

Education, lifetime labor supply, and mortality improvements*

Miguel Sánchez-Romero¹ Hippolyte d’Albis⁴
Alexia Prskawetz^{1,2,3}

^{1,2,3}Wittgenstein Centre (IIASA, VID/ÖAW and WU)

¹Vienna Institute of Demography/Austrian Academy of Science

²Vienna University of Technology

³International Institute for Applied System Analysis

⁴Paris School of Economics - University Paris 1

June 13, 2014

Abstract

This paper presents an analysis of the differential role of mortality for the optimal schooling and retirement age when the accumulation of human capital follows the so-called “Ben-Porath mechanism”. We set up a life-cycle model of consumption and labor supply at the extensive margin that allows for endogenous human capital formation based on Card (2001). This paper makes two important contributions. First, we show that a decrease in mortality may induce a longer education period and an earlier retirement age. Second, using Swedish data for

*We thank Raouf Boucekkine, Uwe Sunde, and participants at the Association for Public Economic Theory (2013), Universidad Autonoma de Madrid, and the International Workshop on Heterogenous Dynamic Models of Economic and Population Systems (Vienna, 2013) for their comments and suggestions. We also thank Kirk Scott and Gustav Oeberg for their reach support providing Swedish data. M. Sánchez-Romero acknowledges the research and financial support of the Max Planck Institute for Demographic Research (Rostock, Germany). H. d’Albis acknowledges financial support from the European Research Council (ERC Stg Grant DU 283953). This project has received funding from the European Union’s Seventh Framework Program for research, technological development and demonstration under grant agreement no. 613247: “Ageing Europe: An application of National Transfer Accounts (NTA) for explaining and projecting trends in public finances”.

cohorts born between 1865 and 2000, we show that our model can match the empirical evidence and permit to decompose the effects into the Ben-Porath mechanism and the lifetime-human wealth effect vs. the years-to-consume effect.

1 Introduction

In many countries, development has been accompanied with significant increases in life expectancy and reductions in labor supply. Since the nineteenth century life expectancy has increased by 40 years at a rate of 3 months per year (Oeppen and Vaupel, 2002; Lee, 2003), while labor supply has decreased at two extensive margins: later entrance in the labor market and earlier retirement. Prior to industrialization, male literacy rates started to increase in the most advanced countries (Cippola, 1969; Cervellati and Sunde, 2005; Boucekkine et al., 2007). This process continued with an expansion of primary education enrollment rates at the end of the nineteenth century and first half of the twentieth century (Benavot and Riddle, 1988). By 1950, the average length of schooling for males was around six years in the most advanced countries and has increased up to twelve years in 2010 (Barro and Lee, 2013). Simultaneously, labor force participation rates for old workers started to fall until recently, even before the introduction of pension systems (Costa, 1998; Schieber and Shoven, 1999). In 1970, the average retirement age was 68 in OECD countries and has declined up to age 63 in 2010 (OECD, 2009).

Existing theoretical models that analyze the effect of mortality on education and retirement, however, find results that contradict the historical empirical evidence. In particular, these models assume a causal positive relation between human capital investments, retirement age, and life expectancy (Boucekkine et al., 2002; Echevarria, 2004; Echevarria and Iza, 2006; Ferreira and Pessoa, 2007; Zhang and Zhang, 2009). While the positive link between human capital investment and life expectancy, through the well-known Ben-Porath (1967)'s mechanism, is undisputed (de la Croix and Licandro, 1999; Kalemli-Ozcan et al., 2000; Zhang et al., 2001, 2003; Cervellati and Sunde, 2005; Soares, 2005; Zhang and Zhang, 2005; Jayachandran and Lleras-Muney, 2009; Oster et al., 2013), except for Hazan and Zoabi (2006). However, studies about the effect of mortality improvements on the decline in retirement age are scarce and offer several complementary explanations. For instance, Kalemli-Ozcan and Weil (2010) suggest that the decline in retirement age might be explained by reductions in the risk of dying before retirement, named "uncertainty effect". More recently, Bloom et al. (2014) point out that positive in-

come effects along the twentieth century might have offset the gains in healthy life after retirement, or “compression of morbidity” effect (Bloom et al., 2007).

In the last two decades, improvements in the understanding of the Ben-Porath mechanism and the link between mortality and retirement have come from two different sources. On the one hand, empirical investigations of the mortality decline over the last two centuries show that mortality does not improve uniformly across age groups (Lee, 1994; Wilmoth and Horiuchi, 1999; Cutler et al., 2006). Early stages of the mortality transition are mainly characterized by reductions of mortality in infants and children, while recent mortality declines occur at older ages. Motivated by this demographic feature, several authors have recently shown that the link between life expectancy and labor supply depends on the age pattern of mortality improvements. In particular, mortality declines during adulthood may cause early retirement, while reductions in mortality at older ages delay retirement (d’Albis et al., 2012; Strulik and Werner, 2012). Also, it was shown that improvements in survival during prime-working ages increase human capital investment (Cervellati and Sunde, 2013). On the other, recent literature on the returns to education show that, despite the rising returns to high education in the US, on average individuals underinvest in human capital due to high psychic costs of school, risk and uncertainty, or unobserved differences in skills (Carneiro et al., 2003; Cunha et al., 2005; Heckman et al., 2006). In other words, individuals do not choose the optimal number of years of education that maximize their lifetime earnings (Becker, 1967; Willis, 1986). To cope with this problem, Card (2001) proposes a model in which individuals have different aptitudes or tastes for schooling relative to work.

In this paper, we explain the differential role of mortality in optimal schooling and retirement age. We set up a life-cycle model of consumption and labor supply at the extensive margin that allows for endogenous human capital formation based on Card (2001). Since the decline in mortality does not occur uniformly across age-groups, following d’Albis et al. (2012) we model the age-specific mortality rates non-parametrically. Thus, using the derivative of a functional (Ryder and Heal, 1973; d’Albis et al., 2012) we study the impact of a mortality decline at any arbitrary age on human capital investment and retirement.

Under this setup, we first derive analytically conditions under which a decrease in mortality may induce higher education and lower labor supply, thereby reconciling the empirical evidence shown by Hazan (2009) with economic theory. We provide the economic intuition of our results by decomposing the differential effect of mortality on schooling and retirement into the “lifetime human wealth” and “years-to-consume” effects. “Lifetime human wealth” ef-

fect stands for the positive impact that a mortality decline has on consumption because it raises the likelihood of receiving a future labor income stream. On the contrary, the “years-to-consume” effect, which is always negative, reflects the overall reduction in consumption due to a longer lifespan. Second, we empirically analyze our model using Swedish schooling, retirement, and mortality data for cohorts born between 1865 and 2000. Our results show two important facts. First, when the life expectancy rises, our model is capable of producing an optimal decline in retirement age and an increase in years of schooling (Hazan, 2009). Second, since in the earlier stage of mortality transition, a decline in mortality belongs mainly to younger people, whereas in the later stage, a decline in mortality has mainly occurred at older ages, we show that the optimal retirement age bottomed out for cohorts born in the 1920s and it is expected to increase from now on.

The rest of the paper is organized as follows: Section 2 introduces the model setup and presents the first-order conditions for optimal consumption, length of schooling, and retirement. Furthermore, the relationship between the optimal length of schooling and retirement is explained. In Section 3, we study –using the Volterra derivative of a functional– the differential role of mortality on the optimal length of schooling and retirement. For a better understanding on the role of mortality on each variable, we distinguish between the direct and indirect impact of mortality on education and retirement, separately. In Section 4, we solve the model numerically and demonstrate, using a simple quantitative exercise, how the mortality transition may increase the length of schooling and reduce the retirement age. Concluding remarks are made in Section 5.

2 The model

We setup a consumer’s problem that consists in choosing the optimal number of years of schooling (S), optimal retirement age (R), and the optimal consumption path ($c(x)$) in order to maximize the expected lifetime utility ($V(S, R, c)$). We assume time is continuous. Agents face lifetime uncertainty, which is represented by the survival function

$$p(x) = e^{-\int_0^x \mu(q) dq}, \quad (1)$$

where $p(x)$ is the (unconditional) probability of surviving to age x , $p(0) = 1$, $p(\omega) = 0$, being $\omega \in (0, \infty)$ the maximum age, and $\mu(x) \geq 0$ is the mortality hazard rate at age x .

Schooling and labor supply are indivisible and the transitions from schooling to working and from working to retirement are irreversible, as in Boucekkinne et al. (2002), Echevarria (2004), Echevarria and Iza (2006), and Cai and Lau (2012). We also assume that agents do not save with a bequest motive in mind and there exists a perfect annuity market, which grants that agents borrow and lend freely at a fixed interest rate. Thereby, consumers optimally choose to purchase annuities (Yaari, 1965). The instantaneous expected utility depends positively on current consumption and negatively on current non-leisure time. The utility of consumption $U(c)$ is an increasing and concave function (i.e. $U_c(\cdot) > 0$, $U_{cc}(\cdot) < 0$).¹ Let $\tilde{\phi}(S, x)$ denote the disutility of non-leisure time at age x of an individual who has completed S years of schooling. Assume $\tilde{\phi}(S, x)$ is a positive and increasing function with respect to age (i.e. $\tilde{\phi}(S, x) > 0$, $\tilde{\phi}_x(S, x) > 0$), which reflects the fact that the disutility of not enjoying leisure is increasing with age (Hazan, 2009; Kalemli-Ozcan and Weil, 2010; d’Albis et al., 2012; Cai and Lau, 2012). After retirement, $\tilde{\phi}(S, x)$ equals zero. Then, assuming that agents discount future utility flows at a subjective discount rate ρ , the expected lifetime utility, conditional on the years of schooling (S), retirement (R), and consumption path (c) is

$$V(S, R, c) = \int_0^\omega e^{-\rho x} p(x) U(c(x)) dx - \int_0^R e^{-\rho x} p(x) \tilde{\phi}(S, x) dx. \quad (2)$$

Following Card (2001), we generalize most existing theoretical papers by assuming that agents may have different tastes for schooling relative to work

$$\tilde{\phi}(S, x) = \begin{cases} \phi(x) + \psi(x) & \text{if } x \leq S, \\ \phi(x) & \text{if } x > S, \end{cases} \quad (3)$$

where $\phi(x) > 0$ (with $\phi_x(x) \geq 0$) is the underlying disutility of non-leisure time and $\psi(x)$ is the relative disutility of school versus work. Factor $\psi(x)$ is positive when the agent prefers work to schooling or negative when schooling is preferred to work. We assume that if $\psi(x)$ is positive, the relative disutility of schooling increases with age (i.e. $\psi_x(x) \geq 0$), whereas if $\psi(x)$ is negative, our agent has a decreasing preference for schooling, or $\psi_x(x) \leq 0$. As a particular case, notice $\psi(x) = 0$ for all $x \in (0, \omega)$ is also included in Card (2001), which is implicitly assumed in Hazan (2009), Kalemli-Ozcan and Weil (2010), Cai and Lau (2012), among many others.

Labor income, denoted by y , is assumed to be proportional to years of schooling and years of post-schooling experience (Mincer, 1974). Then, we

¹We use subscripts to denote the derivative with respect to the variable in the subscript, and apply the same notation for partial derivatives.

write labor income at age x conditional on S years of schooling as $y(S, x) = w(x - S)h(S, x)$, where $w(x - S) > 0$ (with $w(0) > 0$, $w_x(x - S) \geq 0$ and $w_S(x - S) \leq 0$ for $x \leq \underline{x}$, and $w_x(x - S) \leq 0$ and $w_S(x - S) \geq 0$ for $x > \underline{x}$) represents the wage rate per unit of human capital with $x - S$ years of post-schooling experience and $h(S, x)$ is the stock of human capital of an individual at age x with S years of schooling.² Assume the law of motion of human capital of an individual at age x with S years of schooling accumulates according to a Ben-Porath (1967) technology

$$h_x(S, x) = \begin{cases} q(h(S, x)) - \delta h(S, x) & \text{if } x \leq S \\ -\delta h(S, x) & \text{otherwise,} \end{cases} \quad (4)$$

where $q(\cdot)$ is the human capital production function (with $q_h(\cdot) > 0$ and $q_{hh}(\cdot) < 0$), and $\delta > 0$ is the human capital depreciation rate, which is assumed constant across age.³ As a result, the law of motion of financial wealth at age x ($a(x)$) is

$$a_x(x) = \begin{cases} [r + \mu(x)]a(x) + y(S, x) - c(x) & \text{if } S < x < R, \\ [r + \mu(x)]a(x) - c(x) & \text{otherwise,} \end{cases} \quad (5)$$

with boundary conditions $a(0) = 0$ and $a(\omega) = 0$, where r is the (real) interest rate. Integrating (5) with respect to age, subject to the boundary conditions, we obtain the standard lifecycle budget constraint faced by our individual:

$$\int_0^\omega e^{-rx} p(x) c(x) dx = \int_S^R e^{-rx} p(x) y(S, x) dx \equiv W(S, R), \quad (6)$$

where $W(S, R)$ is the lifecycle earnings (measured at age 0) conditional on S years of schooling and retirement age R . For the sake of comparison with the literature on the impact of mortality on retirement and education, notice that we implicitly assume that the only cost of schooling is foregone labor income (Kalemli-Ozcan et al., 2000; Hazan, 2009; Cai and Lau, 2012; Cervellati and Sunde, 2013). Tuition costs, earnings while in school, and taxes are also modeled in the returns to education literature (Willis, 1986; Card, 2001; Heckman et al., 2006).

²This assumption guarantees the usual hump-shape of the wage rate per unit of human capital and thus labor income. Using data from the US Decennial Censuses, the estimated coefficients from Mincer log earnings regressions for white males during the period 1940-90 report an average $\underline{x} - S$ value of 30.1 years (Heckman et al., 2006, Table 2, p. 326).

³The functional form $h(S, S) = e^{\theta(S)}$, used by Hazan (2009), p. 1834, can be obtained assuming either that $\delta = 0$ or that $q(h(S, x))$ is equal to $(\theta_x(x) + \delta)h(S, x)$.

2.1 Optimal consumption, length of schooling, and retirement age

Following d'Albis and Augeraud-Véron (2008), Heijdra and Romp (2009), and d'Albis et al. (2012) we obtain our agent's optimal consumption path, length of schooling, and retirement in two steps. First, we derive the optimal consumption path. We define the optimal consumption at age x , conditional on the length of schooling (S) and retirement age (R), as $c(x, S, R)$. Second, based on the conditional consumption path derived in the first step, we obtain the optimal length of schooling and retirement age. Let us define $\hat{V}(S, R)$ as the expected lifetime utility conditional on the optimal consumption path.

In Proposition 1, we characterize the optimal consumption path, the optimal length of schooling, and the retirement age.

