Mortality decline, productivity increase, and positive feedback between schooling and retirement choices

Zhipeng Cai [*]

Sau-Him Paul Lau [†]

C. Y. Kelvin Yuen [‡]

February 2018

Abstract

The twentieth century has seen phenomenal decline in mortality and increase in productivity level. We study the effects of these two important events in a life-cycle model with schooling and retirement choices. We first show the presence of positive feedback between optimal schooling years and retirement age. Based on this feature, we then obtain interesting results regarding the sign and magnitude of the effects of a mortality or productivity shock on these two endogenous variables. In particular, we show that when compared with an exogenous retirement age model, the impact of mortality decline on schooling years is substantially magnified.

[*] Equity Research, Institutional Securities, Morgan Stanley Asia Limited, Hong Kong. E-mail: Zhipeng.Cai@morganstanley.com

[†] Corresponding author. Faculty of Business and Economics, University of Hong Kong, Hong Kong. E-mail: laushp@hku.hk

[‡] Department of Economics, Washington University in St. Louis, One Brookings Drive, St. Louis, MO 63130, U.S.A. Email: kelvinyuen@wustl.edu

1 Introduction

The twentieth century has seen phenomenal decline in mortality and increase in productivity level. Figure 1 shows the mortality decline and productivity increase in the USA from 1900 to 2000.[1] As shown in panel (a) of Figure 1, the (projected) life expectancy at birth for an average person born in 2000 was 80.9 years, 26 years longer than those born a century ago. Likewise, using real GDP per capita as rough estimates, panel (b) of Figure 1 shows that the productivity level in the USA has increased almost seven times during the same period. Similar magnitude of improvement in life expectancy and productivity is also observed in other developed economies. These two changes have led to a much higher level of expected lifetime wealth for younger generations.

[Insert Figure 1 here.]

Huge demographic and productivity changes, through the effect of expected lifetime wealth, are likely to influence economic decisions, chief among which are the retirement and schooling choices. The impact of mortality decline and/or productivity increase have been widely studied in the literature. Bloom et al. (2014) consider a life-cycle model with endogenous retirement age, as in Bloom et al. (2007) and Kalemli-Ozcan and Weil (2010). They find that optimal retirement age is delayed because of mortality decline, but is reduced by productivity increase. Restuccia and Vandenbroucke (2013) endogenize schooling duration, as in Heijdra and Romp (2009) and Cervellati and Sunde (2013). They find that optimal schooling duration rises over time because of either mortality decline or productivity increase. Boucekkine et al. (2002), Echevarria and Iza (2006) and Sánchez-Romero et al. (2016) consider both schooling and retirement choices, but they focus only on the effects of mortality changes and not those of productivity increase. We observe that while the core issues studied in the above-mentioned papers are similar, the results are quite diverse. For example, a mortality decline leads to a rise in retirement age in Bloom et al. (2014) but may lead to a fall in retirement age in Kalemli-Ozcan and Weil (2010). Moreover, the assumptions made by the researchers are sometimes very different, making it hard to compare the underlying reasons of the different results.

In this paper, we study the effects of mortality decline and productivity increase on optimal schooling years and retirement age. We conduct both

---

[1]Life expectancy data is from the Berkeley Mortality Database (http://www.demog.berkeley.edu/~bmd/), and GDP per capita data is from the Maddison-Project (http://www.ggdc.net/maddison/maddison-project/data.htm).

theoretical and quantitative analysis, but our focus is mainly theoretical. We obtain new results regarding the interaction of the schooling and retirement choices, and provide a useful framework to understand the underlying mechanism determining the effects of a mortality or productivity shock, which is helpful to interpret various diverse results in the literature.

Starting with a careful analysis of the effect of mortality decline or productivity increase on schooling years or retirement age, we find it useful to decompose the effect as the sum of the direct effect (due to the exogenous shock) and the indirect effect (due to feedback from the other endogenous variable). We find that a common feature determining the impact of either an exogenous mortality or productivity shock is *positive feedback* between optimal schooling years and optimal retirement age,[2] and we trace it to the underlying economic factors captured by the model. Intuitively, the optimal choice of schooling years depends positively on the duration that the individual can reap the return of human capital accumulation. This idea can be traced back to the influential work of Ben-Porath (1967). Thus, in response to an (anticipated) change in the retirement age, the agent changes the schooling years in the *same direction*. Similarly, schooling duration and human capital are important in affecting the marginal benefit in extending retirement age, through the effect on the individual's wage profile. As a result, a change in schooling years would also lead to a subsequent change of retirement age in the *same direction*.

While the feature of positive feedback is an important element determining the *sign* of the total effect on the schooling years or retirement age, the other key factor is the sign of the exogenous mortality or productivity shock. We combine these two sets of factors and provide further analysis. We find that if the coefficient of intertemporal elasticity of substitution is smaller than one, then an exogenous increase in productivity will decrease both retirement age and schooling years. Intuitively, an increase in productivity has both income and substitution effects on retirement age. When the intertemporal elasticity of substitution is sufficiently small, the income effect dominates the substitution effect. Combining the direct effect with the positive feedback, this leads to a negative total effect on either retirement age or schooling years. We also obtain analytical results of the effects of mortality decline. We first show that the direct impact of mortality decline on optimal schooling years, holding retirement age fixed, is generally posi-

---

[2]We emphasize that whether the effects of an exogenous shock on schooling years and retirement age are positively or negatively correlated, and whether there is positive or negative feedback between the two endogenous variables, are two different issues. See Section 3.1 (especially footnote 16) and Section 6 about the distinction between these two issues.

tive. We then show that a negative direct effect of a mortality decline on retirement age is a necessary, but not sufficient, condition for a negative total effect of a mortality decline on retirement age. This result implies that the lifetime human wealth channel (d'Albis et al., 2012) is less likely to explain the decreasing retirement age trend if the schooling duration also responds to mortality decline.

The existence of positive feedback also has implications on the *magnitude* of the total effects on the schooling years and retirement age. In particular, we show that the direct effect of an exogenous shock on either schooling years or retirement age is magnified in the presence of the other endogenous variable, when both direct effects are of the same sign or when exactly one of the direct effects is zero. Our analysis extends the result of Ben-Porath (1967) to a setting with endogenous retirement age, and shows that the magnitude of the effect of a mortality shock on schooling years is magnified under such an environment.

We obtain the above results in a baseline model focusing purely on the productivity-enhancing role of schooling (i.e., the Ben-Porath mechanism). There are two advantages in using this model: (a) the analysis, while rather tedious, is still manageable; and (b) the intuition of the results is very transparent. However, there is a major disadvantage when we match the predictions of the model with the data. Even if we allow for various combinations of mortality and productivity shocks, the computational analysis suggests that the model is not able to account for the negative correlation of schooling years and retirement age for the earlier cohorts of the twentieth century. These results are robust to various specifications and parameters values.

One direction to deal with this issue is to improve along the computational dimension. It may also be helpful to include the social security system (as in Gruber and Wise, 1999) to explain the negative correlation of schooling years and retirement age.[3] While we believe it is valuable to pursue further computational analysis, we think that such analysis does not fit well with the approach of this paper, which is mainly theoretical. Instead, we conduct further theoretical analysis by introducing an extra factor: the direct utility benefit of schooling, as in Bils and Klenow (2000) and Restuccia and Vandenbroucke (2013). Using the framework of decomposition between the

---

[3]Note that we do not analyze social security in this paper. The detailed features of social security system (in terms of the payroll tax rate, the level and coverage of pension benefit, eligibility age etc.) differ substantially among countries and for different sub-periods of the twentieth century. Theoretical results are likely to be less sharp in this more complicated environment. Since focusing on positive feedback between schooling and retirement choices and obtaining its implications are the key concerns of this paper, we decide not to include social security.

exogenous shocks and the endogenous feedback, we are able to show that the correlation of schooling years and retirement age may be positive or negative in the extended model. In particular, we show that the extended model is able to explain the negative correlation of schooling years and retirement age, provided that the flow utility of schooling is in some intermediate range.

The paper is organized as follows. In Section 2, we introduce a life-cycle model in which the sole benefit of schooling is its productivity-enhancing effect. Various mechanisms have been emphasized in the literature, and sometimes the underlying reasons in these papers are not very transparent, especially when several mechanisms are mixed together. After a careful investigation, we find it easier to first understand the underlying economic reasons in this simple environment. In Section 3, we provide analytical results regarding the impact of mortality and productivity shocks on schooling years and retirement age. We conduct computational analysis in Section 4. In Section 5, we examine issues related to positive feedback and magnification of the impact of the exogenous shocks, based on theoretical and computational results in earlier sections. Section 6 extends the model to incorporate the direct utility benefit of schooling, so as to achieve a better match between the predictions of the extended model and data. Section 7 concludes.

2   A life-cycle model with schooling and retirement choices

We consider a continuous-time life-cycle model with endogenous schooling years and retirement age. As in Restuccia and Vandenbroucke (2013) and Bloom et al. (2014), as well as many papers in the literature, mortality decline and productivity increase are taken as exogenous and we only investigate the effects, but not the causes, of these changes. Thus, we abstract from any health-enhancing expenditure (as in Chakraborty, 2004) or any feedback of human capital accumulation on economic growth (as in Bils and Klenow, 2000).

As in many existing papers, we ignore changes in infant and child mortality in our model. Assume that individuals in the model begin to make economic decisions at age N. Define "adult age" as the age measuring from age N. Lifetime uncertainty is present in the economic environment that we study. An individual of cohort b faces an age-specific mortality rate function $\mu(x; \theta_b)$, where x is her (adult) age and θ_b is an index of mortality level of this cohort. The function satisfies $\mu(x; \theta_b) \geq 0$ and $\lim_{x \to T} \mu(x; \theta_b) = \infty$, where T is the maximum age in the model. Equivalently, lifetime uncertainty can

be represented by the survival function

$$l(x; \theta_b) = \exp\left[-\int_0^x \mu(t; \theta_b)dt\right],$$

(1)

which is the probability that a cohort-b individual lives for at least x years. In the analysis performed in subsequent sections, the survival functions $l(x; \theta_b)$ of different cohorts shift over time to reflect mortality changes.

Individuals in the model make decisions on three dimensions: the consumption path, education, and working versus retirement. To maintain tractability, we follow many existing papers by assuming individuals stay in school in early stage of the life cycle, and schooling is a discrete choice of either full-time study or no study. It is also assumed that individuals do not return to school after some years of working, and do not take part-time study simultaneously with a full-time job. For the labor-leisure choice, we focus on the extensive, rather than intensive, margin. We also do not consider the event of going back to the job market after a period of (temporary) retirement, consistent with the evidence in Costa (1998, p. 6) that retirement behavior in most cases is "a complete and permanent withdrawal from paid labor." In this environment, an individual spends the first S years of her life in school, joins labor market immediately after graduation, and retires at age R.

As mentioned in the Introduction, we first consider a model in which the sole benefit of schooling is its role to enhance the productivity level of the individual. In the preference side, a cohort-b individual values consumption and dislikes working. She chooses the consumption path, S and R to maximize her expected lifetime utility, which is given by

$$\int_0^T \exp\left(-\rho x\right) l\left(x; \theta_b\right) \frac{c\left(x\right)^{1-\frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}} dx - \int_{\underline{W}}^R \exp\left(-\rho x\right) l\left(x; \theta_b\right) \nu\left(x; \theta_b\right) dx,$$

(2)

where ρ is the subjective discount rate, σ is the coefficient of intertemporal elasticity of substitution, $c(x)$ is the level of consumption at age x, $\nu\left(x; \theta_b\right)$ is the disutility of labor of a cohort-b individual at age x, and $\underline{W}$ is the minimum age such that disutility of labor is positive.[4] It is assumed that $\nu\left(x; \theta_b\right)$ is non-decreasing in age, and may shift down over time to reflect

---

[4]In the literature, disutility of labor (a non-negative term, which may depend on age or health status) is assumed to be important after one reaches some ages around 40 to 50. It captures the cost of delaying retirement age. Before that age, this cost is usually minimal and it is convenient to assume that $\nu(x; \theta_b) = 0$. In the computational analysis, we take $\underline{W}$ as $45 - N$.

the "compression of morbidity" effect (Fries, 1980; Bloom et al., 2007) of exogenous health improvement and mortality decline.

The flow budget constraint is as follows:

$$a'(x) = \begin{cases} [r + \mu(x; \theta_b)]\, a(x) + \phi_b h(S) - c(x) & \text{if } S < x \leq R \\ [r + \mu(x; \theta_b)]\, a(x) - c(x) & \text{if } x \leq S \text{ or } x > R \end{cases}, \quad (3)$$

where r is the real interest rate, $a(x)$ is the level of financial asset at age x, $h(S)$ is the human capital level of the individual, ϕ_b is the index of productivity level of a cohort-b individual, and the boundary conditions

$$a(0) = 0, \ a(T) \geq 0. \quad (4)$$

Under the above specification, the agent has no bequest motive and a perfect annuity market exists to fully insure against mortality risk, similar to Yaari (1965). Therefore, at each age x, the agent can lend or borrow in a perfect financial market with effective (instantaneous) rate of return $r + \mu(t; \theta_b)$.

According to the budget constraint (3), when an individual works (after studying for S years), her wage rate is given by $\phi_b h(S)$. One may think of this specification as consisting of three components: depending on (a) the compensation to raw labor, which is normalized to be 1, (b) one's level of human capital $h(S)$, which is a function of schooling duration,[5] and (c) an index ϕ_b capturing the changing level of productivity of different cohorts, with a person from a more recent cohort benefiting from a higher value of ϕ_b. As in Hazan (2009) and Cervellati and Sunde (2013), we assume that the return to schooling, $\frac{h'(S)}{h(S)}$, is positive but non-increasing in S for $0 \leq S < \overline{S}$. We also assume that $\frac{h'(S)}{h(S)}$ is zero for $S \geq \overline{S}$, where $\overline{S} < \underline{W}$.[6]

We assume there is no social security system in this model, as in Restuccia and Vandenbroucke (2013) and Bloom et al. (2014). Since there is no social security, the marginal benefit of delaying retirement age is the marginal utility of the extra labor income generated. On the other hand, the marginal cost is the disutility of labor.

We consider a model as similar as possible to those in the literature, especially to Restuccia and Vandenbroucke (2013) and Bloom et al. (2014).[7]

---

[5]In this model, human capital is accumulated only through formal schooling, following Bils and Klenow (2000) and Hazan (2009). On the other hand, human capital is also accumulated through on-the-job training in Manuelli et al. (2012).

[6]This technical assumption ensures that optimal schooling years is less than $\overline{S}$ (and thus less than $\underline{W}$.) In the computational analysis, we take $\overline{S} = 30$.

[7]The similarities in these two papers are as follows. They both assume perfect capital market. In their quantitative analyses, they assume a constant growth rate of productivity, and that interest rate equals to the rate of pure discounting.

However, when these two models differ, we choose the assumptions with justification and as standard as possible. We highlight several major features of our model. In terms of labor-leisure choice, we focus on the extensive margin and study retirement age, instead of the intensive margin. Thus, we follow Bloom et al. (2014) in specifying the disutility of labor function based on discrete choice of labor, instead of a utility function consisting of a continuous choice of leisure at any point in time. For the schooling duration choice, we follow Restuccia and Vandenbroucke (2013) to assume that the return to schooling is decreasing in schooling years. However, unlike their paper and Bils and Klenow (2000), we mainly conduct our analysis (in Sections 2 to 5) without relying on a term reflecting direct utility benefit of schooling.[8] We believe that the interaction between schooling and retirement choices is most clearly illustrated in a model based only on productivity-enhancing role of schooling without direct utility benefit of schooling. In terms of the utility function of consumption, Restuccia and Vandenbroucke (2013) use a specification with a log utility function with a subsistence level, but Bloom et al. (2014) assume constant intertemporal elasticity of substitution (CIES) form. We choose the more general CIES specification. It turns out that our results (such as Propositions 2 and 3) depend on the value of the intertemporal elasticity of substitution (σ), which determines the relative importance of the income and substitution effects, but these two effects will cancel out when $\sigma = 1$ (the log case). Besides these three key differences, there is also a difference in the survival function assumed. We follow Bloom et al. (2014) to use the more general non-rectangular survival function. This offers the advantage that the theoretical results hold more generally for different survival functions, and we can also use a realistic survival function in the quantitative analysis.

Since our focus is the impact of mortality decline and productivity increase, we only consider two cohort-specific shocks in the model: θ_b and ϕ_b. Individuals of different cohorts face different productivity levels (indexed by ϕ_b). They also face different survival functions $l(x; \theta_b)$, and different disutility of labor functions $\nu(x; \theta_b)$, with both functions indexed by θ_b. In the remainder of this section, we obtain various choices of a representative individual of a particular cohort, with given θ_b and ϕ_b. (See the Appendix in Section 8 for detailed analysis.) First, conditional on a particular length of the schooling period and retirement age, we obtain the optimal consumption path of a cohort-b individual, defined as $c(x, S, R; \theta_b, \phi_b)$. It can be shown

---

[8]The absence of such a term in our model corresponds to $\zeta = 0$ in (9) of Bils and Klenow (2000) and $\beta = 0$ in (1) of Restuccia and Vandenbroucke (2013). Note that in Section 6, we will extend the model to incorporate direct utility benefit of schooling.

that the (conditional) optimal consumption path is characterized by

$$c\left(x, S, R; \theta_b, \phi_b\right) = \exp\left[\sigma\left(r - \rho\right) x\right] \phi_b c^n\left(0, S, R; \theta_b\right), \qquad (5)$$

where

$$c^n\left(0, S, R; \theta_b\right) = \frac{c\left(x, S, R; \theta_b, \phi_b\right)}{\phi_b} = \frac{h(S) \int_S^R \exp\left(-rx\right) l(x; \theta_b) dx}{\int_0^T \exp\left\{-\left[\left(1 - \sigma\right) r + \sigma\rho\right] x\right\} l(x; \theta_b) dx} \qquad (6)$$

is the initial consumption level normalized by the productivity level. It is clear from (6) that this normalized level is independent of ϕ_b.

