximization problem independently from dynamic considerations.[3]

2.2 Intratemporal Optimization

A consumer's preferences and their implications for intratemporal substitution effects between consumption and leisure are first analyzed in a static setting. Individuals use up their budget, such that (2) holds with equality. Maximizing effort adjusted wage income yields the first-order conditions

$$(1-\tau)\,\theta w = \varphi'(l),$$

$$(1-\tau)\,\theta w l - \varphi(l) - e = \phi'(\delta).$$

Substituting for consumption and exploiting the first-order condition with respect to e yields

$$l = l_0\left[(1-\tau)\,\theta w\right]^{v^W}, \tag{5}$$

$$\delta = v^R \ln\left[\frac{1}{\delta_0}\left((1-\tau)\,\theta w l - \varphi(l) - e\right)\right]. \tag{6}$$

Intensive labor supply only depends on the current wage, productivity, and income tax rate. On the extensive margin, the chosen fraction of time δ during which a job is held reflects the difference between adjusted wages and alternative transfer income. The more generous the transfer income, the less people are inclined to incur the utility cost of e.g. postponed retirement. In the following, we reduce the labor supply for all age groups $a < a^R$ to the intensive dimension only. Age groups $a > a^R$ are defined to be fully retired ($\delta = 0$) and receive income e.

2.3 Intertemporal Optimization

Aggregating period utility over time, we want to allow for the possibility that individuals are risk neutral but have a finite intertemporal elasticity of substitution. This is not possible within the axiom set of the von Neumann-Morgenstern (v) utility theory where the coefficient of relative risk aversion is the inverse of the elasticity of intertemporal substitution. Kreps and Porteus (1978) ( ) show that the application of the v axioms to temporal lotteries amounts to assuming that individuals are indifferent as regards the timing of resolution of uncertainty. This is because

[3]Assuming separable preferences and hence excluding income effects is quite common in the literature, cf. e.g Greenwood, Hercowitz, and Huffman (1988), Heijdra (1998) in intertemporal macroeconomics, and the model of endogenous retirement by Crémer and Pestieau (2003).

v utility theory imposes on temporal gambles the axiom of reduction of compound lotteries which postulates that individuals care only about the compound probability of each offered prize in a compound lottery. If one is willing to drop the axiom of reduction (as in the context of preferences), the separation of the attitude towards risk from intertemporal substitution is possible. We employ a special constant elasticity of intertemporal substitution utility function proposed by Farmer (1990) on the basis of Weil (1990) and Epstein and Zin (1989) that restrict individuals to be risk neutral with respect to income risk, but allow for an arbitrary elasticity of intertemporal substitution.

When we consider an individual's intertemporal decision problem, we write down preferences in recursive form. The analysis is based on the concept of probabilistic aging, as outlined in the appendix. Individuals are assumed not to age deterministically, but to move to the next age-group only with a certain probability. The first subscript, α, is an individual's life-cycle history containing the date of birth along with the dates of past age-group changes; the second subscript, t, is the current date. The superscript a denotes the current age-group. ω^a and γ^a denote the instantaneous probabilities of staying in the current age-group and surviving, respectively. With A_t denoting assets and R_t being the interest rate factor $1 + r_t$, indirect utility is given by[4]

$$V\left(A^a_{\alpha,t}, \theta^a_{\alpha,t}, \delta^a_{\alpha,t}\right) = \max\left[\left(Q^a_{\alpha,t}\right)^{\rho} + \gamma^a\beta\left(\overline{V}^a_{\alpha,t+1}\right)^{\rho}\right]^{1/\rho}, \tag{7}$$

subject to:

$$\begin{aligned}
\overline{V}^a_{\alpha,t+1} &= \omega^a V^a_{\alpha,t+1} + (1-\omega^a)\, V^{a+1}_{\alpha',t+1}, \\
\gamma^a A^a_{\alpha,t+1} &\equiv \gamma^a A^{a+1}_{\alpha',t+1} = R_t\left[A^a_{\alpha,t} + y^a_{\alpha,t} - Q^a_{\alpha,t}\right], \\
y^a_{\alpha,t} &\equiv \delta^a_t w^a_t l^a_t + (1-\delta^a_t)\, e_t, && (8) \\
w^a_t &= (1-\tau)\,\theta^a_t w_t. && (9)
\end{aligned}$$

Assuming that there is no binding time constraint and if the consumption path is deterministic, the parameter $\sigma = 1/(1-\rho)$ is the (constant) elasticity of intertemporal substitution. Under certainty, $\beta = 1/(1+\upsilon)$ is the subjective time discount factor with υ denoting the time preference rate. Unlike in standard expected utility, the discount factor under uncertainty is endogenous with Epstein-Zin preferences – unless $\rho = 0$, in which case it is equal to β as under certainty.

We assume the individuals to intertemporally optimize over their consumption and labor supply and as in Grafenhofer, Jaag, Keuschnigg, and Keuschnigg (2005). This gives rise to the following result.

Result 1 *Optimal consumption policy $C^a_{\alpha,t}$ and indirect utility $V^a_{\alpha,t}$ are*

$$\begin{array}{llcl}
(i) & Q^a_{\alpha,t} & = & \left(1/\Delta^a_t\right)\left(A^a_{\alpha,t} + H^a_{\alpha,t}\right), \quad \sigma = 1/(1-\rho), \\
(ii) & V^a_{\alpha,t} & = & \left(\Delta^a_t\right)^{1/\rho} Q^a_{\alpha,t}, \\
(iii) & \Delta^a_t & = & 1 + \gamma^a\beta^{\sigma}\left(\Omega^a_{t+1}R_{t+1}\right)^{\sigma-1}\Delta^a_{t+1}, \\
(iv) & \Omega^a_{t+1} & = & \omega^a + (1-\omega^a)\left(\Lambda^a_{t+1}\right)^{1-\rho}, \quad \Lambda^a_{t+1} = \left(\Delta^{a+1}_{t+1}/\Delta^a_{t+1}\right)^{1/\rho}, \\
(v) & H^a_{\alpha,t} & = & y^a_t + \gamma^a\overline{H}^a_{\alpha,t+1}/\left(\Omega^a_{t+1}R_{t+1}\right), \\
(vi) & \overline{H}^a_{\alpha,t+1} & = & \omega^a H^a_{\alpha,t+1} + (1-\omega^a)\left(\Lambda^a_{t+1}\right)^{1-\rho} H^{a+1}_{\alpha',t+1}.
\end{array} \tag{10}$$

[4]The case in which $\rho = 0$ can be handled by taking limits and applying de L'Hôpital's rule.

Δ *is the inverse of the marginal propensity to consume and H denotes human capital equal to the present value of future wages and pension benefits.*

Proof. Cf. Grafenhofer, Jaag, Keuschnigg, and Keuschnigg (2005). ■

2.4 Extensive Education Decision

So far, we have considered only one type of skills in the population since labor supplied by variously aged workers is assumed to be perfectly substitutable against each other. The case of two different skill groups has been studied in the literature: Acemoglu (1998) models the discrete skill formation decision by assuming that there are opportunity costs of education; among individuals, there is heterogeneity with respect to the time it takes to become high skilled. In Heckman, Lochner, and Taber (1998), individuals decide whether or not to attend college based on their idiosyncratic cost and the gains in indirect utility from a college education. In this section, we generalize the concept of a discrete education decision to three (and potentially more) differentiated skill groups. Similar to Heckman, Lochner, and Taber (1998), we assume instantaneous education, such that there are no foregone earnings.[5] Here, differences in ability lead to differentiated direct costs of education and thereby to a stratification of workers.