Proposition 1 *For the life-cycle model given by (1)-(5), the optimal consumption path, conditional on a length of schooling S and a retirement age R , is characterized by*

$$U_c(c(x, S, R)) = e^{(\rho-r)x} U_c(c(0, S, R)). \quad (7)$$

Moreover, an interior optimal length of schooling (S^*) satisfies

$$\int_{S^*}^R e^{-r(x-S^*)} \frac{p(x)}{p(S^*)} y_S(S^*, x) dx = y(S^*, S^*) + \frac{e^{(r-\rho)S^*} \psi(S^*)}{U_c(c(0, S^*, R))}, \quad (8)$$

and an interior optimal retirement age (R^*) is given by

$$U_c(c(0, S, R^*)) e^{-rR^*} y(S, R^*) = e^{-\rho R^*} \phi(R^*). \quad (9)$$

Proof. See the proof in Appendix A. ■

Eq. (7) is the standard Euler condition characterizing the consumption path. The left-hand side of Eq. (8) is the marginal benefit of the S^* -th year of schooling, whereas the right-hand side represents the marginal cost of the S^* -th year of schooling. Let us define $f(S, R)$ as the marginal effect of an additional year of schooling (measured at age S) on lifecycle earnings:

$$f(S, R) \equiv \frac{W_S(S, R)}{e^{-rS} p(S)} = \int_S^R e^{-r(x-S)} \frac{p(x)}{p(S)} y_S(S, x) dx - y(S, S), \quad (10)$$

or, equivalently, the marginal benefit of the S -th year of schooling minus the foregone labor income at age S (measured at age S). Provided the log-labor

income is separable in education and experience, from (4) Eq. (10) can be rewritten, after rearranging, as

$$f(S, R) = \frac{W(S, R)}{e^{-rS}p(S)} \left(\frac{q(h(S, S))}{h(S, S)} - \left[\int_S^R e^{-(r+\delta)(x-S)} \frac{p(x)}{p(S)} \frac{w(x-S)}{w(0)} dx \right]^{-1} + \int_S^R \frac{e^{-(r+\delta)x} p(x) w_S(x-S)}{\int_S^R e^{-(r+\delta)u} p(u) w(u-S) du} dx \right), \quad (11)$$

where $q(h(S, S))/h(S, S)$ is the rate of return to education at S (henceforth $r_h(S)$).⁴ The second term inside the parenthesis is the inverse of the present value (measured at age S) of the labor income over the working life relative to the initial labor income, and the third term is the marginal effect of schooling on the wage rate over the working life. For notational convenience, let us denote the sum of the last two terms in (11) as $\bar{r}(S, R)$; that is

$$\bar{r}(S, R) = \left[\int_S^R e^{-(r+\delta)(x-S)} \frac{p(x)}{p(S)} \frac{w(x-S)}{w(0)} dx \right]^{-1} - \int_S^R \frac{e^{-(r+\delta)x} p(x) w_S(x-S)}{\int_S^R e^{-(r+\delta)u} p(u) w(u-S) du} dx. \quad (12)$$

Assuming a flat wage rate, zero interest rate, and no human capital depreciation, the inverse of $\bar{r}(S, R)$ coincides with the expected total labor supply, $\int_S^R \frac{p(x)}{p(S)} dx$, used by Lee (2001), Hazan (2009), and recently by Cervellati and Sunde (2013).⁵ Substituting (10)-(12) in (8), and rearranging, gives

$$r_h(S^*) = \bar{r}(S^*, R) + \frac{e^{-\rho S^*} p(S^*) \psi(S^*)}{W(S^*, R) U_c(c(0, S^*, R))}, \quad (13)$$

or the return to education with S^* years of schooling. Notice that when our individual has a higher preference for working versus schooling ($\psi(S^*) > 0$), we observe an underinvestment in education ($S^* < \bar{S}$) and the return to education is higher than $\bar{r}(S^*, R^*)$. In contrast, when our individual prefers schooling to work ($\psi(S^*) < 0$), we have an over-investment in education ($S^* > \bar{S}$) and hence the return to education is lower than $\bar{r}(S^*, R^*)$. Finally, if education

⁴Notice $q(h(S, S))/h(S, S)$ can be obtained by differentiating the $\log y(S, x)$ with respect to S , which coincides with the rate of return to education $r_h(S^*)$ in the standard Mincer (1974) earnings function: $\log y(S, E + S) = \alpha_0 + r_h S + \beta_0 E + \beta_1 E^2$, where E stands for years of experience.

⁵Recent estimations applied to cohorts born between 1850 to 1990 of US men workers obtain that the expected number of years worked from age 20 has been roughly stable between forty one and forty two years (Lee, 2001).

were considered a pure investment good ($\psi(S^*) = 0$), $r_h(S^*)$ would coincide with $\bar{r}(S^*, R^*)$.

Empirically, the econometric estimations of returns to education report values of $r_h(S^*)$ exceeding those of $\bar{r}(S^*, R^*)$. For example, Card (1999) finds a wide range of rates of returns to education in the US centered around 8% per year, while Heckman et al. (2008) estimate also for the US that the returns to education range between 10 to 15% per year. In contrast, when education is considered a pure investment good, the rate of return to education for an individual with 10 years of education does not exceed 3% per year for a wide range of feasible retirement ages.⁶ Several explanations are suggested in the literature for the positive difference between $r_h(S^*)$ and $\bar{r}(S^*, R^*)$. The most common ones are high “psychic cost” of school, uncertainty, and heterogeneity among individuals (Carneiro et al., 2003; Cunha et al., 2005; Heckman et al., 2006), while credit constraints might be important for going to college decisions (Belley and Lochner, 2007), but not for most students (Carneiro and Heckman, 2002; Heckman et al., 2006). Henceforth, following the literature on returns to education, we assume that $\psi(x) > 0$ for all $x \in (0, S)$.

Eq. (9) is the optimal retirement age condition. Eq. (9) implies that the marginal benefit of continued working at age R^* , which is equivalent to the additional labor income at age R^* measured in utility terms, equals the marginal cost of working at age R^* , or the disutility of continued working at age R^* . This optimal retirement age condition was first derived by Sheshinski (1978).

The first important results one can obtain from Proposition 1 are the effects of an increase in the optimal length of schooling and retirement age on the optimal consumption path. Differentiating (6) and (7) with respect to S , substituting, and using (10) gives

$$\frac{1}{c(0, S, R)} \frac{\partial c(0, S, R)}{\partial S} \Big|_{S=S^*} = \frac{e^{-rS^*} p(S^*) \sigma(c(0, S^*, R)) f(S^*, R)}{\int_0^\omega e^{-rx} p(x) \sigma(c(x, S^*, R)) c(x, S^*, R) dx}. \quad (14)$$

where

$$\sigma(c) = -\frac{U_c(c)}{c \cdot U_{cc}(c)} > 0, \quad (15)$$

is the intertemporal elasticity of substitution (IES) for consumption c . Using

⁶A value of 3% has been calculated based on the wage rate per unit of human capital $\log w(x - S) = \log w(0) + 0.094(x - S) - 0.0013(x - S)^2$ withdrawn from Table 2 (Heckman et al., 2006, p. 326), US death rates of males from the cohort born in year 1900 (BMD), an interest rate of 3%, and no human capital depreciation rate.

(6) and (11)-(13), Eq. (14) becomes⁷

$$\frac{1}{c(0, S, R)} \frac{\partial c(0, S, R)}{\partial S} \Big|_{S=S^*} = \frac{\sigma(c(0, S^*, R))}{\sigma(c(\bar{x}, S^*, R))} (r_h(S^*) - \bar{r}(S^*, R)), \quad (17)$$

where $\sigma(c(\bar{x}, S^*, R))$ is the mean value of the IES over the age range $[0, \omega]$. Assuming a constant IES, Eq. (17) implies that the relative increase in the initial consumption due to an additional year of schooling is equal to the difference between the actual return of schooling and that as if education is considered a pure investment good. Therefore, an additional investment in schooling is efficient when $r_h(S^*) > \bar{r}(S^*, R)$, and inefficient when $r_h(S^*) < \bar{r}(S^*, R)$. Furthermore, the consumption path is maximized if, and only if, $r_h(S^*) = \bar{r}(S^*, R)$ for all $R > S^*$.

To analyze the impact of retirement on the optimal consumption path we differentiate (6) and (7) with respect to R . Substituting and using (9), we have

$$\frac{1}{c(0, S, R)} \frac{\partial c(0, S, R)}{\partial R} \Big|_{R=R^*} = \frac{e^{-rR^*} p(R^*) \sigma(c(0, S, R^*)) y(S, R^*)}{\int_0^\omega e^{-rx} p(x) \sigma(c(x, S, R^*)) c(x, S, R^*) dx}, \quad (18)$$

which is equivalent to

$$\frac{1}{c(0, S, R)} \frac{\partial c(0, S, R)}{\partial R} \Big|_{R=R^*} = \frac{\sigma(c(0, S^*, R)) e^{-rR^*} p(R^*) y(S, R^*)}{\sigma(c(\bar{x}, S^*, R)) W(S, R^*)}. \quad (19)$$

For a constant IES, Eq. (19) states that the relative impact of delaying retirement on the initial consumption is equal to the weight of labor income at age R^* in lifecycle earnings. Thereby, contrary to an increase in the length of schooling, an increase in the retirement age always raises the optimal consumption path because the agent receives an additional labor income at age R^* .

⁷Applying the mean value theorem for integration, we have

$$\begin{aligned} & \int_0^\omega e^{-rx} p(x) \sigma(c(x, S^*, R)) c(x, S^*, R) dx \\ &= \sigma(c(\bar{x}, S^*, R)) \int_0^\omega e^{-rx} p(x) c(x, S^*, R) dx \equiv \sigma(c(\bar{x}, S^*, R)) W(S^*, R), \end{aligned} \quad (16)$$

where $\sigma(c(\bar{x}, S^*, R))$ is the mean value of the IES $\sigma(c(x, S^*, R))$ over the age range $[0, \omega]$.

2.2 Relationship between years of schooling and retirement

In the previous subsection we have shown the first-order conditions for an optimum of S^* and R^* , separately, and how they impact on the optimal consumption path. In this subsection, we turn to a detailed study about the relationship between the optimal years of schooling and the optimal retirement age.

Let us denote the optimal consumption at age 0 conditional on S^* and R^* as c^* . To examine the impact on the optimal length of schooling of a change in the retirement age, we totally differentiate (8) with respect to S^* and R . Taking $\psi(S^*)$ as common factor, and rearranging, we obtain

$$\left. \frac{\partial S^*}{\partial R} \right|_{R=R^*} = \frac{\frac{\partial \hat{V}_S(S^*, R)}{\partial R}}{-\hat{V}_{SS}(S^*, R)} \Big|_{R=R^*} = \frac{\frac{f_R(S^*, R^*)}{f(S^*, R^*)} - \frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{\partial c^*}{\partial R}}{\frac{\psi_S(S^*)}{\psi(S^*)} + \frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{\partial c^*}{\partial S} - \frac{f_S(S^*, R^*)}{f(S^*, R^*)}}. \quad (20)$$

Similarly, applying the implicit function theorem, we totally differentiate (9) with respect to R^* and S to examine the impact on the optimal retirement age of a change in the length of schooling

$$\left. \frac{\partial R^*}{\partial S} \right|_{S=S^*} = \frac{\frac{\partial \hat{V}_R(S, R^*)}{\partial S}}{-\hat{V}_{RR}(S, R^*)} \Big|_{S=S^*} = \frac{-\frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{\partial c^*}{\partial S} + r_h(S^*) + \frac{w_S(R^*-S^*)}{w(R^*-S^*)}}{\frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{\partial c^*}{\partial R} - \frac{w_R(R^*-S^*)}{w(R^*-S^*)} + \delta + \frac{\phi_R(R^*)}{\phi(R^*)}}. \quad (21)$$

Provided (S^*, R^*) is the interior solution of our problem, substituting (10) and (19) in (20)-(21), and using (12), we have

$$\text{sign} \left[\left. \frac{\partial S^*}{\partial R} \right|_{R=R^*} \right] = \text{sign} \left[\left. \frac{\partial R^*}{\partial S} \right|_{S=S^*} \right] = \text{sign} \left[\frac{\bar{r}(S^*, R^*) + \sigma(\bar{c}) \frac{w_S(R^*-S^*)}{w(R^*-S^*)}}{1 - \sigma(\bar{c})} - r_h(S^*) \right], \quad (22)$$

where $\sigma(\bar{c})$ is, hereinafter, the short-hand notation for $\sigma(c(\bar{x}, S^*, R^*))$. Eq. (22) implies that the length of schooling S^* and the retirement age R^* may be either positively or negatively related. On the one hand, looking at Eq. (21), we have that an additional year of schooling after age S^* increases the labor income at age R^* by $r_h(S^*) + \frac{w_S(R^*-S^*)}{w(R^*-S^*)}$, which increases the marginal benefit of working. As a consequence, our individual optimally postpones the retirement age in order to reap the benefits of schooling. On the other hand, the increase in education may also change the marginal utility of consumption, and hence

the marginal benefit of working, by $-\frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{\partial c^*}{\partial S}$.⁸ Thus, the net effect of a change in education on retirement depends upon the strength of the income effect, reflected by the IES (Imrohoroglu and Kitao, 2009; Keane, 2011), and the difference between $r_h(S^*)$ and $\bar{r}(S^*, R^*)$. At the extreme cases, when $\sigma(\bar{c})$ tends to one or $\bar{r}(S^*, R^*) = r_h(S^*)$, we have that the sign of $\frac{\partial S^*}{\partial R^*} \Big|_{R=R^*}$ and $\frac{\partial R^*}{\partial S} \Big|_{S=S^*}$ depend on that of $\bar{r}(S^*, R^*) + \frac{w_S(R^*-S^*)}{w(R^*-S^*)}$, which is always positive.

