WORKING PAPER

A Theory of Education and Health

Titus J. Galama and Hans van Kippersluis

RAND Labor & Population
RAND Bing Center

WR-1094
March 2015
This paper series made possible by the NIA funded RAND Center for the Study of Aging (P30AG012815) and the NICHD funded RAND Population Research Center (R24HD050906).

A Theory of Education and Health

Titus J. Galama* Hans van Kippersluis†

March 5, 2015

Abstract

This paper presents a unified theory of human capital with both health capital and, what we term, skill capital endogenously determined within the model. By considering joint investment in health capital and in skill capital, the model highlights similarities and differences in these two important components of human capital. Health is distinct from skill: health is important to longevity, provides direct utility, provides time that can be devoted to work or other uses, is valued later in life, and eventually declines, no matter how much one invests in it (a dismal fact of life). Lifetime earnings are strongly multiplicative in skill and health, so that investment in skill capital raises the return to investment in health capital, and vice versa. The theory provides a conceptual framework for empirical and theoretical studies aimed at understanding the complex relationship between education and health, and generates several new testable predictions.

Keywords: health investment; lifecycle model; human capital; health capital; optimal control

JEL Codes: D91, I10, I12, J00, J24

*University of Southern California, Dornsife College Center for Economic and Social Research, USA, and RAND Corporation, USA.

†Erasmus School of Economics, Erasmus University Rotterdam (EUR), and Tinbergen Institute, The Netherlands.

1 Introduction

The United States' 20^{th} Century was characterized by unprecedented increases in economic growth, with real per capita income in 2000 five to six times its level in 1900 (Goldin and Katz, 2009). The 20^{th} Century additionally differentiated itself by significant increases in life expectancy, health, and educational attainment. Life expectancy at birth increased by about 30 years, from 46 years in 1900 to 74 in 2000 for white men (Centers for Disease Control and Prevention; cdc.gov), and years of schooling rose from seven years in 1900 to 13 years in 2000 (Bleakley, Costa, and Lleras-Muney, 2013). Similary impressive advances in per capita income, life expectancy, and schooling took place in other developed and increasingly also in developing nations (Deaton, 2013). While increases in life expectancy, health status, and educational attainment appear to contribute to economic growth (Barro, 2001; Bloom and Canning, 2000; Bloom, Canning, and Sevilla, 2004; Goldin and Katz, 2009),[1] it is less clear to what extent, and how, the trends in life expectancy, health, and education are related.

Studying these relations is traditionally guided by human-capital theory, the foundations of which have been laid by the seminal works of Schultz (1961), Becker (1964), Ben-Porath (1967), and Mincer (1974). What the canonical human capital model does not deny, though largely leaves out, is that human capital is multidimensional (Acemoglu and Autor, 2012). Education and health are considered to be the most critical components of human capital (Schultz, 1961; Grossman, 2000), and while they share the defining characteristic of human capital that investing in them makes individuals more productive, there are several important differences between them. Perhaps most importantly, Becker (1964) observes that investments in human capital should decrease with age as the remaining period over which benefits can be accrued decreases. While this is clearly the case for education and training decisions, investments in health generally increase with age, even after retirement when health has lost its importance in generating earnings.[2]

This and other distinctions between health and other types of human capital, identified by, e.g., Mushkin (1962), have led to the development of the so-called health-capital model by Grossman (1972a,b). While the health-capital model has been very influential in health economics, and recognizes the role of education as a productivity-enhancing factor in health investment, it treats both education and longevity as exogenous. In doing so, it leaves out the possibility that individuals jointly optimize health, longevity, and education.[3] As a result, both human-capital theory and health-capital theory fall short

[1]But see Acemoglu and Johnson (2007; 2014) who suggest that gains in life expectancy generate limited or no economic growth.

[2]Investments in health consist of, e.g., medical care, physical exercise, a healthy diet and a healthy lifestyle. Not all such components of health investment necessarily increase with age. For example, the lifecycle profile of exercise is relatively flat (Podor and Halliday, 2012). But medical expenditures (e.g., Zweifel, Felder, and Meiers, 1999) and intake of fruit and vegetables do increase with age (Serdula et al. 2004; Pearson et al. 2005), and smoking rates drop with age (DHHS, 2014).

[3]Ehrlich and Chuma (1990) have included endogenous longevity in the Grossman model, and Galama

of providing a comprehensive framework to study the interactions between education, health, and longevity. As Michael Grossman (2000) put it *" ... Currently, we still lack comprehensive theoretical models in which the stocks of health and knowledge are determined simultaneously ... The rich empirical literature treating interactions between schooling and health underscores the potential payoffs to this undertaking ... "*.

This paper presents an explicit theory of joint investment in skill capital, health capital, and longevity, with three distinct (and endogenous) phases of life: schooling, work, and retirement.[4] Investments in health capital consist of, e.g., medical expenditures and physical exercise, while investments in skill capital consist of, e.g., expenditures on education and (on-the-job) training. Education (or schooling) is a distinct phase of life characterized by large investments in skill and limited or no work, and retirement is a distinct phase of life devoted to leisure and health investment. Individuals make their own decisions and are free to work, i.e., the start of the model corresponds to the mandatory schooling age (around 16 to 18 years for most developed nations) and the decision under consideration is whether to participate in post-mandatory education (or not) and for what duration.

The theory integrates (unifies) the human-capital and health-capital theoretical literatures. We are the first to develop such a comprehensive theory of education and health, the first to investigate such a theory analytically, and the first to employ it to make detailed predictions.[5]

The theory makes two main contributions to the literature. The first contribution is of a fundamental nature: by explicitly modeling joint investment in both skill and in health, the model defines and highlights the similarities and differences in the nature of skill and health. Like skill, health is an investment good that increases individuals' productivity (Grossman, 1972a). Yet, skill and health are different and not interchangeable. In contrast to skill, health provides direct utility (Grossman, 1972a; Murphy and Topel, 2006), and health extends life (Ehrlich and Chuma, 1990). In this paper, we argue for three additional

and Van Kippersluis (2010) have extended the model further by including health behaviors and the decision to accept unhealthy working conditions, which are important causes of ill-health and early mortality. Still, these models treat education as being determined outside of the model.

[4]In order to distinguish health clearly from the traditional notion of human capital, we employ in the remainder of the paper the term "skill capital" to refer to traditional human capital, "health capital" to refer to health, and "skill-capital literature" to refer to the traditional human-capital literature.

[5]Recently, the joint modeling of education and health has gained some traction. Becker (2007) develops a simple two period model of joint decisions regarding health and skill (which he refers to as education), where expenditures on skill and health in the first period increase expected wages and survival in the second period. Strulik (2013) and Carbone and Kverndokk (2014) develop multi-period models, which they numerically simulate and calibrate. Their aim is to simulate the implications of various policies on the association between years of schooling and health and to explain historical patterns. Hai and Heckman (2014) develop and structurally estimate a dynamic lifecycle model of health, skill and wealth that allows for credit constraints and rational addictive unhealthy behavior. Their aim is to quantify causal effects of education and wealth on health and unhealthy behavior and to quantify reverse causality and selection. Our aim is to develop a comprehensive analytical framework for human capital, where health, skill, and education are its most important inputs. Our analytical approach is complementary to these works and allows generating intuitive predictions that transparently follow from economic principles and assumptions.

distinctions. First, skill capital (largely) determines the wage rate, while health capital (largely) determines the time spent working, both within a day by decreasing sick time (as in Grossman, 1972a), but also over the life cycle by affecting retirement and life expectancy (two essential horizons that determine the period over which investments can be recouped). Second, individuals generally start life with a healthy body, but the terminal health state is universally low (for natural causes of death it is the physically frail that eventually face the great reaper). By contrast, individuals generally start life with limited skills, but end life with various degrees of cognitive and mental fitness (some of us have the good fortune to stay mentally sharp till death). In short, skill grows while health declines. Third, skill is valued mostly early in life while health is valued mostly later in life. Thus, investments in skill are high when young, while investments in health are high when old. Hence, despite broadly similar formulations of skill- and health-capital theory, differences in initial conditions, end conditions, and production processes, lead skill and health to exhibit fundamentally different dynamics.

The second contribution of the paper consists of providing a conceptual framework to guide empirical research in human capital. The unified theory provides new insights, makes new predictions, and explains stylized facts that the individual theories of skill capital and health capital on their own cannot. We highlight a few here and discuss these more extensively in section 4.

First, in the empirical literature, with very few exceptions, human capital is operationalized as years of education. Recognizing that health is an essential component of human capital suggests misspecification in many empirical applications of human-capital theory. Examples include, but are not limited to, the importance of health in development accounting efforts of economic growth (e.g., Weil, 2007), and the attribution of the hump-shaped earnings profile over the lifecycle to skill-capital decline (Mincer, 1974; Willis, 1985). Our theory suggests that the hump-shaped earnings profile is in part due to reductions in work as a result of declining health. Indeed Casanova (2013) finds that wages remain flat for two-thirds of workers till retirement, while the remaining one third has flat wages till they transition into partial or full retirement (often involuntarily, e.g., for health reasons).

Second, while the causal effect of education on health outcomes has received much theoretical and empirical attention, the reverse causal effect from health to education has only been studied empirically: it is absent from skill-capital as well as health-capital models.[6] The importance of this channel is illustrated by empirical studies that report a negative effect of childhood ill-health on educational attainment (Perri, 1984; Behrman and Rosenzweig, 2004; Case, Fertig, and Paxson, 2005; Currie, 2009; Bloom and Fink, 2013). Our model not only accounts for an effect of health on educational attainment, but additionally predicts that health raises skill formation beyond the school-leaving age, since (i) health and skill are strongly complementary in generating earnings, so that an increase

[6] However, see Carbone and Kverndokk, 2014, for an exception using a numerically simulated and calibrated model, and Bleakley, 2010a, for an informal discussion of the effect of health on years of schooling.

in health substantially raises the return to investment in skill, (ii) healthy individuals are potentially more efficient producers of skill, and (iii) healthy individuals live longer, increasing the return to skill investment by increasing the period over which its benefits can be reaped. This provides several additional pathways from health to educational attainment and from health to skill that are understudied in empirical as well as theoretical research.

Third, our model predicts a central role for longevity. The ability to utilize resources to postpone death (endogenous longevity) is crucial in explaining observed associations between wealth, skill, and health. Absent ability to extend life (fixed horizon), associations between wealth, skill, and health are absent or small (as in the traditional human- and health-capital literatures). If, however, life can be extended, wealth, skill, and health, are positively associated and the greater the degree of life extension, the greater is their association. The intuition behind this result is that the horizon (longevity) is a crucial determinant of the return to investment in skill and in health. This suggests that in situations where it is difficult to increase life expectancy, associations between wealth, skill, and health, would be weak. This may be the case for a developing nation (where there may be lack of access to basic medical care), for a nation with a high disease burden (where gains from tackling a certain disease may be limited due to the existence of other major diseases in the environment), for the developed world if it were faced with diminishing ability of technology to further extend life, for the developed world before the era of the industrial revolution, or for individuals faced with Huntington's (Oster, Shoulson, and Dorsey, 2013) or other diseases that severely impact longevity.

Fourth, and related, our model highlights that complementarity effects, operating through longevity, reinforce the associations between wealth, skill, health, and technology. That is, the combined effect is greater than the sum of the individual effects. As an example, improvements in the productivity of skill investment and in health investment are reinforced by gains in life expectancy. The United States' 20^{th} Century saw significant improvements in the productivity of health investments (e.g., clean water technologies, introduction of antibiotics) and reductions in the price of skill (e.g., compulsory schooling laws). It has been established that these technological and policy developments led to strong increases in life expectancy (Cutler and Miller, 2005; Lleras-Muney, 2005; Cutler, Deaton, and Lleras-Muney, 2006). Our theory suggests that the combination of (i) a higher productivity of skill and health investment, and (ii) the associated increase in life expectancy may have reinforced each other. Jointly they may have led to high returns to investments in both skill capital and health capital, potentially explaining the unprecedented increases in skill and health during the 20^{th} Century.

Fifth, our model contributes to the debate on whether the 20^{th} Century increase in educational attainment was the result of gains in life expectancy. Recently, this question has attracted attention. For example, Hazan (2009) incorporates a retirement decision in a stylized Ben-Porath model of skill capital to argue that a necessary condition for gains in longevity to increase educational attainment is that it also increases lifetime labor supply. He then shows that in the developed world lifetime labor supply has in fact declined amid significant gains in longevity, to conclude that longevity gains cannot be behind the observed increases in educational attainment. However, in contrast to the skill-capital literature where longevity is treated as exogenous, in our theory it is entirely possible that increases in years of schooling and in skill are accompanied by a reduction in lifetime labor supply since the associated gains in wealth increase the demand for leisure. An increase in lifetime labor supply is thus not a necessary, nor a sufficient, condition for higher life expectancy to induce an increase in the optimal years of schooling (see also Cervellati and Sunde, 2013).

These are just a few examples of how the theory can be used as an analytical framework to study empirical questions and to generate testable predictions. The detailed examples we provide in the paper, and the comparative dynamic analyses we employ to arrive at predictions, provide a template that can be followed by researchers to study their own particular research question of interest.

The paper is organized as follows. In the next section we provide further background on the relation between education, health, and life expectancy. In section 3 we outline our model formulation, contrast it with the existing literature, and present first-order and transversality conditions. In section 4 we discuss the lifecycle trajectories, analyze heterogeneity in these trajectories by employing comparative dynamic analyses, and develop about a dozen predictions. We conclude in section 5.