Second, conditional on the optimal consumption path in (5), we obtain the first-order conditions for the optimal schooling years and retirement age.[9] Conditional on a retirement age (R), the optimal schooling years function, $\widetilde{S}(R)$, is defined implicitly according to[10]

$$\phi_b h'\left(\widetilde{S}(R)\right) \left[\int_{\widetilde{S}(R)}^R \exp\left(-rx\right) l\left(x\right) dx\right] = \phi_b \exp\left(-r\widetilde{S}(R)\right) l\left(\widetilde{S}(R)\right) h\left(\widetilde{S}(R)\right). \qquad (7)$$

The left-hand side of (7) is the marginal benefit of continuing to study, which is measured by the expected present discounted value of the increases in labor income throughout the working years from age $\widetilde{S}(R)$ to age R, due to higher level of human capital. The right-hand side of (7) is the marginal cost (the expected present discounted value of foregone labor income) of postponing the entry into the labor market at age $\widetilde{S}(R)$.

Similarly, conditional on the schooling years (S), the optimal retirement

---

[9] To avoid unnecessarily lengthy expression, we do not specify the dependence of relevant functions on θ_b and ϕ_b in (7) to (10), since we focus on optimal choices for a given cohort in this section. When we consider the comparative static results in later sections (with θ_b and/or ϕ_b changing), the dependence of relevant functions on θ_b and ϕ_b will be specified explicitly.

[10] We could replace R in (7) by R^e, where R^e is the *anticipated* retirement age, if we want to emphasize the role of anticipated retirement age in the optimal schooling years function. We then need to further impose that the actual and anticipated values of retirement age are equal $(R^* = R^e)$ at equilibrium. On the other hand, our simplification by using $\widetilde{S}(R)$ instead of $\widetilde{S}(R^e)$ is consistent with the interpretation that the individual makes schooling and retirement choices simultaneously in this model. Since there is no element of time inconsistency in our model, both specifications give the same results.

age function, $\widetilde{R}(S)$, is defined implicitly according to[11]

$$(\phi_b)^{1-\frac{1}{\sigma}} \exp\left(-r\widetilde{R}(S)\right) h(S) \left[c^n\left(0, S, \widetilde{R}(S)\right)\right]^{-\frac{1}{\sigma}} = \exp\left(-\rho\widetilde{R}(S)\right) \nu\left(\widetilde{R}(S)\right).$$
$$(8)$$

The left-hand side of (8) is the marginal benefit of delaying the retirement age, and the right-hand side is the corresponding marginal cost. Conditional on a given value of schooling duration, the productivity index ϕ_b affects the marginal benefit through two channels. First, an increase in ϕ_b leads to an upward shift of the consumption path and the resulting decrease in marginal utility of consumption. Individuals thus demand more leisure, and retire earlier. Second, a rise in ϕ_b causes increases in labor income at all ages. This rise in the price of leisure causes people to demand less leisure by retiring later. These two effects correspond to the income and substitution effects of a change in productivity level on retirement age.[12]

The optimal choices of schooling years and retirement age, S^* and R^*, are the choices of S and R such that S^* and R^* satisfy both (7) and (8) simultaneously.[13] That is,

$$R^* = \widetilde{R}(S^*), \tag{9}$$

and

$$S^* = \widetilde{S}(R^*). \tag{10}$$

Note that the productivity level ϕ_b directly affects condition (8) for the optimal retirement age, but not condition (7) for optimal schooling years, after cancellation of the common term. However, as will be seen more clearly later, it does affect S^* indirectly through R^*.

---

[11] According to (A4) in the Appendix, when retirement age (R) increases, the marginal cost is given by the expected disutility term $l(R)\nu(R)$, which is then discounted back to age 0 as $\exp(-\rho R) l(R) \nu(R)$. On the other hand, the marginal benefit is given by the discounted expected increase in labor income, which is $\exp(-rR) l(R) \phi_b h(S)$. This is multiplied by the marginal utility $[\phi_b c^n(0, S, R)]^{-\frac{1}{\sigma}}$ to convert it to age-0 utility units. Eq. (8) is obtained after cancelling the common term $l(R)$.

[12] Note that the term $(\phi_b)^{1-\frac{1}{\sigma}}$ in (8) comes from these two effects. One component, ϕ_b, comes from the effect of a change in the productivity level on the opportunity cost of delaying retirement, which is labor income $\phi_b h(S)$, and is associated with the substitution effect. The other component, $(\phi_b)^{-\frac{1}{\sigma}}$, comes from the marginal utility of initial consumption level ($\left[\phi_b c^n\left(0, S, \widetilde{R}(S)\right)\right]^{-\frac{1}{\sigma}}$), and is associated with the income effect.

[13] An alternative way to express the first-order conditions is obtained by substituting (9) and (10) into (7) and (8). This is what we do when we conduct comparative static analysis in Section 3. However, we keep the definitions of $\widetilde{S}(R)$ in (7) and $\widetilde{R}(S)$ in (8), because these two terms are particularly useful in interpreting the results in Section 3.

3 Impact of a mortality or productivity shock

In this paper, we examine the impact on optimal schooling years and retirement age of two exogenous changes: mortality decline and productivity increase. Both analytical and computational approaches are useful in a complementary way to understand these behavior. In this section, we derive comparative static results analytically. We focus on the impact of one exogenous shock at a time, since sharper analytical results are easier to obtain in the absence of the other shock. In Section 4, we will conduct computational analysis regarding the impact of both shocks simultaneously. The analytical results of this section turn out to be not only interesting on its own, but are also helpful in interpreting the computational results.

In the Appendix, we re-write the first-order conditions (7) and (8), evaluated at the optimal choices of $S = S^*$ and $R = R^*$, as (A8) and (A9). Based on (A10) and (A11), it can be shown that when there is only mortality decline, its impact on S^* and R^* are given by

$$\frac{\partial S^*}{\partial \theta_b} = \frac{\partial \widetilde{S}(R^*; \theta_b)}{\partial \theta_b} + \frac{\partial \widetilde{S}(R^*; \theta_b)}{\partial R} \frac{\partial R^*}{\partial \theta_b}, \tag{11}$$

and

$$\frac{\partial R^*}{\partial \theta_b} = \frac{\partial \widetilde{R}(S^*; \theta_b, \phi_b)}{\partial \theta_b} + \frac{\partial \widetilde{R}(S^*; \theta_b, \phi_b)}{\partial S} \frac{\partial S^*}{\partial \theta_b}, \tag{12}$$

where

$$\frac{\partial \widetilde{S}(R^*; \theta_b)}{\partial \theta_b} = \frac{\frac{\int_{S^*}^{R^*} \exp(-rx) \frac{\partial l(x;\theta_b)}{\partial \theta_b} dx}{\int_{S^*}^{R^*} \exp(-rx) l(x;\theta_b) dx} - \frac{\frac{\partial l(S^*;\theta_b)}{\partial \theta_b}}{l(S^*;\theta_b)}}{\frac{2h'(S^*)}{h(S^*)} - \frac{h''(S^*)}{h'(S^*)} - \mu(S^*;\theta_b) - r}, \tag{13}$$

$$\frac{\partial \widetilde{S}(R^*; \theta_b)}{\partial R} = \frac{\frac{\exp(-rR^*) l(R^*;\theta_b)}{\int_{S^*}^{R^*} \exp(-rx) l(x;\theta_b) dx}}{\frac{2h'(S^*)}{h(S^*)} - \frac{h''(S^*)}{h'(S^*)} - \mu(S^*;\theta_b) - r}, \tag{14}$$

$$\frac{\partial \widetilde{R}(S^*; \theta_b, \phi_b)}{\partial \theta_b} = \frac{-\frac{1}{\sigma} \frac{1}{c^n(0,S^*,R^*;\theta_b)} \frac{c^n(0,S^*,R^*;\theta_b)}{\partial \theta_b} - \frac{1}{\nu(R^*;\theta_b)} \frac{\partial \nu(R^*;\theta_b)}{\partial \theta_b}}{r - \rho + \frac{1}{\sigma} \frac{\exp(-rR^*) l(R^*;\theta_b)}{\int_{S^*}^{R^*} \exp(-rx) l(x;\theta_b) dx} + \frac{1}{\nu(R^*;\theta_b)} \frac{\partial \nu(R^*;\theta_b)}{\partial x}}, \tag{15}$$

and

$$\frac{\partial \widetilde{R}(S^*; \phi_b, \theta_b)}{\partial S} = \frac{\frac{h'(S)}{h(S)}}{r - \rho + \frac{1}{\sigma} \frac{\exp(-rR^*) l(R^*;\theta_b)}{\int_{S^*}^{R^*} \exp(-rx) l(x;\theta_b) dx} + \frac{1}{\nu(R^*;\theta_b)} \frac{\partial \nu(R^*;\theta_b)}{\partial x}}. \tag{16}$$

Note that $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R}$ in (14) and $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial S}$ in (16) represent the interaction between the two endogenous variables.

According to (11) and (12), a mortality change affects a particular endogenous variable both directly and indirectly. The underlying reasons of (11) can be traced back to (7) or (A8). Since a mortality change affects the optimal choice of schooling years (S) through the survival function $l\left(.;\theta_b\right)$, the *direct effect* is captured by the term $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial \theta_b}$, evaluated at the original retirement age. Moreover, retirement age (R) appears in (7), and this term is affected by a mortality change and may affect schooling years; thus, the *indirect effect* is represented by the product of $\frac{\partial R^*}{\partial \theta_b}$ and $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R}$. The interpretation of (12) is similar to that in (11), except that the roles of schooling years and retirement age are interchanged.

Similarly, when there is only productivity increase, its impact on S^* and R^* are given by

$$\frac{\partial S^*}{\partial \phi_b} = \frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R}\frac{\partial R^*}{\partial \phi_b}, \tag{17}$$

and

$$\frac{\partial R^*}{\partial \phi_b} = \frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial \phi_b} + \frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial S}\frac{\partial S^*}{\partial \phi_b}, \tag{18}$$

where

$$\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial \phi_b} = \frac{\left(1-\frac{1}{\sigma}\right)\frac{1}{\phi_b}}{r-\rho+\frac{1}{\sigma}\frac{\exp(-rR^*)l(R^*;\theta_b)}{\int_{S^*}^{R^*}\exp(-rx)l(x;\theta_b)dx}+\frac{1}{\nu(R^*;\theta_b)}\frac{\partial\nu(R^*;\theta_b)}{\partial x}}. \tag{19}$$

The interpretation of (17) and (18) is essentially the same as that of (11) and (12). The only exception is that the productivity level does not affect the optimal choice of schooling years, since it appears equally on both sides of (7) and can be cancelled out. As a result, there is only an indirect effect in (17).

In either the system of (11) and (12) or that of (17) and (18), the *total effect* (sum of direct and indirect effects) of an exogenous shock on the two endogenous variables can be obtained by solving the two relevant equations simultaneously. Using ψ to represent either θ_b or ϕ_b, we can solve each of the above two systems as

$$\frac{\partial S^*}{\partial \psi} = M\left[\frac{\partial \widetilde{S}\left(R^*;\theta_b\right)}{\partial \psi} + \left(\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R}\right)\frac{\partial \widetilde{R}\left(S^*;\theta_b,\phi_b\right)}{\partial \psi}\right], \tag{20}$$

and

$$\frac{\partial R^*}{\partial \psi} = M\left[\frac{\partial \widetilde{R}\left(S^*;\theta_b,\phi_b\right)}{\partial \psi} + \left(\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial S}\right)\frac{\partial \widetilde{S}\left(R^*;\theta_b\right)}{\partial \psi}\right], \tag{21}$$

11

where

$$M = \left(\frac{1}{1 - \frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R} \frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial S}} \right) > 0, \qquad (22)$$

because of (A7). Note that the solution of (17) and (18) is also given by (20) and (21) with $\psi = \phi_b$, once we recognize that $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial \phi_b} = 0$.

3.1   Positive feedback between the two optimal choices

It is observed from (20) and (21) that there are similarities as well as differences for the two systems: $\psi = \theta_b$ or $\psi = \phi_b$. In Sections 3.2 and 3.3, we will consider their differences by studying each of them individually.[14] Before that, we first focus on the common elements of these two systems of equations, which are given by the terms $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R}$ and $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial S}$ appearing in both systems. These common elements, which are about the interaction between optimal schooling years and retirement age, exhibit interesting properties, as given in the following Proposition. In all the propositions in this paper, it is assumed that the second-order conditions (A5) to (A7) hold for a meaningful maximization problem. We have checked that they are satisfied computationally in our analysis in Section 4.

Proposition 1. *For the life-cycle model given by (1) to (4),*

(a) anticipating that an exogenous shock will shift up (resp. down) the retirement age function, the agent changes the schooling years in the same direction; and

(b) a rise (resp. fall) in schooling years leads to a subsequent change of retirement age in the same direction.

Proof. See the Appendix.

The intuition of Proposition 1 is as follows. We observe from the first-order condition (7) that retirement age only affects the marginal benefit of increasing schooling years. When retirement age rises (say, in response to an exogenous shock), it shifts up the marginal benefit schedule. With an unchanged marginal cost schedule, the increase in retirement age induces the optimal schooling years to move in the same direction, as given in Proposition 1(a).

---

[14]We decide to consider these two systems separately because of two reasons. First, the underlying economic reasons are different for the two cases. Second, whether there is magnification of the shock or not, and the channels of magnification if they occur, differ slightly for the two shocks, because both direct effects are present for mortality decline but one direct effect is absent for productivity increase. We will elaborate on this point in Section 5.

According to the first-order condition (8), a change in schooling years has two effects on the marginal benefit of continuing working: a term related to human capital function and one related to the normalized consumption level. It can be shown from (6) and (7) that at the optimal schooling years, the effect on normalized consumption level becomes zero.[15] As a result, there is only one effect related to the rate of return of accumulating human capital, as given in (16). Since the rate of return is positive in the relevant region, agents with more schooling will retire later, as given in Proposition 1(b).

According to Proposition 1, $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R} > 0$ and $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial S} > 0$. The changes in the two endogenous variables S^* and R^* (due to a particular exogenous shock, for example) reinforce each other.[16] Positive feedback exists, and this contrasts with the other possibility of negative feedback in which $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R}$ and $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial S}$ are of opposite signs.[17]

3.2 Effects of productivity increase

We first examine how an increase in the productivity parameter (ϕ_b), other things being equal, changes optimal schooling years and retirement age. The analysis is simpler in this case, because productivity increase has no direct effect on schooling years $(\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial \phi_b} = 0)$. The results are summarized in the following Proposition.

Proposition 2. *For the life-cycle model given by (1) to (4). If*

$$0 < \sigma < 1, \tag{23}$$

then $\frac{\partial S^*}{\partial \phi_b} < 0$ and $\frac{\partial R^*}{\partial \phi_b} < 0$.
Proof. See the Appendix.

---

[15] In Section 6, we will further comment on this point for the extended model.

[16] As will be seen in Proposition 6, positive values of $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R}$ and $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial S}$ do not necessarily lead to positive correlation of the two endogenous variables. We need to consider this endogenous interaction component together with another component: the signs of the direct effects of the exogenous shocks.

[17] Positive feedback is perhaps easiest to understand in a dynamic setting in which the responses occur sequentially. For example, according to Vietorisz and Harrison (1973, p. 369), "positive feedback arises when the induced effect—after completion of the cycle—has the same sign as the original effect and thus reinforces it." While we do not emphasize the dynamic process of the interaction of the two endogenous variables in our model, we think the characterization of positive feedback is appropriate because Proposition 1 suggests that the positive response of one endogenous variable to the movement of the other, and vice versa, reinforce each other. The emphasis of "mutually reinforcing elements" is also found in the study of positive feedback by Arthur (1990, p. 99).

The intuition of Proposition 2 is as follows. As observed in (11), (12), (17) or (18), the impact of an exogenous shock (θ_b or ϕ_b) on S^* and R^* can be expressed as the sum of the direct effect (due to the exogenous shock) and the indirect effect (due to the feedback from the other endogenous variable). Moreover, we can solve the system of (11) and (12), or that of (17) and (18), to obtain (20) and (21). According to these two equations, the impact of an exogenous shock on S^* and R^* can be decomposed into two components: the exogenous shock component and endogenous feedback component. Proposition 1, which concerns the feedback (or interaction) term, shows that both $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R}$ and $\frac{\partial \widetilde{R}(S^*;\phi_b,\theta_b)}{\partial S}$ are positive.[18] We call this feature *positive endogenous effect*. Together with second-order condition (A7), we see from (20) and (21), with $\psi = \phi_b$, that when $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial \phi_b} = 0$ (direct effect of the productivity shock on schooling years is zero), both total effects ($\frac{\partial S^*}{\partial \phi_b}$ and $\frac{\partial R^*}{\partial \phi_b}$) are positively related to $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial \phi_b}$ (direct effect of the productivity shock on retirement age). When (23) holds, the income effect dominates the substitution effect, leading to *negative exogenous effect*: $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial \phi_b} < 0$. Combining the negative exogenous effect and positive endogenous effect leads to the negative total effect for the productivity shock in Proposition 2.

For the sake of completeness, we summarize in the following Proposition the remaining two cases about the value of σ. The proof, which is only slightly different from that of Proposition 2, is omitted.[19]

Proposition 3. *(a) When $\sigma = 1$, an increase in productivity level has no effect on both S^* and R^*.*

(b) When $\sigma > 1$, an increase in productivity level leads to increases in both S^* and R^*.

3.3   Effects of mortality decline

We now study how a change in the mortality parameter (θ_b), other things being equal, affects optimal schooling years and retirement age.

Solving (11) and (12) simultaneously gives (20) and (21), with $\psi = \theta_b$. As in Section 3.2, it is helpful to decompose the total effect of a mortality decline on schooling years or retirement age into the exogenous (shock) and endogenous (feedback) components. As seen from (20) and (21), the effects

---

[18]The term $\frac{\partial \widetilde{R}(S^*;\phi_b,\theta_b)}{\partial S}$ is unimportant for the proof of Proposition 2 because $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial \phi_b} = 0$, but is important generally; see, for example, the analysis of the effect of mortality decline in Section 3.3.

[19]The ingredients of the proof of Proposition 3 are very similar to those for Proposition 2. The only difference is that $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial \phi_b} = 0$ (resp. > 0) when $\sigma = 1$ (resp. > 1).