At the beginning of their lifetime, each individual decides which skill level s she wants to acquire according to the decision problem

$$s = \arg\max_{\widetilde{s}\in\{1,\ldots,S\}} \left\{ V_{t,t}^{\widetilde{s},1} - \sum_{i=1}^{\widetilde{s}-1} c^i \right\}$$

where c^i denotes the incremental cost of acquiring the next skill level i and the first subscript t denotes the date of birth. There are S skill levels with 1 denoting the lowest level. We assume heterogeneity in individuals with individual i possessing ability ν^i. The distribution of ν is given by the function $\Gamma(\nu)$. To simplify the analysis, we assume that $\Gamma(\nu)$ has no mass points. We further assume that individual costs of acquiring skill are strictly decreasing in ability such that there is a single crossing property in skill acquisition: If an individual of type ν chooses to attain a certain skill level s, another with $\nu' > \nu$ must at least acquire the same skill level. Therefore, there exists a cutoff level of talent ν^s for every skill level s, such that all $\nu > \nu^s$ get at least skills $s+1$. Assume that the cost of acquiring the next skill level $s+1$ consists only of the type-dependent direct lump-sum cost of education $c^s(\nu) > 0$ with $c^{s\prime}(\nu) > 0$. The individual with talent ν^s needs to be indifferent between belonging to skill level s and $s+1$. Thus, we need $V_{(t),t}^{s} = V_{(t),t}^{s+1,1} - c^s(\nu)$ or $\Delta V_{(t),t}^{s,1} = c^s(\nu)$. The cost of education to be of the form $c^s(\nu) = \xi^s \exp\left(\Gamma(\nu)\big/\varepsilon^D\right)$, where ε^D is the semi-elasticity of the education decision with respect to $\Delta V : \varepsilon^D = -d\nu^s \big/\!\left(d\Delta V^{s,1}\big/\Delta V^{s,1}\right)$. Then, the pivotal value ν_t^s for the choice between the two skill-groups s and $s+1$ is

$$\nu_t^s = \Gamma^{-1}\left(\varepsilon^D\left[\ln \Delta V_{(t),t}^{s,1} - \ln \xi^s\right]\right). \tag{11}$$

An individual's discrete education decision among three skill groups is displayed in figure 2.

[5] Note, however, that the production technology assumed by Heckman, Lochner, and Taber (1998), the two skill classes are equally well substitutable for capital, while the technology assumed here allows for capital-skill complementarity.

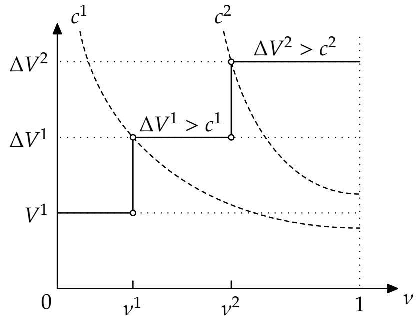

Figure 2

Discrete education decision with three skill groups if ν has support over $[0, 1]$.

We denote the lowest skill group by 1 and the highest skill group by S. With $N^1_{t,t}$ new households in each period, the number of households choosing skill group s will be

$$N^1_{(t),t} p^s_t, \qquad p^s_t = \begin{cases} \Gamma\left(\nu^s_t\right) & \text{for} \quad s = 1, \\ \Gamma\left(\nu^s_t\right)\left[1 - \Gamma\left(\nu^{s-1}_t\right)\right] & \text{for} \quad 1 < s < S, \\ \left[1 - \Gamma\left(\nu^{s-1}_t\right)\right]\left[1 - \Gamma\left(\nu^{s-2}_t\right)\right] & \text{for} \quad s = S. \end{cases}$$

The educational choice is open only to new households, while older ones are locked into their previously chosen skill group. Hence, the law of motion for the number of individuals in age-group 1 with skill level l is

$$N^{s,1}_{t+1} = \gamma^1 \omega^1 N^{s,1}_t + N^1_{(t+1),t+1} p^s_{t+1}.$$

3 Firm Optimization

Usually, life-cycle models with overlapping generations, such as the original Auerbach and Kotlikoff (1987) model assume labor to be homogeneous. Rojas (2005) replaces this assumption with a labor market characterized by imperfect substitutability between less and more experienced workers, hence introducing a vertical differentiation of labor. In our model, the individuals' discrete education choice leads to a partition of the working population into different skill classes, which is mirrored on the production side, where firms' technology uses as inputs differentiated labor and capital, hence giving rise to a horizontal differentiation of labor. For simplicity, we assume that the experience of workers does not alter their type of labor.

3.1 Production Technology

The surge in skill-biased information technology during the 1990s which coincided with a rise in the wages of skilled workers relative to those of unskilled workers suggests that various kinds of labor are of different substitutability with capital. The first evidence of this phenomenon is due to Griliches (1969) who refers to his finding that capital and skilled labor are more complementary as inputs than are capital and unskilled labor as the capital-skill complementary hypothesis. The hypothesis has received broad attention in the macroeconomic literature, where evidence in its favor has important implications on how to specify aggregate production in theory, but also in reassessing the robustness of existing empirical findings. Quantitative implications of the hypothesis are studied e.g. by Duffy, Papageorgiou, and Perez-Sebastian (2004) who find only weak evidence in support of it. The hypothesis has also been tested in many microeconomic studies with firm and industry-level data. For a representative sample of such studies cf. Hamermesh (1993) who concludes that there is strong evidence for capital-skill complementarity. However, he cautions that many of the studies that disaggregate the work force by demographic group exclude capital as a productive input due to the lack of a reliable measure of the capital stock. The assumption of competitive markets allows early studies to proxy variations in the capital stock by variations in the rates of return. More recent studies refrain from this assumption and explicitly include the stock of capital in their analysis (cf. Duffy, Papageorgiuo and Perez, 2004). Goldin and Katz (1998) argue that physical capital and skill have not always been viewed as relative complements: In an earlier era, the transformation from skilled artisan shops to factories involved the substitution of physical capital and unskilled labor for highly skilled labor – precisely the opposite of what is hypothesized to be happening today. This suggests that capital-skill complementarity may be a transitory phenomenon: In a country's development process skilled labor may change from being well substitutable with capital and unskilled labor to being highly complementary to these two inputs. In order to account for this, recent studies consider panel data over long periods of time (e.g. Duff, Papageorgiou and Perez, 2004) or split the full country sample into subsamples of similar development (e.g. Papageorgiou and Chmelarova, 2003). A consistent finding also supported by Galor and Weil (2000) is that the capital-skill complementarity is especially strong at the beginning of a development process and may fade out when a technology becomes widely adopted.

The observed increase in the skill premium during the 1990s in developed countries may be due to skill-biased technological change, as is argued e.g. by Acemoglu (1998) who argues that an increase in the supply of skills reduces the skill premium in the short run, but induces biased technological change and hence increases the skill premium in the long run. Krusell, Ohanian, Rios-Rull, and Violante (2000) challenge this view in an empirical study where they find that with capital-skill complementarity, changes in observed inputs in the aggregate production function alone account for most of the variations in the skill premium.

In order to account for the established difference between labor of various skill levels in their substitutability with capital, we use a nested production function commonly used in the relevant empirical literature; specific to our setting is the inclusion of more than two skill classes. Every differentiated input factor in the production function makes up for an input level in the nesting structure, as shown in figure 3 below. We concentrate on the analysis of three differentiated skill classes, which is a realistic reproduction of the education structure in Switzerland (cf. also section 5.3).