2 The relationship between education, health, and longevity

Typically, either skill-capital theory or health-capital theory is exploited to understand the interactions between education, health, and life expectancy.[7] Skill-capital theory (Becker, 1964; Ben-Porath, 1967) predicts a link between increases in life expectancy and educational attainment – known as the 'Ben Porath' mechanism: longevity (exogenous in the skill-capital literature) increases the return to investment in skill capital by lengthening the horizon over which benefits can be accrued. An extensive economic literature relies on the Ben-Porath mechanism to explain the transition from economic stagnation to growth, and the Ben-Porath mechanism has been invoked in high-level public debates and in policymaking to argue for improving the health and longevity of the poor (Hazan, 2009).

Health-capital theory also predicts a link between educational attainment, health, and longevity, albeit in the other direction. Grossman (1972a,b) emphasizes a causal effect of education (exogenous in the health-capital literature) on health and longevity. The argument is that the higher educated are more efficient producers of health investment

[7]Even though the trends in life expectancy, health, and education have largely moved in tandem over the 20^{th} century, it is a priori not clear that there are strong interactions between education on the one hand and health and life expectancy on the other. For example, concurrent skill-biased technological change (Goldin and Katz, 2009) and advances in medical technology and expanded insurance coverage (Fonseca et al. 2013) may have independently led to increases in educational attainment and improvements in health and mortality, respectively.

through (i) more efficient use of existing inputs (productive efficiency), e.g., better management of their diseases (Goldman and Smith, 2002), (ii) use of a better mix of health investment inputs (allocative efficiency), and (iii) early adoption of new knowledge and new technology (Lleras-Muney and Lichtenberg, 2005; Glied and Lleras-Muney, 2008). Education could also affect working conditions, (time) preferences, rank, and lifestyle choices (e.g., Cutler and Lleras-Muney, 2010), and higher educated individuals place a higher value on health compared to wealth (Galama and Van Kippersluis, 2010).

Empirically, there is supporting evidence for both directions of causality. Several studies have established a causal effect of education on health outcomes (Lleras-Muney, 2005; Conti, Heckman and Urzua, 2010; Van Kippersluis, O'Donnell, and van Doorslaer, 2011), although a number of recent studies find a very small or no effect (Mazumder, 2008; Albouy and Lequien, 2009; Meghir, Palme, and Simeonova, 2012; Clark and Royer, 2014). The Ben-Porath mechanism is also supported by several studies finding a positive effect of life expectancy on skill investment (Soares, 2006; Jayachandran and Lleras-Muney, 2009; Fortson, 2011; Oster, Shoulson, and Dorsey, 2013).

Since each theory emphasizes one particular direction of causality, skill-capital and health-capital theory on their own provide only partial, and often competing, explanations for the relation between education and health. In addition, both skill-capital theory and health-capital theory do not allow for an effect of health on years of schooling, evidence for which is found in many studies (e.g., Case, Fertig, and Paxson, 2005; Currie, 2009; Bloom and Fink, 2013). Childhood health may impact educational attainment through (i) the physical ability to attend school, (ii) associated improved cognitive ability and thereby learning (Grantham-McGregor et al. 2007, Bleakley, 2007; Madsen, 2012), and (iii) incentivizing parents to invest more in their children's education (Soares, 2005). Further, skill and health may be mutually reinforcing. They could exhibit *self-productivity*, where, e.g., a higher stock of skill in one period creates a higher stock of health in the next period (and vice versa), and *dynamic complementarity*, where, e.g., a higher stock of skill in one period makes investments in health in the next period more productive (and vice versa; Cunha and Heckman, 2007).

These various bi-directional causal mechanisms, do not feature in the extant empirical and theoretical skill- and health-capital literatures. A comprehensive framework aimed at studying the interactions between education, skill, health, investments, labor supply, and longevity, requires a framework in which skill and health are jointly determined. We formulate such a theory in the next section.

3 Model formulation and solutions

3.1 Model

The theory merges the human-capital literature with the health-capital literature. We treat health as a form of human capital that is distinct from the component of human capital that individuals invest in through education and training. We loosely refer to the latter as "skill capital" and the former as "health capital". Individuals invest in health (and longevity) through expenditures on (e.g., medical care) and time investments in (e.g., exercise) health; they invest in skill capital through outlays and time investments in skill (e.g., schooling and [on-the-job] training).

Individuals maximize the lifetime utility function

$$\mathbb{U} = \max_{X_C, L, I_E, I_H, S, R, T} \left\{ \int_0^S U[.]e^{-\beta t} dt + \int_S^R U[.]e^{-\beta t} dt + \int_R^T U[.]e^{-\beta t} dt \right\}, \qquad (1)$$

where time $t = 0$ corresponds to the mandatory schooling age (around 16 to 18 years for most developed nations), S denotes years of post-mandatory schooling (endogenous), R denotes the retirement age (endogenous), T denotes total lifetime (endogenous), β is a subjective discount factor and individuals derive utility $U[X_C(t), L(t), H(t)]$ from consumption goods and services $X_C(t)$, leisure time $L(t)$, and health $H(t)$. The utility function is increasing in each of its arguments and strictly concave.

The objective function (1) is maximized subject to the following dynamic constraints for skill capital $E(t)$ and health capital $H(t)$:

$$\frac{\partial E}{\partial t} = F_E\left[I_E(t), E(t), H(t)\right] = f_E\left[I_E(t), E(t), H(t)\right] - d_E(t)E(t), \qquad (2)$$

$$\frac{\partial H}{\partial t} = F_H\left[I_H(t), E(t), H(t)\right] = f_H\left[I_H(t), E(t), H(t)\right] - d_H(t)H(t). \qquad (3)$$

Skill capital $E(t)$ (equation 2) and health capital $H(t)$ (equation 3) can be improved through investments in, respectively, skill capital $I_E(t)$ and health $I_H(t)$, and deteriorate at the biological deterioration rates $d_E(t)$ and $d_H(t)$. Goods and services $X_E(t)$, $X_H(t)$, purchased in the market and own time inputs $\tau_E(t)$, $\tau_H(t)$, are used in the production of investment in skill capital and in health capital $I_E(t)$, $I_H(t)$:

$$\begin{aligned} I_E(t) &= I_E[X_E(t), \tau_E(t)], \\ I_H(t) &= I_H[X_H(t), \tau_H(t)]. \end{aligned}$$

The skill-capital F_E and health-capital F_H production processes are assumed to be increasing and strictly concave in the investment inputs $X_E(t)$, $\tau_E(t)$, and $X_H(t)$, $\tau_H(t)$, respectively.[8] Crucially, this assumption of diminishing returns to investment (concavity) addresses the degeneracy of the solution for investment that plague the health-capital literature as a result of the common assumption of constant returns to scale (see for a discussion Ehrlich and Chuma, 1990; Galama, 2011; Galama and Van Kippersluis, 2013).

The efficiencies of investment in skill capital and health capital are assumed to be functions of the stocks of skill capital $E(t)$ and of health $H(t)$. This allows us to model self-productivity, where skills produced at one stage augment skills at later stages, and

[8] Concavity implies $\partial^2 F_E/\partial X_E^2 < 0$, $\partial^2 F_E/\partial \tau_E^2 < 0$, $\partial^2 F_H/\partial X_H^2 < 0$, $\partial^2 F_H/\partial \tau_H^2 < 0$, $\left(\partial^2 F_E/\partial X_E^2\right)\left(\partial^2 F_E/\partial \tau_E^2\right) > \left(\partial^2 F_E/\partial X_E \partial \tau_E\right)^2$ and $\left(\partial^2 F_H/\partial X_H^2\right)\left(\partial^2 F_H/\partial \tau_H^2\right) > \left(\partial^2 F_H/\partial X_H \partial \tau_H\right)^2$.

dynamic complementarity, where skills produced at one stage raise the productivity of investment at later stages. Self-productivity can be self-reinforcing $\partial F_E/\partial E > 0$, $\partial F_H/\partial H > 0$, and/or cross fertilizing, $\partial F_E/\partial H > 0$, $\partial F_H/\partial E > 0$. Dynamic complementarity too can be self-reinforcing, $\partial^2 F_E/\partial E \partial I_E > 0$, $\partial^2 F_H/\partial H \partial I_H > 0$, and/or cross fertilizing, $\partial^2 F_E/\partial E \partial I_H > 0$, $\partial^2 F_H/\partial H \partial I_E > 0$ (Cunha and Heckman, 2007).

The intertemporal budget constraint for assets $A(t)$ is given by

$$\frac{\partial A}{\partial t} = rA(t) + Y[t, E(t), H(t)] - p_C(t)X_C(t) - p_E(t)X_E(t) - p_H(t)X_H(t). \quad (4)$$

Assets $A(t)$ (equation 4) provide a return r (the rate of return on capital) and increase with income

$$Y[t, E(t), H(t)] = b_S(t) \qquad 0 \leq t < S, \quad (5)$$
$$Y[t, E(t), H(t)] = w[t, E(t)]\tau_w[t, H(t)] \qquad S \leq t < R, \quad (6)$$
$$Y[t, E(t), H(t)] = b_R(R) \qquad R \leq t < T, \quad (7)$$

where during the schooling period (up to S) individuals receive a state (or parental) transfer to fund schooling $b_S(t)$ (e.g., financial aid, conditional on being in school), and during retirement individuals receive a state or private pension annuity $b_R(R)$, typically a fraction (replacement ratio) of average earnings over the individual's working period, and assumed to be a function of retirement age R. During working life (between S and R) earnings consist of the product of the wage rate $w[t, E(t)]$ and the time spent working $\tau_w[t, H(t)]$.[9] Hence, skill capital (largely) determines the wage rate, while health capital (largely) determines the time spent working. Assets decrease with expenditures on investment and consumption goods and services $X_E(t)$, $X_H(t)$ and $X_C(t)$, at prices $p_E(t)$, $p_H(t)$ and $p_C(t)$. Alternatively, or in addition to the schooling subsidy $b_S(t)$, the government may subsidize the cost of skill formation by reducing or fully subsidizing the price $p_E(t)$ of skill investment while in school.[10]

Finally, the total time constraint Ω is given by

$$\Omega = \tau_w(t) + L(t) + \tau_E(t) + \tau_H(t) + s[H(t)]. \quad (8)$$

During working life, the total available time Ω is divided between time spent working $\tau_w(t)$, leisure time $L(t)$, time investments in skill and in health capital $\tau_E(t)$, $\tau_H(t)$, and time lost due to illness $s[H(t)]$ (assumed to be a decreasing function of health). During school years and during retirement individuals do not work, i.e.

$$\tau_w(t) = 0. \quad (9)$$

[9] $\tau_w[t, H(t)]$ is an explicit function of health status through sick time $s[H(t)]$ and an implicit function of health status through the optimal response of time inputs (consumption, skill- and health-capital investment) to health status.

[10] We assume, for simplicity, that if an individual decides to continue her education $S > 0$, she is not allowed to work. In practice there may be attendance requirements and, depending on how stringent these are, students may have varying degrees of time available that they could devote to work for pay.

Thus, we have the following optimal control problem: the objective function (1) is maximized with respect to the control functions $X_C(t)$, $X_E(t)$, $X_H(t)$, $L(t)$, $\tau_E(t)$, $\tau_H(t)$, and the parameters S, R, and T, subject to the constraints (2) to (8), and the following initial and end conditions: $H(0) = H_0$, $H(T) = H_T$, $E(0) = E_0$, $A(0) = A_0$, $A(T) = A_T$, and $E(T) \geq 0$ (and free). Length of life T (Grossman, 1972a;b) is determined by a minimum health level below which an individual dies: $H_T \equiv H_{\min}$.

The Lagrangian (see, e.g., Seierstad and Sydsaeter, 1987; Caputo, 2005) of this problem is:

$$\begin{aligned} \Im &= U[X_C(t), L(t), H(t)]e^{-\beta t} + q_E(t)\frac{\partial E}{\partial t} + q_H(t)\frac{\partial H}{\partial t} + q_A(t)\frac{\partial A}{\partial t} \\ &+ \lambda_{\tau_w}(t) w[t, E(t)]\tau_w(t) + \lambda_{H_{\min}}(t)\left[H(t) - H_{\min}\right], \end{aligned} \tag{10}$$

where $q_E(t)$, $q_H(t)$, and $q_A(t)$ are the co-state variables associated with, respectively, the dynamic equations (2) for skill capital $E(t)$, (3) for health $H(t)$, and (4) for assets $A(t)$, the multiplier $\lambda_{\tau_w}(t)$ is associated with the condition that individuals do not work during school years and retirement (9) ($\lambda_{\tau_w}(t) = 0$ if $\tau_w(t) > 0$ and $\lambda_{\tau_w}(t) > 0$ if $\tau_w(t) = 0$),[11] and $\lambda_{H_{\min}}(t)$ is the multiplier associated with the condition that $H(t) > H_{\min}$ for $t < T$.

The co-state variables $q_E(t)$, $q_H(t)$, and $q_A(t)$ find a natural economic interpretation in the following standard result from Pontryagin

$$q_Z(t) = \frac{\partial}{\partial Z(t)} \int_t^{T^*} U(*)e^{-\beta s} ds, \tag{11}$$

(e.g., Caputo 2005, eq. 21 p. 86) with $Z(t) = \{E(t), H(t), A(t)\}$, and where T^* denotes optimal length of life and $U(*)$ denotes the maximized utility function (i.e., along the optimal paths for the controls, state functions, and for the optimal schooling age, retirement age, and length of life). Thus, for example, $q_E(t)$ represents the marginal value of remaining lifetime utility (from t onward) derived from additional skill capital $E(t)$. We refer to the co-state functions as the "marginal value of skill", the "marginal value of health", and the "marginal value of wealth" (these are also often referred to as the shadow prices of skill capital, of health capital, and of wealth).