14

due to the feedback term are the same as those given in Proposition 1: both $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R}$ and $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial S}$ are positive. Moreover, $1 - \frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R}\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial S} > 0$ because of (A7). On the other hand, the exogenous components correspond to the two direct effects due to mortality decline: $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial \theta_b}$ and $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial \theta_b}$. It is easy to see from (20) and (21) that the total effect on either schooling years or retirement age ($\frac{\partial S^*}{\partial \theta_b}$ or $\frac{\partial R^*}{\partial \theta_b}$) is a linear combination of these two direct effects.

We first examine the direct effect of a mortality change on schooling years ($\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial \theta_b}$). Based on (13), as well as (A5) and (A13) in the Appendix, a positive value of $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial \theta_b}$ is equivalent to

$$\frac{\int_{S^*}^{R^*} \exp\left(-rx\right) l(x;\theta_b) \left[\int_{S^*}^{x} \left(-\frac{\partial \mu(t;\theta_b)}{\partial \theta_b}\right) dt\right] dx}{\int_{S^*}^{R^*} \exp\left(-rx\right) l\left(x;\theta_b\right) dx} > 0. \qquad (24)$$

According to (7), a mortality decline affects optimal schooling years (S^*) directly through higher future income stream (by increasing the survival probabilities from S^* to R^*) in the marginal benefit schedule and foregone current income (by increasing the survival probability at age S^*) in the marginal cost schedule. It is well known that mortality rate at the current age only affects future survival probabilities but not past survival probabilities; see (1) and (A12) also. Since the survival probabilities of age R^* and above do not appear in (7), the effect of a change in θ_b on $\mu(.;\theta_b)$ for ages after R^* is irrelevant. Moreover, the analysis in (A13) shows that the effects of a change in θ_b on $\mu(t;\theta_b)$ for $t \leq S^*$ on the marginal benefit and marginal cost schedules exactly cancel out. Thus, the direct effect of a change in θ_b on optimal schooling years depends only on its impact on $\mu(t;\theta_b)$ for $t \in [S^*, R^*]$.[20] Eq. (24) has a nice interpretation that the linear combination of the effects of a change in θ_b on the survival probabilities from age S^* to age R^*, which only appear in the marginal benefit schedule, is positive.

A more intuitive interpretation can further be obtained in the special case that the mortality decline process causes decreases in mortality rates of the working years from S^* to R^*. In this case,

$$\frac{-\partial \mu(t;\theta_b)}{\partial \theta_b} > 0, \forall t \in [S^*, R^*]. \qquad (25)$$

Since (25) is a sufficient condition for (24), it is easy to see that (24) is satisfied when a change in θ_b decreases mortality rates during the working

---

[20]Cai and Lau (2017, Section 3) provide a proof of this result in a model with endogenous schooling years and exogenous retirement age.

years. Observed mortality decline in the twentieth century usually reduces mortality rates for most, but not necessarily all, ages. Thus, $-\frac{\partial \mu(t;\theta_b)}{\partial \theta_b}$ may be negative for some t, and (25) may not hold. However, the above arguments suggest that, based on the linear-combination interpretation described in the previous paragraph, it is very likely that (24) holds for a lot of empirically relevant mortality decline processes.[21]

We next examine the direct effect of a mortality change on retirement age $\left(\frac{\partial \tilde{R}(S^*;\theta_b,\phi_b)}{\partial \theta_b}\right)$. It is observed from (8) that a mortality decline affects the normalized consumption level on the marginal benefit schedule, and the disutility of labor term on the marginal cost schedule. According to (A14) in the Appendix, the effect of a mortality decline on the consumption level can be decomposed into two effects, which are called the lifetime human wealth effect and years-to-consume effect, following (23) of d'Albis et al. (2012).[22] Combining (15) and (A14), it can be shown that the sign of $\frac{\partial \tilde{R}(S^*;\theta_b,\phi_b)}{\partial \theta_b}$ is the same as the sign of[23]

$$
\frac{1}{\sigma} \frac{\int_0^T \exp\left\{-\left[(1-\sigma)r+\sigma\rho\right]x\right\} l\left(x;\theta_b\right)\left[\int_0^x \left(-\frac{\partial \mu(t;\theta_b)}{\partial \theta_b}\right)dt\right]dx}{\int_0^T \exp\left\{-\left[(1-\sigma)r+\sigma\rho\right]x\right\} l\left(x;\theta_b\right)dx} + \frac{\left(-\frac{\partial \nu(R^*;\theta_b)}{\partial \theta_b}\right)}{\nu\left(R^*;\theta_b\right)}
$$

$$
-\frac{1}{\sigma} \frac{\int_{S^*}^{R^*} \exp\left(-rx\right) l\left(x;\theta_b\right)\left[\int_0^x \left(-\frac{\partial \mu(t;\theta_b)}{\partial \theta_b}\right)dt\right]dx}{\int_{S^*}^{R^*} \exp\left(-rx\right) l\left(x;\theta_b\right)dx}. \tag{26}
$$

Summing up the above analysis, there are three components in the direct effect of a mortality decline on retirement age: the lifetime human wealth effect (by shifting down the marginal benefit schedule), the years-to-consume effect (by shifting up the marginal benefit schedule) and the compression of morbidity effect (by shifting down the marginal cost schedule). If the sum of the years-to-consume effect and compression of morbidity effect is at least as large as the lifetime human wealth effect, then (26) is non-negative and $\frac{\partial \tilde{R}(S^*;\theta_b,\phi_b)}{\partial \theta_b} \geq 0$. In the main model considered by d'Albis et al. (2012) in

---

[21]As will be discussed in Section 5, (24) holds computationally for the USA from 1900 to 2000.

[22]Note that d'Albis et al. (2012) focus on mortality decline *at an arbitrary age* to show that mortality reductions at different ages have systematically different effects on retirement age. On the other hand, our specification allows mortality changes *occurring at all ages*, and use a change in θ_b to capture this more general mortality change. However, both (23) of d'Albis et al. (2012) and (A14) have similar *economic interpretations*.

[23]Similar integral terms appear in (24) and the right-hand side of (26), except that the limits of integration are S^* and x in (24), and 0 and x in (26). In the former case, the lower limit of integration is S^* instead of 0, because the effects of θ_b on $\mu\left(.;\theta_b\right)$ from 0 to S^* on marginal benefit and marginal cost have been cancelled out, as seen from (A13).

which only the lifetime human wealth and years-to-consume effects exist,[24] they argue that when a mortality decline concentrates on old ages, the lifetime human wealth effect is absent and the years-to-consume effect is present, resulting in a delay in retirement. On the other hand, if a mortality decline concentrates on younger ages, then the lifetime human wealth effect may dominate, leading to earlier retirement. Thus, a mortality decline which affects simultaneously mortality rates at different ages will generally have ambiguous effect on retirement age. In the life-cycle model with schooling years being endogenously determined and with an additional compression of morbidity effect, the various effects examined in d'Albis et al. (2012) are also relevant, leading to the same conclusion that the direct effect of a general mortality decline process on retirement age is usually ambiguous.

Combining these two direct effects, we obtain useful results regarding the sign of the total effect of a mortality decline on retirement age, $\frac{\partial R^*}{\partial \theta_b}$. This is given in the following Proposition.

Proposition 4. *Consider the life-cycle model given by (1) to (4). Assume that the direct effect of a mortality decline on schooling years is positive.*

(a) If a mortality decline has a non-negative direct effect on retirement age, then $\frac{\partial R^*}{\partial \theta_b} > 0$.

(b) A necessary, but not sufficient, condition for $\frac{\partial R^*}{\partial \theta_b} < 0$ is a negative direct effect of a mortality decline on retirement age.

The intuition of Proposition 4 is as follows. In a life-cycle model incorporating both schooling and retirement choices, the effect of a mortality decline on retirement age is given by (21), with $\psi = \theta_b$. It is observed that both direct effects $(\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial \theta_b}$ and $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial \theta_b})$ can be important in determining the sign of the total effect $\frac{\partial R^*}{\partial \theta_b}$. Since a mortality decline has a positive direct effect on schooling years $(\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial \theta_b} > 0)$ when (24) holds, and longer schooling duration induces higher retirement age $(\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial S} > 0)$ according to Proposition 1(b), the indirect effect (due to the endogenous change in schooling years) of mortality decline is positive in this case. As a result, the total effect of a mortality decline on retirement age is positive if the direct effect $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial \theta_b}$ is non-negative, as stated in Proposition 4(a). On the other hand, Proposition 4(b) states that $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial \theta_b}$ has to be strongly negative in order to have an overall negative effect.

The above results are related to the debate in the economic demography literature. According to conventional wisdom in this literature, when people

---

[24]d'Albis et al. (2012, Section 4) also consider the compression of morbidity effect and imperfect annuity market, but their argument can be explained most intuitively in their main model.

are expected to live longer, they tend to delay their retirement so as to earn more resources for the post-retirement days. Empirically, however, the average retirement age trend over time is more complicated than the monotonic increasing relationship predicted by conventional theory. As documented in, for example, Costa (1998, Figure 2.1), labor force participation rates of US men aged 65 and over declined from over 60% in 1900 to around 20% at the 1990s. Interestingly, the downward trend of labor force participation rates of men aged 65 and above seems to reverse around the 1990s (Maestas and Zissimopoulos, 2010, Figure 4).[25]

Different reasons to explain the decreasing trend of retirement age for the cohorts born in the late nineteenth century and early twentieth century in developed countries have been offered in the literature. Kalemli-Ozcan and Weil (2010) focus on a decrease in the uncertainty about the age at death, and show that mortality decline may lead to early retirement if this uncertainty effect is very strong. On the other hand, d'Albis et al. (2012) find that a mortality decline may lead to early retirement age if the lifetime human wealth effect dominates the years-to-consume effect, and they clarify that this condition is more likely to hold if the mortality decline concentrates on younger ages. Besides these demographic factors, Gruber and Wise (1998, 1999) examine the role of generous benefits provided by the social security system. Costa (1998) emphasizes the wealth effect associated with sustained economic growth. Bloom et al. (2014) follow up on this idea, and combine mortality decline and increasing wealth to explain declining retirement age in the twentieth century.

Our analysis has a direct contribution to the above debate. The analysis of d'Albis et al. (2012) implies that a *necessary and sufficient condition* for a mortality decline leading to an early retirement age in a life-cycle model with *exogenous schooling years* and the compression of morbidity effect is that the lifetime human wealth effect dominates the sum of the years-to-consume and compression of morbidity effects. In terms of the notations of this paper, the condition is equivalent to a negative value of (26).[26] According to Proposition 4(b), a negative value of (26), which implies that $\frac{\partial \widetilde{R}(S^*; \theta_b, \phi_b)}{\partial \theta_b} < 0$, becomes only a *necessary condition* for a mortality decline leading to earlier retirement when *schooling duration is endogenous* and (24) holds. This result implies that the lifetime human wealth channel emphasized in d'Albis et al. (2012)

---

[25]Note that people who retire in the 1990s correspond roughly to various cohorts born on the 1920s and 1930s.

[26]The point can be seen from (21) with $\psi = \theta_b$. In a model with exogenous schooling years, $\frac{\partial \widetilde{R}(S^*; \theta_b, \phi_b)}{\partial S} = 0$. Thus, a necessary and sufficient condition for $\frac{\partial R^*}{\partial \theta_b} < 0$ is a negative value of the direct effect $\frac{\partial \widetilde{R}(S^*; \theta_b, \phi_b)}{\partial \theta_b}$.

is less likely to explain the declining trend of retirement age in an economic environment in which schooling years also respond to mortality decline. As argued earlier, it is very likely that the direct effect of a mortality decline on schooling years is positive (i.e., (24) holds). In this case, even if the necessary condition of a negative value of the direct effect of a mortality decline on retirement age (i.e., (26) is negative) holds, a mortality decline may not be sufficient to generate a negative total effect on retirement age.

4   Quantitative analysis

In this section, we conduct quantitative analysis to examine the impact of mortality decline and productivity increase on schooling years and retirement age. We will first conduct analysis based on the baseline model, and then perform sensitivity analysis.

4.1   Specifications of the baseline model

As far as possible, the specifications in our baseline model are those commonly used in the literature. We use the Gompertz-Makeham specification for the survival function, as in Heijdra and Romp (2009) and Bloom et al. (2014). The Gompertz-Makeham survival function, which involves three parameters, is given by

$$l\left(x;\theta_b\right) = l^{GM}(x;\theta_b) = \exp\left\{-\mu_{b,0}x - \frac{\mu_{b,1}}{\mu_{b,2}}\left[\exp\left(\mu_{b,2}x\right) - 1\right]\right\}, \qquad (27)$$

where $\mu_{b,i}$ $(i = 0, 1, 2)$ is related to the mortality parameter θ_b defined before, as follows:

$$\mu_{b,i} = \mu_i(\theta_b). \qquad (28)$$

The corresponding age-specific mortality rate function is given by $\mu^{GM}(x;\theta_b) = -\frac{1}{l^{GM}(x;\theta_b)}\frac{\partial l^{GM}(x;\theta_b)}{\partial x} = \mu_{b,0} + \mu_{b,1}\exp\left(\mu_{b,2}x\right)$.[27] Note that the coefficients are cohort-specific.

---

[27] Note that $\mu^{GM}(x;\theta_b)$ does not tend to infinity for finite x, and thus is different from the convenient assumption of a finite maximum age (T) in the theoretical model. However, this discrepancy does not pose any practical problem in our computational analysis, because we assume $T = 110 - N$ or $T = 115 - N$, and the estimated values of $\frac{l^{GM}(x+N;\theta_b)}{l^{GM}(N;\theta_b)}$ are effectively zero for $x + N > 110$. (See Figure 2.)

Following Bloom et al. (2014), the disutility of labor function of cohort-b individuals is assumed to be proportional to that cohort's age-specific mortality rate function:

$$\nu\left(x;\theta_b\right) = \delta\mu^{GM}(x;\theta_b) = \delta\left[\mu_{b,0} + \mu_{b,1}\exp\left(\mu_{b,2}x\right)\right], \tag{29}$$

where $\delta > 0$.

For the human capital function, we assume

$$h\left(S\right) = \exp(\gamma S^\lambda), \tag{30}$$

where $\gamma > 0$ and $0 < \lambda \leq 1$. This functional form is consistent with those in the literature, such as Cervellati and Sunde (2013) and Cai and Lau (2017). According to this specification, the rate of return to schooling is

$$\frac{h'\left(S\right)}{h\left(S\right)} = \gamma\lambda S^{\lambda-1}, \tag{31}$$

which is a decreasing function in S when λ is strictly less than 1.

Finally, we assume a constant growth rate in productivity, as in Restuccia and Vandenbroucke (2013) and Bloom et al. (2014). Normalizing the productivity index for the 1900 cohort as 1, this index for cohort-b is given by

$$\phi_b = \exp\left[g(b-1900)\right], \tag{32}$$

where g $(g > 0)$ is the growth rate of the productivity level.

4.2 Calibration

The values of the parameters in the baseline model are chosen to match those in the literature as far as possible. (We will also use other parameter values in the sensitivity analysis.) In particular, we choose $\rho = 0.03$, $r = 0.03$ and $g = 0.0127$, following Bloom et al. (2014). We set σ, the intertemporal elasticity of substitution, to be 0.6. Following Boucekkine et al. (2003), we assume that $N = 10$.[28] Finally, we assume that $T = 100$, which corresponds to a maximum biological age $(T + N)$ of 110.

To estimate the parameters of the survival functions, we minimize the sum of squared residuals between the survival probability based on the Gompertz-Makeham specification (27) and the data of US men from the Berkeley Mortality Database. For each cohort, we first transform the data to obtain the

---

[28]Note than Bloom et al. (2014) assume $N = 20$, because they focus only on retirement.

survival probabilities conditional on reaching age N, and then choose the three survival parameters ($\mu_{b,0}$, $\mu_{b,1}$ and $\mu_{b,2}$) to minimize

$$SSR^l\left(\mu_{b,0},\mu_{b,1},\mu_{b,2}\right) = \sum_{x=1}^{T}\left[\exp\left\{-\mu_{b,0}x - \frac{\mu_{b,1}}{\mu_{b,2}}\left[\exp\left(\mu_{b,2}x\right)-1\right]\right\} - \frac{l_{N+x}^{data}}{l_N^{data}}\right]^2,$$
(33)

where l_x^{data} is the survival probability data up to age x. Figure 2 shows the comparison of the survival probability data and that of the estimated Gompertz-Makeham model, for the beginning and ending years (1900 and 2000), as well as the mid point (1950) of the USA data in the Berkeley Mortality Database. It is observed that the fit is very good in every case. The fits in other years, which are not shown, are also good.

[Insert Figure 2 here.]

For the estimation of parameters γ and λ in the human capital function $h\left(S\right)$, we also use the non-linear least squares method. Specifically, we choose γ and λ to minimize

$$SSR^S\left(\gamma,\lambda\right) = \sum_{x=0}^{15}\left\{S_{1900+5x}^*\left(\gamma,\lambda\right) - \left[S_{1900+5x}^{data} - \left(N-6\right)\right]\right\}^2,$$
(34)

where $S_b^*\left(\gamma,\lambda\right)$ denotes the optimal schooling years of cohort-b individuals calculated from the model, and S_b^{data} is from the schooling years data set (for US men) used in Goldin and Katz (2008), which starts from 1876 and ends at 1975.[29] There are only 16 observations, for the data in 5-year interval, covered in both the Berkeley Mortality Database and Goldin and Katz (2008). According to the estimation result, $\widehat{\gamma} = 0.0482$, $\widehat{\lambda} = 0.948$, and the root mean squared error ($RMSE$) is 0.302, which is reasonably small. The estimated schooling data fits the actual data reasonably well, as seen in Figure 3.

[Insert Figure 3 here.]

Finally, we calibrate parameter δ of the disutility function such that the optimal retirement age for the 1900 cohort is 65.[30] The parameter values of the baseline case are summarized in Table 1.

---

[29]We are grateful to Diego Restuccia for sending the data set. We use the data in 5-year interval to minimize computational time. Our results are essentially the same when we use the annual data.