As opposed to Heckman, Lochner, and Taber (1998), Blundell and Bond (2000) and Blundell, Dias, and Meghir (2003), we explicitly allow for capital-skill complementarity and hence for a non-constant share of capital in production. There are several ways of nesting the differentiated production factors within a function, of which six allow for full capital-skill complementarity. With a monotonous ordering of labor inputs according to their skill level, there are two nestings:

$$Y = \Upsilon_1\left(H^1, \Upsilon_2\left(H^2, \Upsilon_3\left(H^3, K\right)\right)\right),$$
$$Y' = \Upsilon_1\left(H^3, \Upsilon_2\left(H^2, \Upsilon_3\left(H^1, K\right)\right)\right),$$

where $\Upsilon_1, \Upsilon_2, \Upsilon_3$ are aggregators. Y' implies the elasticity of substitution between H^3 and K to be the same as between H^1 and H^3 by the symmetry restrictions of the functional form. This is at odds with factor elasticity estimates, which suggest that the substitution elasticity between skilled labor and unskilled labor is higher than the substitution elasticity between skilled labor and capital (cf. Hamermesh, 1993, and Krusell, Ohanian, Rios-Rull, and Violante, 2000). Hence, we choose Y which does not seem to be at variance with elasticity estimates. This ordering of the nesting also conforms with the functional form chosen by Duffy, Papageorgiou, and Perez-Sebastian (2004). The according production function writes as

$$
\begin{aligned}
Y &= F(K, L^1, L^2, L^3) \qquad (12)\\
&= Y^0 \left[a^1 \left(L^1\right)^{\kappa^1} + \left(1-a^1\right) \left[a^2 \left(L^2\right)^{\kappa^2} + \left(1-a^2\right) \left[a^3 \left(L^3\right)^{\kappa^3} + \left(1-a^3\right) K^{\kappa^3} \right]^{\frac{\kappa^2}{\kappa^3}} \right]^{\frac{\kappa^1}{\kappa^2}} \right]^{\frac{1}{\kappa^1}}.
\end{aligned}
$$

The most aggregate nest is final output Y which consists of unskilled labor, L^1, and a subnest, which we call Y^2 and which itself consists of medium skilled labor L^2 and a composite Y^3 of high skilled labor L^3 and capital K. a^1, a^2, a^3 are distribution parameters. Note that on each level of the nesting structure only one type of labor enters production. The chosen production technology can be characterized by the price elasticity of demand for the input factors. If labor is paid the value of its marginal product, total compensated demand for labor on aggregation level $s \in \{1,2,3\}$, L_t^s, is given by

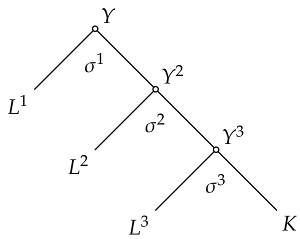

Figure 3 Production technology with three-level nested structure.

$$
L_t^s = Y_t^s (a^s)^{\sigma^s} \left(\frac{C_t^s}{w_t^s} \right)^{\sigma^s} = Y_t \left(w_t^s\right)^{-\sigma^s} (a^s)^{\sigma^s} \left(C_t^1\right)^{\sigma^1} \prod_{n=2}^{s} \left(1 - a^{s-1}\right)^{\sigma^{s-1}} \left(C_t^n\right)^{\sigma^n - \sigma^{n-1}}
$$

where Y_t^s denotes demand, C^s is total cost for the whole nest at level s (cf. figure 3), w_t^s is the unit price of labor of type s and $\sigma^l = \frac{1}{1-\kappa^l}$. By Shephard's lemma, the derivative of total cost with respect to w^i equals the demand for input factor i:

$$
\frac{\partial C^s}{\partial w^i} = \begin{cases} 0 & i < s, \\ L^i & i \geq s. \end{cases}
$$

The compensated own-price elasticity of demand for labor s is

$$
\frac{\partial L^s}{\partial w^s} \frac{w^s}{L^s} = -\sigma^s + \sigma^1 \frac{\partial C^1}{\partial w^s} + \sum_{n=2}^{s} \left(\sigma^n - \sigma^{n-1} \right) \frac{\partial C^n}{\partial w^s}.
$$

Let $k \in \{i, j\}$ denote the deepest price aggregate which contains both w^i and w^j. Taking into account the impact of w^i on $C^{j \leq k}$, the cross price-elasticity is given by

$$
\varepsilon^{i,j} \equiv \frac{\partial L^i}{\partial w^j} \frac{w^j}{L^i} = \sigma^1 \frac{\partial C^1}{\partial w^j} + \sum_{n=2}^{k} \left(\sigma^n - \sigma^{n-1} \right) \frac{\partial C^n}{\partial w^j}.
$$

In analogy, the cross price-elasticity between labor L^s and capital K is

$$
\varepsilon^{L^s,K} \equiv \frac{\partial L^s}{\partial w^K} \frac{w^K}{L^s} = \sigma^1 \frac{\partial C^1}{\partial w^K} + \sum_{n=2}^{s} \left(\sigma^n - \sigma^{n-1} \right) \frac{\partial C^n}{\partial w^K},
$$

where w^K is the cost per unit of capital.

3.2 Optimum Investment and Factor Demand

Given the production technology as described, a representative firm faces the following situation:

$$Y_t \equiv F(\mathbf{L}_t^D, K_t), \quad \mathbf{L}_t^D \equiv \left(L_t^1, \dots, L_t^S\right), \tag{13}$$

$$L_t^s \equiv \sum_{a=1}^{A} \theta_t^{s,a} L_t^{s,a}, \quad L_t^{s,a} = l_t^{s,a} N_t^{s,a}, \tag{14}$$

$$\chi_t = Y_t - \sum_l \sum_a w_t^{s,a} L_t^{s,a} - I_t, \quad w_t^{s,a} = w_t^s \theta_t^{s,a}, \tag{15}$$

$$K_{t+1} = I_t + (1 - \delta^K) K_t. \tag{16}$$

Y_t is output, K_t is capital input, χ_t are net dividends, I_t is gross investment, and δ^K is the rate of depreciation of capital. In order to derive realistic dynamics of investment in a small open economy, investment costs are assumed to increase progressively with gross investment, reflecting adjustment costs of capital. The capital installation technology is assumed to be linear homogeneous, with investment costs $I^C = \psi(I, K)$ and $\psi_I, \psi_{II} > 0$ and $\psi_K < 0$. As a normalization, we specify $\delta^K = \psi\left(\delta^K, 1\right)$, implying that $I^C = I$ in a steady state.

The Bellman equation for the firm's maximization problem is

$$V^K(K_t) = \max_{I_t, \mathbf{L}_t^D} \left\{ \chi_t + \frac{V^K(K_{t+1})}{R_t} \right\} \quad \text{s.t. (13) – (16).} \tag{17}$$

Define $\lambda_t^K \equiv \frac{\partial V_t}{\partial K_t}$ as the marginal value of a unit of capital. Then, the firm's optimality conditions write as:

$$(I_t): \quad \psi_I = \frac{\lambda_{t+1}^K}{R_{t+1}}, \tag{18}$$

$$(L_t^s): \quad w_t^s = F_{L^s}. \tag{19}$$

Note that w_t^s is the wage per efficiency unit, while $l_t^{s,a} w_t^s \theta_t^{s,a}$ is the wage per capita of skill level s and age group a exerting effort $l_t^{s,a}$.

The envelope condition for the firm's problem is

$$(K_t): \quad \lambda_t^K = F_K - \psi_K + (1 - \delta^K) \frac{\lambda_{t+1}^K}{R_t}, \tag{20}$$

which, after using (18), noting that $R = 1 + r$, and in a steady state $I = \delta^K K$, becomes $F_K = r_t + \delta^K$. We can restate the result by Hayashi (1982) as

$$V_t^K \equiv \lambda_t^K K_t. \tag{21}$$

4 Aggregation and General Equilibrium

The laws of motion of population aggregates are given in the appendix. Based on these, aggregate macroeconomic variables are computed as in the following result.

