Schooling, learning on-the-job, earnings and inequality

Luís P. Correia

Discussion Paper No. 06/585

May 2006

Department of Economics
University of Bristol
8 Woodland Road
Bristol BS8 1TN

Schooling, learning on-the-job, earnings and inequality*

Luís P. Correia†

Department of Economics, University of Bristol

and Pembroke College, University of Oxford

May 2006

Abstract

Why might people in poor countries leave school earlier and invest less in learning on-the-job than people in rich ones? How do these human capital decisions impact on inequality? To give quantitative answers to these questions, I build an overlapping generations model with optimal human capital accumulation and a given distribution of abilities. Variation in mortality and population growth rates can generate large variability in schooling decisions, earnings profiles and measures of inequality. High mortality and population growth rates are shown to produce comparatively little investment in human capital, flat earnings profiles and low inequality, both within and across cohorts.

JEL classification numbers: I20, J11, J24, O11, O40; *Keywords*: human capital; earnings profiles; schooling; inequality.

*I am extremely grateful to my Ph.D. dissertation adviser, Jonathan Temple, for insightful discussions. All remaining errors are my responsibility. I acknowledge with gratitude the financial support of the Fundação para a Ciência e a Tecnologia, Ministério da Ciência e do Ensino Superior, Portugal.

†E-mail: Luis.Correia@bristol.ac.uk

1 Introduction

Why do people in poor countries leave school a lot earlier than people in rich ones? Why should we expect people in poor countries to invest less in learning on-the-job than people in rich ones? How do these human capital decisions impact on inequality?

To address these questions, we build on Blanchard's (1985) seminal contribution and develop an overlapping generations framework with schooling and learning on-the-job decisions that can be used to illuminate many issues. In this paper, we use the framework to argue that with fast population growth, high mortality rates and high interest rates, there is little incentive to invest in human capital, so people leave school early and their earnings profiles are relatively flat. We also argue that differences in demographic factors, in conjunction with a stable distribution of abilities, can account for large cross-country differences in inequality, both within and across cohorts.

Our learning-on-the job model is a contribution in its own right. We extend the famous Ben-Porath (1967) human capital model to uncertain lifetimes by assuming that mortality rates do not depend on age. As emphasized by Blanchard (1985), an age-invariable mortality rate is able to capture the finite aspect of lives while at the same time making aggregation tractable. But this assumption also has the unfortunate consequence that people of different ages have *exactly* the same lifetime horizon. This led Blanchard (ibid., p. 224) to believe that his formulation was not suitable to capture "the change in behavior over life". Here we build a model of human capital accumulation on-the-job that, despite the assumption of constant mortality rates, generates an age-dependent human capital investment behaviour and a hump-shaped earnings profile like the ones we observe in the data. This is achieved by assuming that, all else equal, an individual's ability to acquire more human capital on-the-job declines with age. This assumption turns out to be well supported by empirical evidence. At the same time, because constancy of mortality rates is preserved, aggregation across cohorts is still tractable, as in Blanchard (1985).

We attempt not only to quantify the extent to which low life expectancy and high population growth rates may explain low investment rates in human capital, but also the impact of these decisions on within-country inequality. Our explanation for inequality is essentially based on demographic factors. It rests on two assumptions and one indisputable fact. We assume, first, that there is a time-invariant distribution of abilities in the population, second, that there are diminishing returns to schooling years but these set in faster for low ability people than for high ability people. As for the fact, it is the large variability in mortality and population growth rates that has been observed both across time and space.

The intuition for our results can be conveyed in a few paragraphs. Our model of the initial schooling choice is essentially the one proposed by Rosen (1977) and extended by Kalemli-Ozcan, Ryder and Weil (2000) to account for uncertain lives: individuals choose schooling to maximize the present value of (expected) lifetime earnings; at the margin, the return to schooling must be equal to the effective discount rate, the latter being increasing in the interest rate and the mortality rate. Once individuals enter the workforce, they may increase their human capital through on-the-job learning. In this respect, the model is, as stated above, an extension of the Ben-Porath (1967) model that incorporates uncertain lifetimes in a tractable manner. Individuals choose how to allocate their time between production and investment in human capital with the objective of maximizing lifetime earnings.

High mortality and population growth rates lead to high effective discount rates, directly in the case of mortality, indirectly for both via their effect on the interest rate. These demographic forces mean that people rationally choose to invest little in formal education as well as in learning on-the-job. Moreover, because of high discount rates, the dispersion in schooling decisions between high ability individuals and low ability individuals is also small. This is turn generates a relatively small dispersion of the human capital of new workers, and since these workers will invest little on-the-job, experience effects on productivity will also be small. So the picture here is of a relatively equal society in which very able and not so able individuals make similar human capital investment decisions.

Consider what would happen if mortality and population growth rates fell. People would stay longer in school and would also accumulate more human capital at work. Moreover, under some plausible assumptions, the more able would react to these changes more strongly than the less able. This in turn would generate larger variability in schooling choices, in the human capital of new workers, and through the compounding effect of learning on-the-job, these effects could significantly increase the dispersion of earnings in society.

In our model, decreases in mortality, population growth and interest rates should cause both income per capita and inequality to rise. This is because the more able benefit more from the same reduction in mortality and interest rates than the less able. Introducing the government in this framework would allow us to analyze the impact of several redistribution policies. For instance, governments that implement compulsory schooling laws might be able to reduce the dispersion of schooling choices at the lower end of the distribution and reduce overall inequality.

The remainder of this paper is organized as follows. In Section 2 we focus on endogenous investment in human capital while on-the-job. In Section 3 we endogenize the schooling decision. We then proceed to aggregate productive human capital and physical capital (Section 4) and to compute the general equilibrium of the economy (Section 5). Next we introduce heterogeneity in ability and we briefly describe the small changes in the model that this modification entails (Section 6). In Section 7, we calibrate this model to the U.S. balanced growth observations of the second half of the twentieth century. We perform simulations in Section 8. Finally, we conclude in Section 9.

2 On-the-job human capital accumulation

Consider a closed economy populated by a continuum of overlapping generations of agents. Each generation is made of many agents who face a constant probability of death per unit of time. Individuals go through two phases in their lives: first, they go to school for a certain number of years, then they leave school voluntarily and work until they die. During this second phase, they have to decide, at any point in time, how much to consume and how much time they should spend acquiring additional human capital or working, taking as given current and future wage and rental rates. Each newborn is endowed with a certain amount of human capital but no financial assets. Individuals are assumed to have no concerns for their descendants after their death. Finally, they have perfect foresight about all economic variables, except for the fact that they do not know their time of death, only the hazard rate.

In this section we abstract from endogenous schooling by assuming that all agents go to school for the same, exogenously given, number of years. We focus instead on endogenous investment in human capital while on-the-job. The point of departure is the Ben-Porath (1967) model as developed by Heckman (1976). Under the latter's formulation, the individual has to decide how much to consume and how to allocate their time between leisure, work and investment in order to maximize lifetime utility. Heckman (1976) avoids the analysis of corner solutions by assuming that the optimal choice is given by interior solutions throughout the individual's life; there is no initial phase of full time education. His analysis therefore applies to the choices made by a student, in high school or college say, that already works part-time.[1] We build on Ben-Porath (1967) in focusing on the labour-investment trade-off, but we follow Heckman (1976), first, in including the choice of consumption and second, in dealing with interior solutions only.

A brief overview of the remainder of this section may prove useful to the reader. First, we set out the formal problem which is solved by the typical agent. We also take some time to justify the modifications we introduce to the Ben-Porath (1967) formulation. For expositional reasons, we find it convenient to place the agent at the beginning of their working life. At each moment in time, the agent has to decide how much to consume and how to allocate their time between work and investment in human capital. Second, we solve for the optimality conditions of the formal problem (Section 2.1). Among other things, we establish that optimality requires that the marginal returns to time allocated to work and investment must be equal at all ages. Next, we turn to the time paths of human assets and labour income (Section 2.2). Provided some reasonable conditions are satisfied, we show that the profiles for human capital, for productive human capital and for labour income will be hump-shaped. We also study how those time paths vary with the interest rate, the mortality rate and the rate of technical change. Finally, we determine the profiles for consumption and financial assets from birth (Section 2.3). As usual, consumption will grow over time provided the real interest rate is greater than the subjective discount rate. On average (since they can die at any point in time), individuals start out with zero financial assets, incur debt at least until they leave formal education, but then eventually their earnings become bigger than their consumption and interest payments on accumulated debt.

Let us then consider the problem faced by an individual, born at time j, who is about to start their working life at time j', after completing s years of formal education.[2] Since they face an uncertain life, they will maximize their *expected* lifetime utility, where the expectation is taken with respect to the distribution of the duration of life. Yaari (1965) showed that it is possible to write this objective function in the same way as the one under certain lives but with an altered discount factor. Briefly stated, an uncertain life makes an individual discount future consumption

[1] In his words, "There is considerable evidence that the "representative" high school and college student works...one could define a schooling period less arbitrarily as one in which hours of work are low..." (ibid., pp. S14-5)

[2] Hence, $j' = j + s$.

more heavily than the typical infinitely lived agent. If ρ is, as usual, the subjective discount rate, m the instantaneous mortality rate, and the individual stands at time j', then the discount factor attached to the utility obtained at time v ($v \geq j'$) will be equal to $e^{-(\rho+m)(v-j')}$. Hence the individual's objective is to maximize

$$\int_{j'}^{\infty} e^{-(\rho+m)(v-j')} \frac{c(j,v)^{1-\theta}-1}{1-\theta} dv, \tag{1}$$

where θ is the inverse of the intertemporal elasticity of substitution, $\theta > 0$; and $c(j,v)$ denotes consumption at time v for an individual born at time j ($j \leq v$). [3]

Let $w(v)$ and r stand, respectively, for the wage per unit of human capital and the riskless rate of interest at time v. We drop the time argument associated with r from the start because we focus exclusively on balanced growth paths, henceforth BGPs. More specifically, we assume the flow $Y(t)$ of final output at t is given by $Y(t) = K(t)^{\alpha}\left[A(t)H_y(t)\right]^{1-\alpha}$, where $\alpha \in (0,1)$; $A(t)$ is an index of labour augmenting technical progress which grows at the rate g; $K(t)$ is physical capital and $H_y(t)$ is the stock of human capital allocated to production (productive human capital for short). In equilibrium, terms $K(t)$ and $A(t)H_y(t)$ will grow at the same rate and both factors will be paid their marginal products, so r will be constant while $w(v)$ will grow at the rate g per unit of time.

Let $h(j,v)$, $y(j,v)$ and $a(j,v)$ stand, respectively, for human capital, labour income, and (real) financial assets; $s_i(j,v)$ for the fraction of time spent investing in human capital; and $h_i(j,v)$ for the number of efficiency units of human capital used in its production, invested human capital for short. By normalizing the time available at each moment to unity, we have that

$$h_i(j,v) = s_i(j,v)h(j,v) \qquad\qquad y(j,v) = w(v)h(j,v)(1-s_i(j,v))$$

The individual accumulates financial assets $a(j,v)$ according to an almost standard equation:

$$\frac{d}{dv}a(j,v) = y(j,v) - c(j,v) + (r+m)a(j,v) \tag{2}$$

The only difference between a standard equation and this one is the real rate of return that individuals can obtain on their financial wealth. Here the rate of return is the sum of the riskless rate of interest and the probability of death. This assumes that individuals have no concerns for their descendants after their death; that perfect markets for loans and annuities secured by life insurance exist; and that the number of individuals of each cohort is so large that there is no aggregate uncertainty regarding the size of surviving cohorts.[4]

The individual is also able to accumulate human capital on-the-job according to

$$\frac{d}{dv}h(j,v) = E(j,v)^{1-b}h_i(j,v)^b - \delta h(j,v) \tag{3}$$

[3] More generally, $x(j,v)$ stands for the value of the variable x at time v for an individual born at time j, in the main text.

[4] For the pioneering paper, see Yaari (1965). For textbook expositions assuming a constant mortality rate see, for instance, Blanchard and Fischer (1989, pp. 115-26) or Barro and Sala-i-Martin (2004, pp. 179-86).

where $E(j, v)$ is an age-varying efficiency term that is beyond the agent's control from the moment they start working onwards; b is a parameter strictly between zero and one to ensure the gross production function is strictly concave with respect to $h_i(j, v)$; and δ is the rate of depreciation of human capital. Notice how, for simplicity, we are abstracting from the direct costs of acquiring human capital: the only cost of investment in human capital is lost labour income.

If we had instead of term $E(j, v)^{1-b}$ a constant, E_{BP} say, then equation (3) would amount to the Ben-Porath (1967) formulation: $dh(j, v)/dv = E_{BP}h_i(j, v)^b - \delta h(j, v)$, where E_{BP} is an efficiency parameter that can be interpreted as a person's ability to learn. It is plausible, however, to allow E_{BP} to vary positively with the level of schooling as for instance Heckman, Lochner and Taber (1998) have done. This is because formal education may not only increase one's starting productivity but also *permanently* increase one's ability to acquire additional human capital on-the-job. The latter effect will arise if education equips a person with learning methods say that are only, or at least more effectively, acquired in school.

We capture this effect of education on learning efficiency and introduce a new one associated with age – hence the arguments j and v – by assuming that the $E(j, v)$ term of equation (3) takes the following form:

$$E(j, v) = \left[\phi h(j, j')\right] e^{-\lambda(v-j')}. \tag{4}$$

where $\phi < 1$ and $\lambda > 0$. The term in equation (4) in square brackets captures the permanent effect of schooling, the exponential term the effect of age. We therefore propose that one's ability to learn, all else equal, declines with age at the exponential rate λ.

There is a great amount of evidence to back the assumption that people's ability declines with age. Studies in psychology, physiology, neurophysiology, neuroscience, gerontology, ageing, motor behaviour and so on discuss and measure through batteries of tests the extent to which people age, both physically and mentally. That physical ability declines with age from early adulthood is beyond dispute. That the same phenomenon happens to cognitive abilities, the ones that are most relevant for modern economies, may be less obvious but no less real. Naturally, the extent to which people's mental abilities deteriorate is not uniform at all across their brains' functions or capabilities. For example, it has been found that verbal fluency seems to increase until the mid-fifties and that its decline from its peak onwards takes place at very mild rates. By contrast, a person's memory, or the ability to perform mathematical computations, to engage in several activities at the same time, to reason quickly or to perform new tasks, all start an inexorable decline from the early twenties onwards. For more on this issue, see for instance Avolio and Waldman (1994).

Summing up, most learning activities from early adulthood onwards have to make use of abilities that, after controlling for other factors such as initial human capital, will inevitably decline with age. From this it follows naturally that, once again taking other factors as given, the efficiency with which a person *acquires* more human capital must decline with age. This is the crucial modification to the standard Ben-Porath (1967) model that allows us to generate realistic hump-shaped profiles despite age-invariant death rates.

The individual's intertemporal problem at time j' is therefore to maximize (1) by choosing profiles for c and s_i subject to the accumulation constraints (2) and (3), to initial levels of financial assets $a(j,j')$ and human capital $h(j,j')$, and to the following standard transversality conditions

$$\lim_{v\to\infty} a(j,v)\Lambda_a(j,v) = 0 \tag{5}$$
$$\lim_{v\to\infty} h(j,v)\Lambda_h(j,v) = 0, \tag{6}$$

where $\Lambda_a(j,v)$ is the shadow price of one unit of financial assets and $\Lambda_h(j,v)$ is the shadow price of one unit of human capital.