[30]Bloom et al. (2014) also calibrate the disutility parameter of their model such that optimal retirement age for the 1900 cohort is 65. Note that we present the numerical values of various variables in actual age for easier comparison. For example, we add N to the calculated values of R^* (expressed in model age). We also add $N-6$ to the calculated values of S^*, corresponding to the assumption that children go to school from age 6.

[Insert Table 1 here.]

4.3   Effects of mortality decline and productivity increase

In the quantitative analysis, we focus on 21 cohorts of US men: starting from year 1900, and increasing every 5 years up to year 2000. For each cohort, we use the specifications and parameters discussed above, together with the survival probability data of the corresponding cohort. The results of the baseline case are shown in Figure 4, with the optimal retirement ages given in the upper panel and the optimal schooling years given in the lower panel.

[Insert Figure 4 here.]

Optimal retirement age rises in early cohorts (from an imposed value of 65 in 1900) until reaching the peak of 69.0 for the 1950 cohort and falls gradually afterwards, to 67.9 in 2000. It is observed from Figure 4 that the rate of increase in retirement age from 1900 to 1950 is larger in magnitude than the rate of decrease from 1950 to 2000. The path of optimal schooling years, on the other hand, shows a rather different profile as the retirement age. The optimal schooling years series starts from 8.22 years for the 1900 cohort, increases all the way up to 13.71 years for the 1990 cohort, and then decreases slightly to 13.67 years for the 2000 cohort. It is also observed from Figure 4 that the rate of increase in schooling years was quite high in the earlier decades, but the increase slowed down in the middle of the century.

To understand the trend of the optimal retirement age and schooling years paths, we perform decomposition exercises by isolating the effect of two factors individually. Specifically, when analyzing the effect of mortality decline only, we shut down the productivity increase channel by using $g = 0$ in (32). On the other hand, when analyzing the effect of productivity increase only, we use $\mu_{b,j} = \mu_{1900,j}$ ($j = 0, 1, 2$) in (27) and (29). That is, we switch off the mortality decline channel by assuming the survival functions in later decades are the same as the 1900 version. Note that we keep other parameters as those in Table 1 in these decomposition exercises.

We first look at the effect on optimal retirement age in the upper panel of Figure 4. When the productivity level is held constant, retirement age increases monotonically from 65 to 78.1 in response to reductions in mortality rates. It is also observed that the retirement age increases at a decreasing rate, implying that the mortality effect have become weaker over time. This reflects the well-known fact that the rate of mortality decline in developed countries is slowing down in the second half of the twentieth century. As

22

observed in Figure 2, the shifting out of the survival functions from 1950 to 2000 is smaller than that from 1900 to 1950. On the other hand, the optimal retirement age decreases from 65 to 53.0 if there is only productivity increase (and thus, increased lifetime wealth). Note that the change in retirement rate in response to productivity increase is almost linear, unlike the effect of mortality decline.

We observe similar effects on optimal schooling years. When there is only mortality decline, optimal schooling years increase monotonically from 8.22 to 18.45. On the other hand, when there is only productivity increase, the corresponding figure decreases from 8.22 years to 5.29. Again, we observe that the schooling years path is concave when there is only mortality decline, and is rather linear when there is only productivity increase.

The two results on retirement age (positive effect of mortality decline and negative effect of productivity increase) are consistent with those in Bloom et al. (2014). In this sense, their qualitative results, which are developed for a model with exogenous retirement age, are also relevant when retirement age is endogenously determined. On the other hand, we find that introducing endogenous schooling years choice brings new qualitative and quantitative results when we consider the effects of simultaneous changes in mortality and productivity. Bloom et al. (2014) show that the magnitude of the decrease in optimal retirement age in response to productivity increase is about twice that of the increase in retirement age in response to mortality decline. As a result, the optimal retirement age decreases from 1901 to 1951 and also from 1951 to 1996 (Bloom et al., 2014, Table 3). Our analysis, however, shows that optimal retirement age increases in the first half of the century and then decreases in the second half. The intuition of this difference can be traced to the concave shape of the path of schooling years, which is caused by the substantial decline of mortality in the first half of the century and the decreasing rate of return of human capital formation. Bloom et al. (2014) do not consider schooling choice and thus, any indirect effect through schooling years on retirement age is not captured.[31]

---

[31] At first glance, it appears that including schooling choice in the life-cycle model represents a worse move, if our main objective is to explain the decrease in retirement age for the cohorts born in the first half of the twentieth century. However, we think that schooling years, together with retirement age, should be modeled as choice variables in studying the impact of mortality decline and productivity increase. This paper adopts this approach, and has contributions in explaining the decreasing trend of retirement age, even though it mainly examines other issues. As will be shown in Section 6, introducing utility benefit of schooling in addition to its productivity-enhancing role is helpful to explain the decrease in retirement age in a model with both schooling and retirement choices.

Regarding the impact on schooling years, the positive effect of mortality decline is consistent with Restuccia and Vandenbroucke (2013). On the other hand, the negative effect of productivity increase on schooling years is different from the positive effect in Restuccia and Vandenbroucke (2013). We will re-examine this issue in Section 6.

4.4   Sensitivity analysis

We report the results of the sensitivity analysis regarding the parameters listed in Table 1.

First, we are interested to know how sensitive our quantitative results are with respect to the values of interest rate and subjective discount rate. In the baseline case, we have followed Bloom et al. (2014) to assume $r = \rho = 3\%$. The first sensitivity analysis we perform is to assume that $r = \rho = 4\%$ (Case 1a), as in Restuccia and Vandenbroucke (2013).[32] When the values of these two parameters are the same, (5) implies that the consumption path is constant. It is also interesting to consider other parameter values such that consumption is changing over time. We consider the case that $r = 5\%$ and $\rho = 2\%$ (Case 1b), as in Barro et al. (1995). The results for these two cases, together with the baseline case, are presented in panel (a) of Figure 5. It can be observed that there are only minor differences in retirement age and years of schooling. Graphically, the profiles remain mostly unaltered, with a mildly inverted U-shaped graph for the retirement age and a concavely increasing one for schooling years. These analyses show that the baseline results are not sensitive to the choice of r and ρ, at least when they vary within a reasonable range.

[Insert Figure 5 here.]

Second, we vary the value of the growth rate of productivity. In the baseline case, we assume that the growth rate of productivity is 1.27%.[33] We would like to know how sensitive our main results are with respect to this parameter. We consider $g = 1\%$ (Case 2a), lower than the baseline value, and $g = 2\%$ (Case 2b), higher than the baseline value. The results are given in panel (b) of Figure 5. When productivity is growing at a slower rate ($g = 1\%$), the retirement age path is above that of the baseline case, with the

---

[32]Note that in each of the following cases, we re-calibrate the three parameters (γ, λ and δ) according to the procedures described in Section 4.2.

[33]Note that this is the value assumed in Bloom et al. (2014, p. 852), corresponding to the long-run growth rate of real wage in the USA. On the other hand, Restuccia and Vandenbroucke (2013) assume $g = 2\%$.

retirement age for the 2000 cohort at 69.8, which is almost 2 years higher than 67.9 of the baseline case. Intuitively, when the growth rate of productivity is lower, the effect of productivity increase (and thus, wealth) is weaker. Thus, the effect of mortality decline on retirement age becomes relatively more important, resulting in a higher retirement age path. Interestingly, the schooling years profile is not too much different from the baseline case. When productivity is growing faster ($g = 2\%$), the stronger productivity effect drags down the retirement age path substantially, to an extent that the level of retirement age for the 2000 cohort is even lower than the level for the 1900 cohort. The schooling years profile now has an inverted U-shape. For Case 2b, the decrease in retirement age is substantial (4.6 years) from 1945 to 2000, and similarly, the corresponding decrease in schooling years is quite large (1.71 years), offering support regarding the interaction of the effects on these two variables.

We next perform analysis with different values of the intertemporal elasticity of substitution (σ). In the baseline case, we focus on the region $\sigma < 1$ (with the income effect dominating the substitution effect), and assume $\sigma = 0.6$. We consider two cases in this region: $\sigma = 0.5$ (Case 3a), which is assumed in Bloom et al. (2014), and $\sigma = 0.7$ (Case 3b). We also consider two other cases that the income effect does not dominate the substitution effect. In Case 3c, we consider $\sigma = 1$ such that the two effects cancel out. Finally, we consider $\sigma = 1.5$ (Case 3d), in which the substitution effect dominates. The results vary significantly when σ changes. This pattern is particularly clear when we switch off the mortality effect and focus only on the effect of productivity increase. The variation of the relative magnitude of the income and substitution effects with respect to σ is confirmed: retirement age and schooling years decrease over time (with increasing wealth) when $\sigma < 1$, but increase when $\sigma > 1$. (See Table 2.) Quantitatively, the magnitude of the changes in retirement age and schooling years with respect to σ is large. In panel (c) of Figure 5, we focus on the comparison of the baseline case with Cases 3a and 3b, regarding the combined effect of changes in mortality decline and productivity increase. When σ decreases slightly from the baseline value to 0.5, the productivity effect becomes stronger, leading to decreases in the retirement age and schooling years in the second half of the twentieth century. The effect is substantial, and the retirement age for the 2000 cohort is even lower than that for the 1900 cohort. Similarly, the productivity effect becomes weaker when σ increases to 0.7, leading to increases in both retirement age and schooling years over almost the entire period. Again the effect is large, with the retirement age for the 2000 cohort increases to 71.2 years, compared with 67.9 years in the baseline case.

[Insert Table 2 here.]

It is seen from the above analysis that a decrease in g shifts up the retirement age path (and the schooling years in recent years), but a decrease in σ produces opposite effects. In panel (d) of Figure 5, we present the results of $g = 0.85\%$ and $\sigma = 0.5$ (Case 4). In this case, decreases in g and σ (from the baseline values) produce almost completely offsetting effects, resulting in retirement age and schooling years paths similar to the baseline case. We also consider results with different assumptions of the age that individuals begin making economic decisions (N) and maximum age in the model (T). The results of $N = 6$ (Case 5) and $T = 105$ (Case 6) are given in panel (d) of Figure 5. We see no major differences in the retirement age and schooling years, relative to the baseline case. Our results are robust to the choice of N and T.

To summarize, the computational results are robust with respect to r, ρ, N and T, at least when they are within some relevant ranges. They are less robust with respect to g and σ individually, but we also find that simultaneous changes in g and σ can lead to retirement age and schooling years paths close to the baseline case.

For each of the above cases, we also perform the decomposition exercises by focusing on one exogenous change (mortality decline or productivity increase) only. As seen in Table 2, we find that in each case, optimal retirement age and schooling years always change monotonically, and these two variables always move in the same direction.[34] These results are consistent with the theoretical results in Section 3.

5   Positive feedback and magnification

In earlier sections, we show that the impact of either a productivity or mortality shock on schooling years and retirement age depends partially on a common feedback (or interaction) term, which is determined by the interaction between these two endogenous variables ($\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R}$ and $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial S}$). Proposition 1 shows that both derivatives are positive, resulting in positive feedback.

Building on the results in earlier sections, we now examine the link between positive feedback and magnification of the effect of the exogenous shock. Based on the analysis of (20) and (21), it can be concluded that if the two direct effects of an exogenous shock ($\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial \psi}$ or $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial \psi}$) are of

---

[34]This includes the special case that these two variables remain unchanged in response to productivity increase when $\sigma = 1$.

opposite signs, then positive feedback does not necessarily imply magnification.[35] On the other hand, the following Proposition states that when one direct effect is absent or when both direct effects are of the same sign, positive feedback between schooling and retirement choices causes magnification of the two direct effects.

Proposition 5. *Consider the life-cycle model given by (1) to (4). If either*

$$\frac{\partial \widetilde{R}\left(S^*; \theta_b, \phi_b\right)}{\partial \psi} = 0 \ \ and \ \ \frac{\partial \widetilde{S}(R^*; \theta_b)}{\partial \psi} \neq 0, \tag{35}$$

$$\frac{\partial \widetilde{S}(R^*; \theta_b)}{\partial \psi} = 0 \ \ and \ \ \frac{\partial \widetilde{R}\left(S^*; \theta_b, \phi_b\right)}{\partial \psi} \neq 0, \tag{36}$$

or

$$sign\left(\frac{\partial \widetilde{R}\left(S^*; \theta_b, \phi_b\right)}{\partial \psi}\right) = sign\left(\frac{\partial \widetilde{S}(R^*; \theta_b)}{\partial \psi}\right) \neq 0 \tag{37}$$

holds, then (a)

$$\left|\frac{\partial S^*}{\partial \psi}\right| > \left|\frac{\partial \widetilde{S}(R^*; \theta_b)}{\partial \psi}\right|, \tag{38}$$

and (b)

$$\left|\frac{\partial R^*}{\partial \psi}\right| > \left|\frac{\partial \widetilde{R}\left(S^*; \theta_b, \phi_b\right)}{\partial \psi}\right|. \tag{39}$$

As noted earlier, the second-order conditions (A5) to (A7) are assumed to hold. With other ingredients of the model, we conclude that $\frac{\partial \widetilde{S}(R^*; \theta_b)}{\partial R} > 0$ and $\frac{\partial \widetilde{R}(S^*; \theta_b, \phi_b)}{\partial S} > 0$. The proof of Proposition 5, which is straightforward based on analyzing (20) and (21), is omitted.

While magnification of the direct effect occurs when any of the three conditions in Proposition 5 holds, the channels of magnification differ slightly. If (35) or (36) holds, then the non-zero direct effect is *magnified by the feedback* from the other endogenous variable through the term M in (22), because $0 < \frac{\partial \widetilde{S}(R^*; \theta_b)}{\partial R} \frac{\partial \widetilde{R}(S^*; \theta_b, \phi_b)}{\partial S} < 1$. If (37) holds, then the direct effect is *both augmented by the direct effect of the other shock and magnified by the feedback*, as seen in (20) and (21).

Are the theoretical results in Proposition 5 relevant to the two exogenous shocks we consider? First, the impact of productivity increase is magnified

---

[35]The reason is that the two direct effects, when they are of opposite signs, may offset each other and not lead to amplification.

because one direct effect is absent. Second, we have checked that for mortality decline, (24) holds and (26) is positive for all cohorts.[36] As a result, (37) holds and we conclude from Proposition 5 that the total effect on optimal schooling years or retirement age is larger than the corresponding direct effect. Quantitatively, the magnitude of amplification can be seen from Figure 6, in which we compare the impact of mortality decline when both schooling years and retirement age are endogenous with (a) exogenous schooling years or (b) exogenous retirement age. From 1900 to 2000, the increase in retirement age due to the direct effect is 9.6 years. When schooling duration is endogenously determined, there is an extra increase in retirement age of 3.5 years (or 36.5 percent) over the direct effect. The magnification effect on schooling years is even more dramatic. The rise in schooling years due to the direct effect is 4.0 years from 1900 to 2000. When retirement age is endogenously determined, the extra increase in schooling years is 6.3 years (or 158 percent) over the direct effect.

[Insert Figure 6 here.]

Our results are related to the influential work of Ben-Porath (1967). He shows that the optimal human capital investment at the current age (t) is increasing in the working duration, $R - t$, where the compulsory retirement age R in his model is conveniently treated as constant.[37] The underlying reason is that people usually undertake more human capital investment when they are young, because they will have a longer duration in the future to reap the benefit of the investment. The insight of his analysis has then been widely accepted in the profession.

Propositions 1 and 5 contribute to the literature about the Ben-Porath mechanism in two aspects. First, we show in Proposition 1(a) that, in anticipating, for example, an increase in retirement age in the future due to an exogenous shock, the individual chooses longer schooling years to receive the higher benefit of human capital accumulation. Proposition 1(a) extends Ben-Porath's (1967) result, which is proved assuming retirement age is fixed, to an environment in which both schooling years and retirement age are

---

[36]The statement refers to cohorts based on data in 5-year interval. On the other hand, if we use data in 1-year interval, it is found that (26) is always positive but (24) is slightly violated in one instance: with a value of -2.1×10^{-5} in the 1948-49 interval. Fortunately, the violation is small and transitory, and so would not alter our conclusion about the long run trend of schooling years and retirement age, which is the focus of this paper.

[37]Ben-Porath (1967) used the symbol T as the retirement age; see, for example, (11) in his paper.

choice variables.[38] Second, Proposition 5 implies that the magnitude of the Ben-Porath effect is magnified when retirement age is also a choice variable. When condition (37) holds, the impact of a mortality shock on schooling becomes higher when retirement age is endogenously determined, as compared to the magnitude when retirement age is fixed exogenously (and thus, the feedback from anticipated change in retirement age on schooling years is no longer relevant). Quantitatively, the magnification effect is very substantial for mortality decline in USA from 1900 to 2000, as observed in the lower panel of Figure 6.

6 An extension: Including direct utility benefit of schooling

In the previous sections, we have used a life-cycle model focusing purely on the productivity-enhancing role of schooling. While the relative simplicity of the model allows us to obtain useful results and transparent intuition, there is a drawback when we match its predictions with the data. Even if we allow for various combinations of mortality and productivity shocks, it is hard to explain the negative correlation of schooling years and retirement age for those born before the 1930s. Additionally, the effect of a productivity increase on schooling is negative according to the model, which is contradictory to the result in Restuccia and Vandenbroucke (2013) and the conjecture of many researchers.

One way to fix the above inadequacy is to change some specifications or parameter values for the computational analysis. For example, we may consider other specifications of the compression of morbidity term to see whether the direct effect of mortality decline on retirement age is positive or negative. Relaxing the assumption of constant rate of productivity increase to allow for different rates of productivity improvement at different sub-periods may also be helpful. We leave these channels to be explored in future work focusing on quantitative analysis.[39] To keep the theoretical focus of this

---

[38]Ben-Porath (1967) is interested to know why an individual engages more in human capital investment at a young age. His analysis focuses on *an individual of a particular cohort*, and the assumption of a fixed retirement age (of people of the same cohort) is reasonable. On the other hand, we study how mortality decline and productivity increase, by changing expected lifetime wealth, may affect life-cycle choices (including schooling) of *individuals of different cohorts*. In this context, the assumption of an unchanged retirement age (for different cohorts) is less desirable, and it is better to allow both schooling years and retirement age to be endogenously determined.