Result 2 *Aggregate macroeconomic variables are obtained as*

$$C_t \equiv \sum_{s=1}^{S}\sum_{a=1}^{A} C_t^{s,a}, \quad C_t^{s,a} = \sum_{\alpha\in N_t^a} C_{\alpha,t}^{s,a} N_{\alpha,t}^{s,a},$$

$$L_t^{S,s} \equiv \sum_{a=1}^{a^R-1} \theta_t^{s,a} l_t^{s,a} N_t^{s,a} + \delta^{a^R} \theta_t^{s,a^R} l_t^{s,a^R} N_t^{s,a^R}, \tag{22}$$

$$N_t^R = \sum_{a^R+1}^{A} N_t^a + \left(1-\delta^{a^R}\right) N_t^{a^R}, \quad N_t^a = \sum_{s=1}^{S} N_{\alpha,t}^{s,a} \tag{23}$$

Proof. Cf. Grafenhofer, Jaag, Keuschnigg, and Keuschnigg (2005). ■

Closing the model requires the consolidated government budget constraint to hold. Wage taxes must finance public consumption G, which is kept constant per capita of the total population at level g, and pay-as-you-go pension spending, where $L_t^{S,s}$ and N_t^R are given in (22) and (23),

$$\tau \sum_{s=1}^{S} w_t^s L_t^{S,s} = e_t N_t^R + G_t, \quad G_t = g \cdot N_t, \quad N_t = \sum_{a=1}^{A} N_t^a. \tag{24}$$

Finally, we state the capital and labor market clearing conditions. In equilibrium, demand for efficiency units of labor corresponds to aggregate household sector labor supply, scaled by worker productivity. Further, accumulated household sector financial wealth absorbs the value of domestically issued equity, V_{t+1}, and foreign bonds D_{t+1}^f,

$$L_t^s = L_t^{S,s}, \quad A_{t+1} = V_{t+1}^K + D_{t+1}^f. \tag{25}$$

When assets are perfectly substitutable, they must earn an identical rate of return equal to the market interest r_{t+1}. Given optimal household and firm behavior, and with all budget constraints fulfilled, Walras' Law implies the current account

$$D_{t+1}^f = R_{t+1}\left[D_t^f + Q_t - I_t^C - G_t - C_t\right], \tag{26}$$

where the stock of foreign net assets D_t^f is measured at the end of the period.

5 Calibration

We calibrate the model to stylized data of Switzerland in 2000. Table 1 states key parameters characterizing preferences and technology. The choice of the most important parameters is detailed in the following sections by referring to the relevant literature.

5.1 Mortality and Transition Probabilities

For simplicity, we consider a demographic stationary state and ignore time indices in this section. We further assume that aging parameters are equal across skill groups. Setting $\omega^a = 0$ implies that an aging event occurs with probability one in each period, leading to $\widetilde{N}^t = \widetilde{\gamma}^{t-1}\widetilde{N}^{t-1}$. The concept of

Table 5.1
Taste and technology parameter calibration.

Real interest rate r	0.050
Depreciation rate δ^K	0.100
Output elasticity of capital α	0.350
Subjective discount factor β	0.956
Intertemporal elasticity of substitution σ	0.800
Wage elasticity of labor supply ε^L	0.400
Retirement semi-elasticity ε^R	0.600
Skill choice semi-elasticity ε^D	60.000
Capital – low skill elasticity of substitution σ^{K,L^1}	1.800
Capital – medium skill elasticity of substitution σ^{K,L^2}	1.200
Capital – high skill elasticity of substitution σ^{K,L^3}	0.600
Proportional wage tax rate t^W	0.277
Pension replacement rate ζ^1	0.694
Pension replacement rate ζ^2	0.380
Pension replacement rate ζ^3	0.253

Notes: The labor supply elasticity is $\varepsilon^L \equiv \varphi'/(l\varphi'')$, the pension replacement rate $\zeta = e/[(1-t^W)w\theta^{a^R-1}]$, and the production function as in (12).

an age group thus becomes identical with a cohort or vintage where age t is measured by time since birth. The tilde indicates the decomposition in annual cohorts $\widetilde{N}^t$. Taking age-dependent survival rates $\widetilde{\gamma}^t$ from official mortality tables, the cohort composition of the population in a demographic steady state can be constructed. Recursively applying $\widetilde{N}^t = \widetilde{\gamma}^{t-1}\widetilde{N}^{t-1}$ yields the size of cohort t relative to the size of a new cohort. Summing up over all cohorts fixes the size of the new cohort compared to total population size N,

$$\widetilde{N}^t = \widetilde{N}^1 \prod_{s=1}^{t-1} \widetilde{\gamma}^s, \quad \widetilde{N}^1 = N \Bigg/ \sum_{t=1}^{T} \prod_{s=1}^{t-1} \widetilde{\gamma}^s . \tag{27}$$

Taking a total length of life of T years, and based on actual survival rates, we have thus found the stationary decomposition of the population into a total of T cohorts. Using the population decomposition indicated in table 5 in the appendix, each age groups contains several cohorts,

$$N^a = \sum_{t=\alpha_a}^{\alpha_{a+1}} \widetilde{N}^t. \tag{28}$$

Given that the instantaneous probability of staying is $\omega^a\gamma^a$, the expected duration in group a is

$$\alpha_{a+1} - \alpha_a = 1/\left(1 - \omega^a\gamma^a\right). \tag{29}$$

The chosen age group decomposition of the actual population corresponds to a life-cycle history where an agent spends exactly the average duration in each age period. One can thus recover the demographic parameters of the model from our knowledge of aggregated population data N^a and duration in group a. In a demographic steady state, each age group must fulfill the restriction

$$(1-\gamma^a\omega^a)\cdot N^a = \left(\gamma^{a-1} - \gamma^{a-1}\omega^{a-1}\right)\cdot N^{a-1}. \tag{30}$$

At this stage, one knows N^a from aggregated population data and $\gamma^a\omega^a$ from the age group duration as implied by the chosen aggregation. The only unknown in (30) is γ^{a-1} which is easily solved for all groups except the last one. For the last group, γ^A follows directly from (30) on account of the restriction $\omega^A = 1$. For all age groups $a < A$, γ^a can now be recovered recursively.

5.2 Intertemporal Substitution

In the literature, much effort has been devoted to accurately pin down the value of the rate of intertemporal substitution. On the one hand, macroeconomists use a large value in general, reflecting a common view that a high degree of intertemporal substitution is more consistent with aggregate data considering dynamic macroeconomic models. In their seminal paper, Kydland and Prescott (1982) calibrate it to 0.66 and Lucas (1990) argues that an elasticity of 0.5 appears too low when confronted with macro data. Also Weil (1989) finds that in order match growth and business cycle facts the intertemporal rate of substitution is required to be close to unity. In an influential paper Hall (1988) shows that consumption growth is completely insensitive to changes in interest rates and, therefore, intertemporal elasticity is close to zero. Guvenen (2003) argues that this apparent contradiction arises from ignoring heterogeneity across individuals with respect to stock market participation and wealth. The properties of aggregate variables directly linked to wealth, such as investment and output are entirely determined by high-elasticity stockholders, while aggregate consumption, which is more evenly distributed across households, uncovers the low-elasticity of most households. In accordance with the literature on life-cycle modeling, we calibrate the elasticity of intertemporal substitution to 0.4 (cf. eg. Keuschnigg and Keuschnigg, 2004).

5.3 Skill Structure and Labor Supply

With the skill choice given by (11), ξ^s is calibrated using

$$\xi^s = \frac{\Delta V^s}{\exp\left(\frac{\Gamma(v^s)}{\varepsilon^D}\right)},$$

where $\Gamma(v^s)$ can be recovered from the skill distribution of the 20 years old in the population. As to the extensive skill choice, Heckman, Lochner, and Taber (1998) find that an increase in college tuition of \$1000 (or – equivalently – a decrease in the present value of after-tax earnings) decreases the probability of attending college by about 0.08. This estimate is slightly higher than the estimate by Kane (1994) who finds an according decrease by 0.05 (cf. also Dupor, Lochner, Taber, and Wittekind, 1996). The estimated decrease by circa one per mill of net income associated with a 0.06 decrease in the college enrollment probability yields a skill choice semi-elasticity with respect to income of 60. At the beginning of life, with zero assets, indirect utility is proportional to human capital in our model. Therefore, the extensive skill choice semi-elasticity is chosen to be 60 which means that an increase in the difference of indirect utility at the beginning of life between any two consecutive skill classes increases by one percent, the number of individuals choosing the higher skill class increases by 60 percentage points.