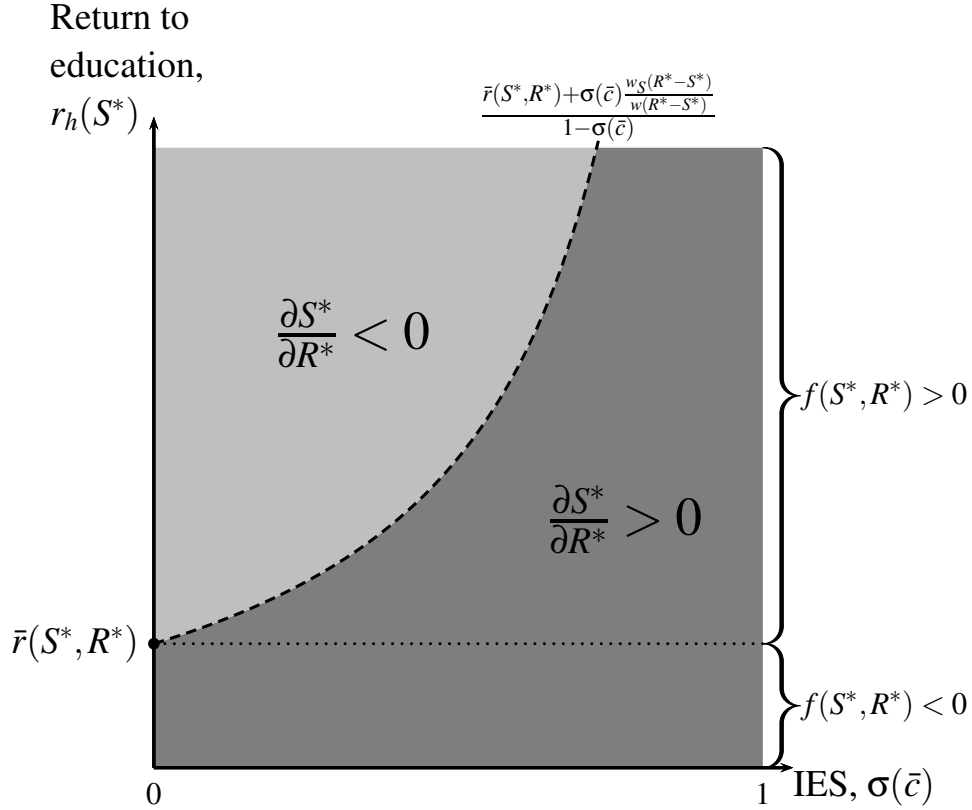


Figure 1: Relationship between S^* and R^* by return to education and intertemporal elasticity of substitution

Figure 1 summarizes the result obtained in Eq. (22). For any given wage

⁸Differentiating (7) with respect to S , we have

$$\frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{\partial c^*}{\partial S} = \frac{1}{\sigma(c(x, S, R^*))} \frac{1}{c(x, S, R^*)} \frac{\partial c(x, S, R^*)}{\partial S} \Big|_{S=S^*} \text{ for all } x \in (0, \omega).$$

rate per unit of human capital, Figure 1 is broken into two shaded areas. A dark gray area that contains the combination of $(r_h(S^*), \sigma(\bar{c}))$ values for which S^* and R^* are positively related, and a light gray area with the combination of $(r_h(S^*), \sigma(\bar{c}))$ values for which S^* and R^* are negatively related. It is clear looking at Figure 1 that S^* and R^* are positively related whenever the return to education is equal to, or lower than, $\bar{r}(S^*, R^*)$ (dark gray area below the horizontal dotted line in Figure 1). In this region, an increase in the retirement age leads to an increase in the optimal length of schooling (Ben-Porath, 1967), as well as an increase in schooling yields an increase in the retirement age (Boucekkine et al., 2002; Echevarria and Iza, 2006). However, when the return to education is higher than $\bar{r}(S^*, R^*)$, S^* and R^* can either be positively or negatively related. The black dashed line in Figure 1 delimits the combination of $(r_h(S^*), \sigma(\bar{c}))$ values at which there isn't a relationship between S^* and R^* ; i.e. $\frac{\partial S^*}{\partial R^*} = 0$. The light gray area, located at the upper-left corner, is characterized by low IES and high return to education. In this area, the income effect dominates. Thus, for a sufficiently high return to education and low IES, when a positive income shock increases the optimal years of schooling, the optimal retirement age decreases, since individuals purchase more leisure time, and the positive effect on years of schooling gets reinforced. The same effect would take place if the positive income shock initially reduces the retirement age. Notice, however, the negative relation between S^* and R^* vanishes as the return to education approaches to the dashed line. On the contrary, in the dark gray area, where the strength of the income effect diminishes –when a positive income shock raises the retirement age– the optimal years of schooling increases and the rise in the retirement age gets also reinforced.

3 Differential impact of mortality decline on optimal schooling years and retirement age

In this Section, we study the impact of a mortality decline at an arbitrary age (x_0) on the optimal length of schooling (S^*) and the optimal retirement age (R^*). For exposition clarity, we make explicit the dependence of the optimal schooling and retirement age on each other and on the underlying mortality schedule; i.e. $S^* \equiv S^*(R; \boldsymbol{\mu})$ and $R^* \equiv R^*(S; \boldsymbol{\mu})$.⁹

Eqs. (23a)-(23b) show how the effect of a mortality decline at an arbitrary

⁹ $\boldsymbol{\mu} = \{\mu(x)\}_{x \in (0, \omega)}$ represents the vector of all mortality hazard rates along the life cycle.

age x_0 is characterized by the sum of a direct and an indirect effect:

$$\left. \frac{-dS^*}{d\mu(x_0)} \right|_{R=R^*} = \left. \frac{-\partial S^*}{\partial \mu(x_0)} \right|_{R=R^*} + \left. \frac{\partial S^*}{\partial R} \right|_{R=R^*} \left. \frac{-dR^*}{d\mu(x_0)} \right|_{S=S^*}, \quad (23a)$$

$$\left. \frac{-dR^*}{d\mu(x_0)} \right|_{S=S^*} = \left. \frac{-\partial R^*}{\partial \mu(x_0)} \right|_{S=S^*} + \left. \frac{\partial R^*}{\partial S} \right|_{S=S^*} \left. \frac{-dS^*}{d\mu(x_0)} \right|_{R=R^*}. \quad (23b)$$

In words, the direct effect is the impact of a mortality decline at x_0 on S^* and R^* –respectively– holding all other variables unchanged, while the indirect effect is the impact of retirement (resp. schooling) on schooling (resp. retirement) that is mediated by a change in mortality. Thus, as shown in (23a) and (23b), the effect of mortality on S^* and R^* are intertwined. However, provided the strict concavity of $\dot{V}(S^*, R^*)$, we can easily derive that the sign of a mortality decline at an arbitrary age x_0 on S^* and R^* are¹⁰

$$\text{sign} \left[\left. \frac{-dS^*}{d\mu(x_0)} \right|_{R=R^*} \right] = \text{sign} \left[\left. \frac{-\partial S^*}{\partial \mu(x_0)} \right|_{R=R^*} + \left. \frac{\partial S^*}{\partial R} \right|_{R=R^*} \left. \frac{-\partial R^*}{\partial \mu(x_0)} \right|_{S=S^*} \right], \quad (24a)$$

$$\text{sign} \left[\left. \frac{-dR^*}{d\mu(x_0)} \right|_{S=S^*} \right] = \text{sign} \left[\left. \frac{-\partial R^*}{\partial \mu(x_0)} \right|_{S=S^*} + \left. \frac{\partial R^*}{\partial S} \right|_{S=S^*} \left. \frac{-\partial S^*}{\partial \mu(x_0)} \right|_{R=R^*} \right]. \quad (24b)$$

According to (24) the total impact of a mortality decline on S^* and R^* can be analyzed, independently, through the analysis of the partial derivatives. For exposition clarity, in Section 3.1 we first focus on the analysis of the direct effects and, in Section 3.2, we study the total impact of a mortality decline on both variables.

¹⁰Solving the system of equations in (23), we have

$$\left. \frac{-dS^*}{d\mu(x_0)} \right|_{R=R^*} = \left(1 - \left. \frac{\partial S^*}{\partial R} \right|_{R=R^*} \left. \frac{\partial R^*}{\partial S} \right|_{S=S^*} \right)^{-1} \left(\left. \frac{-\partial S^*}{\partial \mu(x_0)} \right|_{R=R^*} + \left. \frac{\partial S^*}{\partial R} \right|_{R=R^*} \left. \frac{-\partial R^*}{\partial \mu(x_0)} \right|_{S=S^*} \right),$$

$$\left. \frac{-dR^*}{d\mu(x_0)} \right|_{S=S^*} = \left(1 - \left. \frac{\partial S^*}{\partial R} \right|_{R=R^*} \left. \frac{\partial R^*}{\partial S} \right|_{S=S^*} \right)^{-1} \left(\left. \frac{-\partial R^*}{\partial \mu(x_0)} \right|_{S=S^*} + \left. \frac{\partial R^*}{\partial S} \right|_{S=S^*} \left. \frac{-\partial S^*}{\partial \mu(x_0)} \right|_{R=R^*} \right).$$

Therefore, from (A.10) we obtain

$$\left. \frac{-dS^*}{d\mu(x_0)} \right|_{R=R^*} \propto \left. \frac{-\partial S^*}{\partial \mu(x_0)} \right|_{R=R^*} + \left. \frac{\partial S^*}{\partial R} \right|_{R=R^*} \left. \frac{-\partial R^*}{\partial \mu(x_0)} \right|_{S=S^*},$$

$$\left. \frac{-dR^*}{d\mu(x_0)} \right|_{S=S^*} \propto \left. \frac{-\partial R^*}{\partial \mu(x_0)} \right|_{S=S^*} + \left. \frac{\partial R^*}{\partial S} \right|_{S=S^*} \left. \frac{-\partial S^*}{\partial \mu(x_0)} \right|_{R=R^*}.$$

3.1 Direct effect

Following the same order as the derivation of the first-order conditions, we first study the direct impact that a mortality decline has on the optimal consumption path and, second, we continue with the analysis of the direct effects of a mortality decline on the optimal length of schooling and retirement age.

To study the effect of mortality on our variables of interest, we make use of the derivative of a functional (Ryder and Heal, 1973; d’Albis et al., 2012) to obtain, through (1), that

$$-\frac{\partial p(x)}{\partial \mu(x_0)} = \begin{cases} p(x) & \text{if } x_0 \leq x, \\ 0 & \text{if } x_0 > x. \end{cases} \quad (25)$$

Eq. (25) means that a mortality decline at age x_0 has no effect on the survival probability before age x_0 , but it increases the survival probability at ages above or equal to x_0 . From (25) we derive the impact that a mortality decline at an arbitrary age x_0 has on the optimal consumption path and, in particular, on the initial optimal consumption (c^*). Differentiating (6) and (7) with respect to $-\mu(x_0)$, substituting, and rearranging gives

$$\frac{1}{c^*} \frac{-\partial c^*}{\partial \mu(x_0)} = -\frac{\sigma(c^*)}{\sigma(\bar{c})} \frac{e^{-rx_0} p(x_0) a(x_0)}{\int_0^\omega e^{-rx} p(x) c(x, S^*, R^*) dx} \equiv -\frac{\sigma(c^*)}{\sigma(\bar{c})} \frac{e^{-rx_0} p(x_0) a(x_0)}{W(S^*, R^*)}, \quad (26)$$

where $\sigma(\bar{c})$ is the mean value of the IES $\sigma(c(x, S^*, R^*))$ in $[0, \omega]$. If the IES is constant across the lifecycle, the relative impact of a mortality decline at age x_0 on the initial consumption is minus the ratio between the financial wealth position at age x_0 and lifecycle earnings. Thereby, a decline in mortality at two different ages do not necessarily have the same impact on consumption.¹¹ In addition, according to (26) the impact of a decline in mortality at age x_0 on the optimal consumption path is inversely related to the financial wealth at age x_0 ,

$$\text{sign} \left[\frac{-\partial c^*}{\partial \mu(x_0)} \right] = -\text{sign} [a(x_0)]. \quad (28)$$

Eq. (26) is the extension of Eq. B.5 in d’Albis et al. (2012) to a model with endogenous human capital investment. Like d’Albis et al. (2012) we show that

¹¹The total impact of a mortality decline on the initial consumption is

$$\frac{1}{c^*} \frac{dc^*}{d\mu} = \int_0^\omega \frac{1}{c^*} \frac{-\partial c^*}{\partial \mu(x)} d\mu(x) = -\frac{\sigma(c^*)}{\sigma(\bar{c})} \frac{\int_0^\omega e^{-rx} p(x) a(x) d\mu(x)}{W(S^*, R^*)}. \quad (27)$$

the optimal consumption path increases with a decline in mortality at age x_0 when $a(x_0) < 0$, while the optimal consumption declines when $a(x_0) > 0$, for all $x_0 \in (0, \omega)$. The intuition is simple. On the one hand, a decline in mortality increases the number of years the agent is expected to live. As a consequence, agents compensate a longer lifespan with an overall reduction in consumption. This effect, which is always negative, is named the “years-to-consume” effect. On the other hand, a mortality decline during the working period raises the likelihood of receiving a future labor income stream, which leads to an overall increase in the consumption path. This other effect, which is always positive, is named the “lifetime human wealth” effect. For a better understanding, Proposition 2 gives the net result of these two opposite effects using a CIES utility function.

Proposition 2 *For the life-cycle model given by (1)-(5), if $U_x(x)$ is a power function, the overall result of $\frac{1}{c^*} \frac{-\partial c^*}{\partial \mu(x_0)}$ is the same as that of*

$$g(x_0) = \frac{\int_{S^*}^{R^*} e^{-rx} \left[\frac{-\partial p(x)}{\partial \mu(x_0)} \right] y(S^*, x) dx}{\int_{S^*}^{R^*} e^{-rx} p(x) y(S^*, x) dx} - \frac{\int_{x_0}^{\omega} e^{-[(1-\sigma)r+\sigma\rho]x} p(x) dx}{\int_0^{\omega} e^{-[(1-\sigma)r+\sigma\rho]x} p(x) dx}, \quad (29)$$

where $\sigma \in [0, 1]$ is the intertemporal elasticity of substitution. Moreover, there exists a critical point x_c within the open interval (S^*, R^*) such that

$$\begin{cases} g(x_0) > 0 & \text{for all } x_0 < x_c, \\ g(x_0) = 0 & \text{for all } x_0 = x_c, \\ g(x_0) < 0 & \text{for all } x_0 > x_c. \end{cases} \quad (30)$$

Proof. See Appendix B ■

The first component of (29) is the “lifetime human wealth” effect, while the second component represents the “years-to-consume” effect. An illustration of the shape of both effects across the life-cycle is given in Figure 2. Notice the lifetime human wealth effect dominates the years-to-consume effect up to age $x_c \in (S^*, R^*)$, the year at which the financial wealth is zero, $a(x_c) = 0$. Therefore, a mortality decline early in life leads to an overall increase in consumption. In contrast, a decline in mortality at ages above x_c leads to an overall decline in consumption because the years-to-consume effect dominates the lifetime human wealth effect. Though for simplicity we have not modeled any retirement pension system, our results are robust to the introduction of a more general and realistic framework. Indeed, realize that the introduction of an income during the retirement period will extend the lifetime human wealth effect up to age ω , shifting the age x_c toward older ages.