[39]Another possibility is to include social security and the changes in social security

paper, we decide not to take this route, but instead to extend our model and perform further theoretical analysis.

In this section, we address these issues by adding one ingredient to the model in Section 2, and we consider only the effect of productivity increase.[40] In the extended model, we make two changes. First, we do not consider mortality decline in this section and, simply use the notations $\mu(x)$ and $l(x)$ to represent these cohort-invariant functions. Second, instead of objective function (2), we assume that the individual maximizes

$$\int_0^T \exp(-\rho x) l(x) \frac{c(x)^{1-\frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}} dx + \int_0^S \exp(-\rho x) l(x) \zeta dx$$

$$- \int_{\underline{W}}^R \exp(-\rho x) l(x) \nu(x) dx, \tag{40}$$

where ζ is flow utility of schooling, and is assumed to be non-negative. This modified objective function incorporates direct utility benefit of schooling, as in Bils and Klenow (2000) and Restuccia and Vandenbroucke (2013). The other aspects of the model captured by (1), (3) and (4) remain the same.

It can be shown that the first-order conditions are now given by (8) and

$$\phi_b h'\left(\widetilde{S}(R)\right) \left[\int_{\widetilde{S}(R)}^R \exp(-rx) l(x) dx\right] + \eta\left(\widetilde{S}(R), R; \phi_b\right)$$

$$= \phi_b \exp\left(-r\widetilde{S}(R)\right) l\left(\widetilde{S}(R)\right) h\left(\widetilde{S}(R)\right), \tag{41}$$

where

$$\eta\left(\widetilde{S}(R), R; \phi_b\right) = \frac{\exp\left(-\rho\widetilde{S}(R)\right) l\left(\widetilde{S}(R)\right) \zeta}{(\phi_b)^{-\frac{1}{\sigma}} \left[c^n\left(0, \widetilde{S}(R), R\right)\right]^{-\frac{1}{\sigma}}}. \tag{42}$$

(Detailed derivations for this model are given in an Appendix available upon request.) The first-order condition (8) for retirement age is unaffected by

---

benefits over time, which is particularly relevant for advanced countries in the last few decades. See, for example, Gruber and Wise (1998, 1999).

[40]Some researchers, such as Hansen and Lønstrup (2012) and Sánchez-Romero et al. (2016), have focused on using mortality changes to explain the negative correlation of schooling years and retirement age. While the analysis in Section 3.3 has pointed out that it is possible to do so by specifying various factors in (26) such that the direct effect of a mortality decline on retirement age is strongly negative, we believe it may not be the best approach, partly because of our computational results (in Table 2) that the direct effect of mortality decline on retirement age is always positive. Moreover, such an approach does not explain the positive effect of a productivity increase on schooling years. Our proposed method deals with both issues. Cai (2017, Chapter 3) also examines similar issues.

the introduction of the direct utility of schooling term. On the other hand, compared with (7) of the model in Section 2, we notice that the marginal cost of extending schooling years, given by the right-hand side term of (41), remains the same, but the marginal benefit now consists of two terms. The first term on the left-hand side of (41) represents the productivity-enhancing effect, which is the same as in (7). Additionally, the individual is assumed to have direct utility from schooling. The marginal direct utility due to schooling is given by $\exp\left(-\rho\widetilde{S}(R)\right) l\left(\widetilde{S}(R)\right)\zeta$. Converting this term into monetary units (by dividing it by the marginal utility of consumption) leads to the second term on the left-hand side of (41).

Similar as the analysis in Section 3, we first obtain the relations between the two endogenous variables. It is shown that

$$\frac{\partial\widetilde{S}(R^*;\phi_b)}{\partial R}=\frac{\phi_b h'\left(S^*\right)\exp\left(-rR^*\right)l\left(R^*\right)+\frac{1}{\sigma}\eta\left(S^*,R^*;\phi_b\right)\left[\frac{\exp(-rR^*)l(R^*)}{\int_{S^*}^{R^*}\exp(-rx)l(x)dx}\right]}{\Delta_S},$$
(43)

and

$$\frac{\partial\widetilde{R}(S^*;\phi_b)}{\partial S}=\frac{\frac{h'(S^*)}{h(S^*)}-\frac{1}{\sigma}\frac{1}{c^n(0,S^*,R^*)}\frac{\partial c^n(0,S^*,R^*)}{\partial S}}{r-\rho+\frac{1}{\sigma}\frac{\exp(-rR^*)l(R^*)}{\int_{S^*}^{R^*}\exp(-rx)l(x)dx}+\frac{1}{\nu(R^*)}\frac{\partial\nu(R^*)}{\partial x}},$$
(44)

where

$$\Delta_S=\phi_b\exp\left(-rS^*\right)l\left(S^*\right)h\left(S^*\right)\left[\frac{2h'\left(S^*\right)}{h\left(S^*\right)}-\frac{h''\left(S^*\right)}{h'\left(S^*\right)}-\mu\left(S^*\right)-r\right]$$

$$+\eta\left(S^*,R^*;\phi_b\right)\left[\frac{h''\left(S^*\right)}{h'\left(S^*\right)}+\frac{1}{\sigma}\frac{\eta\left(S^*,R^*;\phi_b\right)}{\phi_b h\left(S^*\right)\int_{S^*}^{R^*}\exp\left(-rx\right)l\left(x\right)dx}+\mu\left(S^*\right)+\rho\right]>0.$$
(45)

The intuition of various terms in (43) and (44) is as follows. A change in retirement age (R) affects the first-order condition (41) of schooling years, both through the productivity-enhancing and direct utility benefit terms. First, it affects the duration that the individual can reap the benefit of schooling. This is represented by the first term on the numerator of (43). We label this the *Ben-Porath effect*. Second, it affects the direct utility term (through the marginal utility of consumption), as given in the second term of the numerator of (43). We label it the *direct utility of schooling effect*. Note that when $\zeta=0$, the direct utility of schooling effect disappears. On the other hand, a change in schooling years (S) affects the first-order condition (8) of retirement age in two ways. First, it affects the wage rate through the term $\frac{h'(S^*)}{h(S^*)}$ related to the human capital formation. This is given in the first

31

term on the numerator of (44). We label it the *return to schooling effect*. Second, a change in schooling years affects the individual's lifetime wealth, as captured by the level of normalized consumption at age 0. This is given in the second term on the numerator of (44). Moreover, it can be shown that $\frac{-1}{c^n(0,S^*,R^*)} \frac{\partial c^n(0,S^*,R^*)}{\partial S} = \frac{\eta(S^*,R^*;\phi_b)}{\phi_b h(S^*) \int_{S^*}^{R^*} \exp(-rx)l(x)dx}$, and it is positive when ζ is positive. We label it the *consumption level effect*. On the other hand, for the model in Section 2, $\frac{1}{c^n(0,S^*,R^*)} \frac{\partial c^n(0,S^*,R^*)}{\partial S} = 0$, because of (7).[41] As a result, the consumption level effect disappears in that model.

It can be concluded from (43) and (44) that $\frac{\partial \widetilde{S}(R^*;\phi_b)}{\partial R} > 0$ and $\frac{\partial \widetilde{R}(S^*;\phi_b)}{\partial S} > 0$. Thus, positive feedback between them continues to hold in the presence of direct utility benefit of schooling.

To examine the total effect of productivity increase on schooling years or retirement age, we first consider the two direct effects. For this model, it can be shown that $\frac{\partial \widetilde{R}(S^*;\phi_b)}{\partial \phi_b}$ is the same as (19) of the model in Section 2, but

$$\frac{\partial \widetilde{S}(R^*;\phi_b)}{\partial \phi_b} = \frac{-\left(1 - \frac{1}{\sigma}\right)\frac{1}{\phi_b}\eta\left(S^*,R^*;\phi_b\right)}{\Delta_S}. \tag{46}$$

If $0 < \sigma < 1$ according to (23), it can be shown that one direct effect is positive ($\frac{\partial \widetilde{S}(R^*;\phi_b)}{\partial \phi_b} > 0$) and the other negative ($\frac{\partial \widetilde{R}(S^*;\phi_b)}{\partial \phi_b} < 0$). In the extended model, the exogenous effect is neither purely positive nor purely negative. Combining this feature with positive feedback between the two endogenous variables (positive endogenous effect), the total effects of productivity increase on schooling years and retirement age depend on the underlying parameters of this model, as given in the following Proposition.

Proposition 6. *Consider the life-cycle model with the direct utility benefit of schooling and an exogenous process of productivity increase, as given by (1), (3), (4) and (40), where θ_b is cohort-invariant. If (23) holds, then*
   (a) $\frac{\partial S^*}{\partial \phi_b} < 0$ *and* $\frac{\partial R^*}{\partial \phi_b} < 0$ *if*

$$0 < \zeta < \left[\frac{\exp\left[-\rho\left(R^* - S^*\right)\right]\frac{l(R^*)}{l(S^*)}\frac{h'(S^*)}{h(S^*)}}{r - \rho + \frac{1}{\nu(R^*)}\frac{\partial \nu(R^*)}{\partial x}}\right]\nu\left(R^*\right); \tag{47}$$

---

[41] The intuition is as follows. When $\zeta = 0$, it can be shown that one objective of the individual is to maximize lifetime wealth. Thus, $\frac{1}{c^n(0,S^*,R^*)} \frac{\partial c^n(0,S^*,R^*)}{\partial S} = 0$. When $\zeta > 0$, there is a non-pecuniary effect, and thus, at the optimal choice, human capital is accumulated "too much" as compared with the model with $\zeta = 0$. Thus, $\frac{1}{c^n(0,S^*,R^*)} \frac{\partial c^n(0,S^*,R^*)}{\partial S} < 0$.

(b) $\frac{\partial S^*}{\partial \phi_b} > 0$ and $\frac{\partial R^*}{\partial \phi_b} < 0$ if

$$\left[\frac{\exp\left[-\rho\left(R^* - S^*\right)\right] \frac{l(R^*)}{l(S^*)} \frac{h'(S^*)}{h(S^*)}}{r - \rho + \frac{1}{\nu(R^*)} \frac{\partial \nu(R^*)}{\partial x}} \right] \nu\left(R^*\right) < \zeta$$

$$< \left[\frac{\exp\left[(r - \rho)\left(R^* - S^*\right)\right] \left[\frac{2h'(S^*)}{h(S^*)} - \frac{h''(S^*)}{h'(S^*)} - \mu\left(S^*\right) - r\right]}{\frac{h'(S^*)}{h(S^*)} - \frac{h''(S^*)}{h'(S^*)} - \mu\left(S^*\right) - \rho} \right] \nu\left(R^*\right); \quad (48)$$

and

(c) $\frac{\partial S^*}{\partial \phi_b} > 0$ and $\frac{\partial R^*}{\partial \phi_b} > 0$ if

$$\zeta > \left[\frac{\exp\left[(r - \rho)\left(R^* - S^*\right)\right] \left[\frac{2h'(S^*)}{h(S^*)} - \frac{h''(S^*)}{h'(S^*)} - \mu\left(S^*\right) - r\right]}{\frac{h'(S^*)}{h(S^*)} - \frac{h''(S^*)}{h'(S^*)} - \mu\left(S^*\right) - \rho} \right] \nu\left(R^*\right). \quad (49)$$

Proposition 6 specifies the conditions which lead to the three possible outcomes. These outcomes are represented graphically in Figure 7.[42] The interpretations of the conditions are as follows. When the extent of utility benefit of schooling (ζ) is small and (47) holds, then the direct effect of productivity increase on schooling years is not strong enough. Thus, the indirect effect dominates and the total effect ($\frac{\partial S^*}{\partial \phi_b}$) is still negative. Similarly, $\frac{\partial R^*}{\partial \phi_b}$ is negative. The model in Section 2, with $\zeta = 0$, could be treated as a limiting case of part (a). When ζ is large enough, the direct effect dominates and the total effect $\frac{\partial S^*}{\partial \phi_b}$ turns positive. Furthermore, we find that when ζ is larger than the first threshold in (48) but not the second one, we have the results $\frac{\partial S^*}{\partial \phi_b} > 0$ and $\frac{\partial R^*}{\partial \phi_b} < 0$, as in part (b). The two thresholds correspond to the upward-sloping dotted lines in Figure 7. When ζ is substantially large to pass the higher threshold, we have $\frac{\partial S^*}{\partial \phi_b} > 0$ and $\frac{\partial R^*}{\partial \phi_b} > 0$, as in part (c).[43]

[Insert Figure 7 here.]

The results in Proposition 6 imply that the extended model is able to explain the negative correlation of schooling years and retirement age, provided that the flow utility of schooling is in the intermediate range given by

---

[42]Note that ζ only affects (41) through (42), but does not affect (8). As a result, different values of ζ in the three cases in Figure 7 are reflected in different positions of $\widetilde{S}(R^*; \phi_2)$ but the position of $\widetilde{R}(S^*; \phi_2)$ remains unchanged.

[43]In case (c), the direct effect $\frac{\partial \widetilde{R}(S^*; \phi_b)}{\partial \phi_b}$ is negative, but the total effect $\frac{\partial R^*}{\partial \phi_b}$ is positive, because ζ is very large and the indirect effect (through the endogenous change in schooling years) is very strong. We believe this case is not empirically important, but we list all three cases in Proposition 6 for the sake of completeness.

(48). Moreover, they have relevance for the results in Restuccia and Vandenbroucke (2013). In that paper, retirement age is assumed to be exogenous. We can see from (20) and (46) that we only need $\zeta > 0$ to guarantee that $\frac{\partial \widetilde{S}(R^*; \phi_b)}{\partial \phi_b} > 0$ in that environment. On the other hand, when retirement age is also a choice variable, we need parameter ζ to be not only positive, but also larger than the lower threshold in (48) in order to explain the positive total effect of productivity increase on schooling years ($\frac{\partial S^*}{\partial \phi_b} > 0$).

7   Conclusion

Mortality decline and productivity increase are two major forces affecting expected lifetime wealth of different cohorts over the twentieth century. In this paper, we study the impact of these two events in a life-cycle model with both schooling and retirement choices. After a careful analysis of the effect of mortality decline or productivity increase on schooling years or retirement age, we find it helpful to decompose the effect in terms of the exogenous (shock) and endogenous (feedback) components. Based on this framework, we obtain three main sets of results, which enhance our understanding of the mechanism determining the effects of these shocks. The results also have implications relevant to the economic demography literature.

First, we show in Proposition 1 that optimal retirement age increases in response to a rise in schooling years, and optimal schooling duration also rises in response to an (anticipated) increase in retirement age. Positive feedback exists between these two endogenous variables, and we further trace it to the underlying economic factors captured in the baseline and extended models.

In the presence of positive feedback between schooling years and retirement age, we then examine, in Propositions 2, 3, 4 and 6, the sign of the effects of either a mortality or productivity shock on these two endogenous variables. In particular, we show in Proposition 4(b) that a negative direct effect of a mortality decline on retirement age (i.e., when (26) is negative), which is the necessary and sufficient condition for a mortality decline to affect retirement age negatively when schooling duration is exogenous, is only a necessary condition when schooling duration is endogenous. This result implies that the lifetime human wealth channel suggested by d'Albis et al. (2012) is less likely to explain the decreasing trend of retirement age when the schooling duration also responds to mortality decline.

We further examine the magnitude of the effects of either a mortality or productivity shock on the schooling years and retirement age. Proposition 5 shows that the direct effect of an exogenous shock on either schooling or retirement choice is magnified in the presence of another endogenous variable,

when exactly one of the direct effects is zero or when both direct effects are of the same sign. Together with Proposition 1(a) and the computational results, our analysis not only extends Ben-Porath's (1967) result to an environment in which both schooling years and retirement age are choice variables, but also shows that the impact of mortally decline on schooling years is magnified substantially in such an environment.

8   Appendix

We derive first-order and second-order conditions of the main model in Section 8.1, and provide detailed analysis for some comparative static exercises in Section 8.2.

8.1   First-order and second-order conditions

We first obtain the individual's optimal consumption path, conditional on schooling years and retirement age. Using standard techniques of dynamic optimization, it is straightforward to obtain (5).

The intertemporal budget constraint at age 0 is given by

$$\int_0^T \exp\left(-rx\right) l(x; \theta_b) c\left(x, S, R; \theta_b, \phi_b\right) dx = \int_S^R \exp\left(-rx\right) l(x; \theta_b) \phi_b h\left(S\right) dx.$$

Differentiating this equation with respect to R and S, respectively, we obtain

$$\int_0^T \exp\left(-rx\right) l(x; \theta_b) \frac{\partial c\left(x, S, R; \theta_b, \phi_b\right)}{\partial S} dx$$

$$= \phi_b \left[h'\left(S\right) \int_S^R \exp\left(-rx\right) l(x; \theta_b) dx - \exp\left(-rS\right) l\left(S; \theta_b\right) h\left(S\right) \right], \quad \text{(A1)}$$

and

$$\int_0^T \exp\left(-rx\right) l(x; \theta_b) \frac{\partial c\left(x, S, R; \theta_b, \phi_b\right)}{\partial R} dx = \exp\left(-rR\right) l(R; \theta_b) \phi_b h\left(S\right).$$
$$\text{(A2)}$$

Conditional on the optimal consumption path (5), we now obtain the first-order necessary conditions for optimal schooling years and retirement age. Substitute (5) into (2) to express the objective function in terms of S and R only. Denote it by

$$U_b\left(S, R\right) = \int_0^T \exp\left(-\rho x\right) l(x; \theta_b) \frac{c\left(x, S, R; \theta_b, \phi_b\right)^{1 - \frac{1}{\sigma}} - 1}{1 - \frac{1}{\sigma}} dx$$

$$- \int_{\underline{W}}^{R} \exp\left(-\rho x\right) l(x; \theta_b) \nu(x; \theta_b) dx.$$

Differentiating $U_b(S, R)$ with respect to S and using (A1) to simplify, we obtain

$$\frac{\partial U_b(S, R)}{\partial S}$$

$$= (\phi_b)^{1-\frac{1}{\sigma}} c^n (0, S, R; \theta_b)^{-\frac{1}{\sigma}} \left[h'(S) \int_S^R \exp\left(-rx\right) l(x; \theta_b) dx - \exp\left(-rS\right) l(S; \theta_b) h(S) \right]$$

$$\equiv a_1 F(S, R), \tag{A3}$$

where $a_1 = (\phi_b)^{1-\frac{1}{\sigma}} c^n (0, S, R; \theta_b)^{-\frac{1}{\sigma}}$ and $F(S, R)$ is the remaining term. Since a_1 is non-zero, the first-order condition for schooling is given by $F(S, R) = 0$, or equivalently, (7).