2.1 Optimality conditions

As stated in the Introduction, we focus on an interior solution. Letting $\mathcal{H}$ stand for the present-value Hamiltonian function of this control problem,

$$\begin{aligned}\mathcal{H} \;=\; & e^{-(\rho+m)j'}\left(c(j,v)^{1-\theta}-1\right)/(1-\theta)+\Lambda_h(j,v)\left\{E(j,v)^{1-b}\left[s_i(j,v)h(j,v)\right]^b-\delta h(j,v)\right\}\\ & +\Lambda_a(j,v)\left\{w(v)h(j,v)\left(1-s_i(j,v)\right)-c(j,v)+(r+m)a(j,v)\right\},\end{aligned}$$

the necessary and sufficient conditions are[5] **(a)** optimal choice of controls: $\partial\mathcal{H}/\partial c = 0$, and $\partial\mathcal{H}/\partial s_i = 0$; **(b)** flow eqs.: $d\Lambda_a(j,v)/dv = -\partial\mathcal{H}/\partial a(j,v)$ and $d\Lambda_h(j,v)/dv = -\partial\mathcal{H}/\partial h(j,v)$; **(c)** boundary conditions: $a(j,j')$ and $h(j,j')$ given, and (5) and (6).

Regarding the choice of controls, we have for all $v \geq j'$,

$$e^{-(\rho+m)(v-j')}c(j,v)^{-\theta} = \Lambda_a(j,v) \tag{7}$$
$$\Lambda_h(j,v)bE(j,v)^{1-b}h_i(j,v)^{b-1}h(j,v) = \Lambda_a(j,v)w(v)h(j,v). \tag{8}$$

Condition (7) says that, for a person born at time j and alive at time j', consumption at each future point in time should be such that its discounted marginal utility is equal to the marginal utility of income. The only difference in relation to the certainty case is that utility is discounted not only because consumers are assumed to have a positive subjective discount rate but also because they take into account that they may not reach time v. Condition (8) equalizes the value of the marginal return to time invested in human capital (l.h.s.) to the value of the marginal return to time at work (r.h.s.).

Turning to the flow equations, the first is just $d\Lambda_a(j,v)/dv = -(r+m)\Lambda_a(j,v)$, and its solution is immediate: $\Lambda_a(j,v) = \Lambda_a(j,j')e^{-(r+m)(v-j')}$. Using equation (8), the second flow equation simplifies to $d\Lambda_h(j,v)/dv = \delta\Lambda_h(j,v) - w(v)\Lambda_a(j,v)$. Now, if the economy is moving along a BGP, $w(v)$ can always be written as $w(v) = w(j')e^{g(v-j')}$, where g stands, as stated previously, for the instantaneous growth rate of wages which must be equal to the rate of labour

[5] It is well known that the Maximum principle only provides necessary conditions for optimality. But it is easy to show that this problem verifies the so-called Arrow sufficiency condition – see Seierstad and Sydsaeter (1987, pp. 107-8). The same cannot be said of other natural formulations, see McCabe (1983) for details.

augmenting technical progress. Define the following discount rates:

$$\begin{aligned} R &= r+m-g+\delta && (9) \\ \widehat{R} &= r+m-g+\lambda && (10) \end{aligned}$$

Then it is easy to show that the relationship that follows is imposed by the transversality condition for human capital:[6]

$$\frac{\Lambda_h(j,v)}{\Lambda_a(j,v)} = \frac{w(v)}{R} \tag{11}$$

We can attempt to provide an intuitive explanation for this condition. Recalling that $\Lambda_h(j,v)$ and $\Lambda_a(j,v)$ are shadow prices and are measured in units of utility, the l.h.s. of equation (11) gives the value, in utils, of an additional unit of human capital relative to the value of an additional unit of financial assets. As for the r.h.s., it gives the return per period, in real terms, of an additional unit of human capital relative to the return of an additional unit of financial assets. Notice that R takes into account not only that financial assets earn interest $r+m$ but also that human capital depreciates at the rate δ and becomes more valuable over time at the rate g. Hence, the condition given in equation (11) says that optimality requires that relative returns in units of utility must be equal to relative returns in real units.

Moreover, it is also straightforward to show that in addition to (11), the transversality condition for human capital imposes that $R>0$ and $\widehat{R}>0$. Both inequalities amount to lower bounds on r and using the definitions of R and $\widehat{R}$ given in (9) and (10) they may be written as follows.

Condition 1 *Two lower bounds on r: (i) $r>g-m-\delta$ and (ii) $r>g-m-\lambda$.*

Now, substituting (11) into the first order condition (8) yields an equality with an intuitive interpretation:[7]

$$\frac{bE(j,v)^{1-b}h_i(j,v)^{b-1}w(v)h(j,v)}{R} = w(v)h(j,v) \tag{12}$$

The numerator of the l.h.s. of (12) is the product of the human capital generated by the marginal unit of time invested in its production and the wage rate. Dividing this by the discount rate R yields the discounted present value of the infinite stream of future labour income earned with that marginal unit of human capital. The r.h.s. of (12) is simply the additional income generated by the marginal unit of time spent working. Therefore, what equation (12) says is that the two marginal returns to time allocation must be equalized at all ages.

2.2 Human assets and labour income

Solving equation (12) for invested human capital we obtain

$$h_i(j,v) = E(j,v)(b/R)^{\frac{1}{1-b}} = h(j,j')e^{-\lambda(v-j')}s_i(R), \tag{13}$$

[6] All derivations involving more than a couple of steps of algebra are omitted from the main text for brevity. Please see the technical appendix for details.

[7] We believe intuition is better provided if we do not cancel the common term $w(v)h(j,v)$ which appears on both sides of the equation.

where $s_i(R)$ stands for the fraction of time that a new worker invests in human capital and is given by

$$s_i(R) = \phi(b/R)^{\frac{1}{1-b}} \tag{14}$$

From equation (13) we can see that the amount of human capital invested while on-the-job decreases monotonically at the rate λ per period over a person's working life. Also, for a given age, this amount depends positively on a person's starting human capital $h(j, j')$, on ϕ and on b, but negatively on R. Given the definition of R – see equation (9) – we can say that, for a given age, the number of efficiency units of human capital invested depends positively on the growth rate of wages g, but negatively on the interest rate, the mortality rate and the depreciation rate of human capital (r, m and δ).

Substituting equation (13) in equation (3) and solving it gives

$$h(j, v) = h(j, j') \left[e^{-\delta(v-j')} + s_i(R) \tfrac{R}{b(\delta-\lambda)} \left(e^{-\lambda(v-j')} - e^{-\delta(v-j')} \right) \right]. \tag{15}$$

The first term inside the square brackets on the r.h.s. of equation (15) accounts for the fact that human capital depreciates at the rate δ per instant of time while the second term accounts for the effect of investment in human capital. By dividing equation (13) by equation (15) we can determine the share of time which is spent acquiring human capital along a BGP, i.e., $s_i(j, v) = h_i(j, v)/h(j, v)$. Finally, the number of efficiency units of human capital spent producing final output, or for short, productive human capital, is obtained by subtracting $h_i(j, v)$ from $h(j, v)$,

$$h_y(j, v) = h(j, j') \left[e^{-\delta(v-j')} + s_i(R) \left\{ \left(\tfrac{R}{b(\delta-\lambda)} - 1 \right) e^{-\lambda(v-j')} - \tfrac{R}{b(\delta-\lambda)} e^{-\delta(v-j')} \right\} \right], \tag{16}$$

while labour income $y(j, v)$ is simply obtained by multiplying $h_y(j, v)$ by the wage-rate per unit of human capital $w(v)$.

The human capital and the productive human capital profiles given in equations (15) and (16) allow for a wide range of theoretical cases. We now show that by imposing two reasonable conditions on parameters we generate human capital profiles $h(j, v)$ and productive human capital profiles $h_y(j, v)$ that are always hump-shaped.

Condition 2 *Interior solution for $s_i(R)$:* $1 > \phi(b/R)^{\frac{1}{1-b}} > 0$.

Condition 3 $h(j, v)$ *grows at the beginning of an agent's working-life:* $\phi(b/R)^{\frac{b}{1-b}} > \delta$.

Condition 4 $\delta > \lambda > 0$

A few comments regarding these conditions are in order. Notice that the part of Condition 2 which reads $s_i(R) > 0$ is trivially guaranteed since we have assumed above that $R > 0$ – see Condition 1. As for Condition 3, it has been derived by evaluating equation (3) at the beginning of an individual's working-life, when $v = j'$, and then forcing the derived expression to take a positive value. As for the part of Condition 4 which reads $\delta > \lambda$, it is not necessary at all to obtain hump-shaped profiles but imposing it halves the number of cases that we need to examine. Moreover, as will be discussed in Section 7, the empirical evidence clearly supports this assumption.

Finally, imposing a positive rate of depreciation of human capital, $\delta > 0$, is necessary to obtain hump-shaped profiles.

The Ben-Porath (1967) model also produces profiles which are very similar to the ones above. But his model differ from ours in two respects. He has an age-invariant constant E_{BP} instead of our $E(j,v)^{1-b}$ term and he posits that people live with certainty until a given, known age T. With our notation, his solution to $h_i(j,v)$ may be written as follows:

$$h_{i,BP}(j,v) = E_{BP}(1 - e^{R(v-j-T)})^{\frac{1}{1-b}}(b/R)^{\frac{1}{1-b}}, \tag{17}$$

where in this case $R = r - g + \delta$. From equation (17) it is clear that it is the fact that a person's lifetime horizon decreases over time that generates an optimal investment plan in human capital that consists of accumulating most of it at young ages in order to be able to reap its benefits – in this case higher earnings – later on. As a matter of fact, if people could live until very old ages, then the term $e^{R(v-j-T)}$ of (17) would be close to 0 throughout most of a person's life and as a consequence $h_{i,BP}(j,v)$ would be approximately constant throughout the same period.

By contrast, in our framework, individuals face a constant probability of death. A person who has been lucky enough to survive until age 80 say has *exactly* the same life-expectancy than a person who has just left school. Now, if λ were 0, i.e., without declining ability, this old person should behave just like a teenager as far as $h_i(j,v)$ is concerned: from equation (13) we can see that the best they can do is to set aside an age-invariant amount of efficiency units of human capital for investment purposes. But this behaviour would have the following consequence: a person's human capital would either always be increasing or decreasing, depending on whether Condition 3 were verified or not – see equation (15). There would be no way of generating a hump-shaped profile.

There is another difference between the two formulations which is worth pointing out. By including the starting human capital $h(j,j')$ in the efficiency term $E(j,v)$ we derived log-earnings experience profiles which are parallel across schooling levels, as in Mincer (1974). In other words, the annual growth rate of labour income is independent of an individual's education. This implication of the model is desirable on its own since it has been broadly corroborated by the data – see for instance Murphy and Welch (1990, p. 207, fig. 2) for the U.S. case. Had we not included $h(j,j')$ in the $E(j,v)$ term, then the profile for $h_i(j,v)$ would be the same irrespective of a new worker's starting human capital. Because of this, individuals who went to school for more and more years should expect their human capital profiles to become flatter and flatter and ultimately be decreasing in experience.[8]

[8] Note also that the fraction of discretionary time invested in human capital, $s_i(j,v)$, is independent of $h(j,j')$: it starts at $s_i(R)$ and then decreases monotonically to $b(\delta - \lambda)/R$.

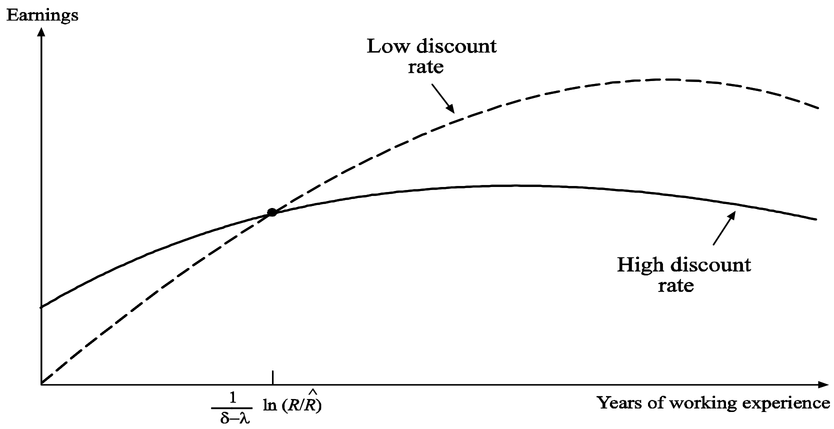

Figure 1: The impact of discount rates on individual earnings.

A case of comparative dynamics of interest to what lies ahead is the following. How do these profiles change with the discount factor R, δ held constant? For simplicity, let agents facing different discount factors start their working lives with the same human capital $h(j,j')$.Using

$$\frac{\partial}{\partial R}\, h_y(j,v)|_{\overline{\delta}} = \frac{h(j,j')s_i(R)}{(1-b)}\left[\frac{e^{-\lambda(v-j')}}{R} - \frac{e^{-\lambda(v-j')} - e^{-\delta(v-j')}}{(\delta-\lambda)}\right],$$

it is straightforward to show that their respective $h_y(j,v)$ profiles can only cross once, at the experience level $\frac{1}{\delta-\lambda}\ln(R/\widehat{R})$. To the left of this point, agents facing high discount factors will have a productive human capital higher than agents facing low discount factors; to the right of $\frac{1}{\delta-\lambda}\ln(R/\widehat{R})$, the opposite will be true. This point is illustrated in Figure 1, where earnings as a function of years of working experience were obtained by multiplying each $h_y(j,v)$ profile by the wage rate per unit of human capital.

2.3 Consumption and financial assets

Substituting the solution for the shadow price of financial assets into the first order condition with respect to consumption gives the solution to its time profile:

$$c(j,v) = c(j,j')e^{\theta^{-1}(r-\rho)(v-j')}, \tag{18}$$

where $c(j,j') = \Lambda_a(j,j')^{-1/\theta}$ is the choice of consumption of the new worker. First, notice that, as first emphasized by Yaari (1965), when life insurance is available, the growth rate of consumption over time is independent of mortality patterns, although the consumption profile under uncertainty may differ markedly from the one under certainty due to different initial consumption choices.