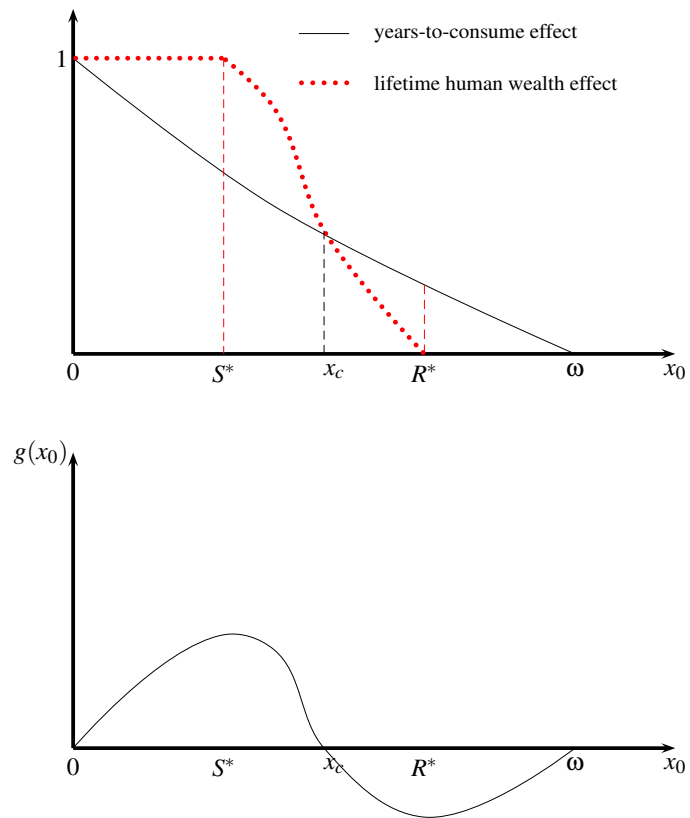


Figure 2: The lifetime human wealth and years-to-consume effect

The impact of a decline in mortality on the length of schooling and retirement age is given in Proposition 3.

Proposition 3 *For the life-cycle model given by (1)-(5),*

(a) *the sign of $\frac{-\partial S^*}{\partial \mu(x_0)}$ is the same as that of*

$$\frac{r_h(S^*) - \bar{r}(S^*, R^*)}{\sigma(\bar{c})} a(x_0) + \int_{S^*}^{R^*} e^{-r(x-x_0)} \frac{-\partial(p(x)/p(x_0))}{\partial \mu(x_0)} y_S(S^*, x) dx \quad (31)$$

and

(b) *the sign of $\frac{-\partial R^*}{\partial \mu(x_0)}$ is the same as that of $a(x_0)$.*

Proof. See Appendix C ■

In Proposition 3 we obtain the same “consumption-leisure” relationship as in d’Albis et al. (2012). That is, given that consumption and leisure are normal goods, Proposition 3(b) implies that if a mortality decline yields an increase in consumption because the lifetime human wealth effect dominates the years-to-consume effect, agents anticipate their optimal retirement age in order to enjoy more leisure time. Similarly, when the decline in mortality implies that the years-to-consume effect dominates the lifetime human wealth effect, agents diminish their consumption and postpone their optimal retirement age.

Proposition 3(a) extends the years-to-consume effect and lifetime human wealth effect reasoning to the accumulation of human capital. In this regard, we obtain unambiguous results concerning the sign on the optimal length of schooling of a mortality decline at ages before the entrance into the labor market, S^* , and after the optimal retirement age, R^* . Specifically, Proposition 3(a) implies that when $r_h(S^*) > \bar{r}(S^*, R^*)$, if a mortality decline yields an increase in consumption because the lifetime human wealth effect dominates the years-to-consume effect, agents reduce their investment in education. Recall that this happens during the schooling period as Figure 2 shows. In contrast, a decline in mortality after the optimal retirement age leads to more years of schooling. Instead, if $r_h(S^*) = \bar{r}(S^*, R^*)$ –as frequently assumed in the literature– a decline in mortality during the schooling period or during retirement period does not have an impact on the optimal length of schooling.

During the working period, Proposition 3(a) shows that a decline in mortality positively affects education through the second term in Eq. (31), which reflects the effect of a mortality decline at age x_0 on the marginal benefit of

schooling (measured at age x_0).¹² Actually, this is the only component driving the effect of mortality on the length of schooling when $r_h(S^*) = \bar{r}(S^*, R^*)$, but it is not the case whenever $r_h(S^*) \neq \bar{r}(S^*, R^*)$.

Next we use the results obtained in Proposition 3 to derive the total effects.

3.2 Total effects

In this section we complete the analysis studying the total impact of a mortality decline at an arbitrary age on the optimal length of schooling and retirement age. To do so, we combine the direct effects, presented in Section 3.1, with the indirect effects (see Eq. (24)).

Proposition 4 gives under the strict concavity of the expected lifetime utility, $\hat{V}(S^*, R^*)$, the sign of a decline in mortality at an arbitrary age x_0 on the optimal length of schooling and retirement age.

Proposition 4 *Assuming the strict concavity of $\hat{V}(S^*, R^*)$, for the life-cycle model given by (1)-(5),*

(a) *the sign of $\left. \frac{-dS^*}{d\mu(x_0)} \right|_{R=R^*}$ is the same as that of*

$$a(x_0) \frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{dc^*}{dS^*} + \int_{S^*}^{R^*} e^{-r(x-x_0)} \frac{-\partial(p(x)/p(x_0))}{\partial\mu(x_0)} y_S(S^*, x) dx, \quad (32)$$

(b) *and the sign of $\left. \frac{-dR^*}{d\mu(x_0)} \right|_{S=S^*}$ is the same as that of*

$$a(x_0) \frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{dc^*}{dR^*} + \left. \frac{\partial S^*}{\partial R} \right|_{R=R^*} \int_{S^*}^{R^*} e^{-r(x-x_0)} \frac{-\partial(p(x)/p(x_0))}{\partial\mu(x_0)} y_S(S^*, x) dx \quad (33)$$

for all $x_0 \in (0, \omega)$.

Proof. See Appendix D ■

Eqs. (32) and (33) show that the total impact of a decline in mortality at an arbitrary age x_0 on the optimal length of schooling and retirement age is given by three factors: (i) the “lifetime human wealth” effect versus the “years-to-consume” effect, which is reflected by the financial wealth at age x_0 ,

¹²Assuming a flat wage rate per unit of human capital, zero interest rate, and zero depreciation of human capital, the integral term in (31) becomes the additional years worked from age x_0 times the return to education. Then, the increase in the expected total years worked caused from a decline in mortality (at all ages) is $\int_{S^*}^{R^*} \int_u^{R^*} \frac{p(x)}{p(u)} dx d\mu(u)$.

i.e. $a(x_0)$; (ii) the total impact of years of schooling and retirement on the initial consumption; and (iii) the effect of mortality on the marginal benefit of schooling. Factors (i) and (ii) only has an impact when $r_h(S^*) \neq \bar{r}(S^*, R^*)$. The first thing to notice is that the increase in the marginal benefit of schooling due to a decline in mortality always has a positive impact on education, but it is not necessarily true on the optimal retirement age since it depends on the relationship between education and retirement. If S^* and R^* are negatively related, agents anticipate their retirement age and enjoy more leisure time when a decline in mortality causes an increase in the marginal benefit of schooling. In contrast, if S^* and R^* are positively related, agents postpone their retirement age in order to reap the benefits of schooling. This is because in the former alternative the income effect dominates over the substitution effect, whereas in the latter the substitution effect dominates over the income effect.

During the schooling period and the retirement period, the effect of mortality on the marginal benefit of schooling is null. Thereby, in these two periods, the sign of the total impact of a decline in mortality on the optimal length of schooling and retirement age solely depend on the lifetime human wealth effect versus the years-to-consume effect and the total impact of years of schooling and retirement on the initial consumption:

$$\text{sign} \left[\frac{-dS^*}{d\mu(x_0)} \Big|_{R=R^*} \right] = \text{sign} \left[a(x_0) \frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{dc^*}{dS^*} \right], \quad (34)$$

$$\text{sign} \left[\frac{-dR^*}{d\mu(x_0)} \Big|_{S=R^*} \right] = \text{sign} \left[a(x_0) \frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{dc^*}{dR^*} \right], \quad (35)$$

for all $x_0 \leq S^*$ and $x_0 \geq R^*$. On the one side, from (5) and Proposition 2 we know that $a(x_0) < 0$ during the schooling period because the lifetime human wealth effect dominates over the years-to-consume effect, while $a(x_0) > 0$ during the retirement period because the years-to-consume effect dominates over the lifetime human wealth effect. On the other side, combining (17), (19)-(21) we have

$$\frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{dc^*}{dS^*} = (1 - \lambda^R) \frac{r_h(S^*) - \bar{r}(S^*, R^*)}{\sigma(\bar{c})} + \lambda^R \left(r_h(S^*) + \frac{w_S(R^* - S^*)}{w(R^* - S^*)} \right), \quad (36)$$

$$\frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{dc^*}{dR^*} = (1 - \lambda^S) \frac{r_h(S^*) - \bar{r}(S^*, R^*)}{\sigma(\bar{c})} + \lambda^S \left(r_h(S^*) + \frac{w_S(R^* - S^*)}{w(R^* - S^*)} \right), \quad (37)$$

where

$$\lambda^R = \frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{\partial c^*}{\partial R^*} \Big/ \left(\frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{\partial c^*}{\partial R^*} - \frac{w_R(R^* - S^*)}{w(R^* - S^*)} + \delta + \frac{\phi_R(R^*)}{\phi(R^*)} \right), \quad (38)$$

$$\lambda^S = \frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{\partial c^*}{\partial S^*} \Big/ \left(\frac{\psi_S(S^*)}{\psi(S^*)} + \frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{\partial c^*}{\partial S^*} - \frac{f_S(S^*, R^*)}{f(S^*, R^*)} \right). \quad (39)$$

From the second-order conditions, and assuming $\psi(S^*) > 0$, we have that $\lambda^R \in (0, 1]$ and $\lambda^S > 0$. Hence, a marginal increase in the optimal length of schooling unambiguously raises consumption since both $\frac{r_h(S^*) - \bar{r}(S^*, R^*)}{\sigma(\bar{c})}$ and $r_h(S^*) + \frac{w_S(R^* - S^*)}{w(R^* - S^*)}$ are positive. Therefore, like Cai and Lau (2012) a decline in mortality during the schooling period has a negative impact on education, while a decline in mortality during the retirement period has a positive impact on education. On the contrary, a marginal increase in the optimal retirement age has an *a priori* ambiguous effect on the initial consumption, and thus the impact of a decline in mortality on the optimal retirement age is also *a priori* ambiguous. From (39), we have that the sign of the derivative of the optimal retirement on the marginal utility of consumption is

$$\begin{cases} > 0 & \text{If } \frac{1}{\lambda^S} > \frac{1 - \sigma(\bar{c})}{r_h(S^*) - \bar{r}(S^*, R^*)} \left(r_h(S^*) - \frac{\bar{r}(S^*, R^*) + \sigma(\bar{c}) \frac{w_S(R^* - S^*)}{w(R^* - S^*)}}{1 - \sigma(\bar{c})} \right), \\ = 0 & \text{If } \frac{1}{\lambda^S} = \frac{1 - \sigma(\bar{c})}{r_h(S^*) - \bar{r}(S^*, R^*)} \left(r_h(S^*) - \frac{\bar{r}(S^*, R^*) + \sigma(\bar{c}) \frac{w_S(R^* - S^*)}{w(R^* - S^*)}}{1 - \sigma(\bar{c})} \right), \\ < 0 & \text{If } \frac{1}{\lambda^S} < \frac{1 - \sigma(\bar{c})}{r_h(S^*) - \bar{r}(S^*, R^*)} \left(r_h(S^*) - \frac{\bar{r}(S^*, R^*) + \sigma(\bar{c}) \frac{w_S(R^* - S^*)}{w(R^* - S^*)}}{1 - \sigma(\bar{c})} \right). \end{cases} \quad (40)$$

Since the term in parenthesis gives the relationship between education and retirement (see Eq. (22)), it is straightforward to obtain from (40) that whenever the relationship between education and retirement is positive, an increase in the optimal retirement age raises consumption. As a consequence, when $\frac{\partial S^*}{\partial R^*} > 0$, a decline in mortality during the schooling period has a negative impact on retirement, while it is positive when the mortality decline occurs during the retirement period. However, when $\frac{\partial S^*}{\partial R^*} < 0$, both alternatives are possible and it is necessary a numerical analysis.

4 Application to Swedish data

In this Section we study numerically the impact of the epidemiological transition on the optimal years of schooling and retirement age when both variables are endogenous. Our analysis delivers two important results. First, when the

life expectancy rises, our model is capable of producing an optimal decline in retirement age and an increase in years of schooling (Hazan, 2009). Second, since in the earlier stage of mortality transition, a decline in mortality belongs mainly to younger people, whereas in the later stage, a decline in mortality decline has mainly occurred at older ages (see Figure 4), we show that the optimal retirement age stops declining after the cohort born in year 1920 and increases thereafter.

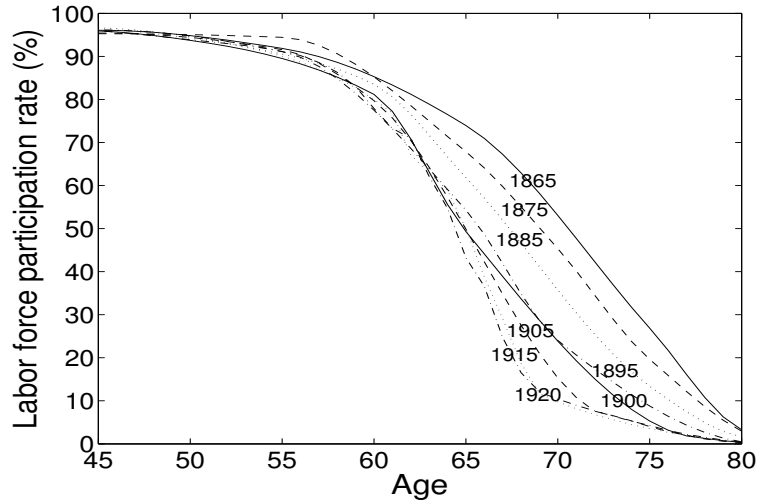
4.1 Data

We are interested in analyzing the marginal effect of the decline in mortality on the length of schooling and retirement age for cohorts born in the 19th and 20th centuries. We implement our analysis to Sweden because there exist reliable economic and demographic information dating back to the 19th century. In particular, we collect information on retirement, schooling, and mortality data for cohorts born between 1865 and 2000.