The three skill classes refer to individuals whose highest level of education is secondary I, secondary II and tertiary, respectively. Figure 4 shows the according [6] levels of our skill classes

[6] is the International Standard Classification of Education.

Figure 4

Skill distribution in the Swiss population, 2000.
Source: (2004)

Skill class		Size	Inc. share
1	1/2	19%	12%
2	3/4	54%	50%
3	5/6	27%	38%

along with their relative population- and labor income shares. The skill class and age group specific productivity parameter θ is calibrated using wage data from Switzerland, 2000. The shapes of individual wage profiles are displayed in figure 5.

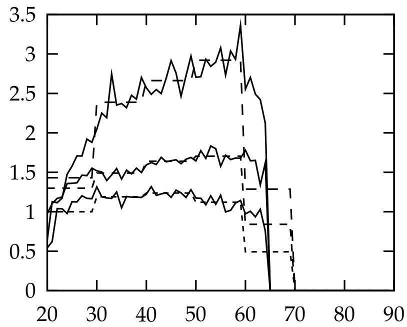

Figure 5

Wageprofiles of high, medium, and low skilled workers.
Source: BFS (2004), own computations

The income share of capital is chosen to be 0.4 (cf. Imrohoroglu, Imrohoroglu, and Joines, 1995, and Conesa and Krueger, 1999, who choose 0.36). The empirical evidence on the static labor supply elasticity is reviewed by Hansson and Stuart (1985). The latest, most advanced studies in their survey yield estimates of the wage elasticity of labor supply between 0.4 and 0.5 (cf. also Trostel, 1993). We calibrate the labor supply elasticity to 0.4.

5.4 Production Technology

For general production functions with more than two inputs, there are multiple possible definitions for the elasticity of substitution between pairs of input. The one most commonly used is the Allen-Uzawa partial elasticity of substitution which measures the percentage change in the ratio of two inputs in response to a change in the ratio of the two input prices, holding all other prices – but not all other inputs – and the output quantity constant. This measure is widely used in the empirical literature, e.g. by Sato (1967), Griliches (1969), and Goldin and Katz (1998). The Allen-Uzawa elasticity of substitution is

$$\sigma^{i,j} \equiv \frac{C^1(\mathbf{w})\, C^1_{i,j}(\mathbf{w})}{C^1_i(\mathbf{w})\, C^1_j(\mathbf{w})}$$

where C^1 denotes total cost of final output and subscripts refer to partial derivatives with respect to inputs i and j.[7] $\mathbf{w}$ is the vector of unit prices of the input factors in $\mathbf{q} = \left\{L^1, L^2, L^3, K\right\}$. Denote by k the deepest aggregate containing both inputs i and j. In the production function (12) with the nest at level 1 producing final output, the elasticities of substitution between any two input

[7] Another measure is the Hicks-Allen direct partial elasticity of substitution which measures the percentage change in the ratio of two inputs in response to a change in the ratio of the two input prices, holding all other prices, inputs, and the output quantity constant. Assume a general production technology $Y = F(\mathbf{q})$. Letting $F_i \equiv \partial F / \partial q^i$, the direct partial elasiticty of substitution between the i$^{\text{th}}$ and the j$^{\text{th}}$ element in $\{\mathbf{q}\}$ is given by $\sigma^{d,i,j} = -\partial \ln\left(q^i/q^j\right) \big/ \partial \ln\left(F_i/F_j\right)$, where Y and all $q^{k \notin \{i,j\}}$ are held constant.

factors i and j write as

$$\sigma^{i,j} = \sigma^1 + C^1 \sum_{n=2}^{k} \frac{\left(\sigma^n - \sigma^{n-1}\right)}{C^n} \tag{31}$$

where C^l is the total cost of a nest at level l and $\sigma^l = \frac{1}{1-\kappa^l}$. Note that the elasticity of substitution between i and j is independent of the subnest elasticity in all nests $s > k$. Applying the concept of capital-skill complementarity implies in the setting of our production technology that

$$\sigma^{K,L^i} < \sigma^{K,L^j} \Leftrightarrow i > j.$$[8]

The relative income shares of the three kinds of labor and capital are known from data. The elasticities of substitution between capital and skilled / unskilled labor is taken from the empirical literature. Using three types of labor, it is clear that high skilled labor is even more complementary with capital than skilled labor in the literature dealing with two kinds of labor only. Hence, we correct the respective elasticities of substitution between capital and skilled / unskilled labor accordingly. The values of $\sigma^1, \sigma^2, \sigma^3$ are computed from (31) recursively:

$$\begin{aligned}
\sigma^1 &= \sigma^{K,H^1}, \\
\sigma^2 &= \frac{C^2}{C^1}\left(\sigma^{K,H^2} - \sigma^1\right) + \sigma^1, \\
\sigma^3 &= \frac{C^3}{C^1}\left(\sigma^{K,H^3} - \sigma^1 - \frac{C^1}{C^2}\left(\sigma^2 - \sigma^1\right)\right) + \sigma^2.
\end{aligned}$$

From national accounting we know the income shares of the different production factors. From these, we can calibrate the distribution coefficients $a_1 = 0.16, a_2 = 0.39, a_3 = 0.37$, and $\kappa_1 = 0.44, \kappa_2 = 0.20, \kappa_3 = -1.3$ (cf. figure 4).

Using a nested production technology with two kinds of labor, Krusell, Ohanian, Rios-Rull, and Violante (2000) estimate the elasticity of substitution between unskilled labor and capital to be 1.67. This is close to the value of 1.5 reported by Johnson (1997). Duffy, Papageorgiou, and Perez-Sebastian (2004) report a range from 1.3 to 10 for this value. Similarly, the estimate of Krusell, Ohanian, Rios-Rull, and Violante (2000) of the elasticity of substitution between skilled labor and capital is 0.67, which is in the range of the estimations by Hamermesh (1993). The estimates by Krusell, Ohanian, Rios-Rull, and Violante (2000) are also used in Lindquist (2004). There are no estimates for three types of labor up to date. In order to account for a third skill class, we set the elasticity of substitution between capital and low, medium, and high skill labor to 1.8, 1.2, and 0.6, respectively.

6 Demographic Scenario

The aging scenario consists of people becoming older on average and potentially living longer. Becoming older means that the mass of people in their 80s becomes larger if more of the younger

[8]Duffy, Papageorgiou, and Perez-Sebastian (2004) show that in the two-level specification of (12) with L^1 denoting high-skilled labor and L^2 low-skilled labor, the capital-skill complementarity holds iff $\sigma^1 < \sigma^2$, regardless of which elasticity measure is used. We are well aware that the Allen-Uzawa elasticity as a measure of substitutability lacks the salient theoretical properties the Hicksian elasticity of substitution exhibits in the two-goods case (cf. Blackorby and Russell, 1989). Since most of the empirical literature deals with that measure, however, we also stick to it.

agents make it to their 80s. Living longer means that the life-time horizon gets longer. To implement the scenario, we raise the survival rates of age groups 5 to 8 by the factors given in table 5 but keep the expected duration in each age group constant. Comparing the columns in tables 2 and 5 shows that the mortality rates $1-\gamma^a$ decline, which necessarily implies that a larger fraction of this group moves to the next one, instead of dying. Since the last group becomes less mortal, $1-\gamma^8$ falling from 0.2 to 0.12, expected duration in that group rises from 5 to $\left(1-\gamma^8\right)^{-1} = 8.3$ years. The representative agent lives longer, namely 93.3 instead of 90 years. This corresponds to the aging scenario in Kalemli-Ozcan, Ryder, and Weil (2000) who discuss the implications of longevity in a model with a single mortality rate.

Table 2

Aging and life-expectancy.