Second, it is well known that with perfect capital markets and no credit constraints, our optimizing problem allows for the complete separation of the consumption/savings decision from investment decisions. This means that the $c(j,j')$ term which appears in equation (18) may in turn be written as being equal to $c(j,j)e^{\theta^{-1}(r-\rho)(j'-j)}$, where $c(j,j)$ is consumption at birth. Hence, the solution to the agent's *lifetime* consumption path can be written as

$$c(j,v) = c(j,j)e^{\theta^{-1}(r-\rho)(v-j)}. \tag{19}$$

What is still unknown in equation (19) is $c(j,j)$. Therefore, we now turn to the determination of this value as well as the profiles for financial assets. During the agent's first s years of life, since they have no labour income, their assets must evolve according to $da(j,v)/dv = -c(j,v) + (r+m)a(j,v)$. The solution for $a(j,v)$, valid for $v \in [j,j']$, is then obtained by making use of equation (19) and condition $a(j,j) = 0$.[9]

On the other hand, by making use of the transversality condition for financial assets – see expression (5) – in the solution of equation (2), valid for $v > j'$, we obtain an equality between the present value of consumption (l.h.s.) and the present value of wealth (r.h.s.), from the perspective of an agent placed at time j':[10]

$$\int_{j'}^{\infty} e^{-(r+m)(x-j')}c(j,x)dx = a(j,j') + \int_{j'}^{\infty} e^{-(r+m)(x-j')}y(j,x)dx. \tag{20}$$

Now, since $a(j,j) = 0$, the value at birth of lifetime labour income is also equal to an agent's wealth, which we denote $\omega_T(j,j)$, $\omega_T(j,j) = \int_{j'}^{\infty} e^{-(r+m)(x-j)}y(j,x)dx$, where the lower limit of integration takes into account that the agent only starts working at time j' ($j+s$). Using $w(x) = w(j)e^{g(x-j)}$ plus the solution for $h_y(j,v)$ given in equation (16), and integrating we obtain

$$\omega_T(j,j) = e^{-(r+m)s}w(j')h(j,j')\left[\frac{1}{R} + \frac{s_i(R)}{\widehat{R}}\frac{1-b}{b}\right]. \tag{21}$$

It is straightforward to see from equation (21) that wealth at birth is increasing in g, but decreasing in r and m. As expected, the detrimental effect that high rates of interest (r), high mortality rates (m) and low rates of technical change (g) have on the incentives to accumulate human capital on-the-job, also causes a reduction in wealth at birth. Using (19), the solution for financial assets when $v = j'$, and (21) in the lifetime constraint (20) gives

$$c(j,j) = [r + m - \theta^{-1}(r-\rho)]\omega_T(j,j), \tag{22}$$

where the following condition must be true:

Condition 5 *Bounded utility:* $r + m > \theta^{-1}(r-\rho)$.

Condition 5 states that the rate of consumption growth cannot be higher than the rate of return obtained on financial assets. If this condition were not verified the consumer would be able to

[9] The solution is $a(j,v) = \frac{c(j,j)}{r+m-\theta^{-1}(r-\rho)}\left[e^{\theta^{-1}(r-\rho)(v-j)} - e^{(r+m)(v-j)}\right]$, $v \in [j,j']$.

[10] The difference between this resource constraint and the standard one is due to uncertain lifetimes. It is easy to show that one can rewrite equation (20) explicitly in terms of the discounted expected values of consumption and income.

achieve unbounded utility in finite time.[11] Notice that the term in square brackets on the r.h.s. of equation (22) is the (marginal) propensity to consume out of wealth. Moreover, although this equation has been set to time j, it is valid for any time t, i.e., $c(j,t) = [r+m-\theta^{-1}(r-\rho)]\omega_T(j,t)$ for $t \geq j$.[12]

Finally, we can use these solutions to get the explicit paths for individual assets $a(j,v)$, before joining the workforce ($v \leq j'$), and after it ($v > j'$). These expressions are necessary when seeking to obtain the aggregate supply of financial resources in the economy.

3 The schooling decision

In this section we make the schooling decision endogenous. Human capital accumulation through formal education may be at least as important as learning on-the-job. To give an example, if the average rate of return to one year of schooling is about 8 percent, a person that spends 12 years in formal education will see their human capital increase approximately 2.5 times. Since this paper focuses on the determinants of human capital accumulation and inequality, it is important to be able to also explain the schooling decision. We do this by drawing on the simple schooling model proposed by Rosen (1977, section 4), as Kalemli-Ozcan et al. (2000) did.[13]

Let us begin the analysis by stating a few assumptions and facts established in Section 2. First, individuals derive utility uniquely from consumption. Second, the growth rate of consumption is equal to $\theta^{-1}(r-\rho)$, while consumption at birth, $c(j,j)$, is proportional to lifetime wealth, $\omega_T(j,j)$, a value which is for convenience rewritten below,

$$\omega_T(j,j) = e^{-\mathrm{R}_s s} w(j) h(j, j+s) \left[\frac{1}{R} + \frac{s_i(R)}{\widehat{R}} \frac{1-b}{b} \right], \tag{23}$$

with j' substituted for $j+s$, $w(j+s)$ for $w(j)e^{gs}$ and where R_s is defined as

$$\mathrm{R}_s = r + m - g.$$

The rate R_s can be considered as the effective discount rate associated with the *schooling* decision, hence the subscript s. If for simplicity we ignore the direct costs of education by assuming they are always zero, then the schooling decision turns out to be a very simple one: s is chosen to maximize the lifetime wealth $\omega_T(j,j)$ which was given above. By maximizing lifetime wealth, the agent is implicitly maximizing consumption at birth and hence attainable consumption (and utility) $c(j,v)$ at each point in time.

So, let the human capital acquired by a person who spent s years in formal education be given by

$$h(j, j+s) = h(j,j)e^{f(s)},$$

[11] Similar conditions are found in the context of infinitely-lived representative agents models (where m equals zero), see for instance Barro (1990, p. S106, eq. 7).

[12] This can be seen by looking at eq. (20): substitute j' for t, denote the r.h.s. by $\omega_T(j,t)$ and solve the integral on the l.h.s.

[13] The Rosen model is also presented in a survey on wage determinants by Willis (1986, section 3.3). For a similar approach which also allows for training, as we do, see Dupor, Lochner, Taber and Wittekind (1996).

with $f'(s) > 0$ and $f''(s) < 0$. The function f measures the total return to s years of education. Notice that a negative $f''(s)$ assumes there are diminishing returns to years of education.[14] A function that captures this is the exponential specification posited by Bils and Klenow (2000) and used in Kalemli-Ozcan et al. (2000): $f(s) = \frac{\pi_1}{1-\pi_2} s^{1-\pi_2}$, with $\pi_1 > 0$ and $\pi_2 \in (0, 1)$. As will be discussed in Section 7, this specification can produce implausible responses of schooling choices to changes in R_s, especially for low levels of R_s. Because of this, we use the following quadratic specification instead, $f(s) = \sigma s - s^2/2\psi$, with $\sigma, \psi > 0$ and $s < \psi\sigma$. The quadratic term $s^2/2\psi$ controls the extent to which returns to education differ from the linear Mincerian specification.

By taking the partial derivative of $\omega_T(j, j)$ with respect to s and equalizing it to zero we get the optimal choice of education: $f'(s) = \mathrm{R}_s$. This condition simply says that optimality requires equalizing the marginal return to schooling (l.h.s.) to the effective discount rate (r.h.s.). Solving explicitly for s gives

$$s^* = \psi(\sigma - \mathrm{R}_s), \tag{24}$$

where we assume $\sigma > \mathrm{R}_s$ and the asterisk is used to denote an optimal choice. The optimal choice of education depends positively on the initial return to schooling, σ, and the rate of technical change, negatively on the interest rate and the mortality rate. Moreover, s^* depends positively on ψ since a higher ψ reduces the extent to which the returns to schooling deviate from the linear Mincerian model.

4 Aggregation

In this section we do the following. First, we sum individual productive human capital across cohorts to obtain the aggregate supply of productive human capital $H_y^s(t)$. We proceed similarly for individual financial assets to obtain the aggregate supply of financial resources in the economy $K^s(t)$. Finally, taking for simplicity the interest rate as exogenous, we briefly discuss two results that can be obtained with the analytical expression for $H_y^s(t)$. The first result says that, under favorable conditions such as low interest rates and low mortality rates, learning on-the-job may more than double the basic productive human capital that results from schooling decisions; the second says that, provided r is greater than the sum of the rate of population growth (n) and the rate of technical change (g), $H_y(t)$ will be decreasing in r but increasing in g.

Let us assume, first, that new individuals are born in each period and second, that the age structure of the population is stable over time. Letting $\widetilde{b}$ ($\widetilde{b} > 0$) stand for the fraction of the population which consists of newborns at each instant of time, and n for the rate of population growth per instant of time, in equilibrium it must be the case that $\widetilde{b} = m + n$.[15].

[14] Evidence for this can be found in Psacharopoulos (1994) or Psacharopoulos and Patrinos (2004).

[15] Let the population at time zero be normalized to 1 and let us assume it has been growing at the rate n ever since. Then the number of newborns at time t will be equal to $\widetilde{b}e^{nt}$. On the other hand, given an age-independent mortality rate, me^{nt} individuals will also die at that instant. Hence the net growth of the population at t must be equal to $(\widetilde{b} - m)e^{nt}$ and this renders $\widetilde{b} - m = n$, as stated in the main text. Notice also that we are allowing for declining populations, but in those cases the assumption $\widetilde{b} > 0$ imposes that $|n| < m$.

Let $L(j,t)$ stand for the number of people born at time j who are still alive at time t and $L(t) = e^{nt}$ stand for the population at time t. Since $L(j,j) = (m+n)e^{nj}$ and since the probability that a person born at time j is still alive at time t is $e^{-m(t-j)}$, we can write

$$L(j,t) = (m+n)e^{-m(t-j)+nj}. \tag{25}$$

Under the assumption that the human capital levels of different workers are perfect substitutes in production, aggregate human capital used for production purposes is given by

$$H_y(t) = \int_{-\infty}^{t-s} h_y(j,t)L(j,t)dj. \tag{26}$$

Now, define the following discount rates,

$$\begin{aligned} \Delta_1 &= m+n+\delta \qquad\qquad \Delta_2 = m+n+\lambda \\ R' &= m+n+g+\theta^{-1}(\rho-r), \end{aligned}$$

term $z(R) = \frac{\phi}{\sigma-\lambda}(b/R)^{\frac{b}{1-b}}$, and for convenience, let us rewrite equation (16) as follows

$$h_y(j,t) = e^{f(s)}\left[\{z(R)-s_i(R)\}\,e^{-\lambda(t-j-s)} + \{1-z(R)\}^{-\delta(t-j-s)}\right], \tag{27}$$

with v substituted for t, j' for $j+s$ and $h(j,j')$ for $e^{f(s)}$. Then $H_y(t)$ can be obtained by substituting equations (25) and (27) into equation (26), and integrating. Aggregating physical capital proceeds along the same lines but it involves more complex expressions since we have to make a distinction between those agents alive at t that are yet to join the workforce and the others. Straightforward algebra produces the following expressions for $H_y(t)$ and $K(t)$:[16]

$$\begin{aligned} H_y(t) &= h(t,t)L(t,t)e^{-(m+n)s+f(s)}\left[\frac{1}{\Delta_1} + \frac{s_i(R)}{\Delta_2}\left(\frac{R}{b\Delta_1}-1\right)\right] && (28) \\ K(t) &= w(t)h(t,t)L(t,t)e^{-(m+n)s+f(s)}\left\{\Psi_1(r)\Psi_2(r) - \Psi_3(r)\right\} && (29) \end{aligned}$$

where $R' > 0$ is necessary in order to guarantee that aggregate physical capital at time t takes a finite value and $\Psi_1(r)$, $\Psi_2(r)$ and $\Psi_3(r)$ are functions of r as well as other parameters and endogenous variables.[17] Since previous conditions guarantee that $\Psi_1(r), \Psi_3(r) > 0$, aggregate physical capital will only be positive if $\Psi_2(r) > 0$. The condition on R' amounts to an upper bound on r. It is also possible to show that when $\delta > \lambda$, then $r > \theta(g-\lambda)+\rho$. We rewrite these two bounds below as two conditions:

[16] We do not use superscripts (s, for supply) in $H_y(t)$ and $K(t)$ since this should not cause any confusion to the reader.

[17] Their expressions are

$$\begin{aligned} \Psi_1(r) &= \left[\frac{1}{R} + \frac{s_i(R)}{\widehat{R}}\frac{1-b}{b}\right] \qquad \Psi_2(r) = \left[e^{(n+g-r)s}\left(\frac{1}{R'} - \frac{1}{n+g-r}\right) + \frac{1}{n+g-r}\right] \\ \Psi_3(r) &= \left[\frac{1}{R\Delta_1} + \frac{s_i(R)}{\Delta_2}\left(\frac{1}{\widehat{R}}\frac{1-b}{b} + \frac{1}{b\Delta_1}\right)\right] \end{aligned}$$

Condition 6 *Upper bound on* r*:* $r < \rho + \theta(m + n + g)$.

Condition 7 *Another lower bound on* r*:* $r > \theta(g - \lambda) + \rho$.

Notice that Condition 6 coincides with $\Psi_2(r) > 0$ whenever $s = 0$, while Condition 7 will in many cases be more stringent than the one imposed by Condition 1 (ii). Before we proceed to determine the equilibrium of this economy, we briefly discuss two partial equilibrium results that are easily obtained by making use of the analytical expression for $H_y(t)$ given above and assuming, for simplicity, that r is exogenously given. The first relates to the importance of learning on-the-job in human capital formation. Under favorable circumstances (e.g., low rates of interest and low mortality rates), learning on-the-job may double or more than double the basic productive human capital that results from schooling decisions; while in less favorable circumstances, it may play a relatively small role in accounting for the level of $H_y(t)$.[18]

The second result is obtained by differentiating $H_y(t)$ with respect to $(r - g)$. First notice, that everything else equal, the choice of schooling years that maximizes society's $H_y(t)$ is given by $f'(s) = m+n$. Provided the interest rate is above $n+g$, the optimal schooling choice made by each individual, s^*, will always lie to the left of that maximum. But since s^* depends negatively on $r - g$ (see Section 3), this means that increases in the interest rate r or decreases in the rate of technical change g will reduce $H_y(t)$ via lower schooling choices. Moreover, and again as long as $r > n + g$, this reduction in $H_y(t)$ will be made more severe by lower investments in learning on-the-job as well.

5 Equilibrium

This section is structured as follows. We begin by solving for the demand for factors $H_y^d(t)$ and $K^d(t)$ by the typical firm. Then we proceed to compute the steady-state general equilibrium of this economy. We define equilibrium as an interest rate r^*; an allocation $\{K^d(t),\, H_y^d(t), Y(t)\}$ for the typical firm; a set of allocations $\{c(j,t), a(j,t)\}$ for the typical individual not in the labour force, $j \in (t\text{-}s, t)$; and a set of allocations $\{c(j,t), s_i(j,t), a(j,t), h_y(j,t)\}$ for the typical worker, $j \leq t\text{-}s$ such that, for all time t: **(a)** firms maximize profits given prices; **(b)** individuals maximize expected lifetime utility subject to human capital and financial assets constraints at given prices; **(c)** the schooling choice s and the population structure are constant; **(d)** all markets clear, $H_y^d(t) = H_y^s(t)$, $K^d(t) = K^s(t)$, $C(t) + I(t) = Y(t)$.

But since the equation that determines the equilibrium interest rate is highly nonlinear and can in principle have many solutions, we next develop a diagrammatic exposition that shows that equilibrium is unique for reasonable parameter configurations and can be seen as the intersection point of two curves, as in the phase diagrams of the Ramsey-Cass-Koopmans model (henceforth simply Ramsey model) and the Blanchard model. Then, to gain intuition for general equilibrium

[18] The effect of learning on-the-job is captured by term $\frac{s_i(R)}{\Delta_2}\left(\frac{R}{b\Delta_1} - 1\right)$ on the r.h.s. of equation (28). For example, when $m + n = .02$; $\delta = .04$; $\lambda = .01$; $R = .1$; $b = .5$; $s_i(R) = .25$, this term is 19.4 whereas the term $1/\Delta_1$ associated with schooling is 16.6; increasing $m+n$ to .04 and decreasing $s_i(R)$ to .20 reduces the learning on-the-job term to 6 and the $1/\Delta_1$ term to 12.5.

effects and comparative statics results, we present two special cases of it. In Section 5.1 we focus on the impact of n, m, g, ρ and θ on r^* when there is no endogenous human capital accumulation (the Blanchard model). In Section 5.2, first, we introduce schooling effects, by making the schooling choice s exogenous (basic extension of the Blanchard model), then by making s endogenous (the Kalemli-Ozcan-Ryder-Weil model, henceforth KRW model); second, we introduce learning on-the-job effects.