Since skill capital $E(T)$ is unconstrained (free), the individual chooses it to have no value at the end of life, $q_E(T) = 0$. However, health capital $H(T)$ and assets $A(T)$ are constrained to their values $H_{\min}$ and A_T, respectively, and as a result they cannot be chosen not to have value at the end of life, and $q_H(T) \geq 0$ and $q_A(T) \geq 0$.

The transversality condition for the optimal length of schooling S, the optimal age of retirement R, and the optimal length of life T, follow from the dynamic envelope theorem

[11] The last term contains the wage rate explicitly to ensure that the dimension of $\lambda_{\tau_w}(t)$ is the same as that of the marginal value of wealth $q_A(t)$, and because time is valued at the wage rate. Ultimately the multiplier is determined by the condition $\tau_w(t) = 0$ and using $\lambda_{\tau_w}(t)$ or $\lambda_{\tau_w}(t)w[t, E(t)]$ has no effect on model solutions.

(e.g., Caputo 2005, p. 293):

$$\frac{\partial}{\partial S}\int_0^T \Im(t)dt = \Im(S_-) - \Im(S_+) + \int_0^T \frac{\partial \Im(t)}{\partial S}dt = 0, \tag{12}$$

$$\frac{\partial}{\partial R}\int_0^T \Im(t)dt = \Im(R_-) - \Im(R_+) + \int_0^T \frac{\partial \Im(t)}{\partial R}dt = 0, \tag{13}$$

$$\frac{\partial}{\partial T}\int_0^T \Im(t)dt = \Im(T) = 0, \tag{14}$$

where S_-, R_- indicate the limit in which S, R are approached from below, and S_+, R_+ the limit in which S, R are approached from above. Conditions (12) and (13) have a natural interpretation in that the optimal length of schooling S and the optimal retirement age R are chosen such that there is no benefit of delaying entrance to the labor market (associated with optimal length of schooling S) and no benefit of delaying retirement R. $\Im(T)$ is the marginal value of life extension T (e.g., Theorem 9.1, p. 232 of Caputo, 2005), and the age at which life extension no longer has value defines the optimal length of life T^*.

3.2 Comparison with the literature

The canonical skill- and health-capital theories (Ben-Porath, 1967; Grossman, 1972a,b, 2000; Ehrlich and Chuma, 1990) models are sub-models of our formulation. The Ben-Porath (1967) model is obtained by removing the schooling decision S, the retirement decision R, leisure time $L(t)$, investment in health capital $I_H(t)$, $X_H(t)$ and $\tau_H(t)$, sick time $s[H(t)]$, and the dynamic equation (3) for health capital $H(t)$ from the model, and assuming fixed length of life T and a skill-capital production process of the form $f_E(t) = A\left[\tau_E(t)E(t)\right]^{\alpha}\left[X_E(t)\right]^{\beta}$ (Ben-Porath neutrality).[12]

The Grossman model is obtained by removing the schooling decision S, the retirement decision R, leisure time $L(t)$, investment in skill capital $I_E(t)$, $X_E(t)$ and $\tau_E(t)$, the dynamic equation (2) for skill capital $E(t)$, and the explicit condition for optimal length of life T (equation 14), and assuming a constant returns to scale health-production process: this consists of the standard assumption made in the health-capital literature of a linear health-production process $f_H(t) = I_H(t)$ and a Cobb-Douglas relation for health investment $I_H(t) = \mu_H(E)X_H(t)^{k_H}\tau_H(t)^{1-k_H}$. The efficiency of health investment $\mu_H(E)$ is allowed to be a function of exogenous skill capital E.

The formulations of Ehrlich and Chuma (1990) and Galama (2011) are as in Grossman (1972a,b; 2000) but assume a health-production process $f_H(t)$ with decreasing returns to scale in investment $I_H(t)$ and add an explicit condition for endogenous longevity (i.e. equation 14).

[12] Even though the Ben-Porath model is formulated as maximizing lifetime earnings $Y(t)$, the characteristics of the model are very similar for a formulation in which the utility of lifetime consumption is maximized (as in this paper), with some exceptions (Graham, 1981).

3.3 First-order conditions and interpretation

In this section we present and discuss the first-order (necessary) conditions and the transversality conditions of the optimal-control problem discussed above. The first-order conditions determine the optimal solutions of the controls skill-capital investment $I_E(t)$, health-capital investment $I_H(t)$, consumption $X_C(t)$, and leisure time $L(t)$, respectively.[13] Appendix A.1 provides detailed derivations. In this section we focus on working ages (i.e. the period between S and R). In section 3.4 we discuss the schooling and retirement phase.

Consumption and leisure The first-order conditions for consumption and leisure are standard

$$\frac{1}{q_A(t)}\frac{\partial U}{\partial X_C} = p_C(t)e^{\beta t}, \tag{15}$$

$$\frac{1}{q_A(t)}\frac{\partial U}{\partial L} = w[t, E(t)]e^{\beta t}. \tag{16}$$

Consumption $X_C(t)$ and leisure time $L(t)$ increase with current wealth $A(t)$ under the standard assumption of diminishing returns to wealth $\partial q_A(t)/\partial A(t) < 0$,[14] and with permanent income (the marginal value of wealth $q_A(t)$ decreases with permanent income).

Consumption and leisure decrease with their respective costs: the price of goods and services $p_C(t)$ (for consumption) and the opportunity cost of time $w[t, E(t)]$ (for leisure). Hence, anticipated increases in wages raise the opportunity cost of time and lead to a reduction in leisure (such a change occurs along the optimal lifecycle trajectory and does not affect the marginal value of wealth $q_A(t)$), but unanticipated (transitory or permanent) increases in wages also raise permanent income (such a change shifts the optimal life cycle trajectory by reducing the marginal value of wealth $q_A(t)$), and may therefore increase leisure if the permanent income effect dominates the opportunity cost of time effect.

Skill-capital investment The first-order condition for skill-capital investment $I_E(t)$ is given by

$$q_{e/a}(t) = \pi_E(t), \tag{18}$$

which equates the marginal benefit of skill-capital investment, given by the ratio of the marginal value of skill capital and the marginal value of wealth $q_{e/a}(t) \equiv q_E(t)/q_A(t)$, or

[13] The first-order conditions for goods and services $X(t)$ are the same as for time inputs $\tau(t)$, as reflected in conditions (21) and (26) (see also Appendix A.1). As a result we have four rather than six controls.

[14] A natural and frequently made assumption is that financial capital (wealth) $A(t)$, skill capital $E(t)$, and health capital $H(t)$, increase remaining lifetime utility (from t onwards), but at a diminishing rate

$$\frac{\partial q_Z(t)}{\partial Z(t)} = \frac{\partial^2}{\partial Z(t)^2}\int_t^{T^*} U(*)e^{-\beta s}ds < 0, \tag{17}$$

with $Z(t) = \{E(t), H(t), A(t)\}$ (see 11).

the *relative marginal value of skill*, for short, to the marginal monetary cost of skill-capital investment $\pi_E(t)$.

The relative marginal value of skill is the solution to the dynamic co-state equation[15]

$$-\frac{\partial q_{e/a}}{\partial t} \quad = \quad \frac{\partial Y}{\partial E} + q_{e/a}(t)\left\{\frac{\partial f_E}{\partial E} - [d_E(t) + r]\right\} + q_{h/a}(t)\frac{\partial f_H}{\partial E}, \tag{20}$$

where the rate at which the relative marginal benefit of skill $q_{e/a}(t)$ depreciates over a short interval of time (left-hand side [LHS]) equals the sum of the direct benefits of skill capital and the contribution of skill capital to enhancing the value of the capital stocks (Dorfman, 1982).[16] Skill capital contributes to wealth by raising earnings $\partial Y/\partial E > 0$ (a production benefit), to skill by raising the efficiency of skill-capital production $\partial f_E/\partial E > 0$ (self-reinforcing self-productivity; valued at the relative marginal value of skill $q_{e/a}(t)$), and to health by raising the efficiency of health-capital production $\partial f_H/\partial E > 0$ (cross-fertilizing self-productivity; valued at the relative marginal value of health $q_{h/a}(t)$). The relative marginal value of skill appreciates with $d_E(t)$ (biological aging depletes the stock of skill, a cost) and the rate of return to capital r (the opportunity cost of investing in skill capital rather than in the stock market), where both costs are valued at the relative marginal value of skill $q_{e/a}(t)$.

The marginal cost of skill-capital investment is defined as

$$\pi_E(t) \equiv \frac{p_E(t)}{\frac{\partial f_E}{\partial I_E}\frac{\partial I_E}{\partial X_E}} = \frac{w[t, E(t)]}{\frac{\partial f_E}{\partial I_E}\frac{\partial I_E}{\partial \tau_E}}. \tag{21}$$

The marginal cost of investment in skill capital increases with the price of investment goods and services $p_E(t)$, and the opportunity cost of not working $w[t, E(t)]$,[17] and decreases in the efficiency of the use of investment inputs in the skill production process, $\partial f_E/\partial X_E$ and

[15] Or, alternatively

$$q_{e/a}(t) = \int_t^T e^{-\int_t^s \left[d_E[x]+r-\frac{\partial f_E}{\partial E}\right]dx}\left(\frac{\partial Y}{\partial E} + q_{h/a}(s)\frac{\partial f_H}{\partial E}\right) ds. \tag{19}$$

Thus the relative marginal value of skill $q_{e/a}(t)$ represents the sum of the lifetime production benefit (earnings) of skill $\partial Y/\partial E$ and the lifetime health-production benefit of skill $\partial f_H/\partial E$, discounted at the rate $d_E(t) + r - \partial f_E/\partial E$, where the discount rate is reduced as a result of the skill-production benefit of skill $\partial f_E/\partial E$.

[16] A unit of skill capital loses value (depreciates) as time passes at the rate at which its potential contribution to lifetime utility becomes its past contribution. This can be understood as follows. The utility that can be derived from skill capital over the remainder of life (from t onward) is a finite amount (analogous to, e.g., the distance a rocket can travel on the remaining fuel). The marginal value of skill $q_E(t)$ is the derivative with respect to the stock $E(t)$ of the remaining lifetime utility (see 11). If the contribution of skill to utility in the current period is high, then the rate of depreciation of the marginal value of skill is high (analogous to the rate at which fuel is burned).

[17] Here too, an anticipated increase in wages raises the opportunity cost of time and hence the marginal cost of skill investment, while an unanticipated increase in wages raises permanent income and the opportunity cost of time, plausibly increasing the relative marginal value of skill.

$\partial f_E/\partial \tau_E$. Because of diminishing returns to scale in skill-capital investment, the marginal cost of skill capital is an increasing function of the level of investment goods / services $X_E(t)$ and investment time inputs $\tau_E(t)$, and hence an increasing function of the level of investment $I_E(t)$ (see 21).[18] Intuitively, due to concavity in investment $I_E(t)$ of the skill production process $f_E(t)$, the higher the level of investment, the smaller the improvement in skill $E(t)$. As a result, the effective cost of investment is higher.

In sum, the decision to invest in skill today (18) weighs the current monetary price and current opportunity cost of time (see 21) with its future benefits (from t to T), consisting of increased earnings, and more efficient skill and health production (see 19).

Health-capital investment Analogous to skill-capital investment, the first-order condition for health-capital investment is given by

$$q_{h/a}(t) \quad = \quad \pi_H(t), \tag{23}$$

where the marginal benefit of health investment $q_{h/a}(t)$ equals the ratio of the marginal value of health and the marginal value of wealth $q_{h/a}(t) \equiv q_H(t)/q_A(t)$, or the *relative marginal value of health*, for short, and $\pi_H(t)$ represents the marginal monetary cost of health-capital investment.