Age groups	1	2	3	4	5	6	7	8
Initial N^a	0.179	0.177	0.175	0.168	0.148	0.107	0.031	0.016
New N^a	0.179	0.177	0.175	0.168	0.148	0.121	0.047	0.058
Factor $\times\gamma^a$	1.000	1.000	1.000	1.000	1.010	1.020	1.050	1.100
Prob. $1-\gamma^a$	0.001	0.001	0.004	0.012	0.018	0.023	0.050	0.120
Prob. $1-\omega^a$	0.099	0.099	0.096	0.089	0.083	0.079	0.158	0.000

Notes: $1-\gamma^a$ mortality rate, $1-\omega^a$ probability of aging.

We compare the aging scenario for different assumptions with regard to international capital mobility. For this purpose, we have calibrated the model with net foreign debt equal to zero. In a small open economy, the real interest rate is fixed on world markets. Any imbalance between domestic savings and investment is thus reflected in a change in the net foreign asset position. In a closed economy, the domestic real interest rate adjusts to keep net foreign assets to zero. This second scenario is interesting, as aging is a worldwide phenomenon that might lead to a decline in the real interest rate via increased savings.[9]

We first turn to the open economy case. Even though the interest rate remains constant, the ratio between capital and the three different kinds of labor varies due to the -production technology. Aging results in a larger number of old people. In the absence of an offsetting decline in fertility, the mass of younger age groups remains almost unchanged, leading to a largely constant workforce and . By assumption, the increased number of retirees erodes pensions per capita to an extent that keeps the pensions to ratio constant. Note, however, that the increase in overall population by about 7% inflates government consumption and thereby necessitates a moderate increase in the wage tax by almost two percentage points.

Effective labor supply of any skill type is affected via three channels. First, active people adjust their labor supply, second, workers in age group 5 defer retirement, and third, individuals sort differently into skill classes. The higher tax erodes the net wage and thereby discourages hours worked at the intensive margin of labor supply. This in itself would lead to lower employment and output. However, the large reduction in per capita pensions makes retirement less attractive compared to continued work. Hence, postponed retirement raises aggregate labor supply in all skill classes at the extensive margin. With the decrease in the use of capital in production, by the capital-skill complementarity, also the use of skilled labor decreases.

[9] For example, computations by Börsch-Supan, Ludwig, and Winter (2004) yield a decline in the return to capital by roughly one percentage point as a result of worldwide aging.

Table 5.3
Long-run impact of aging on key macro variables in percentage changes relative to initial steady state.

Key macro variable			Open	Closed
r	real interest rate*	0.050	0.050	0.018
D^f/Y	debt – ratio*	0.000	0.631	0.000
δ^1	retirement*	0.300	0.330	0.343
δ^2	retirement*	0.400	0.476	0.623
δ^3	retirement*	0.500	0.597	0.781
τ	contribution rate*	0.277	0.296	0.239
e	pension p.c.		-21.507	11.392
$(1-\tau)w^1$	net wage skill type 1		-5.924	28.063
$(1-\tau)w^2$	net wage skill type 2		-3.048	21.019
$(1-\tau)w^3$	net wage skill type 3		-1.154	15.860
l^1	labor supply skill type 1		-2.413	10.400
l^2	labor supply skill type 2		-1.230	7.930
l^3	labor supply skill type 3		-0.463	6.065
N^1	size of skill class 1		7.133	-17.785
N^2	size of skill class 2		-0.667	-0.618
N^3	size of skill class 3		-3.685	13.752
N^W	workers		1.356	4.282
N^R	retirees		25.668	16.472
L^1	effective employment type 1		4.828	-8.883
L^2	effective employment type 2		-1.065	9.918
L^3	effective employment type 3		-3.105	24.380
K	capital stock		-1.746	67.137
Y	gross domestic product,		-1.360	29.741
C	aggregate consumption		1.507	41.033
A	aggregate assets		23.669	67.137

Notes: Pension – ratio and public consumption p.c. constant;
*) absolute values, initial values in first column.

The last line clearly demonstrates the life-cycle savings response. When people must expect a much reduced pension from the pay-as-you-go system (-22%), they must substantially increase their savings for retirement even though the interest rate remains constant. Aggregate assets are therefore 24% higher. In the absence of domestic investment opportunities, this extra wealth is invested abroad and leads to a net foreign asset position of 63% of in the long-run. The additional asset income from abroad allows for an increase in aggregate consumption by roughly one and a half percent, in spite of the decrease in .

Adjustment in the closed economy is much different. Since excess savings cannot be invested abroad, the real interest rate must decline, by 3.2 percentage points in order to balance savings and investment. With production much more capital intensive, wages rise substantially for all skill classes in the long-run and thereby stimulate labor supply both on the intensive and extensive margins. People work more hours, and they postpone retirement as work becomes more attractive relative to pension income. The expansion of effective employment in combination with higher capital intensity boosts the capital stock and aggregate asset holdings by 67% in the long-run, leading to an output gain of almost 30%. Keeping the pension to ratio constant, the economy thus affords an overall increase in pension expenditure and even in pensions per capita. Note that along with an increase in capital, effective employment of high skilled labor increases significantly due to capital-skill complementarity.

Our simulation is insensitive with respect to the semi-elasticity of the skill choice. However, since the shares of employed factors heavily depends on the reaction of the interest rate reaction, we simulate our model with various values of the elasticity of intertemporal substitution as in table 4.

Table 5.4
Closed economy sensitivity to intertemporal substitution.

			$\sigma = 0.2$	$\sigma = 0.4$	$\sigma = 0.8$
r	real interest rate*	0.050	0.017	0.018	0.028
$(1-\tau)w^1$	net wage skill type 1		27.169	28.063	19.606
$(1-\tau)w^2$	net wage skill type 2		21.776	21.019	13.641
$(1-\tau)w^3$	net wage skill type 3		17.757	15.860	9.271
L^1	effective employment type 1		-6.607	-8.883	-8.093
L^2	effective employment type 2		10.214	9.918	7.020
L^3	effective employment type 3		23.833	24.380	17.635
K	capital stock		70.299	67.137	42.859
Y			30.785	29.741	19.992

Note: *) absolute value, initial value in first column.

A disproportionate increase in pension benefits (as in table 3, closed economy) induces individuals to save more for retirement. In a closed economy, the capital stock must thereby increase with a resulting decline in the marginal productivity of capital and the real interest rate. The larger the elasticity of intertemporal substitution, and hence the substitutability of consumption across time, the smaller the increase in additional savings, and therefore the smaller the decrease in the real interest rate. The qualitative results remain basically the same with reasonable values of σ : The higher marginal productivity of labor offsets the effect of higher income tax rates to finance the pension system, which would otherwise be a harmful economic consequence of population aging.

7 Conclusion

We have applied the concept of probabilistic aging with endogenous discrete skill choice to analyze the impact of aging during demographic transition. The model setting allowed for an appraisal of inter- and intragenerational distribution effects by considering eight age groups and three skill classes. The empirically well established complementarity between capital and skilled labor is embraced in a nested production technology. Simulating the effect of aging on key macro variables, we find that the remunerations of labor in different skill classes and accordingly their size are very differently affected. The vanishing attractiveness of early retirement due to a decrease in the relative size of per-capita pensions results in prolonged employment for all skills. In a small open economy, where the interest rate is tied to international capital markets, contracts due to decreased labor supply as a result of lower net wages. In a closed economy setting, aging sparks off a boom via lower interest rates and increased use of capital and skilled labor in production.

8 Appendix

The appendix outlines the concept of probabilistic aging as in Grafenhofer, Jaag, Keuschnigg, and Keuschnigg (2005), which allows for the possibility that an individual keeps her productivity in the next time-period, but also that she may be struck by sudden death. The aging process being the same in all skill classes, for simplicity we omit superscript s throughout the appendix.