As mentioned in Section 2, the flow of final output at t is produced with a Cobb-Douglas production function, $Y(t) = K(t)^{\alpha} [A(t)H_y(t)]^{1-\alpha}$, where we recall that $A(t)$ is an index of labour augmenting technical progress which grows at rate g; $K(t)$ is physical capital, which must be in equilibrium equal to aggregate financial assets; and $H_y(t)$ is the economy's stock of human capital that is allocated to production.

Suppose the economy is moving along a BGP. Then each firm pays $w(t)$ per unit of human capital and $r + \delta_k$ per unit of rented physical capital, where r stands for the riskless market rate of interest and δ_k is the rate of depreciation of physical capital. Since the typical firm faces no adjustment costs, and given that we assume it operates under perfect competition, it will maximize its profits by equalizing at each point in time the marginal rate of technical substitution between $K(t)$ and $H_y(t)$ to the ratio of their prices. This condition can be written as

$$\alpha w(t)H_y(t) = (1 - \alpha)(r + \delta_k)K(t)$$

By substituting equations (28) and (29) into the previous expression, we obtain an equation in r only. This equation is not particularly instructive and must, in general, be solved numerically. This is why we now turn to a diagrammatic exposition.[19] Let $\widehat{x}(t)$ denote aggregate variable $X(t)$ expressed in efficiency units of productive human capital, i.e., $\widehat{x}(t) = X(t)/ [A(t)H_y(t)]$. Equilibrium may then be obtained as the intersection point of two curves on the space $(\widehat{k}, \widehat{c})$, as in the Ramsey model and the Blanchard model. There is a major difference though: the expressions that follow are only valid along a BGP and should *not* be seen as isoclines of a phase diagram on the space $(\widehat{k}, \widehat{c})$, since $H_y(t)$ is endogenous. [20]

First, we have the economy's resource constraint, which says that (gross) output must be equal to consumption plus (gross) investment. In the steady state, $\dot{\widehat{k}} = 0$ and this implies the following relationship between $\widehat{c}$ and $\widehat{k}$:

$$\widehat{c} = \widehat{k}^{\alpha} - (n + g + \delta_k)\widehat{k} = [(r + \delta_k)/\alpha - (n + g + \delta_k)]\, \widehat{k}. \tag{30}$$

This constraint is the same as the one found in the Ramsey model: on the space $(\widehat{k}, \widehat{c})$, the

[19] Notice that an equilibrium with interior solutions, reasonable earnings profiles, etc., is not guaranteed for unrestricted configurations of parameters. For example, we cannot be sure whether or not $s_i(R)$, the time spent acquiring human capital, lies in the unit interval. This issue arises in many similar models. See for instance Benhabib and Perli (1994) for a study of the restrictions on parameter values that deliver a unique interior solution in the Lucas (1988) model.

[20] When one assumes that human capital accumulation decisions, from schooling to the fraction of time used to learn more on-the-job, are exogenously determined, by the government say, then the expressions given in main text can be interpreted as isoclines.

curve defined by $\dot{\widehat{k}} = 0$ is concave, starts at 0 and reaches a maximum when $f'(\widehat{k}) = n + g + \delta_k$ (golden rule). So what makes our model differ from the Ramsey model must be solely attributed to the $\dot{\widehat{c}} = 0$ constraint, to which we now turn. Differentiating aggregate consumption, $C(t) = \int_{-\infty}^{t} c(j,t)L(j,t)dj$, with respect to time gives:

$$\dot{C}(t) = \theta^{-1}(r-\rho)C(t) - mC(t) + c(t,t)L(t,t), \tag{31}$$

The first term on the r.h.s. of equation (31) captures the fact that individual consumption for those already alive grows at the rate $\theta^{-1}(r-\rho)$ per unit of time; the second that a fraction m of the population dies per unit of time; the third that $L(t,t)$ agents are newly born at every instant of time and each one of them consumes $c(t,t)$. It can be shown that the term $c(t,t)L(t,t)$ can be expressed as a function of aggregate variables $C(t)$ and $K(t)$. Using this result in the last equation and writing it in efficiency units, one obtains another relationship between $\widehat{c}$ and $\widehat{k}$ along the BGP (where $\dot{\widehat{c}} = 0$):

$$\widehat{c} = \frac{\left[r + m - \theta^{-1}(r-\rho)\right]\widehat{k}}{1 - R'\left[e^{(r-n-g)s}\Psi_3(r)\Psi_1(r)^{-1} + (e^{(r-n-g)s} - 1)/(r-n-g)\right]} \tag{32}$$

Recall that s is a one-to-one function of r and r is, in turn, a one-to-one function of $\widehat{k}$, so $\widehat{c}$ in equation (32) is a function of $\widehat{k}$ only.[21] Equilibrium on the space $(\widehat{k}, \widehat{c})$ is given by the intersection of the two curves defined by equations (30) and (32). We now proceed to discuss special cases of this general model in order to gain intuition for its general equilibrium properties.

5.1 The Blanchard model

We start with the seminal Blanchard (1985) model, where it is assumed there is no investment in human capital and the population consists only of workers, so $H_y(t) = H(t)$. For expositional reasons, we set $\delta = 0$. In this special case, equation (32) simplifies to[22]

$$\widehat{c} = \frac{\left(r + m - \theta^{-1}(r-\rho)\right)}{1 - R'/\Delta_1}\widehat{k} = \frac{\left(r + m - \theta^{-1}(r-\rho)\right)}{1 - R'/(m+n)}\widehat{k}. \tag{33}$$

Figure 2 illustrates on the space $(\widehat{k}, \widehat{c})$ how the Blanchard equilibrium differs from the Ramsey one (with zero population growth). Equilibrium in the Ramsey model, at point E_R, is given by the intersection of the concave $\dot{\widehat{k}} = 0$ curve with a vertical $\dot{\widehat{c}}_R = 0$ curve. By contrast, when $\theta \geq 1$ (the relevant case according to the empirical evidence, see Section 7), the $\dot{\widehat{c}}_B = 0$ curve of the Blanchard model is convex, starting at the origin and then gradually approaching the vertical asymptote that corresponds to the higher bound for $\widehat{k}$. However, only points which lie to the right of the lower bound for $\widehat{k}$ ($\underline{\widehat{k}}$ in the figure) should be considered as candidates for equilibrium. In Figure 2, the Blanchard equilibrium is represented by point E_B.

[21] We do not make this dependence explicit in what follows to avoid cumbersome notation.

[22] Set to zero the following values: $s_i(R)$, λ, and s.

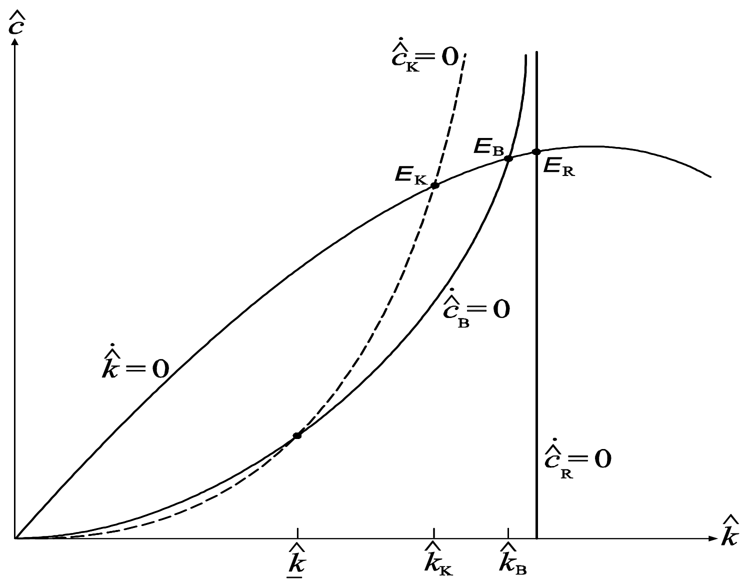

Figure 2: The impact of the arrival of new families (point E_B) and students (point E_K) on the Ramsey equilibrium (point E_R).

As for comparative statics, inspection of equations (30) and (33) shows that increases in n, and g shift the $\dot{\widehat{k}} = 0$ curve downward; while increases in n, g, m, ρ and θ shift the $\dot{\widehat{c}}_B = 0$ curve upward. Hence, the equilibrium interest rate in the Blanchard model depends positively on the following parameters: n, m, g, ρ and θ.

The way parameters g, ρ, θ impact on the equilibrium interest rate is qualitatively similar in the Ramsey model and the Blanchard model. This statement can be made more precise once the equation of motion for $\widehat{c}$ of the Blanchard model is written explicitly:

$$\dot{\widehat{c}} = \left(\theta^{-1}(r-\rho) - g\right)\widehat{c} - (m+n)\left(r + m - \theta^{-1}(r-\rho)\right)\widehat{k} \tag{34}$$

The rate of arrival of consumers per unit of time is equal to the birth rate $m+n$. As this arrival rate tends to zero, the Blanchard model converges to a Ramsey model since the last equation then simplifies to $\dot{\widehat{c}} = \left[\theta^{-1}(r-\rho) - g\right]\widehat{c}$.[23] Therefore, for very low birth rates, the two models produce essentially the same results.

As the arrival rate of consumers moves away from zero, so grows the importance of mortality and population growth rates in the determination of the Blanchard equilibrium. This can be seen by looking at the $\widehat{k}$-term on the r.h.s. of equation (34). On the one hand, new consumers arrive in the economy at the rate $m+n$ per unit of time. On the other hand, the marginal propensity to consume out of wealth for any individual, including newborns, is given by $r + m - \theta^{-1}(r-\rho)$. So what this $\widehat{k}$-term captures is the depressing effect that these new consumers, born without any financial wealth, have on the growth of $\widehat{c}$. This is why, for a given capital stock $\widehat{k}$, $\dot{\widehat{c}}_B = 0$ requires

[23] Recall that we have shown above that the equation of motion for $\widehat{k}$ (the economy's resource constraint) is identical in the two models. The insight that what is crucial for this equivalence is that $m+n \to 0$ is due to Weil (1989).

consumption per effective worker to increase with mortality and population growth rates: a higher m increases both the arrival rate of new consumers and their propensity to consume; a higher n increases the arrival rate of new consumers. Both effects shift the $\dot{\widehat{c}}_B = 0$ curve upward.[24]

5.2 Schooling and learning on-the-job effects

To gain intuition for the effects of schooling, consider first what might be called a KRW model with *exogenous* schooling. That is, relax the assumption that the economy is inhabited by workers only and assume instead that s is positive, although exogenously determined. Now, term R'/Δ_1 of equation (33) should be substituted by a new term that turns out to be increasing in s. The qualitative impact on equilibrium of exogenous schooling is also illustrated in figure 2. Curve $\dot{\widehat{c}}_K = 0$ of this KRW model with exogenous schooling lies, for all admissible values for $\widehat{k}$, everywhere above the $\dot{\widehat{c}}_B = 0$ curve of the Blanchard model. This is what we should expect: for a given capital per effective worker, $\dot{\widehat{c}} = 0$ now requires consumption per effective worker to increase with the fraction of people not in the labour force since students, by drawing their consumption from existing resources, exert downward pressure on the growth of $\widehat{c}$.[25] Equilibrium in this KRW model with exogenous schooling is represented by point E_K in figure 2.

What happens when schooling is endogenous? We have just shown that the equilibrium interest rate in a KRW model with exogenous schooling is an increasing function of s. Let us denote this relationship by $r_1(s)$. On the other hand, optimality in the KRW model of schooling choice requires equalization of the marginal return to schooling and the effective discount rate, .i.e., $f'(s) =$R$_s$.[26] But this equation makes the optimal choice of schooling years a negative function of the interest rate. Denoting the inverse of this last relationship by $r_2(s)$, one can see that the general equilibrium of the KRW model can be represented by the point $\{s^*, r^*\}$ where the two curves $r_1(s)$ and $r_2(s)$ cross.

Figures 3 to 5 show three cases of comparative statics in the KRW model which are also relevant to our more complex model.[27] For all of them, we have used the quadratic specification for $f(s)$ that we proposed in Section 3, i.e., $f(s) = \sigma s - s^2/2\psi$. Hence $r_2(s)$ is simply a line, $r_2(s) = \sigma + g - m - s/\psi$, that shifts rightward as the learning technology improves (higher σ's and ψ's), as technical progress accelerates, and as mortality declines.

Figure 3 illustrates the impact of a higher mortality rate on the long run equilibrium in the KRW model. Initially the economy is at point E. A higher m has a direct negative impact on each individual's choice of schooling since it directly affects their effective discount rate. This is represented by a downward shift of curve $r_2(s)$. But, for a given rate of population growth, a higher m must translate into a higher birth rate (recall that $b = m + n$) and a younger population.

[24] Notice, however, that it is not true that changes in m must have a larger impact on r^* than changes in n: although an increase in m has a double upward impact on the $\dot{\widehat{c}}_B = 0$ curve, it does not change the $\dot{\widehat{k}} = 0$ curve, whereas an increase in n has a downward impact on the $\dot{\widehat{k}} = 0$ curve as well as an upward impact on the $\dot{\widehat{c}}_B = 0$ curve.

[25] However, the difference between term R'/Δ_1 and the new term tends to zero as r approaches its upper bound, i.e., as $R' \to 0$ (this is equivalent to $\widehat{k} \to \underline{\widehat{k}}$).

[26] See Section 3.

[27] Kalemli-Ozcan et al. (2000) do not present a diagrammatic exposition of their model but they do provide an algebraic proof of the existence and uniqueness of the steady-state when $f(s) = \ln(s)$ (ibid, Appendix A, pp. 19-20).

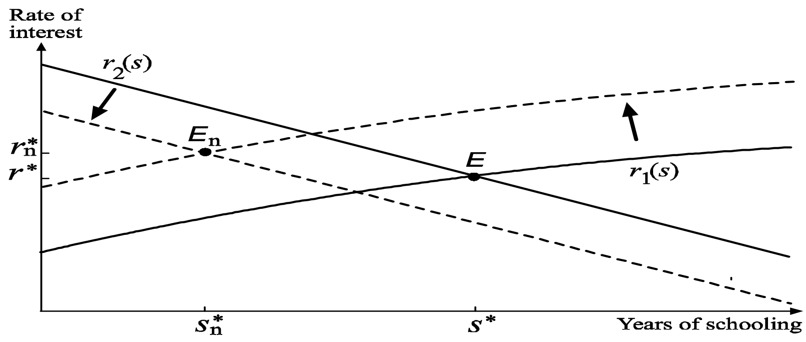

Figure 3: The direct and indirect impact of a higher mortality rate on the KRW equilibrium.

For a given s, the effect of these additional arrivals of young people in the economy is to put upward pressure on the interest rate. This is represented by the upward shift of curve $r_1(s)$. The economy will ultimately converge to the new equilibrium point E_n, with lower schooling than before ($s_n^* < s^*$). However, in theory at least, one cannot say whether or not the new rate of interest lies above (as in the figure) or below the old one.