The relative marginal value of health is the solution to the co-state equation[19]

$$\begin{aligned} -\frac{\partial q_{h/a}}{\partial t} \quad &= \quad \frac{1}{q_A(t)}\frac{\partial U}{\partial H}e^{-\beta t} + \frac{\partial Y}{\partial H} + q_{e/a}(t)\frac{\partial f_E}{\partial H} \\ &+ \quad q_{h/a}(t)\left\{\frac{\partial f_H}{\partial H} - [d_H(t) + r]\right\} + \frac{\lambda_{H_{\min}}(t)}{q_A(t)}, \end{aligned} \tag{25}$$

and the marginal cost of health investment is defined as

$$\pi_H(t) \equiv \frac{p_H(t)}{\frac{\partial f_H}{\partial I_H}\frac{\partial I_H}{\partial X_H}} = \frac{w[t, E(t)]}{\frac{\partial f_H}{\partial I_H}\frac{\partial I_H}{\partial \tau_H}}. \tag{26}$$

[18] Because of concavity of f_E the first derivatives of the production process $\partial f_E/\partial X_E$ and $\partial f_E/\partial \tau_E$, are monotonically decreasing functions of X_E and τ_E, respectively. For example, for the functional form $f_E[I_E(t), E(t), H(t)] \equiv f_E^*[E(t), H(t)]I_E(t)^{\alpha_E}$ (where $0 < \alpha_E < 1$ [diminishing returns]) and a Cobb-Douglass relation between the inputs X_E, τ_E and the output investment $I_E(t)$, $I_E(t) \equiv X_E(t)^{k_E}\tau_E(t)^{1-k_E}$, we have

$$\pi_E(t) = \frac{p_E(t)^{k_E} w[t, E(t)]^{1-k_E}}{\alpha_E f_E^*[E(t), H(t)] k_E^{k_E}(1-k_E)^{1-k_E}} I_E(t)^{1-\alpha_E} \equiv \pi_E^*(t) I_E(t)^{1-\alpha_E}. \tag{22}$$

[19] Or, alternatively

$$\begin{aligned} q_{h/a}(t) &= q_{h/a}(T)e^{-\int_t^T\left[d_H(x)+r-\frac{\partial f_H}{\partial H}\right]dx} \\ &+ \int_t^T e^{-\int_t^s\left[d_H[x]+r-\frac{\partial f_H}{\partial H}\right]dx}\left(\frac{1}{q_A(0)}\frac{\partial U}{\partial H}e^{-(\beta-r)s} + \frac{\partial Y}{\partial H} + q_{e/a}(s)\frac{\partial f_E}{\partial H} + \frac{\lambda_{H_{\min}}(s)}{q_A(s)}\right)ds \end{aligned} \tag{24}$$

Like skill capital, the benefits of health capital consist of increasing earnings $\partial Y/\partial H > 0$ (a production benefit), and potentially raising the efficiency of skill production $\partial f_E/\partial H > 0$ (cross-fertilizing self-productivity; valued at the relative marginal value of skill $q_{e/a}(t)$). Unlike skill capital, health also has a consumption benefit (direct utility) $\partial U/\partial H$, health enables life extension (see 14; Ehrlich and Chuma, 1990), and it is unclear whether health enhances or reduces the efficiency of health production $\partial f_H/\partial H$. The relative marginal value of health appreciates with $d_H(t)$ (biological aging depletes the stock, a cost) and the rate of return to capital r (the opportunity cost of investing in health rather than in the stock market). Both costs are valued at the relative marginal value of health $q_{h/a}(t)$. Further, the constraint that health cannot fall below a minimum level $H_{\min}$ is reflected in an additional term $\lambda_{\mathrm{H}_{\min}}(t)/q_A(t)$ in (25). The term is absent from the relation for the relative marginal value of skill (20), and it increases the rate at which the marginal value of health depreciates. In practice, employing the condition entails restricting solutions to those where the constraint is not imposing.

The marginal cost $\pi_H(t)$ of health investment increases with the price of goods and services in the market $p_H(t)$, the opportunity cost of time $w[t; E(t)]$, and in the level of investment $I_H(t)$ due to decreasing returns to scale. It decreases in the efficiency of the use of investment inputs in the health production process, $\partial f_H/\partial X_H$ and $\partial f_H/\partial \tau_H$.

In sum, similar to investment in skill, the decision to invest in health today (23) weighs the current monetary price and current opportunity cost of time (see 26) with its future benefits (from t to T), consisting of enhanced utility, increased earnings, more efficient skill production, and a longer life (see 24).

Dynamics The dynamic equations for skill (2) and for the relative marginal value of skill (20), together with the initial, end, and transversality conditions, determine the evolution of skill and skill investment over the lifecycle. Likewise, the dynamic equations for health (3) and for the relative marginal value of health (25) determine the evolution of health and health investment.[20] Skill capital and health capital have different initial and terminal conditions. Individuals begin life with limited skills and end life with various degrees of cognitive and mental fitness. This notion is captured in the skill-capital literature by an initial low level of skill E_0 and an end value $E(T)$ that is, apart from being non-negative, unconstrained. Because there is no restriction on the terminal value of the stock of skill $E(T)$, it is chosen such that skill no longer has value at the end of life, $q_E(T) = 0$ (see Heckman, 1976; Chiang, 1992). Thus the relative marginal value of skill capital $q_{e/a}(t)$, and therefore investment in skill, decreases over the life-cycle and approaches zero at the end of life (see 19).

[20] The evolution of assets is given by (4), and the marginal value of assets is determined by its co-state equation (see equation 37 in Appendix A.1):

$$-\frac{\partial q_A}{\partial t} \quad = \quad q_A(t)r. \tag{27}$$

In contrast, most people start adult life with a healthy body, and for natural causes of death the terminal state of health is universally frail. The notion that health cannot be sustained below a certain minimum level is captured in the health-capital literature by the condition $H(T) = H_{\min}$. Health capital eventually decreases over the lifecycle and because the terminal health stock is restricted to $H_{\min}$, it cannot be chosen to have no value, $q_H(T) \geq 0$.

Conjecture 1: Skill capital generally grows as a result of investment but health capital eventually declines.

Conjecture 2: Skill is valued early while health is valued later in life.

Limiting the discussion to adulthood, skill capital is found to increase, at least initially (e.g., Becker 1964, Ben-Porath, 1967), while health capital is found to decrease with age (e.g., Grossman, 1972a;b). Skill-capital investment is thus characterized by a production process that enables improvements in the stock of skill, while health-capital investment is characterized by a production process that (eventually) cannot prevent declining health, no matter how much one invests in it (a dismal fact of life; conjecture 1).

Further, the empirical and theoretical literatures suggest that investments in skill capital tend to decrease with age (e.g., Becker, 1964), while investments in health tend to increase with age (e.g., Zweifel, Felder, and Meiers, 1999). This suggests that the relative marginal value of health $q_{h/a}(t)$ increases with age, while the relative marginal value of skill $q_{e/a}(t)$ decreases with age. Skill is valued early in life while health is valued later in life.[21]

In essence, the two criteria of whether (i) the stock is increasing or decreasing, and (ii) the relative marginal value is increasing or decreasing, define four different regions in the phase diagram. Skill and health occupy distinct regions of phase space. An example is shown in Figure 2 in Appendix A.3 for a simpler version of our theory.

3.4 The schooling, work, and retirement decision

The transition from school to work Our theory emphasizes the difference in nature between years of schooling S, and skill capital $E(t)$. While these terms are often used interchangeably in the literature, we define years of schooling S as the individual's choice of the optimal number of years to attend school, while skill capital $E(t)$ is a stock that builds gradually through investment decisions, of time $\tau_E(t)$ and goods/services $X_E(t)$ devoted to skill formation, that are made in every period.

[21] A decreasing relative marginal value of skill with age suggests high (initial) benefits and slow biological aging of skill (see 20), while an increasing relative marginal value of health with age suggests health is associated with small (initial) benefits and rapid biological aging of health (see 25). Another way of looking at this is that diminishing marginal benefits to skill and health imply that increasing skill decreases benefits ($\partial Y/\partial E$, $\partial f_H/\partial E$, $\partial f_E/\partial E$) while declining health increases benefits ($\partial U/\partial H$, $\partial Y/\partial H$, $\partial f_H/\partial H$, $\partial f_E/\partial H$) over time. Thus, the benefits of skill decline, while the benefits of health grow with age.

In school, the opportunity cost of time investments (e.g., attending class, studying, completing assignments) is low since students do not work and the time that would otherwise be spent working can now be devoted to skill investment $\tau_E(t)$, health investment $\tau_H(t)$, and leisure $L(t)$.[22] Though not exclusively, individuals will use the schooling period (i.e. $t < S$) as a period of life to invest in skill capital $E(t)$, since skill is valued most early in life (see conjecture 2), and because during the schooling phase individuals are encouraged to invest in skill through an education transfer $b_S(t)$, and/or through subsidized schooling (reduced $p_E(t)$; e.g., public schooling).

As individuals gain skill their potential labor earnings increase, and at some point it becomes attractive to join the labor market. Individuals will join the labor market at the age S, the age at which the benefits of work exceed the benefits of staying in school.[23] The transversality condition (12) for the optimal years of schooling S requires us to consider the difference in the Lagrangian right before and right after the individual leaves school, $\Im(S_-) - \Im(S_+)$. Since the Lagrangian does not explicitly depend on S, the last term in (12) is zero. Noting that state and co-state functions are continuous in S, and $\lambda_{\tau_w}(S)w[t, E(t)]\tau_w(S) = 0$, the transversality condition (12) reduces to

$$\begin{aligned} Y(S_+) &= b_S(S) + \frac{1}{q_A(S)}\left[U(S_-) - U(S_+)\right]e^{-\beta S} \\ &+ q_{e/a}(S)\left[f_E(S_-) - f_E(S_+)\right] + q_{h/a}(S)\left[f_H(S_-) - f_H(S_+)\right] \\ &+ p_H(S)\left[X_H(S_+) - X_H(S_-)\right] + p_E(S_+)X_E(S_+) - p_E(S_-)X_E(S_-) \\ &+ p_C(S)\left[X_C(S_+) - X_C(S_-)\right], \end{aligned} \tag{28}$$

where $Y(S_+) = w(S)\tau_w(S_+)$, and we have replaced the limits S_- and S_+ with S for functions that are continuous in S. The optimal school-leaving age S requires the benefits of entering the labor market to equal the benefits of staying in school. The LHS of (28) represents the benefits of entering the labor market consisting of labor income $Y(S_+)$. The right-hand side (RHS) represents the benefit of staying in school, consisting of the schooling subsidy $b_S(S)$ (first term), the monetary value of utility from more leisure time (second term),[24] and the value of higher levels of skill investment and health investment while in

[22] In the absence of earnings from wages, the opportunity cost of time is not determined by the wage rate $w(t)$ but by the constraint (9) that individuals not work $\tau_w(t) = 0$. This situation no longer represents an interior solution but rather a corner solution with $\tau_w(t) = 0$. A simple heuristic argument can be made that the opportunity cost of time is always lower at every age during schooling years (and retirement years), as follows. When individuals are allowed to work they may devote very little time to work, but they will never choose not to work since the decision to work provides an additional margin of adjustment with some benefit. Thus the total time available that can be devoted to leisure, consumption and investment is larger when not working, and thus the opportunity cost of time lower. A comparison of the first-order conditions in Appendix A.1 shows that one can obtain the first-order conditions for non-working ages simply by replacing all occurances of the monetary value of the opportunity cost $q_A(t)w(t;E)$ with $\lambda_{\tau_w}(t)w(t;E)$.

[23] See Haley (1973) for an alternative way of modeling the period of schooling/specialization.

[24] Even if leisure and consumption are substitutes in utility, utility right before the transition from schooling to work is arguably still higher, $U(S_-) - U(S_+) > 0$, as the effect of a change in consumption on utility is a second-order effect (and thus relatively small), operating through the effect that a change

school due to the lower opportunity cost of time (third and fourth term). Further, if time substitutes for goods and services $X_H(t)$ in the production of health investment, then the fifth term on the RHS represents benefits of schooling in terms of reduced expenditures. The sixth term reflects the possibility that the cost of schooling $p_E(t)X_E(t)$ (for $t < S$) may be subsidized, providing another benefit of staying in school. The last term reflects changes in consumption as a result of changes in the marginal utility of consumption due to reduced leisure time while working.

The transition from work to retirement After graduating from school, individuals enter the labor market, and start working. Time previously devoted to skill investment, health investment, and leisure is reduced. As a result, skill increases at a slower pace, and health deteriorates faster. Declining health reduces earnings – through increased sick time and by increasing time devoted to health investment –, and retirement becomes increasingly attractive. Retirement is in part attractive because it lowers the opportunity cost of time. The time otherwise spent working can then be devoted to leisure $L(t)$, time inputs into health-capital investment $\tau_H(t)$, and time inputs into skill-capital investment $\tau_E(t)$.

The optimal retirement age is determined by the transversality condition (13), and requires us to consider the difference in the Lagrangian right before and right after retirement, $\Im(R_-) - \Im(R_+)$ (see also Kuhn et al. 2012). The Lagrangian depends explicitly on the retirement age R, to capture the fact that retirement benefits are often a strong function of the retirement age (and hence the last term of 13 does not vanish).

Noting that state and co-state functions are continuous in R, and $\lambda_{\tau_w}(R)\tau_w(R) = 0$, the transversality condition (13) reduces to

$$\begin{aligned}
Y(R_-) &= \left\{ b(R) - \frac{\partial b(R)}{\partial R}\frac{1}{r}\left[1 - e^{-r(T-R)}\right] \right\} \\
&+ \frac{1}{q_A(R)}\left[U(R_+) - U(R_-)\right] e^{-\beta R} \\
&+ q_{e/a}(R)\left[f_E(R_+) - f_E(R_-)\right] + q_{h/a}(R)\left[f_H(R_+) - f_H(R_-)\right] \\
&+ p_H(R)\left[X_H(R_-) - X_H(R_+)\right] + p_E(R)\left[X_E(R_-) - X_E(R_+)\right] \\
&+ p_C(R)\left[X_C(R_+) - X_C(R_-)\right], \qquad (29)
\end{aligned}$$

where $Y(R_-) = w(R)\tau_w(R_-)$. The optimal age of retirement R requires the benefits of working, consisting of labor income $Y(R_-)$, to equal the benefits of retirement, consisting of the sum of the pension benefit and the cost of delayed retirement at age R (first term on the RHS), the monetary value of additional utility from more leisure time (second term on RHS) and the value of higher skill- and health-capital investments due to the lower cost of time inputs (third and fourth term on RHS). In addition, if time inputs substitute

in leisure time has on the marginal utility of consumption, while the effect of an increase in leisure due to the greater availability of time during schooling has a first-order (direct) effect on utility (and is thus relatively large).

for goods and services there may be an additional benefit of retirement in the form of reduced expenditures on goods and services $X_E(t)$ and $X_H(t)$ (terms five and six on the RHS). The last term reflects changes in consumption as a result of changes in the marginal utility of consumption due to increased leisure time while retired.