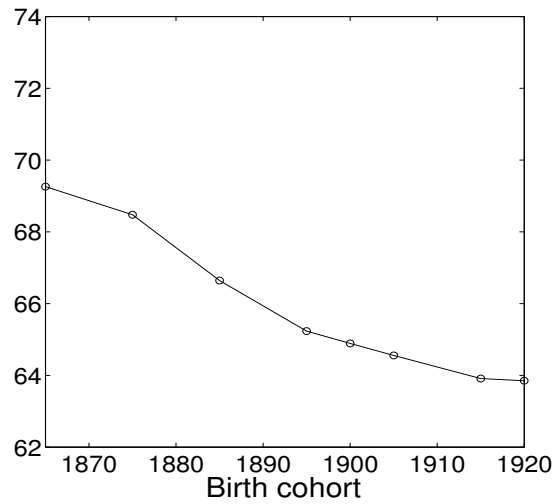
Retirement data. Effective retirement age for men born between years 1865 and 1920 has been calculated using employment rates, based on census data, for the period 1910-1985 and labor force participation rates, taken from labor force surveys, for the period 1975-2004 published by Statistics Sweden.¹³

Figure 3(a) shows that participation rates have fallen at all older ages. For cohort 1865, 85 percent of all sixty-year-olds were in the labor force, for cohort 1900, the participation rates fell to 81 percent, and for cohort 1920 the figures fell to 77 percent, respectively. The decline at older ages has fallen even more sharply. For cohort 1865, 75 percent of all sixty-five-year-olds and 50 percent of all seventy-year-olds were in the labor force, but, for cohort 1920, the participation rates had fallen to 45 and 10 percent, respectively. As a result, the effective retirement age declined from age sixty nine, for cohort 1865, to less than sixty four for the cohort born in year 1920 (see Figure 3(b)). Also, it is worth noticing in Figure 3(b) how the declining trend in retirement age accelerates for cohorts 1875-1895 and decelerates for cohorts born after year 1900.

¹³Employed and working age population in Sweden during the period 1910-1985 is available at <http://www.scb.se/tidsseriehafteforvarvsarbetandefob1910-1985>. Before the 1980s, unemployment rates above age 45 were roughly constant and lower than 2 percent (Ljungqvist and Sargent, 1995; Holmlund, 2009). Therefore, we do not expect a significant error by combining both datasets.



(a) Labour force participation rates by birth cohort (men)



(b) Effective retirement age (men)

Figure 3: Labor force participation rate by age and effective retirement age across cohorts: Sweden, 1865-1920.

NOTE: Labor force participation rates by single year of age are calculated using interpolation techniques.

Schooling data. Information about years of schooling by birth cohort is calculated based on the number of students by educational attainment reported by de la Croix et al. (2008).¹⁴ Their estimates can be decomposed into three groups: (i) primary and lower secondary, (ii) integrated upper secondary, and (iii) tertiary. The average length of schooling has been calculated multiplying the enrollment rate by the expected-years lived in each educational group. To obtain both the enrollment rate and the expected-years lived we make use of the information contained in the Human Mortality Database (2013).

Mortality data. We produce survival probabilities for cohorts born between 1865 and 2000 using Swedish death rates by single-years of age from year 1800 to 2011, taken from the Human Mortality Database (2013). Deaths rates for cohorts born after year 1911 are constructed applying the Lee-Carter model (Lee and Carter, 1992).¹⁵

Figure 4(a) shows the evolution of the conditional survival probabilities from age 6 for cohorts born in 1865, 1900, 1920, 1950, and 2000, whereas Figure 4(b) depicts the corresponding absolute decline in the mortality hazard by age across cohorts. It can be seen in Figure 4(b) how the mortality improvements for cohorts born before 1920 have mainly occurred at young and working ages. The peak for cohorts born between 1865 and 1900 shows the influence of the Spanish flu. However, although mortality improvements for cohorts born after 1920 mainly occur at older ages – in relative terms –, there are still gains in years lived at young ages.

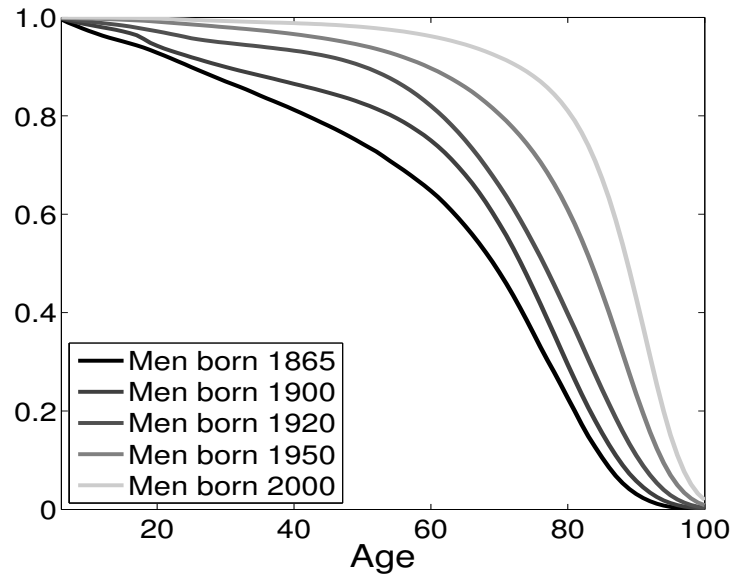
Expected length of work in Sweden. Table 1 summarizes the relevant labor supply statistics for our model based on the information collected for cohorts born between 1865 and 1920. Column two and three report years of schooling and retirement ages across cohorts. Column four shows the significant increase of almost 9 years in the life expectancy (LE) at the age of leaving

¹⁴Information on number of students and enrollment rates by educational attainment in Sweden from year 1768 to 2002 are withdrawn from <http://perso.uclouvain.de/david.delacroix/data/swedish-educ-data.pdf>.

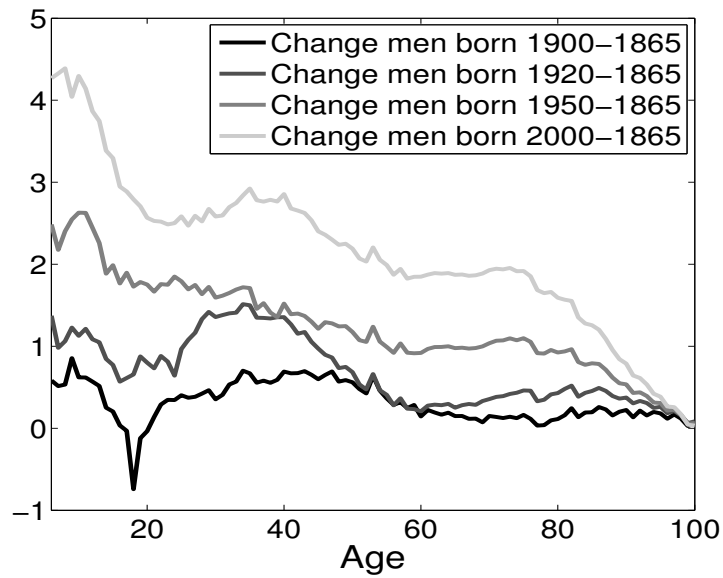
¹⁵Future age-specific death rates from year 2011 are projected applying the Lee-Carter method to actual Swedish period-death rates from year 1945 up to 2011. Thus, it is assumed that the log of the death rate is explained by the following multiplicative process:

$$\log m_{t,x} = a_x + k_t b_x + \epsilon_{t,x}, \text{ where } \epsilon_{t,x} \text{ is } i.i.d.(0, \sigma_\epsilon),$$

where $m_{t,x}$ is the death rate at age x in year t , a_x , b_x are age-specific constants, and k_t is a time-varying index obtained through the singular-value decomposition of a matrix of death rates.



(a) Conditional survival probability at age 6



(b) Change in mortality hazard rate across cohorts

Figure 4: Conditional survival probability at age 6 and changes in the mortality hazard rate across cohorts by age: Sweden, 1865-2000.

school, while column five report the corresponding increase of 6.6 years in the life expectancy from retirement. Like Lee (2001), we obtain that the expected length of work (in years) has been roughly constant and equal to fifty, whereas the last column in Table 1 shows a significant reduction of almost 6 years in the average number of years worked.

Table 1: Estimates of the expected length of work, Sweden, Cohorts: 1865-1920

Birth cohort	Years of schooling (<i>S</i>)	Retirement age (<i>R</i>)	LE from schooling <i>LE</i> (6 + <i>S</i>)	LE from retirement <i>LE</i> (<i>R</i>)	Expected length of work	Number of years worked (<i>R</i>)-(6+ <i>S</i>)
	II	III	IV	V	VI	VII
1865	6.0	69.3	53.4	6.2	50.2	57.3
1875	6.5	68.5	54.1	7.1	50.2	56.0
1875	6.3	66.6	55.0	8.6	49.7	54.3
1895	6.6	65.2	55.9	10.0	49.2	52.6
1900	6.6	64.9	57.1	10.7	49.6	52.3
1905	6.5	64.6	58.6	11.4	50.3	52.1
1915	6.3	63.9	61.2	13.1	51.1	51.6
1920	6.4	63.9	62.1	13.8	51.4	51.5

Source: Own calculations.

Note: LE stands for life expectancy.

4.2 Calibration

To solve the model, we follow Cervellati and Sunde (2013) and consider a constant intertemporal elasticity of substitution (CIES) utility function,

$$U(c) = \frac{c^{1-\frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}}, \text{ with } \sigma \in (0, 1]. \quad (41)$$

The underlying disutility of non-leisure time is assumed constant $\phi(x) = \phi$. For simplicity, we assume a constant relative disutility of school versus work, $\psi(x) = \psi$, for all $x \in (0, S)$. Since our individual devotes her full time to education while she is in school, we use the following simplified version of the Ben-Porath human capital production function

$$q(h(S, x)) = \xi \cdot h(S, x)^\gamma, \text{ with } \xi > 0 \text{ and } \gamma \in (0, 1), \quad (42)$$

where ξ is a scaling factor and γ is the returns to scale in human capital investment. Similar to Hazan (2009) and Cervellati and Sunde (2013) the wage rate is assumed to be constant.

To shed light on the effects of mortality on S^* and R^* , we introduce the next simplifying assumptions. We assume zero discounting, $r = \delta = \rho = 0$, so that the inverse of $\bar{r}(S^*, R^*)$ coincides with the expected length of work (Lee, 2001)

$$\bar{r}(S^*, R^*) = \left[\int_{S^*}^{R^*} \frac{p(x)}{p(S^*)} dx \right]^{-1}. \quad (43)$$

Integrating (4) and using (42) the return to education at age S becomes

$$r_h(S^*) = \xi / (h(0) + (1 - \gamma)\xi S^*). \quad (44)$$

We set $\gamma = 0.65$ as in Cervellati and Sunde (2013). In order to show the importance on the results of the relationship between S^* and R^* , we run six alternative simulations combining three different returns to education function and two different IES. We fix $h(0) = 1$ and set ξ to 1, .25, and 0.075 so as to have a return to education after 13 years in school around 18%, 11.5% (Heckman et al., 2008), and 5.5%.¹⁶ To avoid any spurious result driven by an income effect, in all simulations the labor income faced by the cohort born in 1865 is unitary regardless the return to education function. Thereby, the wage rate per unit of human capital w is set at $1/h(S_{1865}^*, S_{1865}^*)$. To see the role of the IES on our results, we consider two alternative CIES $\sigma = \{0.50, 0.75\}$.¹⁷ Finally, the values of ϕ and ψ are calculated in each simulation so as to have an optimal length of schooling of 6 years and an optimal retirement age of 69.3 years for the cohort born in 1865.

Results. In each of the six simulations we compute the optimal length of schooling (S^*) and optimal retirement age (R^*) for each cohort born between 1865 and 2000. Thereby under this controlled experiment any variation in S^* and R^* across cohorts is solely due to changes in mortality. In addition, to understand the impact of S^* on R^* and of R^* on S^* we also compare the results from the full model to a partial model in which one of the variables is fixed. Hence, differences in S^* , or in R^* , across cohorts between the full model and the partial model are explained by the introduction of an additional endogenous variable.

¹⁶Notice that when $h(0) = 1$, ξ is equal to the initial return to education $r_h(0)$.

¹⁷The importance of the IES value on retirement and human capital investment decisions is frequently stressed in the literature. See Imrohroglu and Kitao (2009) and Rogerson and Wallenius (2009) for an empirical study and Keane (2011) and Keane and Rogerson (2012) for a survey.

Table 2: ABSOLUTE CHANGE IN YEARS OF SCHOOLING AND RETIREMENT AGE FOR SWEDISH MALE COHORTS BORN BETWEEN 1865 AND 2000

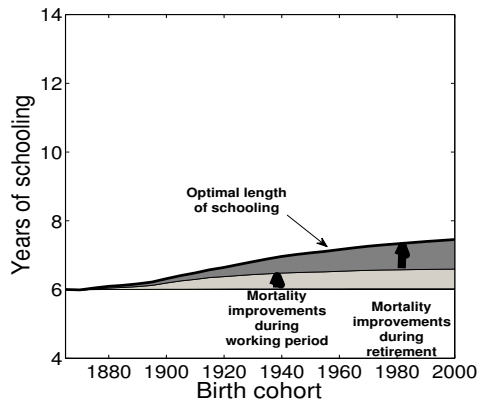
	PARTIAL MODEL		FULL MODEL		
	Endogenous R^*	Endogenous S^*	Endogenous S^* and R^*		
	Fixed S_{1865}^*	Fixed R_{1865}^*	$R_{2000}^* - R_{1865}^*$	$S_{2000}^* - S_{1865}^*$	$L_{2000}^* - L_{1865}^*$
	$R_{2000}^* - R_{1865}^*$	$S_{2000}^* - S_{1865}^*$			
	I	II	III	IV	V=III-IV
<u>CIES ($\sigma = 0.50$)</u>					
$\xi = 1.000$	9.1	1.0	-2.7	1.4	-4.1
$\xi = 0.250$	9.1	2.3	-1.2	2.7	-3.9
$\xi = 0.075$	9.1	7.0	3.4	6.7	-3.3
<u>CIES ($\sigma = 0.75$)</u>					
$\xi = 1.000$	9.1	2.1	1.5	2.0	-0.5
$\xi = 0.250$	9.1	4.1	3.7	3.9	-0.2
$\xi = 0.075$	9.1	8.9	10.5	9.7	0.8

Note: ' S_t^* ' is the optimal length of schooling for cohort born in year t , ' R_t^* ' is the optimal retirement age for cohort born in year t , and L_t^* stands for the number of years worked conditional upon survival by an individual who belongs to cohort t , i.e. $R_t^* - (S_t^* + 6)$. After thirteen years of education, the return to education is 18, 11.5, and 5.5 percent for ξ values 1.000, 0.250, and 0.075, respectively.