Differentiating $U_b(S, R)$ with respect to R and using (A2) leads to

$$\frac{\partial U_b(S, R)}{\partial R} = l(R; \theta_b) \left[(\phi_b)^{1-\frac{1}{\sigma}} \exp\left(-rR\right) h(S) c^n (0, S, R; \theta_b)^{-\frac{1}{\sigma}} - \exp\left(-\rho R\right) \nu(R; \theta_b) \right]$$

$$\equiv a_2 G(S, R), \tag{A4}$$

where $a_2 = l(R; \theta_b)$ and $G(S, R)$ is the remaining term. Since a_2 is non-zero, the first-order condition for retirement age is given by $G(S, R) = 0$, or equivalently, (8).

The corresponding second-order sufficient conditions, evaluated at the optimal choices, are: (a) $\frac{\partial^2 U_b(S^*, R^*)}{\partial S^2} < 0$, which is equivalent to

$$\frac{2h'(S^*)}{h(S^*)} - \frac{h''(S^*)}{h'(S^*)} - \mu(S^*; \theta_b) - r > 0, \tag{A5}$$

(b) $\frac{\partial^2 U_b(S^*, R^*)}{\partial R^2} < 0$, which is equivalent to

$$r - \rho + \frac{1}{\sigma} \frac{\exp\left(-rR^*\right) l(R^*; \theta_b)}{\int_{S^*}^{R^*} \exp\left(-rx\right) l(x; \theta_b) dx} + \frac{1}{\nu(R^*; \theta_b)} \frac{\partial \nu(R^*; \theta_b)}{\partial x} > 0, \tag{A6}$$

36

and (c) $\frac{\partial^2 U_b(S^*,R^*)}{\partial S^2} \frac{\partial^2 U_b(S^*,R^*)}{\partial R^2} - \left[\frac{\partial^2 U_b(S^*,R^*)}{\partial S \partial R}\right]^2 > 0$, which is equivalent to[44]

$$1 - \frac{\partial S^*}{\partial R} \frac{\partial R^*}{\partial S} > 0. \tag{A7}$$

8.2 Comparative static analysis

Combining (7) to (10), the optimal choices S^* and R^* are related by

$$h'(S^*) \left[\int_{S^*}^{R^*} \exp(-rx) l(x; \theta_b) dx\right] = \exp(-rS^*) l(S^*; \theta_b) h(S^*), \tag{A8}$$

and

$$(\phi_b)^{1-\frac{1}{\sigma}} \exp(-rR^*) h(S^*) [c^n(0, S^*, R^*; \theta_b)]^{-\frac{1}{\sigma}} = \exp(-\rho R^*) \nu(R^*; \theta_b). \tag{A9}$$

We differentiate (A8) totally to obtain

$$\left[\frac{2h'(S^*)}{h(S^*)} - \frac{h''(S^*)}{h'(S^*)} - \mu(S^*; \theta_b) - r\right] dS^*$$

$$= \frac{\exp(-rR^*) l(R^*; \theta_b)}{\int_{S^*}^{R^*} \exp(-rx) l(x; \theta_b) dx} dR^* + \left[\frac{\int_{S^*}^{R^*} \exp(-rx) \frac{\partial l(x; \theta_b)}{\partial \theta_b} dx}{\int_{S^*}^{R^*} \exp(-rx) l(x; \theta_b) dx} - \frac{\frac{\partial l(S^*; \theta_b)}{\partial \theta_b}}{l(S^*; \theta_b)}\right] d\theta_b, \tag{A10}$$

and differentiate (A9) totally to obtain

$$\left[r - \rho + \frac{1}{\sigma} \frac{1}{c^n(0, S^*, R^*; \theta_b)} \frac{\partial c^n(0, S^*, R^*; \theta_b)}{\partial R} + \frac{1}{\nu(R^*; \theta_b)} \frac{\partial \nu(R^*; \theta_b)}{\partial x}\right] dR^*$$

$$= \left[\frac{h'(S^*)}{h(S^*)} - \frac{1}{\sigma} \frac{1}{c^n(0, S^*, R^*; \theta_b)} \frac{\partial c^n(0, S^*, R^*; \theta_b)}{\partial S}\right] dS^* + \left(1 - \frac{1}{\sigma}\right) \frac{1}{\phi_b} d\phi_b$$

$$- \left[\frac{1}{\sigma} \frac{1}{c^n(0, S^*, R^*; \theta_b)} \frac{\partial c^n(0, S^*, R^*; \theta_b)}{\partial \theta_b} + \frac{1}{\nu(R^*; \theta_b)} \frac{\partial \nu(R^*; \theta_b)}{\partial \theta_b}\right] d\theta_b. \tag{A11}$$

[44]It can be seen from (A3) and (A4) that the first-order condition for S is defined by $\frac{\partial U_b}{\partial S} = a_1 F(S, R) = 0$, and that for R is defined by $\frac{\partial U_b(S,R)}{\partial R} = a_2 G(S, R) = 0$. Therefore, at the optimal choices S^* and R^*, $\frac{\partial^2 U_b}{\partial S^2} = a_1 \frac{\partial F}{\partial S}$, $\frac{\partial^2 U_b}{\partial R^2} = a_2 \frac{\partial G}{\partial R}$, and $\frac{\partial^2 U_b}{\partial S \partial R} = a_1 \frac{\partial F}{\partial R} = a_2 \frac{\partial G}{\partial S}$. Thus, $\frac{\partial^2 U_b}{\partial S^2} \frac{\partial^2 U_b}{\partial R^2} - \left[\frac{\partial^2 U_b}{\partial S \partial R}\right]^2 > 0$ is equivalent to $1 - \left(\frac{\partial F/\partial R}{\partial F/\partial S}\right)\left(\frac{\partial G/\partial S}{\partial G/\partial R}\right) > 0$, which can be simplified to (A7), because $\frac{\partial S^*}{\partial R} = -\frac{\partial F/\partial R}{\partial F/\partial S}$ and $\frac{\partial R^*}{\partial S} = -\frac{\partial G/\partial S}{\partial G/\partial R}$. Alternatively, we can differentiate $U_b(S, R)$ directly to obtain an expression similar to (A5) and (A6). That expression can be shown to be equivalent to (A7), after using (14) and (16). We prefer (A7) as it is directly useful for subsequent analysis.

Straightforward manipulation of (A10) and (A11) leads to various terms in Section 3.

The following analysis is useful for the impact of mortality decline in Section 3.3. Because of (1), we have

$$\frac{\partial l(x;\theta_b)}{\partial \theta_b} = l(x;\theta_b) \int_0^x \left(-\frac{\partial \mu(t;\theta_b)}{\partial \theta_b}\right) dt. \qquad \text{(A12)}$$

Using (A12), it can be shown that

$$\frac{\int_{S^*}^{R^*} \exp(-rx) \frac{\partial l(x;\theta_b)}{\partial \theta_b} dx}{\int_{S^*}^{R^*} \exp(-rx) l(x;\theta_b) dx} - \frac{\frac{\partial l(S^*;\theta_b)}{\partial \theta_b}}{l(S^*;\theta_b)}$$

$$= \frac{\int_{S^*}^{R^*} \exp(-rx) l(x;\theta_b) \int_0^x \left(-\frac{\partial \mu(t;\theta_b)}{\partial \theta_b}\right) dt dx}{\int_{S^*}^{R^*} \exp(-rx) l(x;\theta_b) dx} - \int_0^{S^*} \left(-\frac{\partial \mu(t;\theta_b)}{\partial \theta_b}\right) dt$$

$$= \frac{\int_{S^*}^{R^*} \exp(-rx) l(x;\theta_b) \left[\int_0^{S^*} \left(-\frac{\partial \mu(t;\theta_b)}{\partial \theta_b}\right) dt + \int_{S^*}^x \left(-\frac{\partial \mu(t;\theta_b)}{\partial \theta_b}\right) dt\right] dx}{\int_{S^*}^{R^*} \exp(-rx) l(x;\theta_b) dx} - \int_0^{S^*} \left(-\frac{\partial \mu(t;\theta_b)}{\partial \theta_b}\right) dt$$

$$= \frac{\int_{S^*}^{R^*} \exp(-rx) l(x;\theta_b) \int_{S^*}^x \left(-\frac{\partial \mu(t;\theta_b)}{\partial \theta_b}\right) dt dx}{\int_{S^*}^{R^*} \exp(-rx) l(x;\theta_b) dx}. \qquad \text{(A13)}$$

We also differentiate $c^n(0, S^*, R^*; \theta_b)$ in (6) with respect to θ_b and use (A12) to obtain

$$\frac{1}{c^n(0, S^*, R^*; \theta_b)} \frac{\partial c^n(0, S^*, R^*; \theta_b)}{\partial \theta_b} = \frac{\int_{S^*}^{R^*} \exp(-rx) l(x;\theta_b) \left[\int_0^x \left(-\frac{\partial \mu(t;\theta_b)}{\partial \theta_b}\right) dt\right] dx}{\int_{S^*}^{R^*} \exp(-rx) l(x;\theta_b) dx}$$

$$- \frac{\int_0^T \exp\{-[(1-\sigma)r + \sigma\rho]x\} l(x;\theta_b) \left[\int_0^x \left(-\frac{\partial \mu(t;\theta_b)}{\partial \theta_b}\right) dt\right] dx}{\int_0^T \exp\{-[(1-\sigma)r + \sigma\rho]x\} l(x;\theta_b) dx}. \qquad \text{(A14)}$$

8.3   Proof of Proposition 1

It is straightforward to conclude that the denominator of (14) is positive because of (A5), and the numerator of (14) is positive. Therefore, $\frac{\partial \widetilde{S}(R^*;\theta_b)}{\partial R} > 0$. This proves (a).

Similarly, the denominator of (16) is positive because of (A6). On the other hand, the return to schooling is positive under reasonable assumptions and thus, the numerator of (16) is positive. Therefore, $\frac{\partial \widetilde{R}(S^*;\theta_b,\phi_b)}{\partial S} > 0$. This proves (b). ∎

8.4 Proof of Proposition 2

We observe from (20) and (21) that when $\psi = \phi_b$, each of the two total effects ($\frac{\partial S^*}{\partial \phi_b}$ or $\frac{\partial R^*}{\partial \phi_b}$) depends on $\frac{\partial \widetilde{R}(S^*; \theta_b, \phi_b)}{\partial \phi_b}$ only, since the other direct effect ($\frac{\partial \widetilde{S}(R^*; \theta_b)}{\partial \phi_b}$) is zero. When (23) holds, it is easy to conclude from (19) and (A6) that

$$\frac{\partial \widetilde{R}(S^*; \theta_b, \phi_b)}{\partial \phi_b} < 0. \tag{A15}$$

According to Proposition 1(a), $\frac{\partial \widetilde{S}(R^*; \theta_b)}{\partial R}$ is positive. Moreover, according to (A7), $1 - \frac{\partial \widetilde{S}(R^*; \theta_b)}{\partial R} \frac{\partial \widetilde{R}(S^*; \theta_b, \phi_b)}{\partial S} > 0$. Combining the above results, it is easy to conclude that $\frac{\partial S^*}{\partial \phi_b}$ and $\frac{\partial R^*}{\partial \phi_b}$ are of the same sign as that of the direct effect $\frac{\partial \widetilde{R}(S^*; \theta_b, \xi_b)}{\partial \phi_b}$, which is negative according to (A15). ∎

References

[1] Arthur, W. B. (1990), Positive feedbacks in the economy. Scientific American 262(2), 92-99.

[2] Barro, R. J., Mankiw, N. G. and Sala-i-Martin, X. (1995), Capital mobility in neoclassical models of growth. American Economic Review 85, 103-115.

[3] Ben-Porath, Y. (1967), The production of human capital and the life cycle of earnings. Journal of Political Economy 75, 352-365.

[4] Bils, M. and Klenow, P. J. (2000), Does schooling cause growth? American Economic Review 90, 1160-1183.

[5] Bloom, D. E., Canning, D., Mansfield, R. K. and Moore, M. (2007), Demographic change, social security systems, and savings. Journal of Monetary Economics 54, 92-114.

[6] Bloom, D. E., Canning, D. and Moore, M. (2014) Optimal retirement with increasing longevity. Scandinavian Journal of Economics 116, 838-858.

[7] Boucekkine, R., de la Croix, D. and Licandro, O. (2002), Vintage human capital, demographic trends and endogenous growth. Journal of Economic Theory 104, 340-375.

[8] Boucekkine, R., de la Croix, D. and Licandro, O. (2003), Early mortality declines at the dawn of modern growth. Scandinavian Journal of Economics 105, 401-418.

[9] Cai, Z. (2017), Essays on demographic changes and life-cycle behavior. PhD thesis, University of Hong Kong.

[10] Cai, Z. and Lau, S.-H. P. (2017), Impact of mortality reductions on years of schooling and expected lifetime labor supply. Journal of Mathematical Economics 72, 134-144.

[11] Cervellati, M. and Sunde, U. (2013), Life expectancy, schooling, and lifetime labor supply: Theory and evidence revisited. Econometrica 81, 2055-2086.

[12] Chakraborty, S. (2004), Endogenous lifetime and economic growth. Journal of Economic Theory 116, 119-137.

[13] Costa, D. L. (1998), The Evolution of Retirement. Chicago: University of Chicago Press.

[14] D'Albis, H., Lau, S.-H. P. and Sánchez-Romero, M. (2012), Mortality transition and differential incentives for early retirement. Journal of Economic Theory 147, 261-283.

[15] Echevarria, C. A. and Iza, A. (2006), Life expectancy, human capital, social security and growth, Journal of Public Economics 90, 2323–2349.

[16] Fries, J. F. (1980), Aging, natural death, and the compression of morbidity. New England Journal of Medicine 303(3), 130-135.

[17] Goldin, C. and Katz, L. F. (2008), The Race Between Education and Technology. Cambridge, MA: Harvard University Press.

[18] Gruber, J. and Wise, D. A. (1998), Social security and retirement: An international comparison. American Economic Review Papers and Proceedings 88(2), 158-163.

[19] Gruber, J. and Wise, D. A. (1999), Introduction and summary, in Gruber, J. and Wise, D. A. (ed.), Social Security and Retirement around the World. Chicago: University of Chicago Press.

[20] Hansen, C. W. and Lønstrup, L. (2012), Can higher life expectancy induce more schooling and earlier retirement? Journal of Population Economics 25, 1249-1264.

[21] Hazan, M. (2009), Longevity and lifetime labor supply: Evidence and implications. Econometrica 77, 1829-1863.

[22] Heijdra, B. J. and Romp, W. E. (2009), Human capital formation and macroeconomic performance in an ageing small open economy. Journal of Economic Dynamics and Control 33, 725-744.

[23] Kalemli-Ozcan, S. and Weil, D. (2010), Mortality change, the uncertainty effect, and retirement. Journal of Economic Growth 15, 65-91.

[24] Maestas, N. and Zissimopoulos, J. (2010), How longer work lives ease the crunch of population aging. Journal of Economic Perspectives 24(1), 139-160.

[25] Manuelli, R. E., Seshadri, A. and Shin, Y. (2012), Lifetime labor supply and human capital investment. Working paper (Washington University St. Louis).

[26] Restuccia, D. and Vandenbroucke, G. (2013), A century of human capital and hours. Economic Inquiry 51, 1849-1866.

[27] Sánchez-Romero, M., d'Albis, H. and Prskawetz, A. (2016), Education, lifetime labor supply, and longevity improvements. Journal of Economic Dynamics and Control 73, 118-141.

[28] Vietorisz, T. and Harrison, B. (1973), Labor market segmentation: Positive feedback and divergent development. American Economic Review 63(2), 366-376.

[29] Yaari, M. E. (1965), Uncertain lifetime, life insurance, and the theory of the consumer. Review of Economic Studies 32, 137-150.