Intuitively, if utility $U(t)$, consumption, and investments in skill and health capital were continuous in R, and the state pension annuity $b(R) = b$ were independent of the age of retirement, the decision to retire would simply be determined by the age R at which earnings in retirement b_R exceeded, for the first time, earnings from work $Y(t)$. Generous retirement (e.g., social security) benefits b_R and low labor income $Y(t)$ (e.g., due to worsening health and declining skill capital with age) encourage early retirement.

In practice, the pension benefit $b(R)$ is a function of the age of retirement R, typically at least initially increasing in R. It is then attractive for individuals to delay retirement in order to receive higher benefits $b(R)$ per period. However, this comes at the cost of a shortened horizon $T - R$ over which these benefits are received, as reflected in the term $(\partial b(R)/\partial R)\,(1/r)\left[1 - e^{-r(T-R)}\right]$. Further, individuals value the utility from additional leisure in retirement, they value the additional investment in skill capital and health capital due to the lower opportunity cost of time, and there are potential reductions in expenditures on consumption and skill and health-capital investment goods and services. As a result, individuals do not need to be compensated dollar for dollar in income, and retirement occurs while pension benefits are less than labor income.[25]

4 Model predictions

In this section we summarize results, analyze the dynamics of the model, and make predictions. In section 4.1 we discuss life-cycle trajectories, and in section 4.2 we present comparative dynamic analyses to explore cross-sectional heterogeneity in these profiles.

4.1 Lifecycle trajectories

The characteristics of the solutions are visually represented in Figure 1, where S is years of schooling, R is the age of retirement, and T denotes total lifetime.[26]

The top panel presents the life-cycle profile of skill investment $I_E(t)$ (left), and skill $E(t)$ (right). Since skill determines wages $w(t, E)$, they show a similar pattern (for this reason wages are not separately shown). The center panel presents the life-cycle profile of health investment $I_H(t)$ (left) and health $H(t)$ (right). The bottom-left panel presents several time uses: leisure $L(t)$ (thick dotted line), sick time $s[H(t)]$ (small dotted line), time devoted to skill investment $\tau_E(t)$ (dash dotted line), time devoted to health investment $\tau_H(t)$ (dashed line), and time devoted to work $\tau_w(t)$ (solid line). The dotted horizontal

[25] Net pension replacement rates for OECD countries are on average 72 percent and range between 41 percent (Japan) and 112 percent (Iceland) for the median male earner (OECD, 2011).

[26] Note that these are based on analytical reasoning, not on a numerical simulation.

line at the top of the graph represents Ω, the total time available per unit of time (let's say, a day). Last, the bottom-right panel presents earnings $Y(t)$.

The various benefits of skill investment (in the production of earnings, skill and health) are high early in life as the horizon over which benefits can be accrued is long, and due to diminishing returns to skill (see 18, 20, and 21) as individuals have low levels of skill early in life. During schooling, the use of time inputs in investment in skill is encouraged, as the opportunity cost of time is low when one is not allowed to work (i.e., mimimum schooling ages reduce the cost of time). Further, individuals potentially receive a transfer $b_S(t)$ and/or schooling is subsidized, i.e. small $p_E(t)$. This further encourages investment. Thus skill investment is high early in life (top-left panel), in particular during the schooling phase, and skill increases rapidly during the schooling phase (top-right panel).

As skill increases, the benefit of work (earnings) increases, individuals leave school and start working. Less time will be devoted to skill investment during working life because of the higher opportunity cost of time (hence the drop in the level of investment $I_E(t)$ at $t = S$, top-left panel), and the rate at which skill is produced slows (hence the downward change in the slope for $E(t)$ at $t = S$, top-right panel).[27] Skill $E(t)$ (top-right panel) may eventually decline, as biological deterioration outweighs declining skill-capital investment (see 2).

After retirement, time spent working is zero. The greater availability of time encourages individuals to invest more time in skill, hence the jump upward in skill investment $I_E(t)$[28] (top-left) and the upward change in the slope of the skill $E(t)$ profile (top-right) at $t = R$.[29] While skill is no longer useful in the production of earnings $\partial Y/\partial E = 0$, during retirement it still delivers an important health-production benefit $\partial f_H/\partial E > 0$. This results in the following prediction.

[27] Investment in skill $I_E(t)$ is shown, for illustrative purposes, to jump down and to decrease more rapidly during working life. The conditions for which the more rapid decrease holds are discussed in Appendix section A.7. The jump is due to the increased opportunity cost of time. One might argue that on-the-job training is not associated with an opportunity cost of time as training simply happens on the job, so that skill investment does not necessarily exhibits a discontinuous jump upon leaving school. However, as Becker (1964, Chapter 3) argues, the firm would not be willing to pay for perfectly general training (as it benefits are also useful to other firms) while individuals would be willing to pay for it (as it raises their earnings). It is thus not firms but individuals that pay for general training by accepting lower wages. The opposite holds for perfectly specific training (the benefits of which are useful only to the specific firm that employs the worker but the worker cannot use it elsewhere). Thus effectively there is an opportunity cost of time for on-the-job training and its size depends on the extent to which the training is general or firm specific.

[28] This result is somewhat counterintuitive since after retirement there is no longer a production benefit of skill, $\partial Y/\partial E$, and so one might be inclined to view skill as less valuable after retirement. The co-state variables $q_E(t)$ and $q_A(t)$, however, are continuous at $t = R$, so there is no discontinuity in the relative marginal value of skill $q_{e/a}(t)$ (skill remains equally valuable). Given the equilibrium condition $q_{e/a}(R) = \pi_E(R)$ (see 18), and since the cost of time is reduced, there is however a discontinuous increase in investment $I_E(t)$ at $t = R$ (see 21).

[29] The conditions for which this holds are discussed in Appendix section A.7.

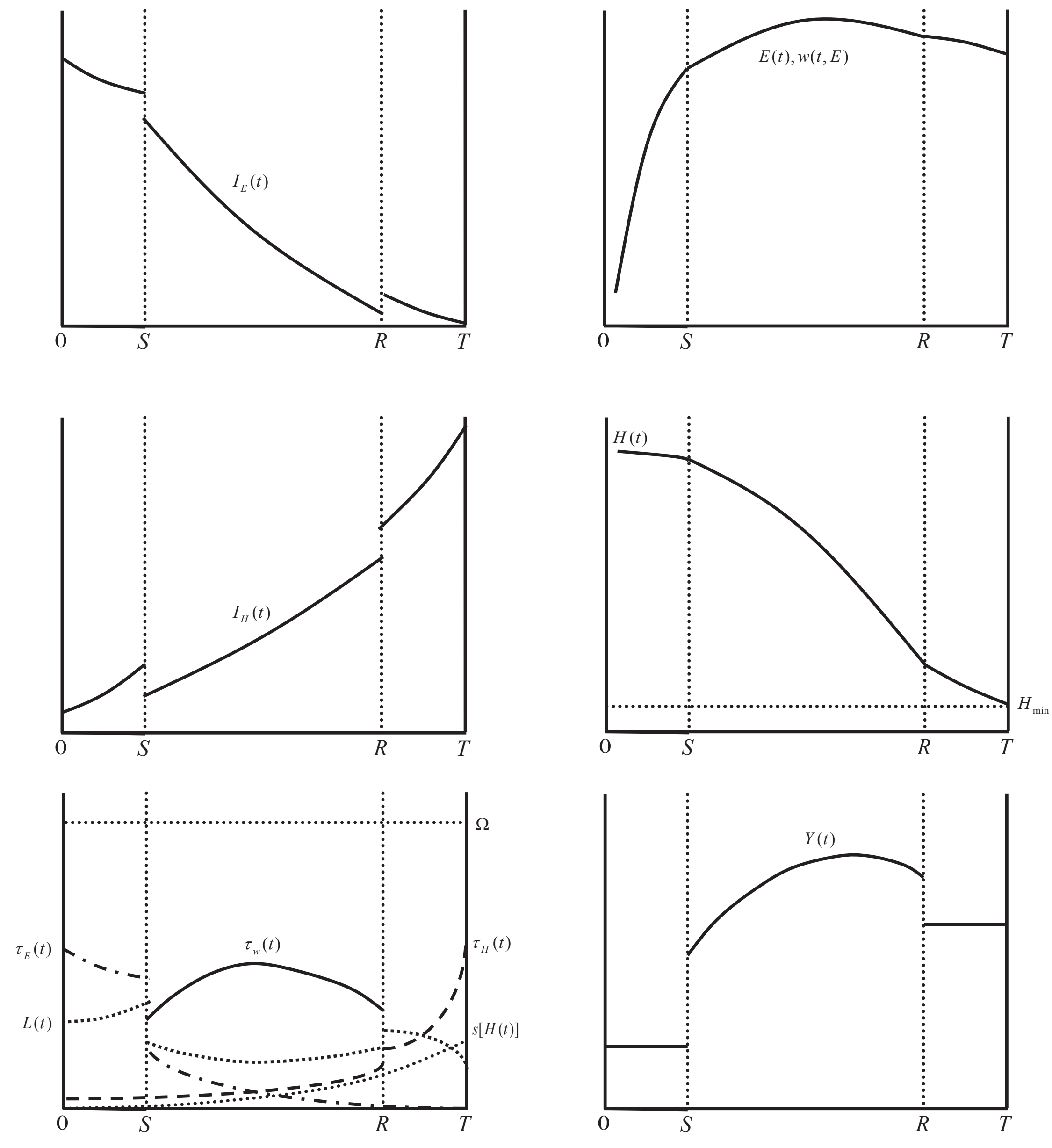

Figure 1: *Illustration of the time paths for skill investment $I_E(t)$ (top left), skill capital $E(t)$ and the wage rate $w[t, E(t)]$ (top right), health investment $I_H(t)$ (center left), health $H(t)$ (center right), time use (bottom left), and earnings $Y(t)$ (bottom right).*

Prediction 1: Individuals will continue to invest in skill capital after retirement. Even though skill no longer contributes to earnings, it still provides important health benefits. Note that this prediction is in sharp contrast to the skill-capital literature (e.g., Becker, 1964; Ben-Porath, 1967), which predicts no skill investment after retirement.[30]

Eventually, investment in skill capital declines to zero (top-left panel), since individuals place no marginal value on the terminal stock of skill $q_E(T) = 0$ (see 19 and the discussion in section 3.3).

At early stages in the life cycle, the individual is endowed with a large stock of health $H(t)$. As a result, the benefit of health, and therefore its relative marginal value, is relatively low due to diminishing returns, and it may be optimal to devote resources to skill capital instead (low health investment $I_H(t)$, center-left panel). As the individual ages, the stock of health declines monotonically (center-right panel).[31] Health investments then become essential to counteract biological aging and to extend life. With declining health, the relative marginal value of health increases. As a result, health investment, in contrast to skill investment but in line with empirical evidence (Zweifel, Felder, and Meiers, 1999; De Nardi, French and Jones, 2010), increases over the life cycle (center-right panel).[32] Retirement further encourages health investment due to the reduced cost of time inputs. Combining this with the earlier discussion for skill, we obtain the following prediction.

Prediction 2: The schooling period is primarily used to invest in skill, and the retirement period is primarily used to invest in health.

Like skill capital, health capital does not contribute to earnings during retirement and therefore the production benefit $\partial Y/\partial H$ is zero. However, health still provides an important home-production benefit as better health reduces sick time, time that can be devoted to leisure $L(t)$, investment in skill $\tau_E(t)$, and investment in health $\tau_H(t)$. Unlike skill capital, health capital also provides direct utility in retirement, providing additional incentives to invest in health after retirement.[33] The health stock eventually reaches a

[30] Some elderly enroll in education programs during retirement. Perhaps skill provides, besides the health benefit, additional benefits, such as a home-production benefit (e.g., cognition enables individuals to remain independent) or a consumption benefit.

[31] Similar to skill, the rate of health decline changes at S and at R due to changes in the opportunity cost of time, and associated changes in the level of health investment.

[32] In part, this is because the terminal level of health is constrained to $H_{\min}$. As a result, the relative marginal value of health at the end of life $q_{h/a}(T)$ does not have to be zero. In contrast, the end condition for skill, $E(T)$ free, does not allow for solutions where skill investment keeps growing till the end of life, since it implies $q_E(T) = 0$ and hence $I_E(T) = 0$ (see the discussion in section 3.3). Thus, a crucial difference between skill and health is the notion that individuals end life in universally poor health but with varying levels of skill.

[33] Analogous to skill, the center-left panel shows a slowing of the rate of growth in health investment during working life. The conditions for which this holds are discussed in Appendix section A.7.

minimum level $H_{\min}$ at the end of life T (indicated by the dotted horizontal line).