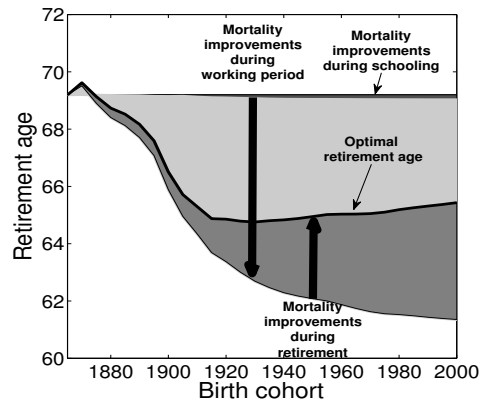
A summary of the impact of the mortality transition on S^* and R^* is given in Table 2. Specifically, Table 2 reports for the full model and the partial models the absolute change in S^* and R^* for male cohorts born in Sweden between year 1865 and 2000. Looking at Table 2 we obtain the following results. First, in a model with fixed length of schooling, the effect of a decline in mortality on retirement is always positive and constant regardless the return to education function and the IES (see column I). Second, in a model with fixed retirement age, the length of schooling increases in all simulations. Nevertheless, smaller differences between returns to education (i.e. $r_h(S^*) - \bar{r}(S^*, R^*) \uparrow 0$) and higher IES boost the positive effect of mortality on the length of schooling. Third, in a model with endogenous schooling and retirement decisions, the impact of mortality on the optimal retirement age changes depending upon the return to education and the IES (see column III in Table 2). Fourth, the results on optimal length of schooling do not significantly change between a model with and without endogenous retirement decisions (c.f. columns II and IV). Fifth, in a full model with endogenous education and retirement decisions, the impact of a decline in mortality on the number of years worked becomes negative whenever the difference between returns to education are high and the IES is low, and vice versa (see column V).

Finally, we conclude the empirical analysis by plotting in Figure 5 the contribution of mortality improvements at different stages of life on the optimal years of schooling and optimal retirement age. In this simulation, we divide the lifespan in three periods: childhood (ages 6-15), adulthood (ages 16-65), and retirement (ages 66+); and calculate the optimal schooling and retirement considering exclusively the gains in survival during either childhood, or adulthood, or retirement across cohorts *ceteris paribus* the survival probability in the other stages of life.

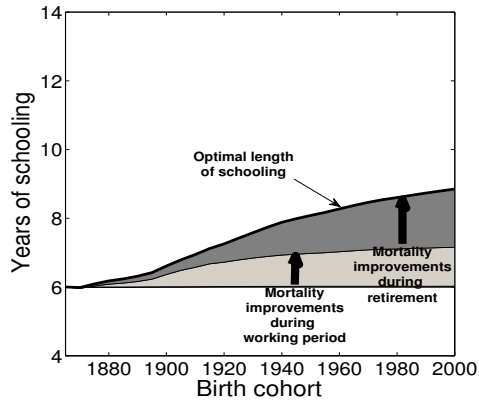
The figures on the left-hand side show how mortality improvements during both the working period (light gray area) and the retirement period (dark gray area) raise the optimal years of schooling, as it is pointed out by the upward arrows. According to Proposition 4, in the first case, a decline in mortality during the working period raises the marginal benefit of schooling because the likelihood of receiving a future labor earning increases –also known as the Ben-Porath’s mechanism–, while, in the second case, a decline in mortality during retirement leads agents to continue studying in order to finance the consumption after retirement through an increase in lifetime earnings –that is, the years-to-consume effect dominates over the lifetime human wealth effect–. The sum of these two positive effects of mortality on the optimal length of



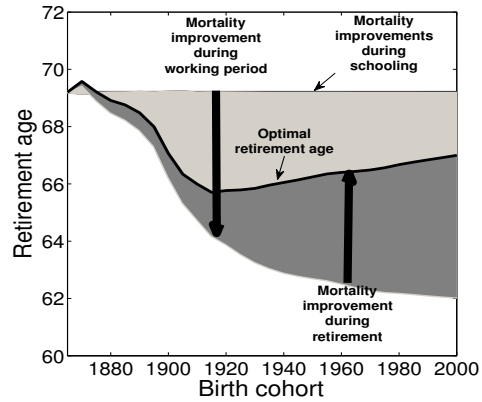
(a) Years of schooling, $\xi = 1.000$



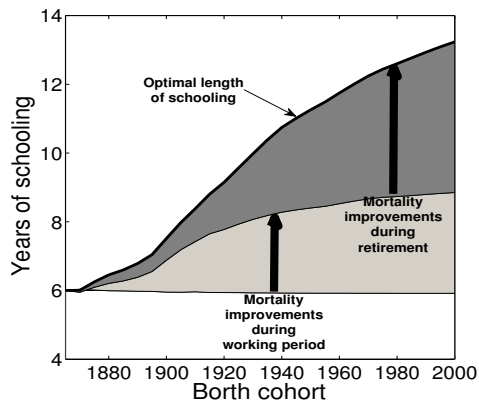
(b) Retirement age, $\xi = 1.000$



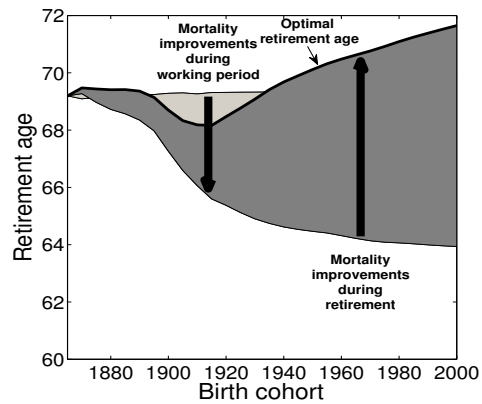
(c) Years of schooling, $\xi = 0.250$



(d) Retirement age, $\xi = 0.250$



(e) Years of schooling, $\xi = 0.075$



(f) Retirement age, $\xi = 0.075$

Figure 5: Contribution of mortality decline by stage of life on the optimal years of schooling and retirement age across cohorts by average return to education, Swedish males cohorts born in 1865-2000

schooling is represented by the black solid line.¹⁸

The figures on the right-hand side show how mortality improvements during childhood, working period (light gray area), and the retirement period (dark gray area) affect the optimal retirement age. On the one hand, since the parameter values assumed imply that the income effect dominates over the substitution effect (i.e. $\frac{\partial S^*}{\partial R^*} < 0$), a decline in mortality during the working period leads to early retirement (see the second term in Eq. 33). Individuals use the additional income to increase consumption and enjoy more leisure time. The extent to which mortality improvements during the working period reduce the optimal retirement age for each cohort is indicated by the downward arrow in Figures 5(b), 5(d), and 5(f). Note that the length of this arrow decreases with lower returns to education. This is because the negative relation between education and retirement vanishes when $r_h(S^*)$ tends to $\bar{r}(S^*, R^*)$ (see Figure 1). Consequently, if the relation between S^* and R^* were positive, a decline in mortality during the working period would lead to a delay in the retirement age. On the other hand, mortality declines late in life delay the optimal retirement age because individuals need more earnings to finance the additional consumption due to a longer retirement period. This positive effect on retirement is indicated by the upward arrow, while the strength is represented by the dark gray area. The net effect of mortality improvements at different stages of life on the optimal retirement age is represented by the black solid line. Notice that in all cases we observe a turning point in the evolution of the optimal retirement age after the cohorts born in the 1920s. This is because mortality improvements for cohorts born before the 1920s mainly occurred at young and working ages, whereas the improvements in mortality for more recent cohorts mainly occur at older ages. Therefore, this finding suggests that the recent increase in retirement in some developed countries might be partially explained by the decline in mortality at older ages as well as by the negative relation between education and retirement.

5 Conclusion

Existing theoretical models derive a causal positive relation between increasing life expectancy and human capital investments and retirement age. However, one salient feature of the economic development during the last two centuries is the negative relation between life expectancy and labor supply (Hazan, 2009). To reconcile the empirical evidence with economic theory, we develop

¹⁸The contribution of mortality improvements during childhood on the optimal length of schooling is almost negligible. For this reason, we opted for not representing the effect.

a lifecycle model with endogenous human capital investment and labor supply in which mortality declines may cause higher schooling and early retirement.

This article makes two important contributions. First, we derive the differential impact of the mortality decline at any arbitrary age on education and retirement. Thereby, our results extend the findings of d’Albis et al. (2012) and Cervellati and Sunde (2013) in a unified framework. We show that a mortality decline at older ages always result in higher human capital investment and may or may not result in early retirement, whereas a mortality decline at young ages leads to lower investments in human capital and may cause late retirement.

Provided data on the Swedish mortality transition for cohorts born between year 1865 and 2000, the second important contribution is to show that the decline in labor supply along with an increasing life expectancy stems from the negative relation between schooling and retirement. The intuition is as follows. Higher human capital investment cause two opposite effects. First, an increase in earnings at all ages which raises the marginal benefit of working. Second, an increase in consumption and hence a decline in the marginal utility of consumption that reduces the marginal benefit of working. The net effect on the optimal retirement age depends upon the strength of the income effect and may, or may not, result in early retirement.

Although our model abstracts from realistic features like the existence of a pension system, the intervention of governments in the access to all levels of education, the introduction of mandatory years of schooling and retirement ages, among many others, our results offer an explanation to the empirical evidence, collected during the last centuries in Sweden, on the evolution of education and retirement. In addition, our model is robust to the introduction of such features. For instance, if education and retirement are negatively related, positive spillovers from education and publicly provided education will increase the marginal benefit of schooling and will reduce even further the retirement age. Similarly, the existence of pension incentives for early retirement and the overall increase in the labor-augmenting technological progress during the last century would have induced earlier retirement ages and higher increases in the marginal benefit of schooling. Therefore, our results suggest some interesting directions for future research. In particular, first, a logical extension of our framework is the consideration of realistic pension systems. Second, the implementation of the model in a computable general equilibrium setting in order to analyze the effect of changes in wages and interest rates. The implementation of these issues can provide researchers and policy-makers a better understanding of the effect of changes in the population on modern economic growth.

A Proof of Proposition 1

We first derive the optimal length of schooling condition (S^*) and optimal retirement age (R^*) introduced in Eqs. (8)-(9). Second, we study the conditions for a maximum in S^* and R^* . Substituting the conditional optimal consumption, $c(x, S, R)$, into (2) and differentiating it with respect to S , we obtain

$$V_S(S, R) = \int_0^\omega e^{-\rho x} p(x) U_c(c(x, S, R)) \frac{\partial c(x, S, R)}{\partial S} dx - e^{-\rho S} p(S) \psi(S). \quad (\text{A.1})$$

Substituting (7) in (A.1), and rearranging, gives

$$V_S(S, R) = U_c(c(0, S, R)) \int_0^\omega e^{-rx} p(x) \frac{\partial c(x, S, R)}{\partial S} dx - e^{-\rho S} p(S) \psi(S). \quad (\text{A.2})$$

Differentiating (6) with respect to S , and simplifying, we have

$$\begin{aligned} \int_0^\omega e^{-rx} p(x) \frac{\partial c(x, S, R)}{\partial S} dx \\ = e^{-rS} p(S) \left(\int_S^R e^{-r(x-S)} \frac{p(x)}{p(S)} y_S(S, x) dx - y(S, S) \right). \end{aligned} \quad (\text{A.3})$$

Substituting (A.3) in (A.2), and taking $U'(c(0, S, R))e^{-rS}p(S)$ as common factor, we obtain

$$\begin{aligned} V_S(S, R) = U_c(c(0, S, R)) e^{-rS} p(S) \cdot \\ \left(\int_S^R e^{-r(x-S)} \frac{p(x)}{p(S)} y_S(S, x) dx - y(S, S) - \frac{e^{(r-\rho)S} \psi(S)}{U_c(c(0, S, R))} \right). \end{aligned} \quad (\text{A.4})$$

Setting $V_S(S, R)$ to zero and simplifying, the first-order condition for an optimal retirement age is given by (8).

Applying a similar approach, the derivative of (2) with respect to R becomes

$$V_R(S, R) = p(R) \left(U_c(c(0, S, R)) e^{-rR} y(S, R) - e^{-\rho R} \phi(R) \right). \quad (\text{A.5})$$

Then, setting $V_R(S, R)$ to zero and simplifying, the first-order condition for an optimal retirement age is given by (9).

Let $\hat{V}(S, R)$ be the expected lifetime utility conditional on the optimal consumption path. Also, let c^* be optimal initial consumption condition on

$S = S^*$ and $R = R^*$. Since \hat{V} is strictly concave, it satisfies at $S = S^*$ and $R = R^*$

$$\hat{V}_{SS} < 0, \hat{V}_{RR} < 0, \quad (\text{A.6a})$$

$$\begin{vmatrix} \hat{V}_{SS} & \hat{V}_{SR} \\ \hat{V}_{RS} & \hat{V}_{RR} \end{vmatrix} > 0. \quad (\text{A.6b})$$

Substituting (10) in (A.4), differentiating with respect to S , using (15), and simplifying gives

$$\begin{aligned} \hat{V}_{SS}(S, R^*) \Big|_{S=S^*} &= U_c(c^*) e^{-rS^*} p(S^*) \\ &\cdot \left(f_S(S^*, R^*) - f(S^*, R^*) \left(\frac{\psi_S(S^*)}{\psi(S^*)} - \frac{U_{cc}(c^*)}{U_c(c^*)} \frac{\partial c^*}{\partial S} \right) \right). \end{aligned} \quad (\text{A.7})$$

Since the $\text{sign} \left[\frac{\psi_S(S^*)}{\psi(S^*)} \right]$ and $\text{sign} \left[\frac{\partial c^*}{\partial S} \right]$ is the same as the sign of $f(S^*, R^*)$, a necessary and sufficient condition for $S^* < R^*$ to be a maximum of $\hat{V}(S, R^*)$ is

$$f_S(S^*, R^*) < f(S^*, R^*) \left(\frac{\psi_S(S^*)}{\psi(S^*)} - \frac{U_{cc}(c^*)}{U_c(c^*)} \frac{\partial c^*}{\partial S} \right). \quad (\text{A.8})$$

For $f(S^*, R^*) > 0$ notice that $\frac{f_S(S^*, R^*)}{f(S^*, R^*)} - \frac{\psi_{S^*}(S^*)}{\psi(S^*)} < -\frac{U_{cc}(c^*)}{U_c(c^*)} \frac{\partial c^*}{\partial S}$. It is also worth noticing that for $f(S^*, R^*) = 0$, $f_S(S^*, R^*) < 0$.