In order to reproduce realistic wage profiles, individuals must undergo some kind of aging. To preserve aggregability, aging is absent in Blanchard-style models of "perpetual youth" with overlapping generations. Since we are interested in modeling life-cycle wage-profiles in the context of an model of the Blanchard style, while still being able to find closed-form aggregate solutions, we use the concept of probabilistic aging to model aging as shown in figure 6. In this section, we concentrate on individual decisions once their skill level s is determined and fixed. We therefore temporarily drop superscript s. Each individual is a member of one of A age- or productivity-groups $a \in \{1, \ldots, A\}$. Individuals are born in age group $a = 1$ as workers. The affiliation of a variable at time t to an individual currently belonging to age-group a is denoted by $(\cdot)_t^a$.

Households differ not only by their date of birth, but also by their diverse life-cycle histories. An agent's life-cycle history is her biography of aging events that have happened since birth. At date t, the set of possible histories of a household that belongs to age group a is

$$\mathcal{N}_t^a \equiv \{(\alpha_1, \ldots, \alpha_a) : \alpha_1 < \ldots < \alpha_a \leq t\}. \tag{32}$$

A particular life-cycle history is represented by a vector $\alpha \in \mathcal{N}_t^a$. The element α_i, $i \in \{1, \ldots, a\}$, denotes the date at which the household who was formerly in age group $i-1$, became a member of group i. In denoting the unborn by a virtual age group zero, the element α_1 lists the date of birth when an agent switches from the group of unborn to the first age group. We say that a member of group $a = 1$ aged only once with no further aging since birth. Nevertheless, different persons of the first age group are heterogeneous since they were born at different moments in the past. The set of possible biographies is $\mathcal{N}_t^1 = \{(\alpha_1) : t \geq \alpha_1\}$. With this notation, the vectors α describing the biography of people in group a have a elements since such persons have aged a times in total.

The individual biographies are updated when a person is subject to an aging event. Suppose a person is in age group $a-1$ and is identified by a given biography $\alpha = (\alpha_1, \ldots, \alpha_{a-1})$. When the next aging event occurs in period t, she arrives in group a next period. Her biography is appended by the entry $t+1$ and will thus read $(\alpha_1, \ldots, \alpha_{a-1}, t+1)$. Accordingly, the set $\mathcal{N}_t^a$ of biographies of age group a will be augmented next period by all the biographies $\alpha' \in \mathcal{N}_t^{a-1} \times (t+1)$ that have $t+1$ as their last entry and refer to people who currently switch from group $a-1$ to a. Hence,

$$\mathcal{N}_{t+1}^a = \mathcal{N}_t^a \cup \mathcal{N}_t^{a-1} \times (t+1), \quad a \in \{1, \ldots, A\}. \tag{33}$$

To model demographics, we allow for mortality among younger age groups. When an individual with an arbitrarily given life-cycle history plans for next period, she faces the risk of aging and dying (cf. figure 6). She must thus reckon with three possible events: (i) she dies with probability $1-\gamma^a$; (ii) she survives without aging and remains in the same age group with probability $\gamma^a \omega^a$, and (iii) she survives and ages and belongs to age group $a+1$ next period with probability $\gamma^a (1-\omega^a)$. Individuals in the last age group have exhausted the aging process and remain in this group with probability one, $\omega^A = 1$. They may either survive with probability γ^A within group A or die with probability $1 - \gamma^A$. The last age group behaves according to the mortality and demographic assumptions of Blanchard's (1985) perpetual youth model.

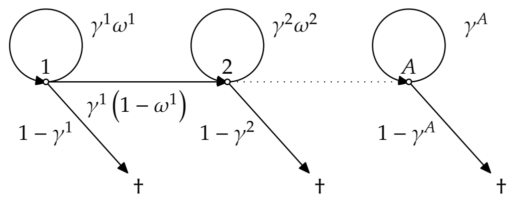

Figure 6

Aging and mortality hazard of individuals in the model.

Since the characteristics of people such as their earnings potential differ across age groups, an agent's consumption, assets and other economic variables will generally depend on her particular life-cycle history. For example, assets depend on the agent's past earnings history which, in turn, is linked to her aging trajectory. To keep track of the population's heterogeneity, one must thus identify each agent by her age group as well as her aging biography α which also includes the date of birth. The number of agents at date t, in state of life a and with aging history α is given by $N^a_{\alpha,t}$. Within this group, agents are identical and face the same independent probability of moving to one of the alternative states. With stochastically independent risks, the law of large numbers implies that the above stated individual probabilities correspond to the fraction of people that are subject to the respective event. Consequently, the group is divided into three subgroups next period: (i) those who die, (ii) those who survive within the same age group a, and (iii) those who are hit by an aging event, switch to the next higher age group. The age group and the biography α remains unchanged in case (ii) while switching to the next higher age group $a+1$ in case (iii) adds another event in a person's life-cycle history α and thereby results in a new biography α':

$$\begin{array}{llcll} (i) & N^{\dagger}_{\alpha^{\dagger},t+1} & = & N^a_{\alpha,t}\cdot(1-\gamma^a), & \text{death,} \\ (ii) & N^a_{\alpha,t+1} & = & N^a_{\alpha,t}\cdot\gamma^a\omega^a, & \text{no aging,} \\ (iii) & N^{a+1}_{\alpha',t+1} & = & N^a_{\alpha,t}\cdot\gamma^a(1-\omega^a), & \text{aging.} \end{array} \tag{34}$$

Individuals enter our model at age 20. We exclude ages $0-20$ from the analysis for two reasons: First, in our model setting, the extensive schooling decision as described in section 2.4 is considered to make up the initial condition for on-the-job training and second, the actual survival function cannot be easily approximated as the mortality hazard is non-monotonous.

Table 5 shows our decomposition of the population from ages 20 to 90. The chosen aging pattern determines the representative agent's life-cycle biography $\alpha = (20, 30, 40, 50, 60, 70, 80, 85)$, which serves as an aggregation key. It allows for a statement about the size and law of motion of the various age groups in result 3.

Result 3 *With the number of individuals in age group a given by*

$$N^a_t \equiv \sum_{\alpha\in N^a_t} N^a_{\alpha,t}, \tag{35}$$

(a) the aggregate law of motion for age group $1 < a \leq A$ *is*

$$N^a_{t+1} = \gamma^a\omega^a\cdot N^a_t + \gamma^{a-1}\left(1-\omega^{a-1}\right)\cdot N^{a-1}_t, \quad \omega^A = 1, \tag{36}$$

(b) age group 1 *evolves according to*

$$N^1_{t+1} = \gamma^1\omega^1\cdot N^1_t + N^1_{(t+1),t+1}, \tag{37}$$

Table 5.5
Demographic and life-cycle parameters.

Age groups	1	2	3	4	5	6	7	8
Cohorts	20-29	30-39	40-49	50-59	60-69	70-79	80-84	85-89
Data N^a/N	0.168	0.222	0.192	0.168	0.120	0.089	0.025	0.016
Model N^a/N	0.179	0.177	0.175	0.168	0.148	0.107	0.031	0.016
Labor prod. $\theta^{1,a}$	1.000	1.189	1.237	1.121	0.492	0.000	0.000	0.000
Labor prod. $\theta^{2,a}$	1.299	1.490	1.641	1.705	0.841	0.000	0.000	0.000
Labor prod. $\theta^{3,a}$	1.431	2.388	2.663	2.921	1.286	0.000	0.000	0.000
Prob. $1-\gamma^a$	0.001	0.001	0.004	0.012	0.028	0.042	0.096	0.200
Prob. $1-\omega^a$	0.099	0.099	0.096	0.089	0.074	0.061	0.115	0.000
Factor Ω^a	1.058	1.087	1.122	1.154	1.176	1.278	1.311	1.000
Propens. $1/\Delta^a$	0.056	0.062	0.070	0.083	0.101	0.129	0.181	0.236

Notes: θ^a life-cycle labor productivity determines wage $w^a = w\theta^a$, $1-\gamma^a$ probability of dying, $1-\omega^a$ probability of aging, Ω^a magnification interest factor reflecting increase in mortality, $1/\Delta^a$ marginal propensity to consume.
Data sources: BFS (2001, 2004) and own calculations.