Following the discussion above, and illustrated in the bottom-left panel, time inputs into skill-capital investment $\tau_E(t)$ (dash-dotted line) and goods purchased $X_E(t)$ (not shown) are high during the schooling period, then decrease in line with decreasing investment, eventually reaching zero at T. In contrast, health investment is characterized by growing time inputs $\tau_H(t)$ (dashed line) and goods purchased $X_H(t)$ (not shown), and extra time can be devoted after individuals retire. Sick time $s[H(t)]$ (small dotted line) increases with declining health $H(t)$. Leisure time $L(t)$ (large dotted line) decreases and subsequently increases somewhat during working life, reflecting the high opportunity cost $w[t, E(t)]$ (top-right panel) of not working during the prime working ages. Upon retiring, leisure time is higher, yet as a result of competition from increasing sick time $s[H(t)]$ and increasing time devoted to health $\tau_H(t)$, leisure could decline with age. The remaining time $\tau_w(t)$ (thick line) is devoted to work, and first increases with age as accumulated skill makes it attractive to work, but later on it declines as increased sick time and time devoted to health investment prevent the individual from working, making retirement increasingly attractive.

The product of time devoted to work $\tau_w(t)$ and the wage rate $w[t, E(t)]$ (top right) provides the earnings profile $Y(t)$ (bottom-right panel), where income may be provided by the state during schooling and individuals receive a pension during retirement. Since earnings are the product of the wage rate and time spent working, earnings will decrease more rapidly than wages, as a result of declining health. This leads to the following prediction.

Prediction 3: The observed hump-shaped earnings profile is partially due to reduced time spent working as a result of declining health.[34]

4.2 Cross-sectional variation in the life-cycle trajectories

The life-cycle trajectories discussed in section 4.1 can be viewed as representing the average individual in a representative sample. We are also interested in understanding cross-sectional heterogeneity in these profiles. Our theory describes the entire lifecycle, and is highly dynamic, limiting the use of comparative static analyses. To gain further insight into the characteristics of the theory, we have to resort to comparative dynamic analyses, which allow analyzing variation in the lifecycle profiles with respect to the three types of resources an individual possesses, financial capital (wealth), skill capital, and health capital, as well as other model parameters of interest.

Following Ben-Porath (1967) and Heckman (1976) we can make some convenient assumptions to arrive at a tractable version of our general theory that permits derivation of analytical expressions for the comparative dynamic results. The simpler model retains

[34] In practice, the wage rate is potentially a function of health as well. Still the prediction holds since earnings capture the effect of reductions in labor supply as well as in wages.

the essential characteristics of the general theory. There are some costs associated with the simplifications, which we discuss in detail in Appendix section A.4, but the benefits arguably outweigh the costs. Most importantly, the assumptions enable obtaining analytical expressions for the comparative dynamic analyses. We find that the predictions of the simpler model also hold for the general model with some nuanced differences (which are discussed in detail in Appendix A.6). Since our approach does not solely rely on the simpler model we obtain robust comparative dynamic results.

We start by introducing the simplified theory.

4.2.1 A simpler tractable model

Individuals maximize a constant relative risk aversion (CRRA) lifetime utility function

$$U(t) = \frac{1}{1-\rho}\left(X_C(t)^{\zeta}\left\{L(t)[E(t)+H(t)]\right\}^{1-\zeta}\right)^{1-\rho}, \tag{30}$$

with ζ the "share" of consumption and $1-\zeta$ the "share" of leisure in utility, and $1/\rho$ the elasticity of substitution. Consumption $X_C(t)$ and "effective" leisure time $L(t)[E(t)+H(t)]$ are complements in utility if $\rho < 1$ and substitutes in utility for $\rho > 1$. Leisure time $L(t)$ is multiplied by $E(t)+H(t)$, reflecting the notion that human capital (consisting of the sum of skill and health capital, $E(t)+H(t)$) augments the agent's consumption time (Heckman, 1976). The utility function is maximized subject to the same dynamic constraints for skill capital (2), for health capital (3), and for assets (4), as in the general framework.

We assume no sick time $s[H(t)]$, that earnings consist of the product of human capital, $E(t)+H(t)$, and the fraction of time available for work

$$Y[E(t),H(t)] \quad = \quad [E(t)+H(t)]\left[1-\tau_E(t)-\tau_H(t)-L(t)\right], \tag{31}$$

and, last, that the production functions of skill capital and of health capital are of a Cobb-Douglas form,

$$\begin{aligned} f_E[\tau_E(t),X_E(t),E(t),H(t)] &= \theta_E(t)\left\{\tau_E(t)\left[E(t)+H(t)\right]\right\}^{\alpha_E} X_E^{\beta_E}, \\ &= \mu_E(t)q_{e/a}(t)^{\frac{\gamma_E}{1-\gamma_E}}, \end{aligned} \tag{32}$$

$$\begin{aligned} f_H[\tau_H(t),X_H(t),E(t),H(t)] &= \theta_H(t)\left\{\tau_H(t)\left[E(t)+H(t)\right]\right\}^{\alpha_H} X_H^{\beta_H}, \\ &= \mu_H(t)q_{h/a}(t)^{\frac{\gamma_H}{1-\gamma_H}}, \end{aligned} \tag{33}$$

where $\theta_E(t)$ and $\theta_H(t)$ denote the technologies of production of skill investments and health investments, respectively, $\gamma_E = \alpha_E + \beta_E < 1$, and $\gamma_H = \alpha_H + \beta_H < 1$ (diminishing returns to scale).[35] The functions $\mu_E(t)$ and $\mu_H(t)$ are generalized productivity factors

$$\mu_E(t) \equiv \left[\frac{\alpha_E^{\alpha_E}\beta_E^{\beta_E}\theta_E(t)}{p_E(t)^{\beta_E}}\right]^{\frac{1}{1-\gamma_E}}, \tag{34}$$

$$\mu_H(t) \equiv \left[\frac{\alpha_H^{\alpha_H}\beta_H^{\beta_H}\theta_H(t)}{p_H(t)^{\beta_H}}\right]^{\frac{1}{1-\gamma_H}}. \tag{35}$$

The technologies of production $\theta_E(t)$, $\theta_H(t)$, and the generalized productivity factors $\mu_E(t)$, $\mu_H(t)$, can be considered as being determined by technology as well as biology.

The begin and end conditions H_0, $H(T) = H_{\min}$, E_0, A_0, $A(T) = A_T$, and the transversality conditions $q_E(T) = 0$, and $\Im(T) = 0$, also apply here. The analytical solutions of the simpler model are presented in Appendix A.2.

4.2.2 Comparative dynamics

Comparative dynamic analyses allow us to analyze differences in behavior as a function of model parameters. We start with an analysis of endowed wealth, health, and skill.

Consider a generic control, state, or co-state function $g(t)$, and a generic variation δZ_0 in an initial condition or model parameter. The effect of the variation δZ_0 on the optimal path of $g(t)$ can be broken down into variation for fixed longevity T and variation due to the resulting change in the horizon T

$$\frac{\partial g(t)}{\partial Z_0} = \left.\frac{\partial g(t)}{\partial Z_0}\right|_T + \left.\frac{\partial g(t)}{\partial T}\right|_{Z_0}\frac{\partial T}{\partial Z_0}. \tag{36}$$

The comparative dynamic effects of a small perturbation in initial wealth δA_0, initial skill δE_0, and initial health δH_0 are summarized in Table 1.[36] Detailed derivations are provided in Appendix section A.5.[37]

We distinguish between two cases, one in which length of life is fixed (exogenous), and one in which length of life can be freely chosen (endogenous).

Prediction 4: Absent ability to extend life T, associations between wealth, skill and health are absent or small.

[35] Proof that the skill f_E and health f_H production functions can be expressed in terms of the relative marginal value of skill $q_{e/a}(t)$, and of health $q_{h/a}(t)$, is provided in Appendix A.2 (see equations 54 to 57).

[36] Note that we can restart the problem at any time t, taking $A(t)$, $E(t)$, and $H(t)$, as the new initial conditions. Thus the comparative dynamic results derived for variation in initial wealth δA_0, initial skill δE_0, and initial health δH_0, have greater validity, applying to variation in wealth $\delta A(t)$, skill $\delta E(t)$, and health $\delta H(t)$, at any time $t \in [0, T)$.

[37] See equations (84), (85), and (86) for initial wealth A_0, equations (87) to (91) for initial skill E_0, and equations (93) to (97) for initial health H_0.

Table 1: Comparative dynamic effects of initial wealth A_0, initial skill E_0, and initial health H_0, on the state and co-state functions, control functions and the parameter T.

	δA_0		δE_0		δH_0	
Function	T fixed	T free	T fixed	T free	T fixed	T free
$E(t)$	0	> 0	> 0	> 0	0	> 0
$q_{e/a}(t)$	0	> 0	0	> 0	0	> 0
$X_E(t)$	0	> 0	0	> 0	0	> 0
$\tau_E(t)\,[E(t) + H(t)]$	0	> 0	0	> 0	0	> 0
$H(t)$	0	> 0	0	> 0	≥ 0	> 0
$q_{h/a}(t)$	0	> 0	0	> 0	< 0	+/-
$X_H(t)$	0	> 0	0	> 0	< 0	+/-
$\tau_H(t)\,[E(t) + H(t)]$	0	> 0	0	> 0	< 0	+/-
$A(t)$	≥ 0	+/-	+/-	+/-	+/-	+/-
$q_A(0)$	< 0	$< 0^\dagger$	< 0	$< 0^\dagger$	< 0	$< 0^\dagger$
$X_C(t)$	> 0	$> 0^\dagger$	> 0	$> 0^\dagger$	> 0	$> 0^\dagger$
$L(t)\,[E(t) + H(t)]$	> 0	$> 0^\dagger$	> 0	$> 0^\dagger$	> 0	$> 0^\dagger$
T	n/a	> 0	n/a	> 0	n/a	> 0

Notes: 0 is used to denote 'not affected', $+/-$ is used to denote that the sign is 'undetermined', n/a stands for 'not applicable', and † is used to denote that 'the sign holds under the plausible assumption that the wealth effect dominates the effect of life extension'. This is consistent with the empirical finding (Imbens, Rubin and Sacerdote, 2001; Juster et al. 2006; Brown, Coile, and Weisbenner, 2010) that additional wealth leads to higher consumption, even though the horizon over which consumption takes place is extended (see section A.5 for further detail).

Prediction 4 highlights a key feature of the Ben-Porath model: absent ability to increase the horizon over which benefits can be accrued (fixed length of life T), additional wealth does not lead to more skill investment and health investment, leaving skill and health unchanged (rows 1 to 8 for T fixed). The additional wealth is primarily used to finance additional consumption and leisure (rows 11 and 12). Both skill capital and health capital are forms of wealth, in the sense that they increase wages and therefore lifetime wealth (reducing the marginal value of initial wealth $q_A(0)$). Thus a positive variation in skill δE_0 and in health δH_0 operates in a manner similar to a positive variation in wealth δA_0 (see columns 3 and 5 in the Table, 88 and 93), with some differences: endowed skill E_0 leads to greater skill, endowed health H_0 leads to greater health, and endowed health H_0 reduces the relative marginal value of health $q_{h/a}(t)$ and thereby the demand for health investment. Thus also for additional skill and additional health there are no additional investments made, and for additional health, health investments are even reduced.

This lack of an association between skill and wealth (and in our case also health) in the Ben-Porath model has been noted before (Heckman, 1976; Graham, 1981). While Graham (1981) suggests it is due to the fact that in the Ben-Porath model individuals maximize lifetime earnings, and not utility, we find it also holds for a model in which individuals maximize lifetime utility. It is instead the result of the assumptions of "Ben-Porath neutrality" (see Appendix section A.4) and fixed length of life. These assumptions ensure that the relative marginal value of skill $q_{e/a}(t)$ is not a function of wealth, skill, and health (see 60). Thus additional wealth, skill or health does not lead to greater levels of investment in skill.

These results, however, have greater validity, as length of life is a crucial determinant of the return to investments. For fixed length of life, in both the general and simpler model, any additional health investment needs to be compensated by eventual lower investment in order for health to reach $H_{\min}$ at $t = T$. The response to additional resources will therefore be muted (see Appendix A.6 for detail). As a result, there are no strong associations between wealth, skill, and health for fixed T, and $\partial g(t)/\partial Z_0|_T$, the first term on the RHS of (36), is generally small for variation δZ_0 in any model parameter of interest.

Now consider the case where T is free.

Prediction 5: Wealthy, skilled, and healthy individuals live longer.

The bottom row of Table 1 shows that positive variations in endowments, in the form of wealth, skill or health, lead to a longer life span. For variation in initial wealth A_0, the intuition is as follows. At high values of wealth (and hence consumption and leisure), individuals prefer investing in health to consuming and leisure. While additional consumption or leisure per period would yield only limited utility due to diminishing utility of consumption and leisure, investments in health extend life, increasing the period over which they can enjoy the utility benefits of leisure and consumption (see also Becker, 2007; Hall and Jones, 2007). With sufficient wealth one starts caring more about other goods, in particular health. Hence, wealth increases health investment and thereby health, $\partial H(t)/\partial A_0 > 0 \quad \forall t$, and extends life $\partial T/\partial A_0 > 0$. Endowments in skill and health are also forms of wealth, in the sense that skill and health increase earnings and therefore lifetime wealth. Hence, a similar reasoning can be applied to variations in initial skill and initial health. Prediction 5 also holds for the general model (see Appendix A.6 for detail).

Prediction 6: Wealthy and healthy individuals value skill more, invest more in skill, and are more skilled at every age. Individuals with more endowed skill are more skilled at every age, but potentially value skill less.