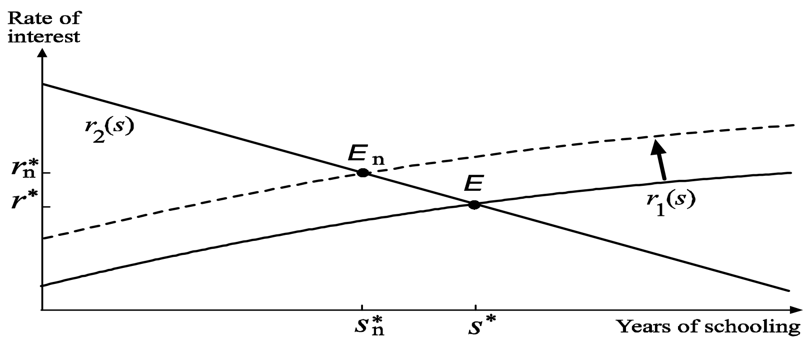

Figure 4: The indirect impact of a higher population growth rate on the KRW equilibrium.

Figure 4 illustrates the impact of a higher population growth rate. A higher n does not have a direct impact on schooling choices, hence $r_2(s)$ does not shift. The negative impact of a higher n on schooling is entirely accounted for by its upward pressure on the economy's interest rate. In the new equilibrium, schooling years will be lower than before because the equilibrium rate of interest has gone up.

Figure 5 illustrates the impact of an improvement in the learning technology as captured by an increase in σ. In the new equilibrium, both schooling years and the rate of interest will be bigger than before.

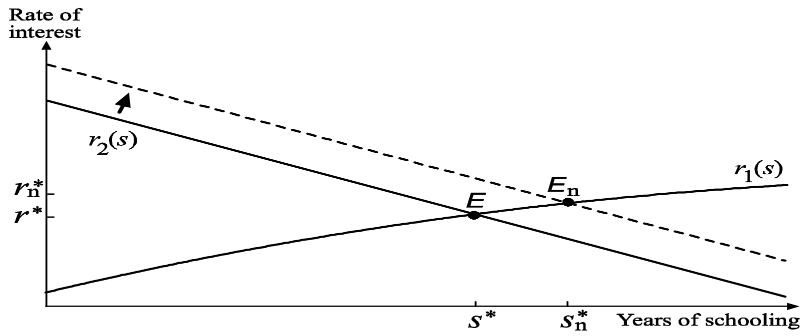

Figure 5: The positive direct effect of a better learning technology on the KRW equilibrium.

We now turn to endogenous learning on-the-job effects. Assume the population consists entirely of workers, so that $s = 0$. It is possible to show that equilibrium in this economy will be unique for all reasonable parameter configurations by using a similar diagram to the ones given in figures 3 to 5. Notice that the crucial human capital decision in this case is $s_i(R)$, the fraction of time that a new worker invests in human capital. Now, on the one hand, r^* is an increasing function of the fraction of exogenously given investment $s_i(R)$ because of the following: when investment on-the-job increases, there is an additional fraction of existing human capital that is not producing but still consuming; this puts downward pressure on the growth of consumption per effective worker, so now the $\dot{\widehat{c}} = 0$ curve will lie, once again for all admissible values for $\widehat{k}$, everywhere above the $\dot{\widehat{c}}_B = 0$ of the simple Blanchard model. On the other hand, the optimal choice of $s_i(R)$ is a negative function of the interest rate. So equilibrium is given by the intersection point of these two relationships on the space $\{s_i(R), r\}$. Moreover, a higher m will have both a direct and an indirect negative impact on on the optimal choice of $s_i(R)$, whereas a higher n will only have an indirect negative impact on it.

Finally, we summarize this discussion by briefly commenting on a few characteristics of the model with endogenous schooling and learning on-the-job. First, the ways mortality rates and population growth rates impact on equilibrium in this model and the Blanchard one are very similar. Increases in these rates translate into higher arrival rates of consumers in the economy per unit of time. These new consumers, born without any financial wealth, will use resources that could have been used to increase capital per effective worker. Moreover, the lower the lifetime horizon (high mortality rates), the stronger this depressing effect on capital per effective worker, since new consumers will have higher marginal propensities to consume. Second, increases in schooling and/or learning on-the-job following an (exogenous) increase in the efficiency of learning will be partially crowded out by higher interest rates.

6 Heterogeneous ability

While the importance of the number of years of formal education and working experience as predictors of earnings was the focus of Mincer (1974, ch. 5), there is now a voluminous literature that extends his human capital earnings function to take into account other factors. The reasons for this are well known: individuals that are apparently equivalent in terms of education, experience, occupation may earn very different amounts. It is also widely agreed that this residual variability in earnings may to some extent be explained by differences in cognitive skills (e.g., IQ or aptitude test scores) and family background (e.g., education of a parent or sibling). Here, we aggregate these differences into one index which we call ability and we take differences in this ability index as being exogenously determined. We also assume that individuals learn their ability at birth.

Let a person's ability ε be a random draw from a time-invariant distribution $G_\varepsilon(\varepsilon)$ with support $[\varepsilon_L, \varepsilon_H]$ on $\mathbb{R}$ and probability density function $g_\varepsilon(\varepsilon)$. Regarding the impact of ability on schooling, we assume that ψ is a positive function of ability, i.e., $f(s,\varepsilon) = \sigma s - s^2/2\psi(\varepsilon)$, with $\psi'(\varepsilon) > 0$. This gives $s^*(\varepsilon) = \psi(\varepsilon)(\sigma - \mathrm{R}_s)$. This formulation implies that the change in s^* in response to a change in the effective discount rate R_s is increasing in ability while at the same time the elasticity of s^* to R_s is invariant to ability: if R_s increases by 1 percent, the schooling choice of both low and high ability people will decrease by $-\mathrm{R}_s/(\sigma - \mathrm{R}_s)$ percent. For transparency, we rewrite the crucial condition below:

Condition 8 *Spence-Mirrlees condition:* $\psi'(\varepsilon) > 0$.

The name given to the condition reflects the fact that the marginal return to education is increasing in ability ($\partial^2 f(s,\varepsilon)/\partial s \partial \varepsilon > 0$). We posit the simplest possible function for $\psi(\varepsilon)$, $\psi(\varepsilon) = \psi\varepsilon$, with $\psi > 0$. With this specification, optimal schooling turns out to be linear in ability:

$$s^*(\varepsilon) = \psi(\sigma - \mathrm{R}_s)\varepsilon \tag{35}$$

Moreover, the total return to $s^*(\varepsilon)$ years of education is also linear in ability, $f(s^*(\varepsilon)) = \frac{\psi}{2}(\sigma^2 - \mathrm{R}_s^2)\varepsilon$. So the starting human capital of the new worker with ability ε, $h(j, j + s^*(\varepsilon)) = h(j,j)e^{f(s^*(\varepsilon))}$, may be seen as a r. v., denoted H^* for short, which is a monotonically increasing function of the r. v. ε. Likewise, wealth at birth for an individual with ability ε may be seen as a r. v., denoted $\boldsymbol{\omega}_T^*$ say, which is a function of the r. v. ε. It is easy to show that we can obtain expressions for the densities of H^* and $\boldsymbol{\omega}_T^*$ from the density of ε.

We now turn to the changes to the model that result from heterogeneity. In a nutshell, they are minimal. First, the analyses of Sections 2 and 3 should be seen as applying to a type-ε individual. To reflect this, a new argument should be added to several variables, such as in $h(j, v, \varepsilon)$, $h_y(j, v, \varepsilon)$ or $a(j, v, \varepsilon)$. For example,

$$h_y(j,v,\varepsilon) = \\ h(j,j)e^{f(s^*(\varepsilon))}\left[e^{-\delta(v-j'(\varepsilon))} + s_i(R)\left\{\left(\frac{R}{b(\delta-\lambda)} - 1\right)e^{-\lambda(v-j'(\varepsilon))} - \frac{R}{b(\delta-\lambda)}e^{-\delta(v-j'(\varepsilon))}\right\}\right],$$

with $j'(\varepsilon) = j + s^*(\varepsilon)$. Second, aggregation must be done across ages *and* ability levels. Since we have assumed the distribution $G_\varepsilon(\varepsilon)$ and therefore its density $g_\varepsilon(\varepsilon)$ are stable over time, the number of type-ε individuals born at time j who are still alive at time time t is given by

$L(j,t,\varepsilon) = L(j,t)g_\varepsilon(\varepsilon)$. So, for instance, aggregate productive human capital is

$$H_y(t) = \int_{\varepsilon_L}^{\varepsilon_H} g_\varepsilon(\varepsilon) \int_{-\infty}^{t-s^*(\varepsilon)} h_y(j,t,\varepsilon) L(j,t) dj d\varepsilon,$$

where the inside integral is identical to the $H_y(t)$ of equation (28), except that s should be substituted by $s^*(\varepsilon)$.

7 Calibration

In this section, we calibrate the model with heterogeneous agents for the U.S. growth observations of the second half of the twentieth century. Tables 1 to 3 give the values chosen for each parameter: Table 1 presents general/macro parameters; Table 2 focuses on the parameters associated with learning on-the-job; while Table 3 covers the parameters associated with the schooling decision as well as with heterogeneous ability.[28]

Turning to Table 1, we rely on the King and Rebelo (1999, pp. 953-4) calibration of the basic real business cycle model for the parameters r, g, α, and δ_k. There is no consensus as to the appropriate value for θ. Surveying many microeconomic studies, Browning, Hansen and Heckman (1999, p. 552) conclude that if constancy of this elasticity is imposed across the population, then there is no strong evidence against the view that θ is slightly above 1. We therefore use logarithmic utility ($\theta = 1$) and $\theta = 2$.[29] As for ρ, the subjective discount rate, it is a free parameter. It is the last value to be set and it is chosen to ensure the model produces the interest rate given above. The two values given in the table correspond to the two values for θ.

Parameter	**Description**	**Value**
r	interest rate	$.065$
n	population growth rate	$.012$
m	mortality rate	$1/70$
g	rate of technical change	$.016$
θ	inverse of the elasticity of substitution	$1; 2$
ρ	subjective discount rate	$.0335; .002$
α	share of income received by physical capital	$1/3$
δ_k	depreciation rate of physical capital	$.1$

Table 1: Calibration - General Parameters

As for demographics, the U.S. population was about 157.8 million in 1950 and 285 million in 2000. Assuming continuous compounding, n solves $285 = 157.8e^{50n}$. The other parameter value that we have to determine is the mortality rate. Given that a constant mortality rate implies a life expectancy of $1/m$, we can use our knowledge of the latter to calculate the former. The

[28] A more detailed discussion of the calibration procedure is available from the author upon request.

[29] The value 2 has been used several times as the benchmark case – see Trostel (1993, p. 336, note 10) for references.

chosen value reflects the fact that our model ignores pre-schooling years.[30]

Choosing the parameter values of the human capital production function – see Table 2 – requires some judgement. Firstly, there are no estimates of our model since it includes an age effect through λ which has never been taken into account as well as permanent effect of schooling through $\phi h(j, j')$ – see equation (4). There are a few estimates though of the simpler specification $dh(j, v)/dv = E_{BP} h_i(j, v)^b - \delta h(j, v)$ in a partial equilibrium setting which also estimate the interest rate – see Heckman (1976, p. S33, Table 1), Haley (1976, p. 1233, Table III) and the U.S. Bureau of the Census 1960 study reported in Browning et al. (1999, p. 585, Table 2.3). Secondly, the chosen parameterization must allow the discount rate R to change within a reasonable range without generating implausible human capital profiles. A value for b of $.5$ accomplishes this while still being very close to the lower estimate for b found in Haley (1976, p. 1233, Table III) as well as the estimate of the U.S. Bureau of the Census as reported by Browning et al. (1999, p. 585, Table 2.3). The value for δ is set as in Trostel (1993).

Parameter	**Description**	**Value**
b	captures the extent of diminishing returns to $h_i(j, v)$	$.5$
δ	depreciation rate of human capital	$.04$
λ	rate that captures the effect of ageing on learning ability	$.01$
ϕ	efficiency parameter associated with learning	$.01497$

Table 2: Calibration - Learning on-the-job Parameters

As for λ, we draw on the work of Avolio and Waldman (1994) on variations on ability across the working life span that may be related to factors such as race, education and occupational type. Avolio and Waldman (1994) present in a table and figures the mean General Aptitude Test Battery scores (collected from 1970 to 1984) on nine ability factors for black, hispanic and white Americans by six age groupings. This produces twenty-seven age ability profiles. The major conclusion that we can draw from their study is that ability factors start declining from the early twenties onwards but they do so at mild rates. Using the above mentioned profiles, we obtained 1 percent per year as a rough estimate for λ.

The only variable that remains to be determined in order to produce human capital profiles is ϕ. We draw on estimates by Krueger and Pischke (1992) that use the March 1989 Current Population Survey (CPS) to obtain a value for ϕ for two reasons. First, their estimates produce profiles which are broadly similar to the actual profiles calculated by Murphy and Welch (1990) using CPS's from 1964 to 1988, second, Krueger and Pischke (1992) also produce estimates of the return to schooling, a number which we also make use of in this exercise.

These authors estimate a standard Mincerian earnings regression of the type $\ln y = \alpha_1 + \alpha_2 s + \gamma_1 x + \gamma_2^2 x + \eta$, where y is weekly earnings, s is the number of years in full-time education, x is the number of potential years of working experience, and η is an error term.[31] Their point

[30] Life expectancy at birth for the U.S. was 68.2 years in 1950 and 77 years in 2000. The round number we have chosen is approximately equal to life expectancy at ages 5 or 6 for the period from the mid-80s onwards – see Arias (2004, p. 22, Table 12).

[31] They also add a dummy indicating gender but that is immaterial for this discussion (ibid., Table 2a, column 5).

estimates are $\alpha_1 = .093$, $\gamma_1 = .032$ and $\gamma_2 = -.00048$. A value for ϕ equal to $.01497$ was found to be the one which produced both a peak age and peak value closer to those implied by the quadratic estimates of Krueger and Pischke (1992).[32]

Turning to the education parameters – see Table 3. As an estimate of the average number of school years of the typical American we use, for consistency, the median value for the year 1989, the same year used to estimate the experience earnings profile, hence $\overline{s} = 12.7$.[33] As for ψ and σ, first, let $\overline{\varepsilon}$ stand for the average ability of the American population and set it, without any loss of generality, to unity. Then if we take the $\overline{s}$ schooling years given above as resulting from equalizing $f'(s)$ to R_s, as shown in Section 3, we must have $\sigma - \overline{s}^*/\psi = \mathrm{R}_s$. This amounts to one equation in two unknowns, ψ and σ. The way we proceeded was to draw on the estimates of Bils and Klenow (2000) to obtain a reasonable value for ψ and then we used it in the last equation to get the value of σ.

Parameter	**Description**	**Value**
σ	rate of return to the first schooling year	$.1264$
ψ	captures the extent to which returns to edu. decrease with y. of school.	200
$\overline{s}$	average years of school completed by a worker	12.7
$\overline{\varepsilon}$	average ability	1
ε_L	lowest ability	$.5$
ε_H	highest ability	1.5
κ_1	coefficient of ε^2 in the density function of ability $g(\varepsilon)$	-6
κ_2	coefficient of ε in $g(\varepsilon)$	12
κ_3	constant term in $g(\varepsilon)$	-4.5

Table 3: Calibration - Schooling and Ability Parameters

Figure 6 represents the same optimal schooling choice under our specification and the Bils and Klenow (2000) one. The horizontal line gives the value of R_s; the black downward sloping line shows the relationship between the number of schooling years and the marginal return to schooling implied by our specification, $f'(s) = .1264 - s/200$; while the dashed curve shows the one implied by a Bils and Klenow (2000) specification, $f'(s) = .276s^{-.58}$. It should be obvious from the graph that, at low levels of R_s, the second specification implies schooling responses to changes in R_s that may be implausible.