Differentiating (A.5) with respect to R , at $S = S^*$ and $R = R^*$, gives

$$\begin{aligned} \hat{V}_{RR}(S^*, R) \Big|_{R=R^*} &= U_c(c^*) e^{-rR^*} p(R^*) y(S^*, R^*) \\ &\cdot \left(\frac{U_{cc}(c^*)}{U_c(c^*)} \frac{\partial c^*}{\partial R} + \frac{y_R(S^*, R^*)}{y(S^*, R^*)} - \frac{\phi_R(R^*)}{\phi(R^*)} \right). \end{aligned} \quad (\text{A.9})$$

Provided $y_R(S^*, R^*) \leq 0$ for all $R^* \geq \underline{x}$ and since $\phi_x(x) \geq 0$, $\hat{V}_{RR}(S^*, R) \Big|_{R=R^*} < 0$ is strictly negative. Finally, the second-order conditions of the maximum in $S = S^*$ and $R = R^*$ is satisfied, if (A.8) holds, $y_R(S^*, R^*) \leq 0$, and

$$\frac{\hat{V}_{SR}(S, R) \Big|_{S=S^*, R=R^*}}{-\hat{V}_{SS}(S, R) \Big|_{S=S^*, R=R^*}} \cdot \frac{\hat{V}_{RS}(S, R) \Big|_{S=S^*, R=R^*}}{-\hat{V}_{RR}(S, R) \Big|_{S=S^*, R=R^*}} < 1. \quad (\text{A.10})$$

B Proof of Proposition 2

If $U_x(x) = x^{-\frac{1}{\sigma}}$, where σ is the intertemporal elasticity of substitution, from (7) we have

$$c(x, S, R) = c(0, S, R) e^{\sigma(r-\rho)x} \text{ for all } x \in (0, \omega). \quad (\text{B.1})$$

Substituting (B.1) in (6), and simplifying, we obtain

$$c(0, S, R) = \frac{\int_S^R e^{-rx} p(x) y(S, x) dx}{\int_0^\omega e^{-[(1-\sigma)r+\sigma\rho]x} p(x) dx}. \quad (\text{B.2})$$

Taking logarithms at both sides of (B.2) and differentiating with respect to $-\mu(x_0)$ gives

$$\begin{aligned} \frac{1}{c(0, S, R)} \frac{-\partial c(0, S, R)}{\partial \mu(x_0)} &= \frac{\int_S^R e^{-rx} \left[\frac{-\partial p(x)}{\partial \mu(x_0)} \right] y(S, x) dx}{\int_S^R e^{-rx} p(x) y(S, x) dx} - \frac{\int_{x_0}^\omega e^{-[(1-\sigma)r+\sigma\rho]x} p(x) dx}{\int_0^\omega e^{-[(1-\sigma)r+\sigma\rho]x} p(x) dx}, \end{aligned} \quad (\text{B.3})$$

Note that the right-hand side of (B.3) at $S = S^*$ and $R = R^*$ is (29). Now, from (29), we obtain $g(S^*) > 0$, $g(R^*) < 0$,

$$g'(x_0) = -\frac{e^{-rx_0} p(x_0) y(S^*, x_0)}{\int_{S^*}^{R^*} e^{-rx} p(x) y(S^*, x) dx} + \frac{e^{-[(1-\sigma)r+\sigma\rho]x_0} p(x_0)}{\int_0^\omega e^{-[(1-\sigma)r+\sigma\rho]x} p(x) dx}, \quad (\text{B.4})$$

and

$$\begin{aligned} g''(x_0) &= \frac{\left[r + \mu(x_0) - \frac{y_{x_0}(S^*, x_0)}{y(S^*, x_0)} \right] e^{-rx_0} p(x_0) y(S^*, x_0)}{\int_{S^*}^{R^*} e^{-rx} p(x) y(S^*, x) dx} \\ &\quad - \frac{[(1-\sigma)r + \sigma\rho + \mu(x_0)] e^{-[(1-\sigma)r+\sigma\rho]x_0} p(x_0)}{\int_0^\omega e^{-[(1-\sigma)r+\sigma\rho]x} p(x) dx}, \end{aligned} \quad (\text{B.5})$$

for any x_0 within the interval (S^*, R^*) . Since $g(\cdot)$ is a continuous function in (S^*, R^*) , $g(S^*) > 0$, and $g(R^*) < 0$ imply that there exists at least a critical age x_c within the interval (S^*, R^*) such that $g(x_c) = 0$.

In order to prove that x_c is unique, we show that there exists only one local optimum in the interval (S^*, R^*) . At a local optimum (denoted by \tilde{x}_0 , with $g'(\tilde{x}_0) = 0$), from (B.4) and (B.5) we obtain

$$g''(\tilde{x}_0) = \frac{\left[\sigma(r - \rho) - \frac{y_{x_0}(S^*, x_0)}{y(S^*, x_0)} \right] e^{-r\tilde{x}_0} p(\tilde{x}_0) y(S^*, \tilde{x}_0)}{\int_{S^*}^{R^*} e^{-rx} p(x) y(S^*, x) dx}. \quad (\text{B.6})$$

Let \tilde{x}_0^i and \tilde{x}_0^{ii} be two possible candidates, which satisfy that $g'(\tilde{x}_0^i) = 0$ and $g'(\tilde{x}_0^{ii}) = 0$ with $\tilde{x}_0^i < \tilde{x}_0^{ii}$. Provided that $y(S^*, x)$ is strictly concave within the interval (S^*, R^*) , \tilde{x}_0 is unique ($\tilde{x}_0^i = \tilde{x}_0^{ii}$) either because $\sigma(r - \rho) - \frac{y_{x_0}(S^*, x_0)}{y(S^*, x_0)} < 0$ or $\sigma(r - \rho) - \frac{y_{x_0}(S^*, x_0)}{y(S^*, x_0)} > 0$ for all $\tilde{x}_0 \in (S^*, R^*)$, or \tilde{x}_0^i is a local maximum and \tilde{x}_0^{ii} a local minimum, which proves that x_c is unique.

C Proof of Proposition 3

Totally differentiating $S^*(R^*; \boldsymbol{\mu})$ with respect to a mortality decline at an arbitrary age x_0 , $-\mu(x_0)$, and the optimal retirement age R^* , using (A.7), and simplifying, we obtain

$$\frac{-\partial S^*}{\partial \mu(x_0)} = \frac{\frac{-\partial f(S^*, R^*)}{\partial \mu(x_0)} - \frac{f(S^*, R^*)}{\sigma(c^*)} \frac{1}{c^*} \frac{-\partial c^*}{\partial \mu(x_0)}}{f(S^*, R^*) \left(\frac{\psi_S(S^*)}{\psi(S^*)} + \frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{\partial c^*}{\partial S} \right) - f_S(S^*, R^*)}. \quad (\text{C.1})$$

Thus, from (A.8) we have

$$\text{sign} \left[\frac{-\partial S^*}{\partial \mu(x_0)} \right] = \text{sign} \left[\frac{-\partial f(S^*, R^*)}{\partial \mu(x_0)} - \frac{f(S^*, R^*)}{\sigma(c^*)} \frac{1}{c^*} \frac{-\partial c^*}{\partial \mu(x_0)} \right]. \quad (\text{C.2})$$

Note from (25) that the first term on the right-hand side of (C.2) is zero whenever $x_0 \leq S^*$ or $x_0 \geq R^*$. Substituting (11)-(12) and (26) in (C.2), and taking $e^{-r(x_0 - S^*)} \frac{p(x_0)}{p(S^*)}$ as common factor gives

$$\text{sign} \left[\frac{-\partial S^*}{\partial \mu(x_0)} \right] = \text{sign} \left[\frac{r_h(S^*) - \bar{r}(S^*, R^*)}{\sigma(\bar{c})} a(x_0) \right]. \quad (\text{C.3})$$

when $x_0 \leq S^*$ and $x_0 \geq R^*$, and

$$\text{sign} \left[\frac{-\partial S^*}{\partial \mu(x_0)} \right] = \text{sign} \left[\frac{r_h(S^*) - \bar{r}(S^*, R^*)}{\sigma(\bar{c})} a(x_0) + \int_{x_0}^{R^*} e^{-r(x-x_0)} \frac{p(x)}{p(x_0)} y_S(S^*, x) dx \right]. \quad (\text{C.4})$$

when $x_0 \in (S^*, R^*)$. This proves Proposition 3(a).

Similarly, totally differentiating $R^*(S^*; \boldsymbol{\mu})$ with respect to a mortality decline at an arbitrary age x_0 , $-\mu(x_0)$, and the optimal length of schooling S^* , and using (A.9), we obtain

$$\frac{-\partial R^*}{\partial \mu(x_0)} = \frac{-\frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{-\partial c^*}{\partial \mu(x_0)}}{\frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{\partial c^*}{\partial R} - \frac{w_R(R^* - S^*)}{w(R^* - S^*)} + \delta + \frac{\phi_R(R^*)}{\phi(R^*)}}. \quad (\text{C.5})$$

From (A.9), we have

$$\text{sign} \left[\frac{-\partial R^*}{\partial \mu(x_0)} \right] = -\text{sign} \left[\frac{-\partial c^*}{\partial \mu(x_0)} \right]. \quad (\text{C.6})$$

Now, in order to prove Proposition (3)(b) we show that the sign of $\frac{-\partial R^*}{\partial \mu(x_0)}$ is that of $a(x_0)$.

Differentiating (6) at $S = S^*$ and $R = R^*$ with respect to $-\mu(x_0)$, and rearranging, gives

$$\begin{aligned} \int_{x_0}^{\omega} e^{-rx} p(x) c(x, S^*, R^*) dx + \int_0^{\omega} e^{-rx} p(x) \frac{-\partial c(x, S^*, R^*)}{\partial \mu(x_0)} dx \\ = \int_{S^*}^{R^*} e^{-rx} \frac{-\partial p(x)}{\partial \mu(x_0)} w(x - S^*) h(S^*, x) dx. \end{aligned} \quad (\text{C.7})$$

The intertemporal budget constraint at age x_0 can be expressed as

$$\begin{aligned} e^{-rx_0} p(x_0) a(x_0) \\ = \begin{cases} \int_{x_0}^{\omega} e^{-rx} p(x) c(x, S^*, R^*) dx & \text{if } x_0 \leq S^*, \\ \int_{x_0}^{\omega} e^{-rx} p(x) c(x, S^*, R^*) dx - \int_{x_0}^{R^*} e^{-rx} p(x) w(x - S^*) h(S^*, x) dx & \text{if } S^* < x_0 < R^*, \\ \int_{x_0}^{\omega} e^{-rx} p(x) c(x, S^*, R^*) dx & \text{if } x_0 \geq R^*. \end{cases} \end{aligned} \quad (\text{C.8})$$

Substituting (C.8) into (C.7), we obtain

$$\int_0^{\omega} e^{-rx} p(x) \frac{-\partial c(x, S^*, R^*)}{\partial \mu(x_0)} dx = -e^{-rx_0} p(x_0) a(x_0). \quad (\text{C.9})$$

Differentiating (7) at $S = S^*$ and $R = R^*$ with respect to $-\mu(x_0)$, and simplifying, gives

$$\frac{1}{c(x, S^*, R^*)} \frac{-\partial c(x, S^*, R^*)}{\partial \mu(x_0)} = \frac{\sigma(c(x, S^*, R^*))}{\sigma(c^*)} \frac{1}{c^*} \frac{-\partial c^*}{\partial \mu(x_0)}. \quad (\text{C.10})$$

Substituting (C.10) into (C.9), and rearranging, we get (26) and thus (28). Finally, substituting (28) in (C.6), we have

$$\text{sign} \left[\frac{-\partial R^*}{\partial \mu(x_0)} \right] = -\text{sign} \left[\frac{-\partial c^*}{\partial \mu(x_0)} \right] = \text{sign} [a(x_0)], \quad (\text{C.11})$$

which proves Proposition 3(b).

D Proof of Proposition 4

At $S = S^*$ and $R = R^*$, Eqs. (23a) and (23b) are equivalent to

$$\frac{-dS^*}{d\mu(x_0)} = \frac{\frac{-\partial \hat{V}_S(S^*, R^*)}{\partial \mu(x_0)}}{-\hat{V}_{SS}(S^*, R^*)} + \frac{\hat{V}_{SR}(S^*, R^*)}{-\hat{V}_{SS}(S^*, R^*) - \hat{V}_{RR}(S^*, R^*)} \frac{\frac{-\partial \hat{V}_R(S^*, R^*)}{\partial \mu(x_0)}}{-\hat{V}_{RR}(S^*, R^*)}, \quad (\text{D.1})$$

$$\frac{-dR^*}{d\mu(x_0)} = \frac{\frac{-\partial \hat{V}_R(S^*, R^*)}{\partial \mu(x_0)}}{-\hat{V}_{RR}(S^*, R^*)} + \frac{\hat{V}_{RS}(S^*, R^*)}{-\hat{V}_{RR}(S^*, R^*) - \hat{V}_{SS}(S^*, R^*)} \frac{\frac{-\partial \hat{V}_S(S^*, R^*)}{\partial \mu(x_0)}}{-\hat{V}_{SS}(S^*, R^*)}. \quad (\text{D.2})$$

Differentiating (A.4) and (A.5) with respect to $-\mu(x_0)$, respectively, we have

$$\frac{-\partial \hat{V}_S(S^*, R^*)}{\partial \mu(x_0)} = U_c(c^*) e^{-rS^*} p(S^*) \left(\frac{-\partial f(S^*, R^*)}{\partial \mu(x_0)} - \frac{f(S^*, R^*)}{\sigma(c^*)} \frac{1}{c^*} \frac{-\partial c^*}{\partial \mu(x_0)} \right), \quad (\text{D.3})$$

$$\frac{-\partial \hat{V}_R(S^*, R^*)}{\partial \mu(x_0)} = U_c(c^*) e^{-rR^*} p(R^*) y(S^*, R^*) \left(-\frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{-\partial c^*}{\partial \mu(x_0)} \right). \quad (\text{D.4})$$

We first derive the total impact of a decline in mortality on the optimal length of schooling. Substituting (D.3)-(D.4) in (D.1), taking $\frac{U_c(c^*) e^{-rS^*} p(S^*)}{-\hat{V}_{SS}(S^*, R^*)}$ as common factor, and rearranging gives

$$\begin{aligned} \frac{-dS^*}{d\mu(x_0)} &= \frac{U_c(c^*) e^{-rS^*} p(S^*)}{-\hat{V}_{SS}(S^*, R^*)} \\ &\cdot \left(\frac{-\partial f(S^*, R^*)}{\partial \mu(x_0)} - \frac{f(S^*, R^*)}{\sigma(c^*)} \frac{1}{c^*} \frac{-\partial c^*}{\partial \mu(x_0)} \left(1 + \frac{dR^*}{dS} \frac{e^{-rR^*} p(R^*) y(S^*, R^*)}{e^{-rS^*} p(S^*) f(S^*, R^*)} \right) \right). \end{aligned} \quad (\text{D.5})$$

From (17)-(19) we have $\frac{\partial c^*}{\partial R} / \frac{\partial c^*}{\partial S} = \frac{e^{-rR^*} p(R^*) y(S^*, R^*)}{W(S^*, R^*) (r_h(S^*) - \bar{r}(S^*, R^*))}$, taking $\left(\frac{\partial c^*}{\partial S}\right)^{-1}$ as common factor, and rearranging gives

$$\begin{aligned} \frac{-dS^*}{d\mu(x_0)} &= \frac{U_c(c^*) e^{-rS^*} p(S^*)}{-\hat{V}_{SS}(S^*, R^*)} \\ &\cdot \left(\frac{-\partial f(S^*, R^*)}{\partial \mu(x_0)} - \frac{f(S^*, R^*)}{\sigma(c^*)} \frac{1}{c^*} \frac{-\partial c^*}{\partial \mu(x_0)} \left(\frac{dc^*}{dS} / \frac{\partial c^*}{\partial S} \right) \right). \end{aligned} \quad (\text{D.6})$$

Substituting (11), (17), (26), and (A.7) in (D.6), and taking $e^{-r(x_0 - S^*)} \frac{p(x_0)}{p(S^*)}$ as common factor gives

$$\begin{aligned} \frac{-dS^*}{d\mu(x_0)} &= e^{-r(x_0 - S^*)} \frac{p(x_0)}{p(S^*)} \\ &\frac{a(x_0) \frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{dc^*}{dS} + \int_{S^*}^{R^*} e^{-r(x-x_0)} \frac{-\partial}{\partial \mu(x_0)} \left[\frac{p(x)}{p(x_0)} \right] y_S(S^*, x) dx}{f(S^*, R^*) \left(\frac{\psi_S(S^*)}{\psi(S^*)} - \frac{U_{cc}(c^*)}{U_c(c^*)} \frac{\partial c^*}{\partial S} \right) - f_S(S^*, R^*)}. \end{aligned} \quad (\text{D.7})$$

Therefore, the sign of $\frac{-dS^*}{d\mu(x_0)}$ gives (32).