The first four rows of Table 1 show that positive variations in the form of endowed wealth, skill, or health, lead to a higher marginal value of skill $q_{e/a}(t)$, higher levels of investment inputs $X_E(t)$, $\tau_E(t)[E(t) + H(t)]$, and greater skill $E(t)$ (T free). A longer horizon (prediction 5) increases the return to investment in skill, such that wealthy individuals value skill more. As a result they invest more in skill and are more skilled at every age $\partial E(t)/\partial A_0 > 0 \quad \forall t$. Endowments in skill and health are also forms of wealth, so that similar reasoning can be applied here. These results also hold for the general model, with the exception that in the general model skilled individuals could potentially value skill less $\partial q_{e/a}(t)/\partial E_0 < 0$, invest less in skill, but still be more skilled at every age (see Appendix A.6 for detail).

Table 2 in Appendix A.5 presents the comparative dynamics for variation in the generalized productivity of skill $\mu_E(t)$ – a composite measure that is increasing in the productivity of skill $\theta_E(t)$, and decreasing in the price of skill $p_E(t)$ (see 34). The signs of the effect of variation in the generalized productivity of skill $\delta\mu_E(t)$ are the same as those for the effect of variation in skill, δE_0.[38] The result also holds for the general model (see Appendix A.6 for detail).

Prediction 7: Wealthy and skilled individuals value health more, invest more in health, and are healthier at every age. Individuals with more endowed health are healthier at every age, but potentially value health less.

The fifth to eight rows of Table 1 show that positive variation in the form of greater endowed wealth and skill, lead to a higher marginal value of health $q_{h/a}(t)$, higher levels of investment inputs $X_H(t)$, $\tau_H(t)[E(t) + H(t)]$, and greater health $H(t)$ (T free). Yet, while endowed health does lead to greater health at every age $\partial H(t)/\partial H_0 > 0 \quad \forall t$, it does not unambiguously lead to a higher marginal value of health $q_{h/a}(t)$, and thereby higher levels of investment $X_H(t)$, $\tau_H(t)[E(t)+H(t)]$. The easiest way to understand this is for fixed length of life. In contrast to skill capital, where the terminal value $E(T)$ is unconstrained, the terminal value of health is $H_{\min}$. If the horizon T is fixed, additional health $H_0 + \delta H_0$ needs to be offset by lower health investment throughout life in order to reach $H_{\min}$ at $t = T$. For free length of life T, if the increase in length of life $\partial T/\partial H_0$ is sufficiently large, health investment is higher throughout life. If it is small, health investment is lower throughout life. These results also hold for the general model (see Appendix A.6 for detail).

Table 2 in Appendix A.5 presents the comparative dynamics for variation in the generalized productivity of health $\mu_H(t)$ – a composite measure that is increasing in the productivity of health $\theta_H(t)$, and decreasing in the price of health $p_H(t)$ (see 35). The signs of the effect of variation in the generalized productivity of health $\delta\mu_H(t)$ are identical to the effect of variation in health, δH_0.[39] The result also holds for the general model (see Appendix A.6 for detail).

Prediction 8: Gains in life expectancy reinforce the associations between wealth, skill, health, and technology.

Gains in life expectancy play a powerful role in generating the associations between

[38]This is quite intuitive, since a higher productivity of skill investment (higher $\theta_E(t)$) or a lower price of skill investment $p_E(t)$ encourage skill investment, leading to a higher stock of skill. An important assumption in obtaining this result is that permanent effects dominate temporary effects.

[39]An important assumption in obtaining this result is that permanent effects dominate temporary effects.

wealth, skill, health, and technology. Equation (36) illustrates this. From prediction 4 we obtain that the first term on the RHS $\partial g(t)/\partial Z_0|_T$ is generally small for variation δZ_0 in any model parameter of interest. From predictions 5, 6 and 7, we have $\partial T/\partial Z_0 > 0$ for $Z_0 = \{A_0, E_0, H_0, \mu_E(t), \mu_H(t)\}$. Finally, for the simpler model we have $\partial g(t)/\partial T|_{Z_0} > 0$ for $g(t) = \{E(t), H(t)\}$ (see 83).[40]

Thus, the size of the effect of Z_0 on $g(t)$ increases with the degree of life extension $\partial T/\partial Z_0$. In other words, if resources, biology, medical technology, institutional, environmental and/or other factors, do not allow for life extension ($\partial T/\partial Z_0$ small), then the effect more closely resembles that of the fixed T case. As in the fixed T case, there would be a small association between wealth, skill, health, and technology (see prediction 4; small $\partial g(t)/\partial Z_0|_T$). In contrast, if additional resources afford considerable life extension ($\partial T/\partial Z_0$ large), the horizon over which the benefits of skill investment and health investments can be reaped is larger. Further, utility from leisure and consumption can be enjoyed with additional years of life. Together, these various benefits of life extension substantially raise investment in skill and in health, thereby improving skill and health. The prediction also holds for the general model, with some differences (see Appendix A.6 for detail).

Prediction 9: Variations in two model parameters, both of which have a positive effect on longevity, generate complementarity effects and reinforce the associations between wealth, skill, health, and technology.

The analytical comparative dynamic expressions of the simpler model can be employed to not only study the sign of the comparative dynamic effects, but also to study complementarities between different model parameters. Is, for example, the effect of endowed skill on health formation $\partial H(t)/\partial E_0$ greater or smaller for the wealthy? Exploring this question requires combining equations (83), (90), and (91) in Appendix A.5. There are many such relationships, given the many permutations possible. It is therefore impossible to discuss all of them, and their expressions can become quite involved. But the interested reader can use Appendix A.5 to delve further into relationships of interest.

The general lesson from this type of analysis is that variations in two parameters, Z_0 and W_0, both of which have a positive effect on longevity, often reinforce each other (complementarity), i.e. that the total effect on a model outcome $g(t)$ is greater than the sum of the individual effects (see also Fonseca et al. 2013). Equation (36) shows (noting that $\partial g(t)/\partial Z_0|_T$ is small; prediction 4) that this could be either due to complementarities between Z_0 and W_0 in their effect on life expectancy (if $\partial T/\partial Z_0$ is increasing in W_0), or it could be due to complementarities between Z_0 and W_0 in the effect of life expectancy on the model outcome $g(t)$ (if $\partial g(t)/\partial T|_{Z_0}$ is increasing in W_0). Longevity is thus essential in generating complementarity effects. See Appendix A.6 for more detail (though in this

[40] The sign for $\partial A(t)/\partial T|_{Z_0}$ is ambiguous because the additional resources (endowments in wealth, skill, or health, or technological improvement) have to be spread over a longer horizon, but the longer horizon at the same time encourages greater investment in skill and in health, which in turn accumulates wealth.

case no full proof could be established for the general model).

As a concrete example, consider the effect of life expectancy T on skill capital $E(t)$ (83 in Appendix A.5), an effect that has attracted much attention in both the theoretical and empirical literatures (e.g., Ben-Porath, 1967; Hazan, 2009; Jayachandran and Lleras-Muney, 2009; Fortson, 2011; Oster, Shoulson, and Dorsey, 2013). Expression (83) shows that the effect of life expectancy on skill capital is reinforced by higher productivity of skill-capital investment $\mu_E(t)$. This productivity factor increases in the technology of skill investment production $\theta_E(t)$ and decreases in the price of skill-capital investment $p_E(t)$ (see 34). This leads to the following prediction.

Prediction 9b (example): Life expectancy and skill-capital productivity reinforce each other in generating skill.

If skill-capital investment is relatively unproductive (e.g., low quality teachers, children infected with worms, or malaria), or if the cost of skill-capital investment is high (e.g., high tuition, long distance to schools, crops that need to be collected), then the effect of gains in life expectancy on skill-capital formation is predicted to be modest. By contrast, when skill investment is productive, and affordable, the effect of life expectancy on skill capital is strong. This suggests there could be important heterogeneity in the effect of longevity gains on skill-capital formation; an insight that is particularly important since the variation that is used to identify the effect of life expectancy gains on skill capital typically derives from developing countries, with potentially low productivity of skill-capital investment (e.g., Jayachandran and Lleras-Muney, 2009; Cutler et al. 2010; Fortson, 2011).

Conjecture 3a: Wealthy individuals have shorter working lives, staying in school longer and retiring earlier.

Conjecture 3b: The relationship between skill and years of employment follows an inverse U-shape, with low- and high-skilled individuals having shorter working lives, and medium-skilled individuals longer working lives.

Conjecture 3c: Healthy individuals stay in school longer and retire later.

Predictions 4 to 7 allow for comparative dynamic analyses of the effects of wealth, skill, and health, on the optimal school-leaving age S, and the optimal retirement age R, using the transversality conditions for schooling (equation 28) and for retirement (equation 29) of the general model. A detailed discussion is provided in Appendix section A.7; here we summarize results.[41]

[41] An important assumption we make in deriving our results is that factors affecting skill formation are most important to the schooling decision and factors affecting health formation are most important to the retirement decision. Early in life the marginal value of skill is high and investment in skill is high while

Individuals with large endowed wealth, ceteris paribus, plausibly have the shortest working lives (conjecture 3a): they leave school later, and retire earlier. This is because they can afford schooling and early retirement (a wealth effect), because they value the increased time during schooling for skill investment (through a higher value of skill, prediction 6), because they value the increased time during retirement for health investment (through a higher value of health, prediction 7), and because wealth, through enhanced skill and health, raises the benefit of skill production and of health production. The only effect that encourages later retirement is that wealthy individuals live longer (prediction 5). The assumption therefore is that this longevity effect is small relative to the forces encouraging early retirement.[42]

While there are many competing effects for skill, a pattern is discernible. Casual observation suggests the skilled, not the unskilled, value continued schooling, suggesting that self-productivity and dynamic complementarity (raising skill production f_E), and life extension enabled by endowed skill (raising the marginal value of skill), are potentially important mechanisms in encouraging continued schooling. As a result, the skilled leave school later. Additional schooling increases wealth and the skilled value the time available for health investment during retirement (through a higher value of health, prediction 6), encouraging early retirement. However, the skilled also live longer, and their pre-retirement wages are higher, providing incentives to postpone retirement. We then expect an inverse U-shape in the relationship between education and retirement in countries, such as, e.g., the United States and those of Europe, where a minimum level of income is guaranteed in old age (e.g., social security). Lower educated individuals retire early as social security makes retirement affordable, they have low pre-retirement wages (low benefit of continued work), live shorter lives, and attach a high value to the increased time for health investment during retirement since they are generally unhealthier.[43] The highest educated accumulate enough wealth and can afford to retire early. The mid-level educated group, in contrast, is relatively healthy, earns reasonable wages, but does not accumulate sufficient wealth to afford early retirement. Therefore, they may end up retiring last and have longer working lives (conjecture 3b).

Empirical evidence (see section 2) suggests that healthy individuals stay longer in school, suggesting that higher permanent income (wealth effect), a lower value of health (prediction 7), and the benefits of greater skill (due to greater health, prediction 6) in enhancing skill and health production outweigh the benefit of enhanced earnings in joining

---

the marginal value of health is low and investment in health is low, so that factors affecting skill formation are plausibly more important than factors regarding health formation in affecting the schooling decision (the opposite holds for retirement).

[42] As an extreme example, the children of the very rich, those for which income from inherited wealth $rA(t)$ substantially outweighs income from labor $Y(t)$, may never work. For them the only benefit of schooling is the enhanced productivity of health from skill. They will not go to school and never work, enjoying a life of leisure (effectively, very early retirement). In practice, the very wealthy may demand skill to manage their wealth (such a benefit of schooling is not included in our theory).

[43] In developing nations, in absence of social security schemes or family provision of income during retirement, poor individuals may not be able to retire.

the labor force. The effect on the retirement decision is harder to predict. Healthy individuals are more capable of accumulating wealth, encouraging early retirement, but also live longer and value the additional time in retirement for health investment less, both encouraging late retirement. It is well established that adverse health shocks lead to labor force withdrawal (e.g., Currie and Madrian, 1999; Garcia-Gomez et al. 2013), suggesting that a health shock reduces the monetary benefits of employment (through increased sick time and higher demand for time devoted to health), and increases the value of retirement as a time to recover from the shock. In sum, the theory, informed by the above empirical stylized facts, suggests that healthy individuals leave school later, but also retire later (conjecture 3c).

5 Discussion

This paper presents a theory of joint investment in skill capital, health capital, and longevity, with three distinct phases of life: schooling, work, and retirement. Investments in health capital consist of, e.g., medical expenditures and physical exercise, while investments in skill capital consist of, e.g., expenditures on education and (on-the-job) training. The theory brings together (or unifies) the skill- and health-capital literatures, encompassing canonical skill-capital theories such as those developed by Becker (1964) and Ben-Porath (1967), and canonical health-capital theories such as those developed by Grossman (1972a;1972b; 2000) and Ehrlich and Chuma (1990).

The contribution of this paper is twofold. First, by unifiying the skill- and health-capital literatures our theory provides new insight into the distinct characteristics of skill and health that have heretofore not been uncovered. Second, the theory provides a framework for explaining stylized facts and for deriving new predictions that can be explored in future research.

Human capital is multidimensional (Acemoglu and Autor, 2012), and skill and health are potentially its most important dimensions (Schultz, 1961; Grossman, 2000; Becker, 2007). Both skill and health are human-capital stocks that depreciate over time, and investing in them can (partially) counteract their deterioration. Skill and health share the defining characteristic of human capital that they make individuals more productive. Despite their similarities, there are some notable differences. Grossman (1972a; 1972b; 2000) has argued that health, in contrast to skill, provides a consumption benefit $\partial U/\partial H$ (direct utility) in addition to a production benefit $\partial Y/\partial H$. Ehrlich and Chuma (1990) have emphasized that health is also distinct from skill in that maintaining health extends life (see 14).