[32] In order to see whether or not the year 1989 was representative of the calibration period, we also used the estimates for γ_1 (.042) and γ_2 (.00061) which Krueger and Pischke (1992, Table 2b) obtained for a sample of men only. The men's profile peaks at about the same x (33.6) as the two-sexes' profile given above but it reaches a higher value for $h_y(x^*)$, about 2. By inspection of the profiles obtained in the Murphy and Welch (1990, p. 207, fig. 2) study of (white) men, we conclude that our estimate is reasonable.

[33] This is not significantly different from the mean. Calculations by Jones (2002, p. 225) produce a *mean* of 12.5 years for 1993.

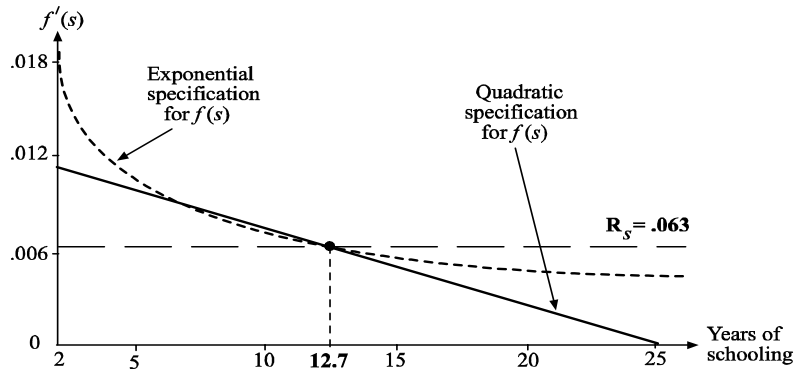

Figure 6: Optimal schooling choice with the exponential specification and the quadratic specification.

As for the extreme values for ability, ε_H is chosen to produce an optimal s of 19 years, the doctoral level say. This gives $\varepsilon_H \approx 1.5$. Then we impose symmetry around average ability ($\overline{\varepsilon} = 1$), so that $\varepsilon_L \approx .5$. The resulting optimal schooling choice for ε_L is 6.35 years. For simplicity, we posit a quadratic in ε: $g(\varepsilon) = \kappa_1\varepsilon^2 + \kappa_2\varepsilon + \kappa_3$.[34] This function has to verify the following conditions:(i) $g(\varepsilon) \geq 0$ for all $\varepsilon \in [\varepsilon_L, \varepsilon_H]$ and $\int_{\varepsilon_L}^{\varepsilon_H} g(\varepsilon)d\varepsilon = 1$; (ii) $\overline{\varepsilon} = \int_{\varepsilon_L}^{\varepsilon_H} \varepsilon g(\varepsilon)d\varepsilon = 1$; (iii) $g(\varepsilon_H) = 0$. Condition (i) guarantees that $g(\varepsilon)$ is a density, condition (ii) that average ability is the one arbitrarily chosen, while condition (iii) captures the idea that the percentage of people close to the extremes of the distribution should be small. The three together pin down the values of κ_1, κ_3 and κ_3.

8 Simulations

In this section we analyze the impact of demographic factors on levels of endogenous variables as well as on measures of inequality along the BGP. Tables 4 and 5 report the effects of, respectively, mortality rates and population growth rates. The notation of the first column of each table requires a few explanations. We assume throughout that we stand at time $t = 0$, so $A(t)$, the index of technical change, is equal to 1, and we drop term t in the expressions below. Recall also that the population at time 0 has been normalized to unity. Variables $\overline{h}_y$, $\overline{k}$ and $\overline{y}$ are, respectively, productive human capital, physical capital and output, all in per worker terms. The next four variables have already been introduced before: $\overline{\mathrm{s}}^*$ is the average school attainment; $s_i^*(R)$ is the fraction of time that a new worker invests in human capital; $\overline{\mathrm{H}}_y^*$ is the average productive human capital of a new worker and $\overline{\boldsymbol{\omega}}_T^*$ is their corresponding wealth at birth.[35] Finally, SD stands for standard deviation, while CV stands for the coefficient of variation (standard deviation over mean), a measure of inequality that is Lorenz-consistent – see for instance Ray (1998, ch. 6).

[34]Positing a density function $g(\varepsilon)$ which is a polynomial in ε makes integration across ability straightforward: calculations amount to integrating a function of the type $e^{a\varepsilon}g(\varepsilon)$ with respect to ε (a being a constant).

[35]We use an asterisk to indicate that these variables relate to optimal decisions.

We begin with the impact of mortality rates – see table 4. The values shown in the second column are those obtained for the calibrated U.S. economy. The values reported in the third and fourth columns, respectively, are the new equilibrium values resulting from mortality rates of $1/55$ and $1/40$, respectively, when $\theta = 1$. The numbers in parenthesis measure changes in relation to the benchmark economy (second column). So, for instance, a value for $\overline{y}$ equal to -10 shown in parenthesis means that income per worker in this case is 10% lower than the one of the calibrated U.S. economy. The fifth and sixth columns repeat the simulations of the same mortality changes for the case $\theta = 2$.

Variable		$\boldsymbol{\theta = 1}$		$\boldsymbol{\theta = 2}$	
	$\mathbf{m}: \frac{\mathbf{1}}{\mathbf{70}}$	$\frac{\mathbf{1}}{\mathbf{55}}$	$\frac{\mathbf{1}}{\mathbf{40}}$	$\frac{\mathbf{1}}{\mathbf{55}}$	$\frac{\mathbf{1}}{\mathbf{40}}$
r	.065	.0673 (+4)	.0708 (+9)	.0695 (+7)	.0758 (+17)
R_s	.063	.0695(+10)	.0798 (+27)	.0717 (+14)	.0848 (+35)
$\overline{h}_y$	3.18	2.89 (-9)	2.43 (-24)	2.79 (-12)	2.21 (-31)
$\overline{k}$	9.12	8.13 (-11)	6.62 (-27)	7.68 (-16)	5.78 (-36)
$\overline{y}$	4.51	4.08 (-10)	3.39 (-25)	3.91 (-13)	3.05 (-32)
$\overline{\mathrm{S}}^*$	12.7	11.5 (-9)	9.4 (-26)	11 (-13)	8.4 (-34)
$s_i^*(R)$	.35	.31 (-11)	.26 (-26)	.30 (-14)	.24 (-31)
$\overline{\mathrm{H}}_y^*$	2.25	2.19 (-3)	2 (-11)	2.15 (-4)	1.88 (-16)
$\overline{\boldsymbol{\varpi}}_T^*$	21.75	18.19 (-16)	14.04 (-35)	17.14 (-21)	12.59 (-42)
$SD(\mathrm{S}^*)$	2.84	2.56 (-10)	2.1 (-26)	2.46 (-13)	1.88 (-34)
$CV(\mathrm{S}^*)$	.22	.22 (0)	.22 (0)	.22 (0)	.22 (0)
$SD(\mathrm{H}_y^*)$	.61	.55 (-10)	.43 (-30)	.53 (-13)	.37 (-39)
$CV(\mathrm{H}_y^*)$	.27	.25 (-7)	.22 (-19)	.24 (-11)	.20 (-26)
$SD(\boldsymbol{\omega}_T^*)$	1.96	1.34 (-32)	.69 (-65)	1.17 (-43)	.50 (-74)
$CV(\boldsymbol{\omega}_T^*)$	.09	.07 (-22)	.05 (-44)	.068 (-24)	.04 (-56)
$SD(h_y)$	1.05	.87 (-17)	.61 (-42)	.815 (-22)	.51 (-51)
$CV(h_y)$	.33	.30 (-9)	.25 (-24)	.29 (-12)	.23 (-30)

Table 4: The impact of mortality on human capital decisions and inequality

The first thing to notice about table 4 is that the effective discount rate R_s increases significantly with mortality rates. This arises because m impacts on R_s both directly, and indirectly through the equilibrium interest rate. This explains why R_s when m equals $1/40$ is about 30 percent higher than in the benchmark case. Second, these differences translate into significant differences in schooling and investment on-the-job choices. For instance, for the case $\theta = 2$, a reduction in life expectancy of 30 years (from 70 to 40 years) reduces average school attainment by 34 percent or 4.3 years in absolute terms; it also reduces the fraction of time that new workers spend investing in their human capital by 31 percent. Third, high mortality rates have a strong negative impact on output per worker because they reduce both productive human capital and physical capital per worker.

We now turn to the impact of mortality rates on inequality within cohorts. Recall that the schooling choices of high ability individuals are more responsive (in absolute terms, not in percentage terms) to changes in the effective discount rate than those of low ability individuals. It follows that the standard deviation of S^*, for a given cohort, is decreasing in R_s. This effect is illustrated in Figure 7. The three downward sloping lines represent the marginal returns to years of schooling of the lowest ability (L), median ability (M) and highest ability individuals (H) in the economy, whereas the two horizontal dashed lines show the effective discount rates that they would face under conditions of high mortality (High R_s) and low mortality (Low R_s).

As can be seen from the picture, high mortality rates, by increasing R_s, reduce inequality in schooling choices. This demographic effect on inequality is large in our simulations. For example, when life expectancy drops to 40 years, $SD(\mathrm{S}^*)$ decreases by 26 percent (case $\theta = 1$) or 34 percent (case $\theta = 2$). But if we compute the unit free measure $CV(\mathrm{S}^*)$, mortality changes turn out to have no impact on inequality. This is because our specification implies that $CV(\mathrm{S}^*) = CV(\varepsilon)$ and the distribution of ability is unchanged.

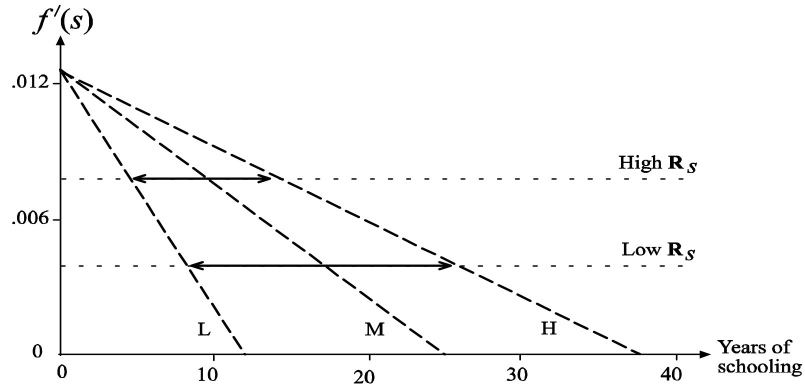

Figure 7: The impact of high mortality rates on inequality in terms of years of schooling.

However, both the SD and the CV of the productive human capital of new workers H_y^* as well as of their wealth at birth $\boldsymbol{\omega}_T^*$ decrease significantly with mortality rates. When these are high, the dispersion in S^* is relatively small. Since high ability individuals make schooling choices that are not very different from those made by low ability individuals, the dispersion in their starting productive human capital is also relatively small. This explains why high mortality reduces inequality in H_y^*. High mortality has an even stronger impact on inequality in wealth at birth because cash-flows earned at different stages in a person's life are now more heavily discounted.

Earnings inequality across cohorts also decreases considerably with mortality rates – see the last two rows of table 4. There are two main reasons for this. First, as stated above, high mortality reduces the dispersion in H_y^*. Second, high mortality reduces directly, and indirectly through its effect on r, the incentives to invest in learning on-the-job. Consequently, it reduces the impact of working experience on productivity.

Variable		**$\theta = 1$**		**$\theta = 2$**	
	n : .012	**0**	**.024**	**0**	**.024**
r	.065	.0575 (-12)	.0717 $(+10)$	.0496 (-24)	.0781 $(+20)$
R_s	.063	.056 (-12)	.07 $(+10)$	.048 $(76-24)$	.076 $(+20)$
$\overline{h}_y$	3.18	3.27 $(+3)$	2.87 (-10)	3.69 $(+16)$	2.58 (-19)
$\overline{k}$	9.12	10.08 $(+11)$	7.77 (-15)	12.28 $(+35)$	6.61 (-28)
$\overline{y}$	4.51	4.76 $(+6)$	4 (-11)	5.51 $(+22)$	3.53 (-22)
$\overline{\mathrm{S}}^*$	12.7	14.2 $(+12)$	11.36 (-11)	15.78 $(+24)$	10.08 (-21)
$s_i^*(R)$	.35	.41 $(+17)$	.31 (-11)	.48 $(+37)$	.28 (-20)
$\overline{\mathrm{H}}_y^*$	2.25	2.26 (0)	2.18 (-3)	2.14 (-5)	2.07 (-8)
$\overline{\boldsymbol{\omega}}_T^*$	21.75	27.7 $(+27)$	17.94 (-18)	37.17 $(+71)$	15.23 (-30)
$SD(\mathrm{S}^*)$	2.84	3.18 $(+12)$	2.54 (-11)	3.53 $(+24)$	2.25 (-21)
$CV(\mathrm{S}^*)$	.22	.22 (0)	.22 (0)	.22 (0)	.22 (0)
$SD(\mathrm{H}_y^*)$	.61	.65 $(+7)$	.54 (-11)	.66 $(+8)$	.47 (-23)
$CV(\mathrm{H}_y^*)$	.27	.29 $(+7)$	.25 (-7)	.31 $(+15)$	.23 (-15)
$SD(\boldsymbol{\omega}_T^*)$	1.96	3.12 $(+59)$	1.29 (-34)	5.18 $(+164)$	.87 (-56)
$CV(\boldsymbol{\omega}_T^*)$	.09	.11 $(+22)$	.07 (-22)	.14 $(+56)$	.06 (-33)
$SD(h_y)$	1.05	1.45 $(+38)$	.83 (-21)	1.7 $(+62)$	.68 (-35)
$CV(h_y)$	.33	.44 $(+33)$	.29 (-12)	.46 $(+39)$	.26 (-21)

Table 5: The impact of population growth on human capital decisions and inequality

The effects of population growth are presented in table 5. For clarity, the second column reproduces the benchmark values already shown in table 4. The third column considers a zero population growth scenario, while the fourth column doubles the growth rate of population relative to the benchmark case. The last two columns repeat these simulations for $\theta = 2$. The first point we make is that the impact of fast population growth is somewhat less severe than the impact of high mortality rates. This reflects the fact that population growth has a negative impact on human capital decisions indirectly only, via the interest rate. Compare for instance the high mortality scenario of table 4 (when $m = 1/40$) with the high population growth scenario of table 5 ($n = .024$): although r is very similar in the two cases, R_s is almost 1 percentage point higher in the first scenario. Still, the negative impact of a fast growing population on human capital decisions and inequality is quite large. For example, for the case $\theta = 2$, a doubling of n decreases average school attainment by 21 percent or 2.6 years in absolute terms; it also reduces earnings inequality as measured by $SD(h_y)$ by 35 percent.

The second point we make is that the impact of lower population growth on human capital decisions and inequality measures is higher when n starts at a lower value. So, for instance, for the case $\theta = 2$, decreasing n from 2.4 percent to 1.2 percent results in an increase in $\overline{\mathrm{s}}^*$ of about 2.6 years; but decreasing it from 1.2 percent to 0 percent results in an increase in $\overline{\mathrm{s}}^*$ of about 3.1 years. What is happening here is that the effect of n on r is nonlinear. The same effect is also responsible for the larger increases in inequality measures that take place when n goes from 1.2 percent to 0 in comparison to those arising when n goes from 2.4 percent to 1.2 percent.