Figure 1. Mortality Decline and Productivity Increase in the USA

(a) Life expectancy at birth, USA
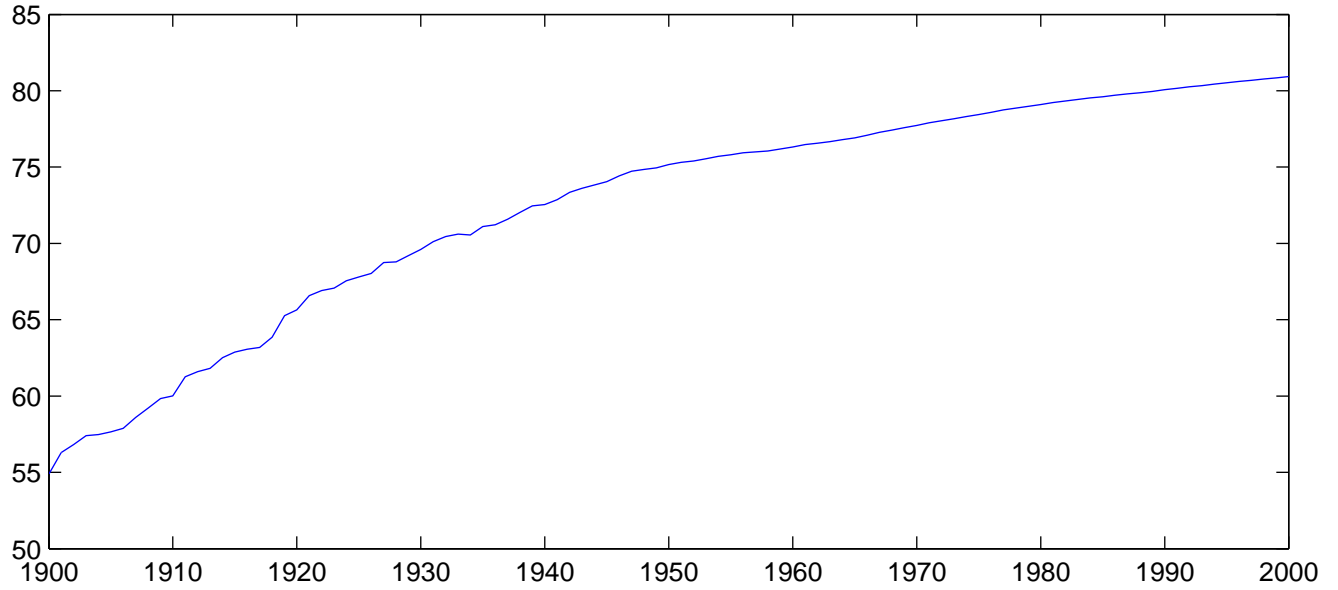
(b) GDP per capita, USA (in thousand 1990 International Geary–Khamis dollars)
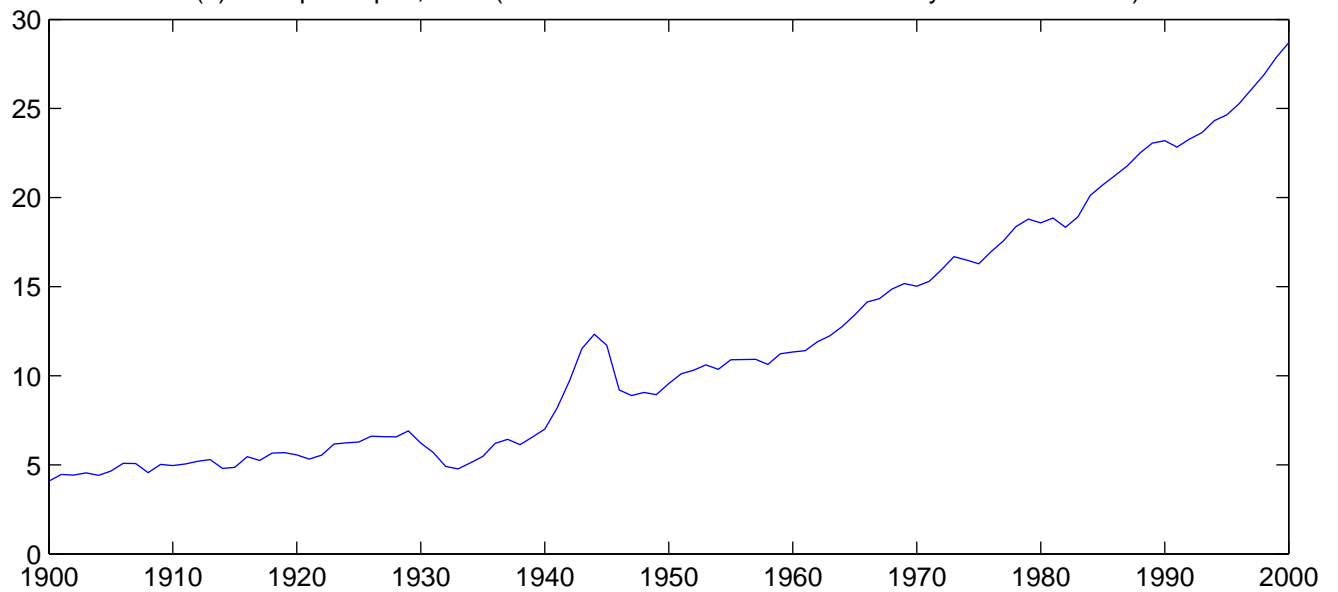
Figure 2. Survival Probability: Data and Estimated Gompertz–Makeham Survival Functions

Legend:
- 1900 Data
- 1900 Est.: $\mu_0 = 0.00286$, $\mu_1 = 0.000368$, $\mu_2 = 0.0794$
- 1950 Data
- 1950 Est.: $\mu_0 = 0.00123$, $\mu_1 = 0.000104$, $\mu_2 = 0.0918$
- 2000 Data
- 2000 Est.: $\mu_0 = 0.000641$, $\mu_1 = 7.4\text{e}{-005}$, $\mu_2 = 0.0932$

Age
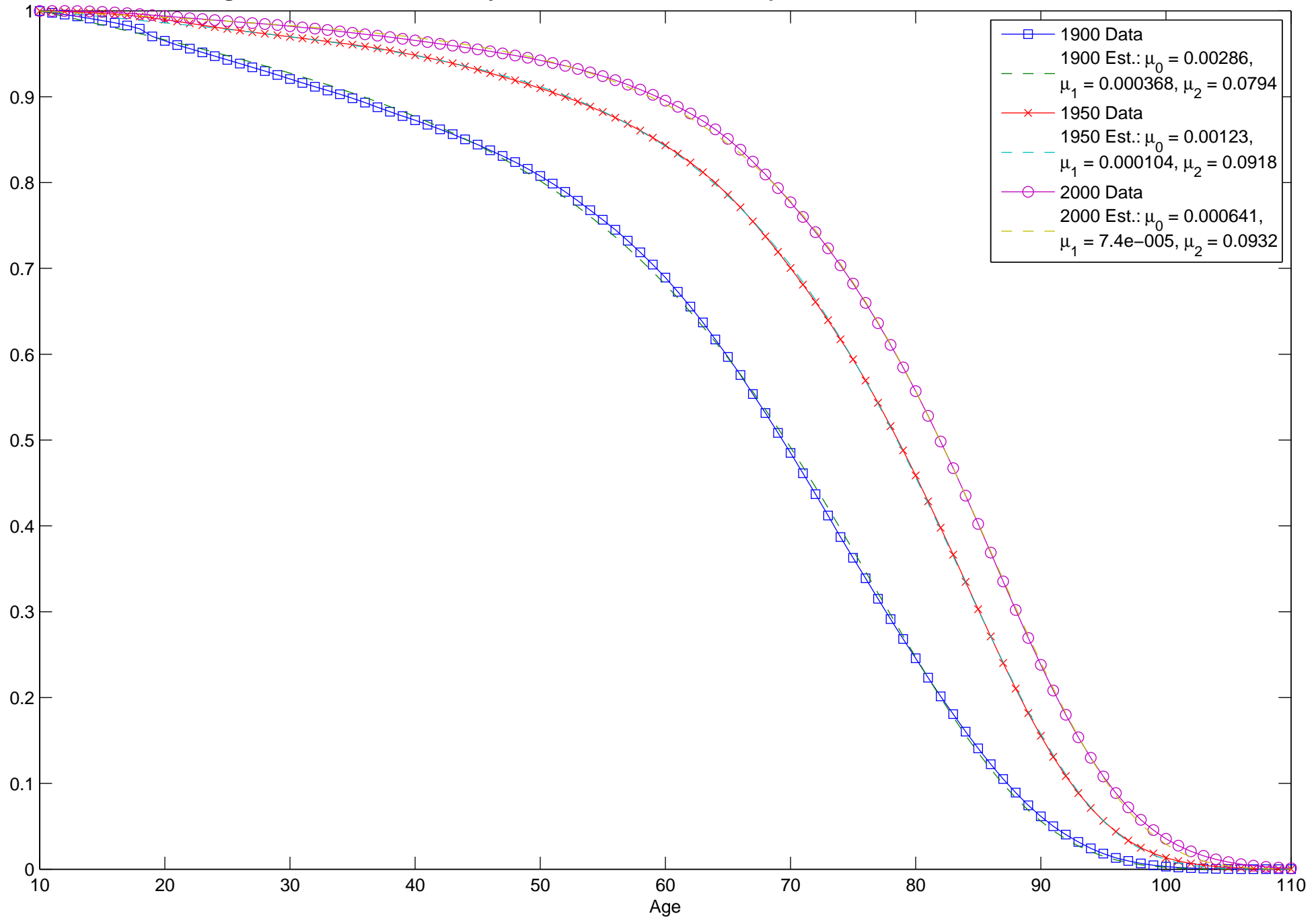
Figure 3. Schooling Years
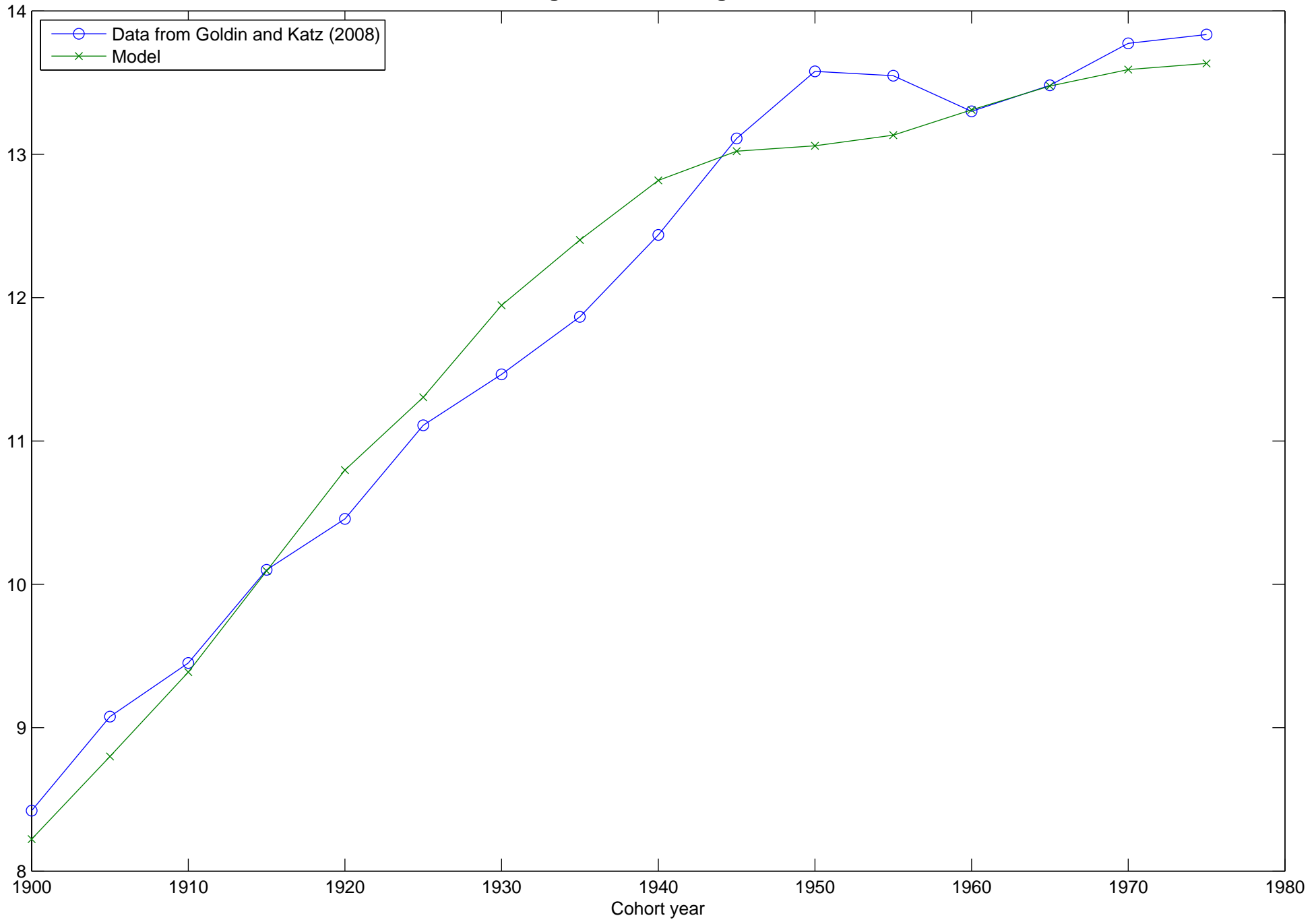
Figure 4. Baseline Results

Retirement age (R* + N)

Legend:
- Both changes
- Mortality decline only
- Increased productivity only

Schooling years (S* + N − 6)

Legend:
- Both changes
- Mortality decline only
- Increased productivity only
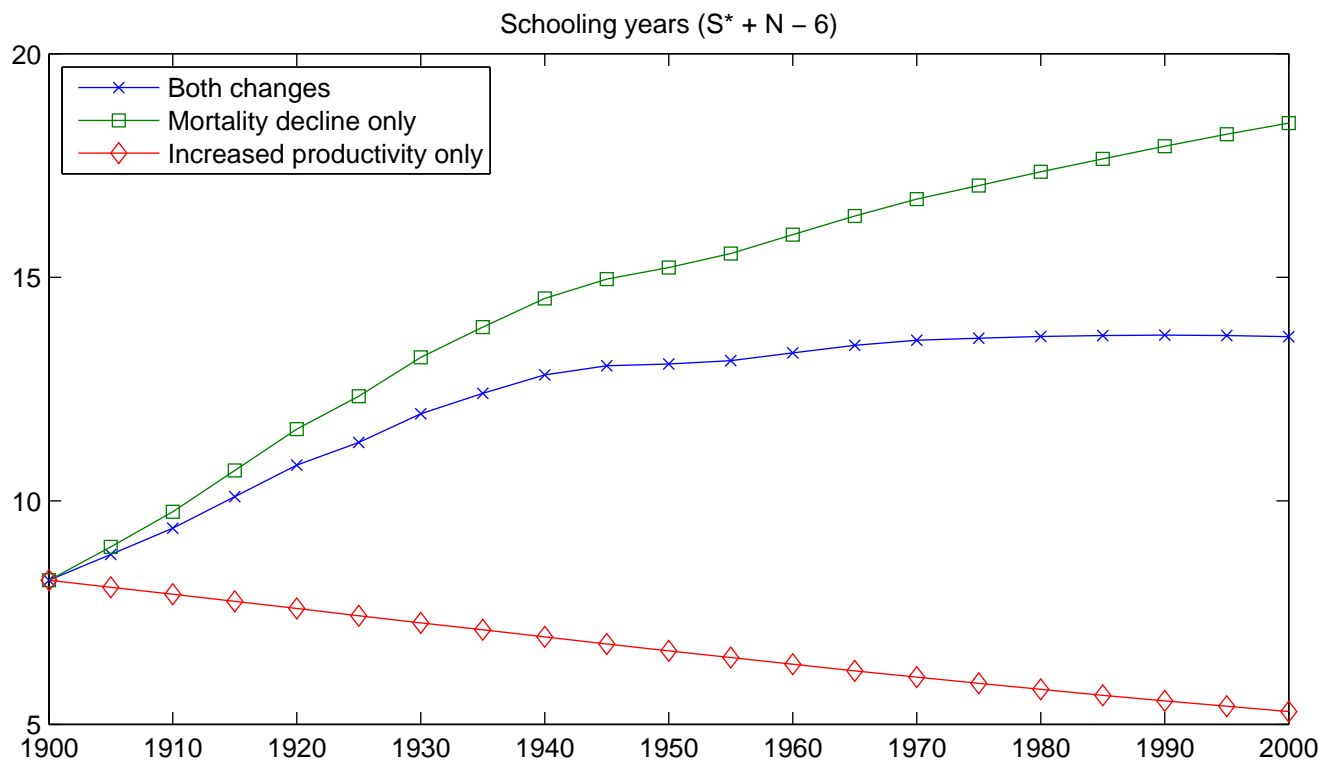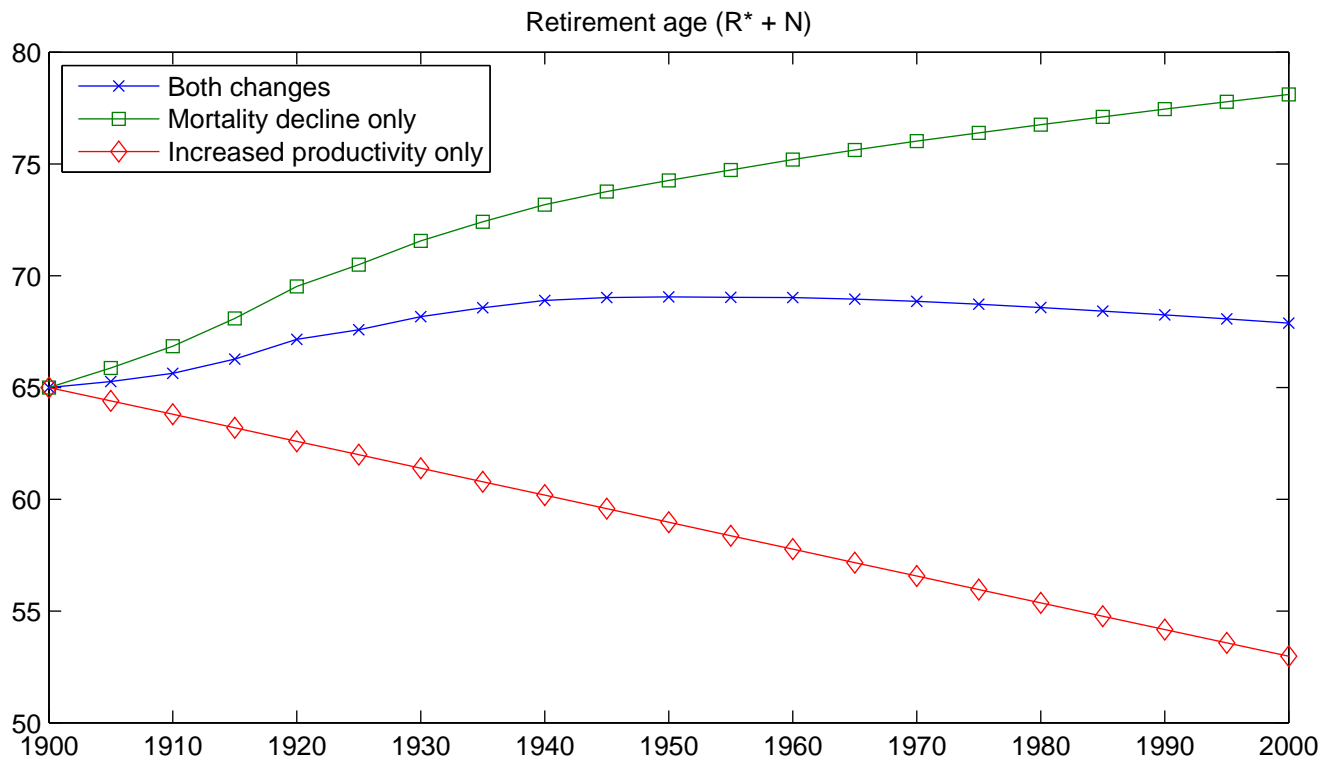
Figure 5. Sensitivity Analysis (With $\sigma < 1$)