(c) total population grows by

$$N_{t+1} = N_t + N^1_{(t+1),t+t} - \sum_{a=1}^{A} (1-\gamma^a) N^a_t, \quad N_t \equiv \sum_{a=1}^{A} N^a_t. \tag{38}$$

Proof. Cf. Grafenhofer, Jaag, Keuschnigg, and Keuschnigg (2005). ■

References

A , D. (1998): "Why Do New Technologies Complement Skills? Directed Technical Change And Wage Inequality," *Quarterly Journal of Economics*, 113(4), 1055–1089.

A , D., A. J. A , L. J. K , K. A. S , J. W (2001): "Simulating Fundamental Tax Reform in the United States," *American Economic Review*, 91(3), 574–595.

A , A. J., L. J. K (1987): *Dynamic Fiscal Policy*. Cambridge University Press, Cambridge, .

B , G. S. (1962): "Investment in Human Capital: A Theoretical Analysis," *Journal of Political Economy*, 70(5), 9–49.

B -P , Y. (1967): "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 75(4), 352–365.

(2001): *Szenarien zur Bevölkerungsentwicklung der Schweiz 2000-2060*. Bundesamt für Statistik, Neuchâtel.

——— (2004): *Die Schweizerische Arbeitskräfteerhebung ( )*. Bundesamt für Statistik, Neuchâtel.

B , C., R. R. R (1989): "Will the Real Elasticity of Substitution Please Stand Up? (A Comparison of the Allen-Uzawa and Morishima Elasticities," *American Economic Review*, 79(4), 882–888.

B , O. J. (1985): "Debt, Deficits and Finite Horizons," *Journal of Political Economy*, 93(2), 223–247.

B , R., S. B (2000): " Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews*, 19(2), 321–340.

B , R., M. C. D , C. M (2003): "The Impact of Wage Subsidies: A General Equilibrium," *mimeo*.

B¨ -S , A. (2003): "Labor Market Effects of Population Aging," *Review of Labor Economics and Industrial Relations*, 17(Special Issue), 5–44.

B¨ -S , A., A. L , J. W (2004): "Aging, Pension Reform, and Capital Flows," *mimeo, University of Mannheim ( ).*

C , J. C., D. K (1999): "Social Security Reform with Heterogeneous Agents," *Review of Economic Dynamics*, 2(4), 757–795.

C ´ , H., P. P (2003): "The Double Dividend of Postponing Retirement," *International Tax and Public Finance*, 10(4), 419–434.

D , J., C. P , F. P -S (2004): "Capital-Skill Complementarity? Evidence from a Panel of Countries," *Review of Economics and Statistics*, 86(1), 327–344.

D , B., L. L , C. T , M. B. W (1996): "Some Effects of Taxes on Schooling and Training," *American Economic Review*, 86(2), 340–346.

E , L. G., S. E. Z (1989): "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57(4), 937–969.

F , R. E. A. (1990): "Rince Preferences," *Quarterly Journal of Economics*, 105(1), 43–60.

G , O., D. N. W (2000): "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond," *American Economic Review*, 90(4), 806–828.

G , M. (1999): "Government Debt and Social Security in a Life-Cycle Economy," *Carnegie-Rochester Conference Series on Public Policy*, 50, 61–110.

G , C., L. K (1998): "The Origins of Technology-Skill Complementarity," *Quarterly Journal of Economics*, 113(3), 693–732.

G , D., C. J , C. K , M. K (2005): "Probabilistic Aging," *University of St. Gallen Department of Economics Discussion Paper 2005-8*.

G , J., Z. H , G. W. H (1988): "Investment, Capacity Utilization, and the Real Business Cycle," *American Economic Review*, 78(3), 402–417.

G , Z. (1969): "Capital-Skill Complementarity," *Review of Economics and Statistics*, 51(4), 465–468.

G , F. (2003): "Reconciling Conflicting Evidence on the Elasticity of Intertemporal Substitution: A Macroeconomic Perspective," *University of Rochester, Working Paper 491*.

H , R. E. (1988): "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 96(2), 339–357.

H , D. S. (1993): *Labor Demand*. Princeton University Press, Princeton, .

H , I., C. S (1985): "Tax Revenue and the Marginal Cost of Public Funds in Sweden," *Journal of Public Economics*, 27(3), 331–353.

H , F. (1982): "Tobin's Marginal Q and Average Q: A Neoclassical Interpretation," *Econometrica*, 50(1), 213–224.

H , J. J., B. E. H ´ (1990): "The Empirical Content of the Roy Model," *Econometrica*, 58(5), 1121–1149.

H , J. J., L. L , C. T (1998): "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1(1), 1–58.

H , B. J. (1998): "Fiscal Policy Multipliers: The Role of Monopolistic Competition, Scale Economies, and Intertemporal Substitution in Labour Supply," *International Economic Review*, 39(3), 659–696.

İ ˘ , A., S. İ ˘ , D. H. J (1995): "A Life Cycle Analysis of Social Security," *Economic Theory*, 6(1), 83–114.

——— (1999): "Social Security in an Overlapping Generations Economy with Land," *Review of Economic Dynamics*, 2(3), 638–665.

J , G. E. (1997): "Changes in Earnings Inequality: The Role of Demand Shifts," *Journal of Economic Perspectives*, 11(2), 41–54.

K -O , S., H. E. R , D. N. W (2000): "Mortality Decline, Human Capital Investment, and Economic Growth," *Journal of Development Economics*, 62(1), 1–23.

K , T. J. (1994): "College Entry by Blacks Since 1970: The Role of College Costs, Family Background, and the Returns to Education," *Journal of Political Economy*, 105(5), 878–911.

K , C., M. K (2004): "Aging, Labor Markets and Pension Reform in Austria," *Finanzarchiv*, 60(3), 359–392.

K , D. M., E. L. P (1978): "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46(1), 185–200.

K , P., L. E. O , J.-V. R -R , G. L. V (2000): "Capital-Skill Complemetarity and Inequality: A Macroeconomic Analysis," *Econometrica*, 68(5), 1029–1053.

K , F. E., E. C. P (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50(6), 1345–1370.

L , M. J. (2004): "Capital-Skill Complementarity and Inequality over the Business Cycle," *Review of Economic Dynamics*, 7(3), 519–540.

L , R. E. (1990): "Supply-Side Economics: An Analytical Review," *Oxford Economic Papers*, 42(2), 293–316.

M , J. A. (1974): *Schooling, Experience and Earnings*. Columbia University Press, New York.

——— (1997): "The Production of Human Capital and the Life Cycle of Earnings: Variations on a Theme," *Journal of Labor Economics*, 15(1), S26–S47.

P , C., V. C (2005): "Nonlinearities in Capital-Skill Complementarity," *Journal of Economic Growth*, 33(1), 59–89.

R , F. (1928): "A Mathematical Theory of Saving," *Economic Journal*, 38(152), 534–559.

R , J. A. (2005): "Life-Cycle Earnings, Cohort Size Effects and Social Security: A Quantitative Exploration," *Journal of Public Economics*, 89(2-3), 465–485.

R , A. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3(2), 135–146.

S , K. (1967): "A Two-Level Constant-Elasticity-of-Substitution Production Function," *Review of Economic Studies*, 34(2), 201–218.

T , P. A. (1993): "The Effect of Taxation on Human Capital," *Journal of Political Economy*, 101(2), 327–350.

W , P. (1989): "The Equity Premium Puzzle and the Risk-Free Rate Puzzle," *Journal of Monetary Economics*, 24(3), 401–421.

——— (1990): "Nonexpected Utility in Macroeconomics," *Quarterly Journal of Economics*, 105(1), 29–42.

W , Y. (1986): "The Determinants of Life Cycle Earnings: A Survey," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and R. Layard, vol. I, chap. 11, pp. 603–639. Elsevier, Amsterdam.

W , R. J., S. R (1979): "Education and Self-Selection," *Journal of Political Economy*, 87(5), 7–36.