For clarity, the simulation results reported in tables 4 and 5 only change a demographic variable at a time. It should be obvious that, in terms of human capital accumulations decisions, the worst possible combination is high mortality and high population growth rates. For instance, for the case $\theta = 2$, when $m = 1/40$ and $n = 2.4$ percent, average school attainment drops 49 percent, while $s_i^*(R)$ drops 40 percent, in relation to the benchmark economy.[36]

9 Concluding remarks

In this paper we constructed an overlapping generations model of a closed economy which assumes, as in Becker (1964)'s pioneering analysis, that economic agents make human capital investment decisions in the same way that firms make their investment decisions: agents compare the monetary cost of an investment today with the discounted value of the future cash-flows that their decision will produce. Not surprisingly, mortality rates and the economy's interest rate, which can both be seen as intertemporal prices, turn out to be crucial factors in their choice. Population growth rates, to the extent that they have an impact on the equilibrium interest rate, also affect human capital decisions.

Moreover, by positing the existence of a time-invariant distribution of abilities in the population and that diminishing returns to schooling years set in faster for low ability people than for high ability people, we offered a new mechanism by which mortality rates, population growth rates and interest rates may impact on inequality, both within and across cohorts.

For simplicity, there are no frictions in this model economy. We assume the existence of perfect credit markets. Agents may differ in their endowment of ability, but they are all born with zero assets and are all free to incur debt for at least the duration of their studies. This assumption is certainly not realistic, as most poor families that are credit constrained, because they lack collateral, can testify. But it helps us to focus on the connection between demographic variables and human capital only.

The model can easily be adapted or extended to consider many other issues. For example, we have used an open economy version of it with realistic age-specific mortality rates to study the

[36] Space constraints preclude the discussion of the way other exogenous variables may interact in nonlinear ways with mortality and fertility in the determination of human capital and inequality.

impact on income and expected lifetime utility of technological shocks and demographic shocks such as the HIV-AIDS epidemic.

References

Arias, Elizabeth, "United States Life Tables, 2002," *National Vital Statistics Reports*, 2004, *53* (6), 1–39.

Avolio, Bruce J. and David A. Waldman, "Variations in Cognitive, Perceptual, and Psychomotor Abilities Across the Working Life Span: Examiming the Effects of Race, Sex, Experience, Education, and Occupational Type," *Psychology and Aging*, 1994, *9* (3), 430–42.

Barro, Robert J., "Government Spending in a Simple Model of Endogeneous Growth," *Journal of Political Economy*, 1990, *98* (5), S103–25.

____ **and Xavier Sala-I-Martin**, *Economic Growth*, second ed., Cambridge (MA): MIT Press, 2004.

Becker, Gary S., *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, New York: Columbia University Press, 1964.

Ben-Porath, Yoram, "The Production of Human Capital and the Life Cycle of Earnings," *Journal of Political Economy*, 1967, *75* (4), 352–365.

Benhabib, Jess and Roberto Perli, "Uniqueness and Indeterminacy: On the Dynamics of Endogenous Growth," *Journal of Economic Theory*, 1994, *63* (1), 113–42.

Bils, Mark and Peter J. Klenow, "Does Schooling Cause Growth?," *American Economic Review*, 2000, *90* (5), 1160–1183.

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

____ **and Stanley Fischer**, *Lectures on Macroeconomics*, Cambridge, MA: MIT Press, 1989.

Browning, Martin, Lars Peter Hansen, and James J. Heckman, "Micro Data and General Equilibrium Models," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Vol. 1A, Amsterdam: North Holland, 1999, chapter 8.

Dupor, Bill, Lance Lochner, Christopher Taber, and Mary Beth Wittekind, "Some Effects of Taxes on Schooling and Training," *American Economic Review*, 1996, *86* (2), 340–6.

Haley, William J., "Estimation of the Earnings Profile from Optimal Human Capital Accumulation," *Econometrica*, 1976, *44* (6), 1223–38.

Heckman, James J., "A Life-Cycle Model of Earnings, Learning, and Consumption," *Journal of Political Economy*, 1976, *84* (4), S11–44.

——— **, Lance Lochner, and Christopher Taber**, "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, 1998, *1* (1), 1–58.

Jones, Charles I., "Sources of U.S. Economic Growth in a World of Ideas," *American Economic review*, 2002, *92* (1), 220–239.

Kalemli-Ozcan, Sebnem, Harl E. Ryder, and David N. Weil, "Mortality Decline, Human Capital Investment, and Economic Growth," *Journal of Development Economics*, 2000, *62* (1), 1–23.

King, Robert G. and Sérgio T. Rebelo, "Resuscitating Real Business Cycles," in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Vol. 1B, Amsterdam: North-Holland, 1999, chapter 14, pp. 927–1007.

Krueger, Alan B. and Jörn-Steffen Pischke, "A Comparative Analysis of East and West German Labor Markets: Before and After Unification," 1992. NBER Working Paper No. 4154.

Lucas, Robert E. Jr., "On the Mechanics of Economic Development," *Journal of Monetary Economics*, 1988, *22* (1), 3–42.

McCabe, Peter J., "Optimal Leisure-Effort Choice with Endogenously Determined Earnings," *Journal of Labor economics*, 1983, *1* (3), 308–29.

Mincer, Jacob A., *Schooling, Experience and Earnings*, New York: Columbia University Press, 1974.

Murphy, Kevin M. and Finis Welch, "Empirical Age-Earnings Profiles," *Journal of Labor Economics*, 1990, *8* (2), 202–229.

Psacharopoulos, George, "Returns to Investment in Education: A Global Update," *World development*, 1994, *22* (9), 1325–43.

——— **and Harry Anthony Patrinos**, "Returns to Investment in Education: A Further Update," *Education Economics*, 2004, *12* (2), 111–34.

Ray, Debraj, *Development Economics*, New Jersey: Princeton University Press, 1998.

Rosen, Sherwin, "Human Capital: Relations Between Education and Earnings," in Michael D. Intriligator, ed., *Frontiers of Quantitative Economics*, Vol III-B, Amsterdam: North-Holland, 1977, chapter 17, pp. 731–53.

Seierstad, Atle and Knut Sydsaeter, *Optimal Control Theory with Economic Applications*, North-Holland, 1987.

Trostel, Philip A., "The Effect of Taxation on Human Capital," *Journal of Political Economy*, 1993, *101* (2), 327–50.

Weil, Philippe, "Overlapping Families of Infinitely-Lived Agents," *Journal of Public Economics*, 1989, *38* (2), 183–198.

Willis, Robert J., "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions," in Orley Ashenfelter and Richard Layard, eds., *Handbook of Labor Economics*, Vol. 1, Amsterdam: North-Holland, 1986, chapter 10, pp. 525–602.

Yaari, Menahem E., "Uncertain Lifetime, Life Insurance, and the Theory of the Consumer," *Review of Economic Studies*, 1965, *32* (2), 137–50.

A Technical Appendix

A.1 Derivation of eq. (11) and Condition 1

From the first order condition with respect to s_i – eq. (8) – we obtain the optimal quantity of efficiency units of human capital used for investment:

$$h_i(j,v) = b^{1/(1-b)} E(j,v) \left[\tfrac{\Lambda_h(j,v)}{w(v)\Lambda_a(j,v)}\right]^{1/(1-b)} \tag{36}$$

Using the solution $\Lambda_a(j,v) = \Lambda_a(j,t)e^{-(r+m)(v-t)}$ where $v \geq t \geq j'$, and $w(v) = w(t)e^{g(v-t)}$ in the second flow eq. $d\Lambda_h(j,v)/dv = \delta\Lambda_h(j,v) - w(v)\Lambda_a(j,v)$, we get, after defining $R = r + m + \delta - g$ and assuming that $R \neq 0$,

$$\begin{aligned}\Lambda_h(j,v) &= e^{\delta(v-t)}\left(\Lambda_h(j,t) - w(t)\Lambda_a(j,t)\int_t^v e^{-(r+m+\delta-g)(s-t)}ds\right)\\ &= e^{\delta(v-t)}\left(\left[\Lambda_h(j,t) - \tfrac{w(t)\Lambda_a(j,t)}{R}\right] + e^{-R(v-t)}\tfrac{w(t)\Lambda_a(j,t)}{R}\right)\end{aligned}$$

On the other hand, the solution for human capital is

$$h(j,v) = e^{-\delta(v-t)}\left(h(j,t) + \int_t^v \Psi(j,s)e^{\delta(s-t)}ds\right)$$

with $\Psi(j,s) = E(j,s)^{1-b}h_i(j,s)^b$. Hence,

$$\begin{aligned}h(j,v)\Lambda_h(j,v) &= \left(\underbrace{h(j,t) + \int_t^v \Psi(j,s)e^{\delta(s-t)}ds}_{(I)}\right)\times\\ &\quad\left(\underbrace{\left[\Lambda_h(j,t) - \tfrac{w(t)\Lambda_a(j,t)}{R}\right] + e^{-R(v-t)}\tfrac{w(t)\Lambda_a(j,t)}{R}}_{(II)}\right)\end{aligned} \tag{37}$$

Now, term (I) of eq. (37) is always positive because the integral is a sum of positive terms. If the integral in (I) converges as $v \to \infty$, this first term (in parenthesis) of eq. (37) will tend to a positive constant, otherwise it will explode to infinity. Necessary conditions for $\lim_{v\to\infty} h(j,v)\Lambda_h(j,v) = 0$ are then:[37]

$$\Lambda_h(j,t) = R^{-1}w(t)\Lambda_a(j,t) \tag{38}$$

$$R > 0 \tag{39}$$

Eq. (38) is eq. (11) given in main text, with t substituted for v; eq. (39) is Condition $1(i)$.

However, these conditions are not sufficient because it is still possible for the integral of term (I) of eq. (37) to grow faster than its term (II) converges to 0. In order to deal with sufficiency, we need to compute the integral of term (I). Given condition (38) and eq. (36), we

[37] When $R = 0$, this product would not converge either as $v \to \infty$.

have $\Psi(j,s) = E(j,s)(b/R)^{b/(1-b)}$. Finally, assuming that $E(j,s) = \phi h(j,j)e^{-\lambda(s-j)}$, eq. (37) may be rewritten as

$$\begin{aligned} h(j,v)\Lambda_h(j,v) &= e^{-R(v-t)}\Lambda_h(j,t)\left[h(j,t) - \frac{\phi h(j,j)(b/R)^{b/(1-b)}e^{-\lambda(t-j)}}{\delta-\lambda}\right] \\ &\quad + e^{-\widehat{R}(v-t)}\frac{\Lambda_h(j,t)\phi h(j,j)(b/R)^{b/(1-b)}e^{-\lambda(t-j)}}{\delta-\lambda} \end{aligned}$$

where $\widehat{R} = r+m+\lambda-g$, and we assume $\widehat{R} \neq 0$. Now, as $v \to \infty$, the first term of the previous expression will tend to zero because of condition (39); but the second term will only converge to zero if $\widehat{R} > 0$, and this is Condition $1(ii)$.

A.2 Derivation of $K(t)$

A prerequisite to aggregation is to have the expressions for individual assets. In footnote (9) we gave the expression for $a(j,v)$ when $v \leq j'$:

$$a(j,v) = \tfrac{c(j,j)}{r+m-\theta^{-1}(r-\rho)}\left[e^{\theta^{-1}(r-\rho)(v-j)} - e^{(r+m)(v-j)}\right] \tag{40}$$

By rewriting the last eq. using eqs. (21) and (22), we have, for $v \leq j'$,

$$a(j,v) = w(j')h(j,j')\left[\tfrac{1}{R} + \tfrac{s_i(R)}{\widehat{R}}\tfrac{1-b}{b}\right]\left[e^{-(r+m)s+\theta^{-1}(r-\rho)(v-j)} - e^{(r+m)(v-j')}\right]. \tag{41}$$

On the other hand, solving eq. (2) for $v > j'$ gives

$$a(j,v) = a(j,j')e^{(r+m)(v-j')} - \int_{j'}^{v} e^{(r+m)(v-x)}c(j,x)dx + \int_{j'}^{v} e^{(r+m)(v-x)}y(j,x)dx.$$

Now compute the first two terms of the last eq. using eqs. (19), (40) and (22):

$$a(j,v) = \omega^T(j,j)\left[e^{\theta^{-1}(r-\rho)(v-j)} - e^{(r+m)(v-j)}\right] + \int_{j'}^{v} e^{(r+m)(v-x)}y(j,x)dx \tag{42}$$

Now, $y(j,x) = w(x)h_y(j,x)$. Substituting $w(x)$ for $w(j')e^{g(x-j')}$ and using eq. (16), with v substituted for x, we have:

$$y(j,x) = w(j')h(j,j')\left[e^{(g-\delta)(x-j')} + s_i(R)\left\{\left(\tfrac{R}{b(\delta-\lambda)} - 1\right)e^{(g-\lambda)(x-j')} - \tfrac{R}{b(\delta-\lambda)}e^{(g-\delta)(x-j')}\right\}\right],$$

hence,

$$\int_{j'}^{v} e^{(r+m)(v-x)}y(j,x)dx = w(j')h(j,j')\times$$
$$\int_{j'}^{v} e^{(r+m)(v-x)}\left[e^{(g-\delta)(x-j')} + s_i(R)\left\{\left(\tfrac{R}{b(\delta-\lambda)} - 1\right)e^{(g-\lambda)(x-j')} - \tfrac{R}{b(\delta-\lambda)}e^{(g-\delta)(x-j')}\right\}\right]dx,$$

or, after integrating,

$$\int_{j'}^{v} e^{(r+m)(v-x)} y(j,x)dx = w(j')h(j,j')\times$$
$$\left\{ \begin{array}{c} \dfrac{e^{(r+m)(v-j')} - e^{(g-\delta)(v-j')}}{R} \\ +s_i(R)\left\{\left(\frac{R}{b(\delta-\lambda)} - 1\right)\dfrac{e^{(r+m)(v-j')} - e^{(g-\lambda)(v-j')}}{\widehat{R}} - \frac{R}{b(\delta-\lambda)}\dfrac{e^{(r+m)(v-j')} - e^{(g-\delta)(v-j')}}{R}\right\} \end{array} \right\},$$

which can be rearranged to give

$$\int_{j'}^{v} e^{(r+m)(v-x)} y(j,x)dx = w(j')h(j,j')\times$$
$$\left\{\left[\tfrac{1}{R} + \tfrac{s_i(R)}{\widehat{R}}\tfrac{1-b}{b}\right] e^{(r+m)(v-j')} + \left[\tfrac{s_i(R)}{b(\delta-\lambda)} - \tfrac{1}{R}\right] e^{(g-\delta)(v-j')} - \left[\tfrac{s_i(R)}{\widehat{R}}\left(\tfrac{R}{b(\delta-\lambda)} - 1\right)\right] e^{(g-\lambda)(v-j')}\right\}.$$

Finally, substitute the last eq. and eq. (21) into eq. (42) to obtain, for $v > j'$,

$$a(j,v) = w(j')h(j,j')\left\{ \begin{array}{c} \left[\tfrac{1}{R} + \tfrac{s_i(R)}{\widehat{R}}\tfrac{1-b}{b}\right] e^{-(r+m)s+\theta^{-1}(r-\rho)(v-j)} \\ + \left[\tfrac{s_i(R)}{b(\delta-\lambda)} - \tfrac{1}{R}\right] e^{(g-\delta)(v-j')} - \left[\tfrac{s_i(R)}{\widehat{R}}\left(\tfrac{R}{b(\delta-\lambda)} - 1\right)\right] e^{(g-\lambda)(v-j')} \end{array} \right\}. \tag{43}$$