(a) r & ρ

(b) g
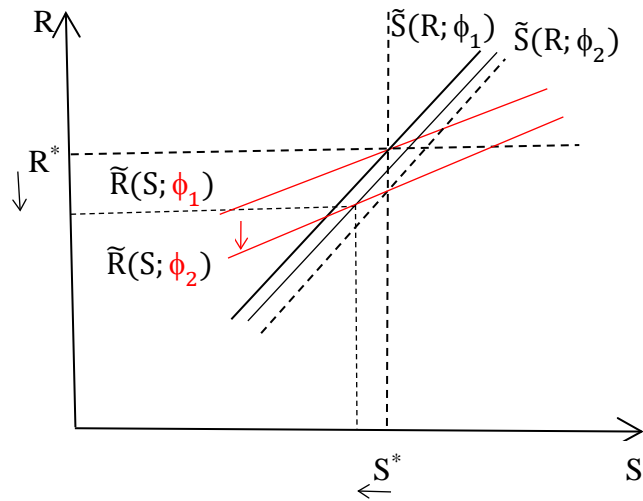
(c) σ
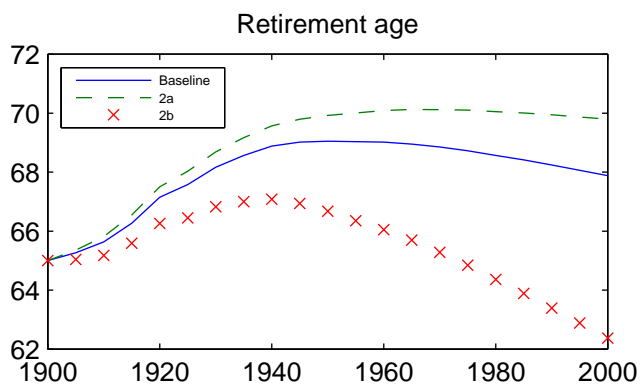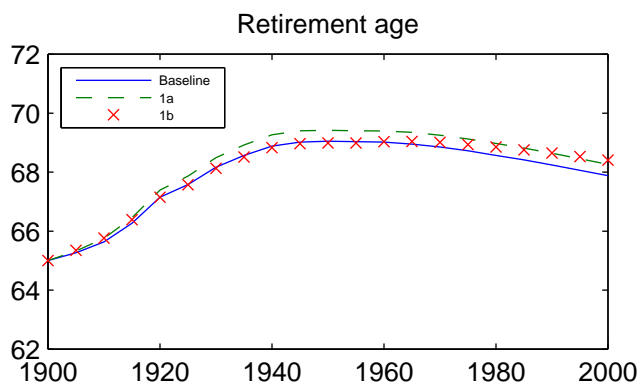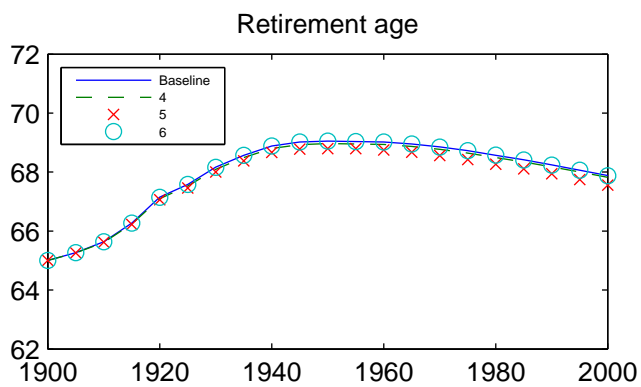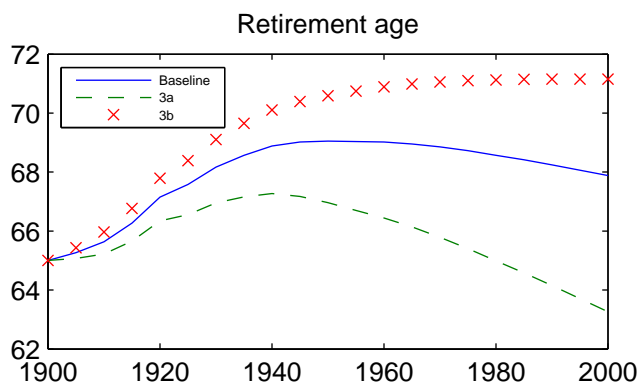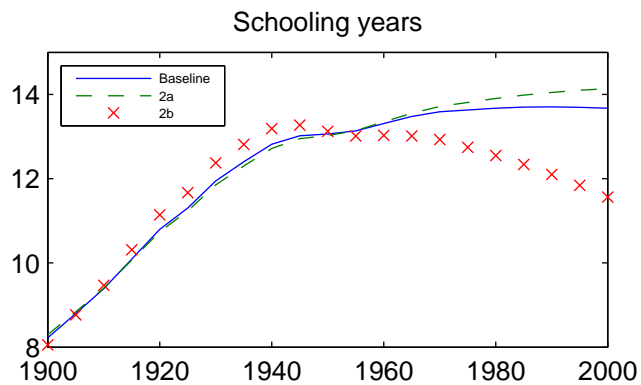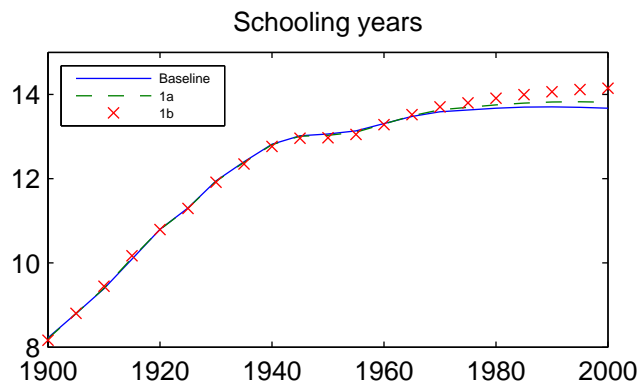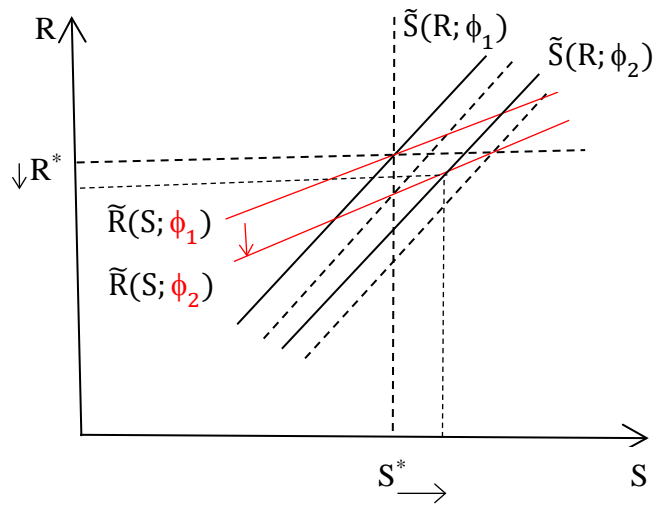
(d) Others
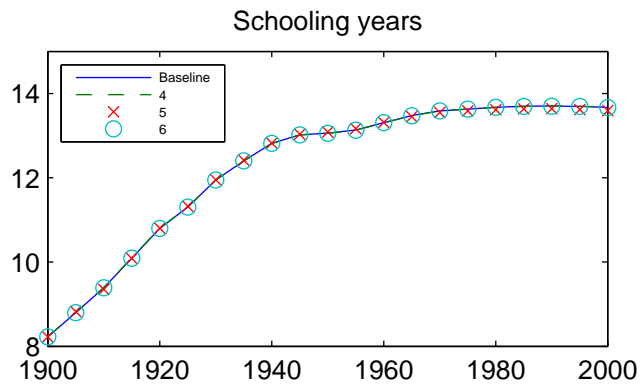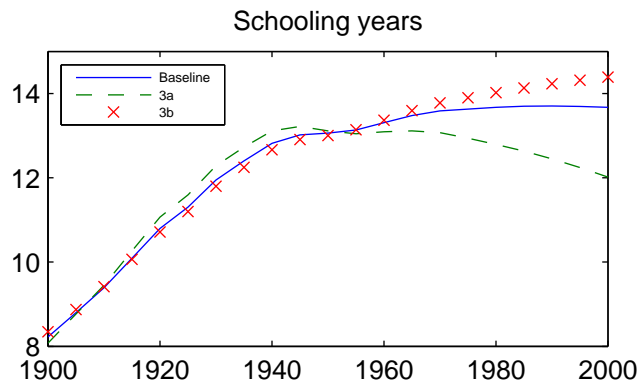
Figure 6. Magnification Effect Under Mortality Decline

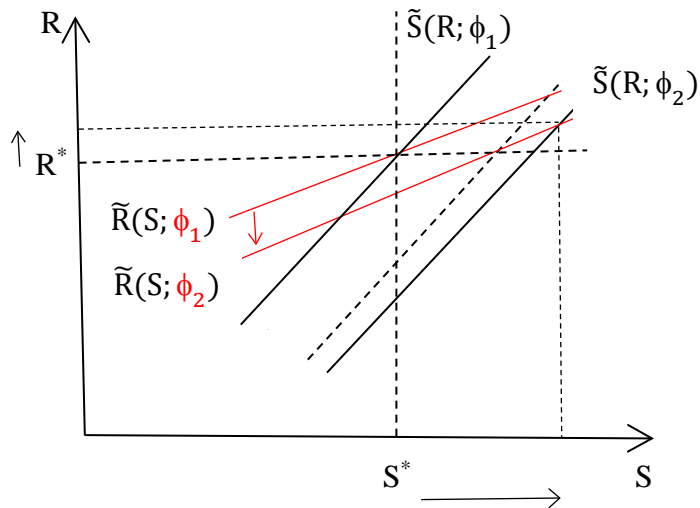## Retirement age

Schooling years
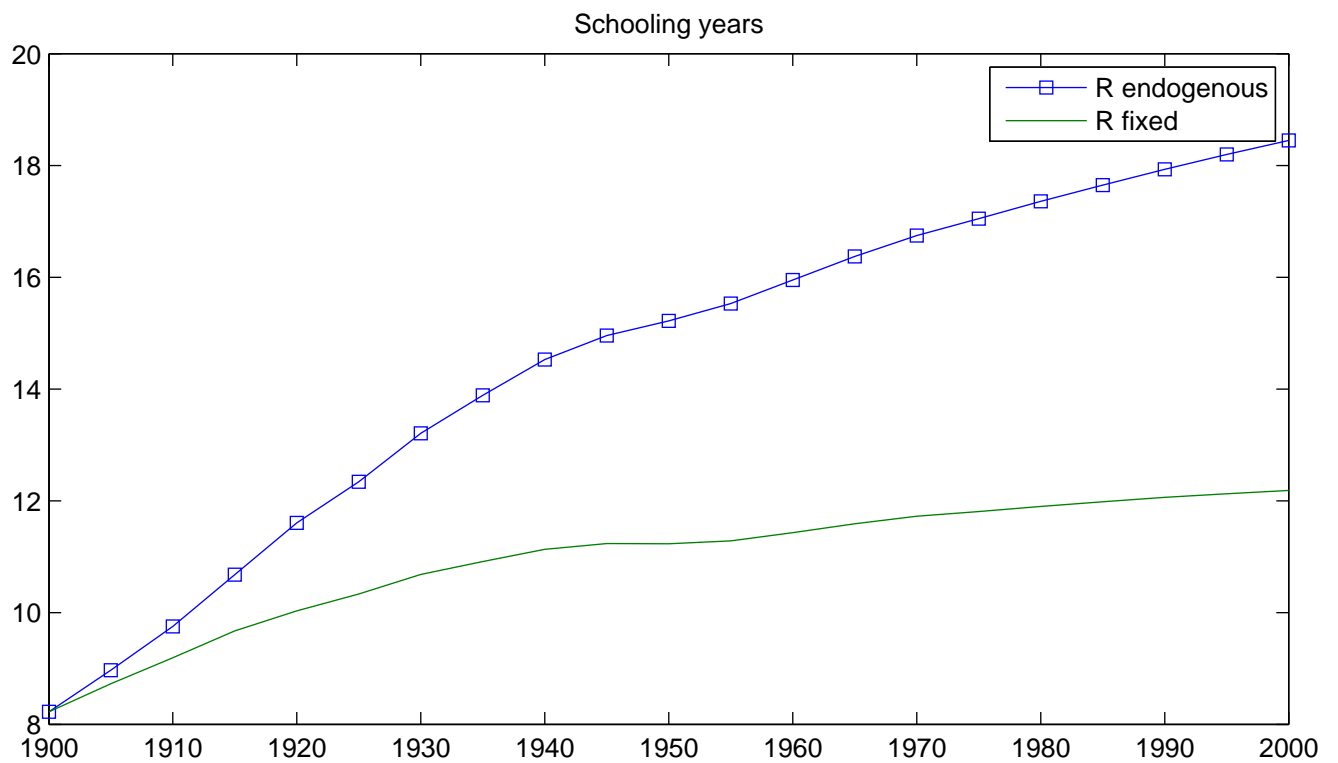
Figure 7. The Impact of Productivity Increase for the Extended Model

(a) Flow utility of schooling is small: (47) holds

(b) Flow utility of schooling is within the intermediate range: (48) holds

(c) Flow utility of schooling is large: (49) holds

Parameter	Meaning	Value
ρ	Subjective discount rate	3%
r	Real interest rate	3%
g	Growth rate of productivity	1.27%
σ	Intertemporal elasticity of substitution	0.6
N	The age that individuals begin making economic decisions	10
T	Maximum age in the model	100
γ	Free parameter in human capital function	0.0482
λ	Free parameter in human capital function	0.948
δ	Free parameter in disutility function	40.2

Table 1: Parameters of the Baseline Model

Cases	Values	RMSE		Mortality decline & productivity increase		Mortality decline only		Productivity increase only	
				R*+N	S*+N-6	R*+N	S*+N-6	R*+N	S*+N-6
Baseline	/	0.302	1900	65.0	8.22	65.0	8.22	65.0	8.22
			Max. value	69.0	13.71	78.1	18.45	65.0	8.22
			(Year)	(1950)	(1990)	(2000)	(2000)	(1900)	(1900)
			2000	67.9	13.67	78.1	18.45	53.0	5.29
1a	$r = 4\%$; $\rho = 4\%$	0.303	1900	65.0	8.21	65.0	8.21	65.0	8.21
			Max. value	69.4	13.82	78.9	18.03	65.0	8.21
			(Year)	(1950)	(1995)	(2000)	(2000)	(1900)	(1900)
			2000	68.3	13.81	78.9	18.03	52.3	5.47
1b	$r = 5\%$; $\rho = 2\%$	0.305	1900	65.0	8.17	65.0	8.17	65.0	8.17
			Max. value	69.0	14.15	76.4	17.14	65.0	8.17
			(Year)	(1965)	(2000)	(2000)	(2000)	(1900)	(1900)
			2000	68.4	14.15	76.4	17.14	56.0	6.38
2a	$g = 1\%$	0.263	1900	65.0	8.30	65.0	8.30	65.0	8.30
			Max. value	70.1	14.14	77.7	17.46	65.0	8.30
			(Year)	(1965)	(2000)	(2000)	(2000)	(1900)	(1900)
			2000	69.8	14.14	77.7	17.46	53.1	5.56
2b	$g = 2\%$	0.615	1900	65.0	8.05	65.0	8.05	65.0	8.05
			Max. value	67.1	13.27	79.5	21.88	65.0	8.05
			(Year)	(1940)	(1945)	(2000)	(2000)	(1900)	(1900)
			2000	62.4	11.56	79.5	21.88	52.6	4.43
3a	$\sigma = 0.5$	0.541	1900	65.0	8.08	65.0	8.08	65.0	8.08
			Max. value	67.3	13.22	78.9	21.23	65.0	8.08
			(Year)	(1940)	(1945)	(2000)	(2000)	(1900)	(1900)
			2000	63.3	12.02	78.9	21.23	47.6	4.10
3b	$\sigma = 0.7$	0.250	1900	65.0	8.34	65.0	8.34	65.0	8.34
			Max. value	71.2	14.39	77.6	16.90	65.0	8.34
			(Year)	(1995)	(2000)	(2000)	(2000)	(1900)	(1900)
			2000	71.1	14.39	77.6	16.90	57.2	6.58
3c	$\sigma = 1$	0.252	1900	65.0	8.48	65.0	8.48	65.0	8.48
			Max. value	77.1	15.06	77.1	15.06	/	/
			(Year)	(2000)	(2000)	(2000)	(2000)	/	/
			2000	77.1	15.06	77.1	15.06	65.0	8.48
3d	$\sigma = 1.5$	0.263	1900	65.0	8.52	65.0	8.52	65.0	8.52
			Max. value	81.8	15.22	77.0	14.22	71.0	9.37
			(Year)	(2000)	(2000)	(2000)	(2000)	(2000)	(2000)
			2000	81.8	15.22	77.0	14.22	71.0	9.37
4	$g = 0.85\%$; $\sigma = 0.5$	0.302	1900	65.0	8.22	65.0	8.22	65.0	8.22
			Max. value	69.0	13.71	77.9	18.42	65.0	8.22
			(Year)	(1950)	(1990)	(2000)	(2000)	(1900)	(1900)
			2000	67.8	13.68	77.9	18.42	47.8	4.50
5	$N = 6$	0.304	1900	65.0	8.22	65.0	8.22	65.0	8.22
			Max. value	68.8	13.64	77.7	18.04	65.0	8.22
			(Year)	(1950)	(1990)	(2000)	(2000)	(1900)	(1900)
			2000	67.6	13.59	77.7	18.04	52.4	4.25
6	$T = 105$	0.302	1900	65.0	8.22	65.0	8.22	65.0	8.22
			Max. value	69.0	13.70	78.1	18.45	65.0	8.22
			(Year)	(1950)	(1990)	(2000)	(2000)	(1900)	(1900)
			2000	67.9	13.67	78.1	18.45	53.0	5.29

Table 2: Sensitivity Analysis