Second, we derive the total impact of a decline in mortality on the optimal retirement age. We apply the same steps as in $\frac{-dS^*}{d\mu(x_0)}$. Substituting (D.3)-(D.4)

in (D.2), taking $\frac{U_c(c^*)e^{-rR^*}p(R^*)y(S^*,R^*)}{-\hat{V}_{RR}}$ as common factor, and rearranging gives

$$\begin{aligned} \frac{-dR^*}{d\mu(x_0)} = & \frac{U_c(c^*)e^{-rR^*}p(R^*)y(S^*,R^*)}{-\hat{V}_{RR}(S^*,R^*)} \left(-\frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{-\partial c^*}{\partial \mu(x_0)} \left(\frac{dc^*}{dR} \Big/ \frac{\partial c^*}{\partial R} \right) \right. \\ & \left. + \frac{\partial S^*}{\partial R} \frac{e^{-rS^*}p(S^*)}{e^{-rR^*}p(R^*)y(S^*,R^*)} \frac{-\partial f(S^*,R^*)}{\partial \mu(x_0)} \right). \quad (\text{D.8}) \end{aligned}$$

Substituting (11), (26), and (A.9) in (D.8), taking $\frac{e^{-rx_0}p(x_0)}{W(S^*,R^*)}$ as common factor, and rearranging gives

$$\begin{aligned} \frac{-dR^*}{d\mu(x_0)} = & \frac{\frac{e^{-rx_0}p(x_0)}{W(S^*,R^*)}}{\left(-\frac{U_{cc}(c^*)}{U_c(c^*)} \frac{\partial c^*}{\partial R} - \frac{y_R(S^*,R^*)}{y(S^*,R^*)} + \frac{\phi_R(R^*)}{\phi(R^*)} \right)} \left(a(x_0) \left(\frac{1}{\sigma(\bar{c})} \frac{dc^*}{dR} \Big/ \frac{\partial c^*}{\partial R} \right) \right. \\ & \left. + \frac{\partial S^*}{\partial R} \frac{W(S^*,R^*)}{e^{-rR^*}p(R^*)y(S^*,R^*)} \int_{S^*}^{R^*} e^{-r(x-x_0)} \frac{-\partial}{\partial \mu(x_0)} \left[\frac{p(x)}{p(x_0)} \right] y_S(S^*,x) dx \right). \quad (\text{D.9}) \end{aligned}$$

Using (19) in (D.9), and taking $\left(\frac{\partial c^*}{\partial R}\right)^{-1}$ as common factor, and after rearranging we obtain

$$\begin{aligned} \frac{-dR^*}{d\mu(x_0)} = & e^{-r(x_0-R^*)} \frac{p(x_0)}{p(R^*)} \\ & \cdot \frac{a(x_0) \frac{1}{\sigma(c^*)} \frac{1}{c^*} \frac{dc^*}{dR} + \frac{\partial S^*}{\partial R} \int_{S^*}^{R^*} e^{-r(x-x_0)} \frac{-\partial}{\partial \mu(x_0)} \left[\frac{p(x)}{p(x_0)} \right] y_S(S^*,x) dx}{y(S^*,R^*) \left(-\frac{U_{cc}(c^*)}{U_c(c^*)} \frac{\partial c^*}{\partial R} - \frac{y_R(S^*,R^*)}{y(S^*,R^*)} + \frac{\phi_R(R^*)}{\phi(R^*)} \right)}. \quad (\text{D.10}) \end{aligned}$$

where the sign of (D.10) is (33).

References

- Barro, R. J., Lee, J. W., 2013. A new data set of educational attainment in the world, 1950-2010. *Journal of Development Economics* 104, 184–198.
- Becker, G. S., 1967. *Human capital and the personal distribution of income*. Ann Arbor Michigan: University of Michigan Press.
- Belley, P., Lochner, L., 2007. The Changing Role of Family Income and Ability in Determining Educational Achievement. *Journal of Human Capital* 1 (1), 37–89.
- Benavot, A., Riddle, P., 1988. The expansion of primary education, 1870-1940: Trends and Issues. *Sociology of Education* 61 (3), 191–210.
- Ben-Porath, Y., 1967. The production of human capital and the life cycle of earnings. *The Journal of Political Economy* 75 (4), 352–365.
- Bloom, D. E., Canning, D., Mansfield, R. K., Moore, M., 2007. Demographic change, social security systems, and savings. *Journal of Monetary Economics* 54, 92–114.
- Bloom, D. E., Canning, D., Moore, M. 2014. Optimal Retirement with Increasing Longevity. *The Scandinavian Journal of Economics*. doi: 10.1111/sjoe.12060
- Boucekkine, R., de la Croix, D., Licandro, O., 2002. Vintage human capital, demographic trends, and endogenous growth. *Journal of Economic Theory* 104, 340–375.
- Boucekkine, R., de la Croix, D., Peeters, D., 2007. Early literacy achievements, population density and the transition to modern growth. *Journal of the European Economic Association* 5:183–226.
- Cai, Z., Lau, P., 2012. Does longevity improvement always raise the length of schooling through the longer-horizon mechanism? *Society for Economic Dynamics*, 2013 Meeting Papers, number 292.
- Card, D., 1999. The causal effect of education on earnings. In O. C. Ashenfelter and D. Card (Eds.), *Handbook of labor economics* (1st ed., Vol. 3, pp. 1801–1863) Amsterdam: Elsevier.
- Card, D., 2001. Estimating the return to education: Progress on some persistent economic problems. *Econometrica* 69 (5), 1127–1160.

- Carneiro, P., Heckman, J. J., 2002. The evidence on credit constraints in post-secondary schooling. *Economic Journal* 112 (482), 705–734.
- Carneiro, P., Hansen, K. T., Heckman, J. J., 2003. 2001 Lawrence R. Klein Lecture Estimating Distributions Of Treatment Effects With An Application To The returns to education And Measurement Of The Effects Of Uncertainty On College Choice. *International Economic Review* 44 (2), 631–422.
- Cervellati, M., Sunde, U., 2005. Human capital formation, life expectancy, and the process of development. *American Economic Review* 95, 1653–1672.
- Cervellati, M., Sunde, U., 2013. Life expectancy, schooling, and lifetime labor supply: Theory and evidence revisited. *Econometrica* 81 (5), 2055–2086.
- Cipolla, C., 1969. *Literacy and Development in the West*. Baltimore: Penguin Books.
- Costa, D. L., 1998. The evolution of retirement. In *The Evolution of Retirement: An American Economic History, 1880-1990*. University of Chicago Press, 6–31.
- Cunha, F., Heckman, J., Navarro, S., 2005. Separating Uncertainty From Heterogeneity In Life Cycle Earnings. *Oxford Economic Papers* 2005 57 (2), 191–261.
- Cutler, D., Deaton, A., Lleras-Muney, A., 2006. The determinant of mortality. *Journal of Economic Perspectives* 20 (3), 97–120.
- d’Albis, H., Augeraud-Véron, E., 2008. Endogenous retirement and monetary cycles. *Mathematical Population Studies* 15, 214–229.
- d’Albis, H., Lau, P., Sanchez-Romero, M., 2012. Mortality transition and differential incentives for early retirement. *Journal of Economic Theory* 147 (1), 261–283.
- de la Croix, D., Licandro, O., 1999. Life expectancy and endogenous growth. *Economics Letters* 65, 255–263.
- de la Croix, D., Lindh, T., Malmberg, B., 2008. Swedish economic growth and education since 1800. *Canadian Journal of Economics* 41 (1), 166–185.
- Echevarria, C. A., 2004. Life expectancy, schooling time, retirement, and growth. *Economic Inquiry* 42 (4), 602–617.

- Echevarria, C. A., Iza, A., 2006. Life expectancy, human capital, social security and growth. *Journal of Public Economics* 90, 2323–2349.
- Ferreira, P. C., Pessoa, S. A., 2007. The effects of longevity and distortions on education and retirement. *Review of Economic Dynamics* 10 (3), 472–493.
- Hazan, M., Zoabi, H., 2006. Does longevity cause growth? A theoretical critique. *Journal of Economic Growth*, 11 (4), 363–376.
- Hazan, M., 2009. Longevity and lifetime labor supply: Evidence and implications. *Econometrica* 77 (6), 1829–1863.
- Heckman, J. J., Lochner, L. J., Todd, P. E., 2006. Earnings Functions, Rates of Return and Treatment Effects: The Mincer Equation and Beyond. In: E. Hanushek and F. Welch, Editor(s), *Handbook of the Economics of Education* 1, 307–458.
- Heckman, J. J., Lochner, L. J., Todd, P. E., 2008. Earnings Functions and Rates of Return. *Journal of Human Capital* 2(1), 1–31.
- Heijdra, B. J., Romp, W. E., 2009. Retirement, pensions, and ageing. *Journal of Public Economics* 93, 586–604.
- Holmlund, B., 2009. The Swedish unemployment experience. *Oxford review of economic policy*, 25 (1), 109–125.
- Hu, S. C., 1999. Economic growth in the perpetual-youth model: Implications of the annuity market and demographics. *Journal of Macroeconomics* 21 (1), 107–124.
- Human Mortality Database, 2013. University of California, Berkeley (USA), and Max Planck Institute for Demography Research (Germany). Available at www.mortality.org or www.humanmortality.de (data download on February 2013).
- Imrohoroglu, S., Kitao, S., 2009. Labor supply elasticity and social security reform. *Journal of Public Economics* 93, 867–878.
- Jayachandran, S., Lleras-Muney, A., 2009. Life expectancy and human capital investments: Evidence from maternal mortality declines. *The Quarterly Journal of Economics* 124 (1), 349–397.
- Kalemli-Ozcan, S., Weil, D., 2010. Mortality change, the uncertainty effect, and retirement. *Journal of Economic Growth* 15 (1), 65–91.

- Kalemli-Ozcan, S., Ryder, H. E., Weil, D. N., 2000. Mortality decline, human capital investment, and economic growth. *Journal of Development Economics* 62, 1–23.
- Keane, M. P., 2011. Labor supply and taxes: a survey. *Journal of Economic Literature* 49 (4), 961–1075.
- Keane, M.P., Rogerson, R., 2012. Micro and macro labor supply elasticities: a reassessment of conventional wisdom. *Journal of Economic Literature* 50 (2), 464–476.
- Lee, R. D., 1994. *Demography of Aging*. National Academy Press, Washington, D.C., Ch. The formal demography of population aging, transfers, and the economic life cycle, pp. 8–49.
- Lee, C., 2001. The expected length of male retirement in the United States. *Journal of Population Economics* 14, 641–650.
- Lee, R., 2003. The demographic transition: three centuries of fundamental change. *The Journal of Economic Perspectives*, 17 (4), 167–190.
- Lee, R. D., Carter, L. R., 1992. Modeling and forecasting U. S. mortality. *Journal of the American Statistical Association* 87 (419), 659–671.
- Ljungqvist, L., Sargent, T. J., 1995. The Swedish unemployment experience. *European Economic Review*, 39 (5), 1043–1070.
- Ludwig, A., Vogel, E., 2010. Mortality, fertility, education and capital accumulation in a simple OLG economy. *Journal of Population Economics* 23, 703–735.
- Maddison, A., 2007. *The world economy volume 1: A millennial perspective volume 2: Historical statistics*. Academic Foundation.
- Mincer, J., 1974. *Schooling, Experience and Earnings*. New York: Columbia University Press.
- OECD (2009). *Society at a Glance 2009: OECD Social Indicators*. OECD Publishing.
- Oeppen, J., Vaupel, J. W., 2002. Broken limits to life expectancy. *Science*, 296 (5570), 1029–1031.

- Oster, E., Shoulson, I., Dorsey, E., 2013. Limited life expectancy, human capital and health investments. *The American Economic Review* 103 (5), 1977–2002.
- Rogerson, R., Wallenius, J., 2009. Micro and Macro Elasticities in a Life Cycle Model With Taxes. *Journal of Economic Theory*, 144, 2277–2292.
- Ryder, H.E., Heal, G.M., 1973. Optimal growth with intertemporally dependent preferences. *Review of Economic Studies* 40, 1–31.
- Sheshinski, E., 1978. A model of social security and retirement decisions. *Journal of Public Economics* 10, 337–360.
- Sheshinski, E., 2008. *The Economic Theory of Annuities*. Princeton, NJ: Princeton University Press.
- Schieber, S. J., Shoven, J. B. 1999. *The Real Deal: The History and Future of Social Security*. London: Yale University Press.
- Soares, R. R., 2005. Mortality reductions, educational attainment, and fertility choice. *American Economic Review* 95 (3), 580–601.
- Strulik, H., Werner, K., 2012. Life expectancy, labor supply, and long-run growth: Reconciling theory and evidence, university of Goettingen, Cege Discussion Paper 140.
- Willis, R., 1986. Wage determinants: A survey and reinterpretation of human capital earnings functions. In *Handbook of Labor Economics*, ed. by Orley Ashenfelter and Richard Layard. Amsterdam and New York: North Holland.
- Wilmoth, J. R., Horiuchi, S., 1999. Rectangularization revisited: Variability of age at death within human populations. *Demography* 36, 475–495.
- Yaari, M. E., 1965. Uncertain lifetime, life insurance, and the theory of the consumer. *Review of Economic Studies* 32, 137–150.
- Zhang, J., Zhang, J., 2005. The effect of life expectancy on fertility, saving, schooling and economic growth: Theory and evidence. *Scandinavian Journal of Economics* 107, 45–66.
- Zhang, J., Zhang, J., 2009. Longevity, retirement, and capital accumulation in a recursive model with an application to mandatory retirement. *Macroeconomic Dynamics* 13 (3), 327–348.

Zhang, J., Zhang, J., Lee, R., 2001. Mortality decline and long-run economic growth. *Journal of Public Economics* 80, 485–507.

Zhang, J., Zhang, J., Lee, R., 2003. Rising longevity, education, savings, and growth. *Journal of Development of Economics* 70, 83–101.