We now proceed to aggregate these individual assets. We have to make a distinction between those agents alive at t that are yet to join the workforce and the others. Hence, aggregate physical capital is

$$K(t) = \int_{-\infty}^{t-s} a(j,t)L(j,t)dj + \int_{t-s}^{t} a(j,t)L(j,t)dj$$

Using eqs. (41) and (43), with v substituted for t, j' for $j+s$, and term $w(j')h(j,j')$ substituted for $w(0)e^{g(j+s)+f(s)}$, we have:

$$\begin{aligned} \frac{K(t)e^{mt}}{m+n} &= \int_{-\infty}^{t} w(0)e^{(m+n+g)j+f(s)}\left[\tfrac{1}{R} + \tfrac{s_i(R)}{\widehat{R}}\tfrac{1-b}{b}\right] e^{(g-r-m)s+\theta^{-1}(r-\rho)(t-j)}dj \\ &+ \int_{-\infty}^{t-s} w(0)e^{(m+n+g)j+gs+f(s)}\times \\ &\left\{\left[\tfrac{s_i(R)}{b(\delta-\lambda)} - \tfrac{1}{R}\right] e^{(g-\delta)(t-j-s)} - \left[\tfrac{s_i(R)}{\widehat{R}}\left(\tfrac{R}{b(\delta-\lambda)} - 1\right)\right] e^{(g-\lambda)(t-j-s)}\right\} dj \\ &- \int_{t-z-s}^{t} w(0)e^{(m+n+g)j+f(s)}\left[\tfrac{1}{R} + \tfrac{s_i(R)}{\widehat{R}}\tfrac{1-b}{b}\right] e^{(r+m)(t-j)+(g-r-m)s}dj, \end{aligned}$$

Defining $R' = m+n+g+\theta^{-1}(\rho - r)$ and imposing $R' > 0$ to guarantee that $K(t)$ takes a finite value, the last expression may be integrated to give[38]

$$\frac{K(t)}{(m+n)w(0)e^{(n+g)t}} = \frac{1}{R'}e^{(g-r-m)s+f(s)}\left[\frac{1}{R}+\frac{s_i(R)}{\widehat{R}}\frac{1-b}{b}\right] + e^{-(m+n)s+f(s)}\left\{\frac{1}{\Delta_1}\left[\frac{s_i(R)}{b(\delta-\lambda)}-\frac{1}{R}\right]-\frac{1}{\Delta_2}\left[\frac{s_i(R)}{\widehat{R}}\left(\frac{R}{b(\delta-\lambda)}-1\right)\right]\right\} - e^{(g-r-m)s+f(s)}\frac{1}{n+g-r}\left[\frac{1}{R}+\frac{s_i(R)}{\widehat{R}}\frac{1-b}{b}\right]\left[1-e^{(r-n-g)s}\right],$$

or, substituting $(m+n)w(0)e^{(n+g)t}$ for $w(t)h(t,t)L(t,t)$, and rearranging

$$\begin{aligned}K(t) &= w(t)h(t,t)L(t,t)e^{(g-r-m)s}e^{f(s)}\left[\frac{1}{R}+\frac{s_i(R)}{\widehat{R}}\frac{1-b}{b}\right]\left[\frac{1}{R'}-\frac{1}{n+g-r}+\frac{e^{(r-n-g)s}}{n+g-r}\right] \\ &- w(t)h(t,t)L(t,t)e^{-(m+n)s}e^{f(s)}\left[\frac{1}{R\Delta_1}+\frac{s_i(R)}{\Delta_2}\left(\frac{1}{\widehat{R}}\frac{1-b}{b}+\frac{1}{b\Delta_1}\right)\right] \\ &= w(t)h(t,t)L(t,t)e^{-(m+n)s+f(s)}\left\{\Psi_1(r)\Psi_2(r)-\Psi_3(r)\right\},\end{aligned}$$

with $\Psi_1(r)$, $\Psi_2(r)$ and $\Psi_3(r)$ given below – and also in footnote (17):

$$\begin{aligned}\Psi_1(r) &= \left[\frac{1}{R}+\frac{s_i(R)}{\widehat{R}}\frac{1-b}{b}\right] \qquad \Psi_2(r) = \left[e^{(n+g-r)s}\left(\frac{1}{R'}-\frac{1}{n+g-r}\right)+\frac{1}{n+g-r}\right] \\ \Psi_3(r) &= \left[\frac{1}{R\Delta_1}+\frac{s_i(R)}{\Delta_2}\left(\frac{1}{\widehat{R}}\frac{1-b}{b}+\frac{1}{b\Delta_1}\right)\right]\end{aligned}$$

A.3 Derivation of Condition 7

In the main text, we have established that those agents alive at t that are yet to join the workforce will all have negative financial wealth. But this means that the agents who work must, in the aggregate, have positive financial assets, otherwise $K(t)$ would necessarily be negative. Let us then compute the financial assets of the *working* population. In what follows, we will use the first four conditions given in main text. They are $R > 0$ and $\widehat{R} > 0$; $0 > s_i(R) > 1$; $e^{\phi}(b/R)^{\frac{b}{1-b}} > \delta$ and $\delta > \lambda > 0$.

Use (43) with v substituted for t, j' for $j+s$, and the product $w(j')h(j,j')$ for $w(0)e^{gj+gs+f(s)}$, plus eq. (25) to get

$$\int_{-\infty}^{t-s} a(j,t)L(j,t)dj = (m+n)e^{-mt}\int_{-\infty}^{t-s} w(0)e^{(g+m+n)j+gs+f(s)}\times \left\{\begin{array}{c}\left[\frac{1}{R}+\frac{s_i(R)}{\widehat{R}}\frac{1-b}{b}\right]e^{-(r+m)s+\theta^{-1}(r-\rho)(t-j)} \\ +\left[\frac{s_i(R)}{b(\delta-\lambda)}-\frac{1}{R}\right]e^{(g-\delta)(t-j-s)}-\left[\frac{s_i(R)}{\widehat{R}}\left(\frac{R}{b(\delta-\lambda)}-1\right)\right]e^{(g-\lambda)(t-j-s)}\end{array}\right\}dj,$$

[38]We assume that $r \neq n+g$ just to avoid having to deal with two cases.

or, integrating and then substituting $(m+n)w(0)e^{(n+g)t}$ for $w(t)h(t,t)L(t,t)$,

$$\int_{-\infty}^{t-s} a(j,t)L(j,t)dj = w(t)h(t,t)L(t,t)e^{-(m+n)s+f(s)}\times \left\{ \begin{array}{c} \frac{1}{R'}\left[\frac{1}{R}+\frac{s_i(R)}{\widehat{R}}\frac{1-b}{b}\right]e^{-[r+m-\theta^{-1}(r-\rho)]s} \\ +\frac{1}{\Delta_1}\left[\frac{s_i(R)}{b(\delta-\lambda)}-\frac{1}{R}\right]-\frac{1}{\Delta_2}\left[\frac{s_i(R)}{\widehat{R}}\left(\frac{R}{b(\delta-\lambda)}-1\right)\right] \end{array} \right\} \tag{44}$$

We know that the expression in eq. (44) inside the curly brackets must be positive, otherwise the whole integral will be negative. But because of Condition 5, which reads $r+m>\theta^{-1}(r-\rho)$, the exponential term in the first line of the expression inside the curly brackets must necessarily be less than one. Hence, the expression inside the curly brackets has to be smaller than the following term[39]

$$\digamma = \frac{1}{R'}\left[\frac{1}{R}+\frac{s_i(R)}{\widehat{R}}\frac{1-b}{b}\right]+\frac{1}{\Delta_1}\left[\frac{s_i(R)}{b(\delta-\lambda)}-\frac{1}{R}\right]-\frac{1}{\Delta_2}\left[\frac{s_i(R)}{\widehat{R}}\left(\frac{R}{b(\delta-\lambda)}-1\right)\right],$$

which can be rearranged and simplified to give[40]

$$\digamma = \underbrace{\frac{1}{R}\left[\frac{1}{R'}-\frac{1}{\Delta_1}\right]}_{(A)}+s_i(R)\left[\underbrace{\frac{1-b}{b}\frac{1}{\widehat{R}}\left[\frac{1}{R'}-\frac{1}{\Delta_2}\right]}_{(B)}-\underbrace{\frac{1}{b\Delta_1\Delta_2}}_{(C)}\right]$$

Now, if $\digamma$ is not positive, $K(t)$ will necessarily be negative. In order to analyse the effects of this restriction, let us assume that r is equal to $\theta(g-\delta)+\rho$. Then $R'=\Delta_1$, term (A) is zero and $\digamma$ simplifies to

$$\digamma = -\frac{s_i(R)}{b\Delta_1\Delta_2\widehat{R}}\left[\widehat{R}+(1-b)(\delta-\lambda)\right]$$

But since $\widehat{R}>0$ and $\delta>\lambda$, it must necessarily be true that $\widehat{R}+(1-b)(\delta-\lambda)>0$. But in this case $\digamma$ would be negative. We have therefore shown that r cannot be equal to $\theta(g-\delta)+\rho$.

On the other hand, Condition 1 (i) in main text reads $r>g-m-\delta$. If $g-m-\delta>\theta(g-\delta)+\rho$, then the inequality $r>\theta(g-\delta)+\rho$ is trivially verified. So consider the case where $g-m-\delta<\theta(g-\delta)+\rho$.[41] What happens if $g-m-\delta<r<\theta(g-\delta)+\rho$? Then $R'>\Delta_1$ and term (A) is negative. Moreover, since $\Delta_1>\Delta_2$, term (B) is also negative. But in this case $\digamma$ would be negative. Hence so far we have established the following: $r>\theta(g-\delta)+\rho$, $R'<\Delta_1$, and term (A) must be positive.

But can $\theta(g-\lambda)+\rho\geq r>\theta(g-\delta)+\rho$? Then in this case $\Delta_2\leq R'<\Delta_1$, $\frac{1}{\Delta_2}\geq\frac{1}{R'}>\frac{1}{\Delta_1}$, $\frac{\delta-\lambda}{\Delta_1\Delta_2}\geq\frac{1}{R'}-\frac{1}{\Delta_1}>0$, and $0\geq\frac{1}{R'}-\frac{1}{\Delta_2}>\frac{\lambda-\delta}{\Delta_1\Delta_2}$. So term (B) would be non-positive and

$$\frac{1}{R}\left[\frac{1}{R'}-\frac{1}{\Delta_1}\right]-s_i(R)\frac{1}{b\Delta_1\Delta_2}\leq\frac{1}{R}\frac{\delta-\lambda}{\Delta_1\Delta_2}-s_i(R)\frac{1}{b\Delta_1\Delta_2}$$

[39] Notice that we could have obtained this number directly from equation (29) by setting s to 0. This would have amounted to assume that the population consists entirely of workers.

[40] To get the expression to the right of $s_i(R)$, we substituted R for $\widehat{R}+\delta-\lambda$ and $(1/\Delta_1-1/\Delta_2)$ for $-(\delta-\lambda)/(\Delta_1\Delta_2)$.

[41] This is necessarily true when $\theta\geq 1$.

or

$$\tfrac{1}{R}\left[\tfrac{1}{R'}-\tfrac{1}{\Delta_1}\right]-s_i(R)\tfrac{1}{b\Delta_1\Delta_2}\leq\tfrac{1}{\Delta_1\Delta_2}\left[-\tfrac{\lambda}{R}+\tfrac{\delta}{R}-\tfrac{s_i(R)}{b}\right] \tag{45}$$

Now, given Condition 3, which amounts to say that $\frac{s_i(R)}{b}>\frac{\delta}{R}$, the l.h.s. of inequality (45) must be negative. But in this case $\digamma$ would be negative. We can therefore conclude that, when $\delta>\lambda$, $R'<\Delta_2<\Delta_1$, (A) and (B) will both be positive, and a necessary condition for equilibrium is that $r>\theta(g-\lambda)+\rho$.

A.4 Derivation of $c(t,t)L(t,t)$ as a function of $C(t)$ and $K(t)$

Let $y^{pv}(j,t)$ stand for the present value of labour income at t for a individual born at time j. We have $y^{pv}(j,t)=\int_t^\infty e^{-(r+m)(x-t)}y(j,x)dx$, where $y(j,x)=w(x)h_y(j,x)$. We need to make a distinction between those agents alive at t who already work and those who do not. Regarding those not yet working, $t\leq j'$, [42] we can use eq. (16) to get:

$$y^{pv}(j,t)=\int_{j'}^\infty e^{-(r+m)(x-t)}y(j,x)dx=w(t)h(t,t)e^{f(s)-\mathrm{R}_s(j+s-t)}\Psi_1(r),$$

where we recall that $\Psi_1(r)=\left[\frac{1}{R}+\frac{s_i(R)}{\widehat{R}}\frac{1-b}{b}\right]$. As for those already in the workforce, $t>j'$, and once again making use of eq. (16), we obtain

$$y^{pv}(j,t)=w(t)h(t,t)e^{f(s)}\left[\frac{e^{\delta(j+s-t)}}{R}+s_i(R)\left\{\left(\frac{R}{b(\delta-\lambda)}-1\right)\frac{e^{\lambda(j+s-t)}}{\widehat{R}}-\frac{e^{\delta(j+s-t)}}{b(\delta-\lambda)}\right\}\right]$$

Defining $Y^{pv}(t)$ as aggregate human wealth, $Y^{pv}(t)=\int_{-\infty}^t y^{pv}(j,t)L(j,t)dj$, we have

$$\begin{aligned}Y^{pv}(t)&=\int_{-\infty}^{t-s}y^{pv}(j,t)L(j,t)dj+\int_{t-s}^{t}y^{pv}(j,t)L(j,t)dj\\&=w(t)h(t,t)L(t,t)e^{f(s)}\left[\Psi_3(r)e^{-(m+n)s}+\Psi_1(r)\frac{e^{-\mathrm{R}_s s}-e^{-(m+n)s}}{n+g-r}\right]\end{aligned}$$

or

$$Y^{pv}(t)=w(t)h(t,t)L(t,t)e^{-(m+n)s+f(s)}\left[\Psi_3(r)+\Psi_1(r)\frac{e^{(n+g-r)s}-1}{n+g-r}\right] \tag{46}$$

But aggregate wealth $W(t)$ is equal to aggregate human wealth $Y^{pv}(t)$ plus aggregate financial wealth, $K(t)$. Moreover, aggregate consumption is proportional to aggregate wealth. So we also have

$$Y^{pv}(t)=\frac{C(t)}{r+m-\theta^{-1}(r-\rho)}-K(t) \tag{47}$$

On the other hand, $c(t,t)=\left[r+m-\theta^{-1}(r-\rho)\right]\omega_T(t,t)$. Using eq. (23), we have

[42] We recall that $j'=j+s$.

$\omega_T(t,t) = w(t)h(t,t)e^{-\mathrm{R}_s s + f(s)}\Psi_1(r)$, hence

$$c(t,t)L(t,t) = w(t)h(t,t)L(t,t)e^{-\mathrm{R}_s s + f(s)}\left[r + m - \theta^{-1}(r-\rho)\right]\Psi_1(r)$$

Finally, using eqs. (46) and (47) we can cancel term $w(t)h(t,t)L(t,t)$ of the last eq. to get:

$$c(t,t)L(t,t) = \frac{C(t) - \left[r + m - \theta^{-1}(r-\rho)\right]K(t)}{e^{(r-n-g)s}\Psi_3(r)\Psi_1(r)^{-1} + (e^{(r-n-g)s} - 1)/(r-n-g)}$$

This is the expression which is then substituted in eq. (31) in order to obtain eq. (32).