This paper suggests three important additional differences between skill and health. First, we argue that skill capital largely determines the rate of return per period (the wage rate), while health capital largely determines the period itself, determining the amount of time that can be devoted to work and other uses, not just within a day (as in Grossman, 1972a; 1972b) but over the entire lifecycle, determining the duration of the schooling, work, and retirement phases of life.

Second, skill is valued early in life, while health is valued later in life (conjecture 2), and, third, skill formation is governed by a production process in which investment in skill increases the level of skill, at least initially, while health formation is governed by a production process where health eventually declines, no matter how much one invests in it (conjecture 1). An implication is that individuals will use the schooling period primarily to invest in skill, while retirement is mostly devoted to health investment and leisure (prediction 2).

We explored the lifecycle trajectories and the comparative dynamics of the model and derived a number of predictions from these analyses. We find, perhaps not surprisingly, that greater endowed wealth, skill, and health, and improvements in technology, lead to higher investment, greater skill, better health, and a longer life (predictions 5, 6, and 7).[44] We also find that individuals will continue to invest in skill after retirement (prediction 1), a result that runs counter to the skill-capital literature, which predicts no investment. In our theory this is because skill provides an important health-production benefit. One could easily imagine other home-production benefits of skill during retirement, or even a consumption benefit, potentially explaining why some elderly participate in educational programs.

There are several implications and possible uses of the theory that we elaborated upon in the introduction and shortly summarize here. First, recognizing that health is an essential component of human capital suggests misspecification in many empirical applications of human-capital theory. Examples include, but are not limited to, the importance of health in development-accounting efforts of economic growth (e.g., Weil, 2007), and the attribution of the hump-shaped earnings profile over the lifecycle to skill-capital decline (prediction 3). Second, in contrast to the skill-capital and health-capital literatures, our model suggests several novel pathways from health to educational attainment and from health to skill that are understudied in empirical as well as theoretical research. Third, our model predicts a central role for longevity. Additional resources (e.g., wealth, skill, health, permanent income) lead to more health and skill investment only if they are accompanied by an increase in longevity (predictions 4 and 8). Finally, the theory predicts heterogeneity in the effect of longevity gains on skill and health, as a result of differences in institutions and environment (prediction 9; 9b). Responses to longevity gains are predicted to be small if the returns to education are small (e.g., in a society where an extractive and exclusive elite controls the nation's wealth; Deaton, 2013). This may explain why some studies have found effects of longevity gains on skill formation (e.g., Bleakley, 2007; Jayachandran and Lleras-Muney, 2009; Bleakley, 2010b; Fortson, 2011), while others have not (e.g., Acemoglu and Johnson, 2007; Cutler et al. 2010).

These are just a few examples of how the theory can be used as an analytical framework

[44] Though healthy individuals may have lower health investment, if better health does not substantially improve longevity (prediction 7).

to study empirical research questions and to generate testable predictions. Empirical tests of the model could either take the form of reduced form analyses investigating specific predictions, or more structural analyses involving calibrated simulations and/or estimation. The theory is rich, and it is impossible to produce an exhaustive list of its possible uses. We hope the theory will aid researchers in studying their own particular questions of interest. For example, the analytical comparative dynamic expressions of the simpler model can be employed to not only study the sign of the comparative dynamic effects but also to provide information on its determinants (see Appendix section A.5), and then following a similar logic to that applied in Appendix A.6, one can assess to what extent results also hold for the general model. The discussion of prediction 9b provides an illustration of the potential of this type of analysis.

Acknowledgements

Research reported in this publication was supported by the National Institute on Aging of the National Institutes of Health under Award Numbers K02AG042452, R01AG037398, and through the USC Roybal Center for Health Policy Simulation, P30AG024968. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Hans van Kippersluis thanks the Netherlands Organization of Scientific Research for financial support (NWO Veni grant 016.145.082). Titus Galama is grateful to the School of Economics of Erasmus University Rotterdam for a Visiting Professorship in the Economics of Human Capital. We thank Marco Angrisani, Hans Bloemen, Michael Caputo, Isaac Ehrlich, Michael Grossman, James Heckman, Jennifer Kohn, Adriana Lleras-Muney, Peter Savelyev, Darjusch Tafreschi, Tom Van Ourti, participants at the Netspar Theme Workshop, Erasmus University, the Netherlands, at ASHEcon Minneapolis 2012 and ASHEcon Los Angeles 2014, at the International Conference on Health, Education and Retirement over the Prolonged Life Cycle, Vienna, Austria, 2013, at the International Health Economics Association (iHEA) meeting, Sydney, Australia, 2013, and seminar participants at the RAND Corporation, the department of Economics at Vanderbilt, and the University of Southern California's (USC) Schaeffer Center, the USC Center for Economic and Social Research (CESR), the USC Department of Economics, and the USC Marshall school of Business, for helpful comments.

References

Acemoglu, D., and D. Autor (2012), "What Does Human Capital Do? A Review of Goldin and Katz's *The Race between Education and Technology*", NBER Working Paper 17820.

Acemoglu, D., and S. Johnson (2007), "Disease and Development: The Effect of Life Expectancy on Economic Growth", *Journal of Political Economy* 115(6): pp. 925-985.

Acemoglu, D., and S. Johnson (2014), "Disease and Development: A Reply to Bloom, Canning, and Fink", NBER Working Paper, 20064.

Albouy, V. and L. Lequien (2009), "Does compulsory education lower mortality?", *Journal of Health Economics*, 28(1): 155-168.

Barro, R.J. (2001), "Human Capital and Growth," In: Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association, *The American Economic Review*, 91(2): pp. 12-17.

Becker, G. (1964), "Human Capital: A theoretical and empirical analysis, with special reference to education", New York: National Bureau of Economic Research.

Becker, G. (2007), "Health as human capital: synthesis and extensions", Oxford Economic Papers: 59, 379410.

Behrman, J.R. and M.R. Rosenzweig (2004), "Returns to Birthweight", *Review of Economics and Statistics*, 86(2): 586-601.

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

Bleakley, H. (2007), "Disease and Development: Evidence from Hookworm Eradication in the American South", *Quarterly Journal of Economics* 122(1): 73-117.

Bleakley, H. (2010a), "Health, Human Capital, and Development", *Annual review of economics*, 2: pp. 283 - 310.

Bleakley, H. (2010b), "Malaria Eradication in the Americas: A Retrospective Analysis of Childhood Exposure", *American Economic Journal: Applied Economics*, 2(2): pp. 1-45.

Bleakley, H., Costa, D. and Lleras-Muney (2013), "Trends in health, education and income in the United States, 1820-2000", *NBER Working Paper Series*, 19162.

Bloom, D.E. and D. Canning (2000), "The health and wealth of nations", *Science*, 287: 1207-1209.

Bloom, D.E., Canning, D. and Sevilla, J. (2004), "The effect of health on economic growth: a production function approach", *World Development* 32(1): 1-13.

Bloom, D.E., and Fink, G. (2013), "The Economic Case for Devoting Public Resources to Health", *Manson's Tropical Diseases: Expert Consult-Online*, 23.

Brown, Jeffrey R., Courtney C. Coile, and Scott J. Weisbenner (2010). "The Effect of Inheritance Receipt on Retirement", *Review of Economics and Statistics*, 92(2): pp. 425-434.

Carbone, J. and Kverndokk (2014), "Individual investments in education and health", University of Oslo health economics research network, working paper 2014:1.

Case A., A. Fertig, and C. Paxson (2005), "The lasting impact of childhood health and circumstance", *Journal of Health Economics*, 24(2): pp. 365-389.

Caputo, M.R. (2005), *Foundations of Dynamic Economic Analysis*, Cambridge University Press.

Casanova, M. (2013), "Revisiting the Hump-Shaped Wage Profile", UCLA Working Paper, August 2013.

Cervellati, Mattero and Sunde, Uwe (2013), "Life expectancy, schooling and lifetime labor supply: theory and evidence revisited", *Econometrica* 81(5), 20552086.

Chiang, Alpha C. (1992), *Elements of dynamic optimization*, McGraw-Hill Book Co, Singapore.

Clark, D. and H. Royer (2014), "The Effect of Education on Adult Mortality and Health: Evidence from the United Kingdom" *American Economic Review*.

Conti, G., J. J. Heckman, and S. Urzua (2010), "The Education-Health Gradient" *American Economic Review* 100: 234-238.

Cunha, F. and J. Heckman (2007), "The Technology of Skill Formation", *American Economic Review*, 97(2): 31-47.

Currie, J., and Madrian, B.C. (1999), "Health, Health Insurance and the Labor

Market", In *Handbook of Labor Economics*, Volume 3, ed. O. Ashenfelter and D. Card, 3309407.Amsterdam, North Holland.

Currie, J. (2009), "Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development", *Journal of Economic Literature*, 47(1): pp. 87-122.

Cutler, D., and Miller, G. (2005). "The role of public health improvements in health advances: the twentieth-century United States", *Demography*, 42(1): pp. 1-22.

Cutler, D.M., A. Deaton, and A. Lleras-Muney (2006), "The determinants of mortality", *Journal of Economic Perspectives*, 20(3): pp. 97-120.

Cutler, D.M. and A. Lleras-Muney (2008), "Education and Health: Evaluating Theories and Evidence", in Robert F. Schoeni et al., eds., Making Americans Healthier: Social and Economic Policy As Health Policy (National Poverty Center Series on Poverty and Public Policy).

Cutler, David M., and Adriana Lleras-Muney (2010). "Understanding differences in health behaviors by education", *Journal of Health Economics*, 29(1): 1-28.

Cutler, D., Fung, W., Kremer, M., Singhal, M., and Vogl, T. (2010). "Early-life malaria exposure and adult outcomes: Evidence from malaria eradication in India", *American Economic Journal: Applied Economics*, 72-94.

Deaton, A. (2013). "The Great Escape. Health, Wealth, and the Origins of Inequality", Princeton University Press.

De Nardi, M., E. French, and J.B. Jones (2010), "Why do the elderly save? The role of medical care expenses", *Journal of Political Economy*, 118(1): 39-75.

DHHS (2014), "Current cigarette smoking among adults United States, 2005-2012", US Department of Health and Human Services, MMWR 63(2): 29-46.

Ehrlich, I. and H. Chuma (1990), "A Model of the Demand for Longevity and the Value of Life Extension", *Journal of Political Economy*, 98(4): 761-782.

Fonseca, Raquel, Michaud, Pierre-Carl, Kapteyn, Arie and Titus Galama (2013), "Accounting for the Rise of Health Spending and Longevity" IZA Discussion Paper No. 7622.

Fortson, J. (2011), "Mortality Risk and Human Capital Investment: The Impact of

HIV/AIDS in Sub-Saharan Africa," *Review of Economics and Statistics*, 93(1): pp. 1-15.

Galama, T.J. and H. van Kippersluis (2010). "A theory of socioeconomic disparities in health", RAND Working Paper, WR-773.

Galama, T.J. (2011), "A contribution to Health Capital Theory", RAND Working Paper WR-831.

Galama, T.J. and H. van Kippersluis (2013). "Health Inequalities through the Lens of Health Capital Theory: Issues, Solutions, and Future Directions", In: Research on Economic Inequality, Ed: John Bishop. Volume 21: Health and Inequality, Eds: Pedro Rosa Dias, Owen O'Donnell.

Garcia-Gomez, P., H. van Kippersluis, O. ODonnell, and E. van Doorslaer (2013), "Long-term and spillover effects of health shocks on employment and income", *Journal of Human Resources* 48(4): 873-909.

Glied, S. and A. Lleras-Muney (2008), "Health Inequality, Education and Medical Innovation", *Demography* 45(3): 741-761.

Goldin, C., and L. Katz (2009), "The Race between Education and Technology", Cambridge, MA: Harvard University Press.

Goldman, D.P. and J.P. Smith (2002), "Can Patient Self-Management Help Explain the SES Health Gradient?", *Proceedings of the National Academy of Science*, 99: 10929-10934.

Graham, J.W. (1981), "An Explanation for the Correlation of Stocks of Nonhuman Capital with Investment in Human Capital", *American Economic Review*, 71(1):248-255.

Grantham-McGregor, S., Cheung, Y.B., Cueto, S., Glewwe, P., Richter, L. et al. (2007), "Developmental Potential in the First 5 years for Children in Developing Countries", *Lancet* 369: 60-70.

Grossman, M. (1972a), *The Demand for Health-A Theoretical and Empirical Investigation*, New York: National Bureau of Economic Research.

Grossman, M. (1972b), "On the Concept of Health Capital and the Demand for Health", *Journal of Political Economy*, 80(2): 223-255.

Grossman, M. (2000), "The Human Capital Model", in Handbook of Health Economics, Volume 1, 347-408, A.J. Culyer and J.P. Newhouse (Eds.), Elsevier Science.

Hai, R., and Heckman, J. (2014),"A Dynamic Model of Health, Education and Wealth with Credit Constraints and Rational Addiction", Mimeo, University of Chicago.

Haley, William J. (1973), "Human Capital: The Choice between Investment and Income", *American Economic Review* 63(5): 929-944.

Hazan, Moshe (2009),"Longevity and lifetime labor supply: Evidence and implications." *Econometrica* 77(6): 1829-1863.

Heckman, J.J. (1976), "A lifecycle model of earnings, learning, and consumption", *Journal of Political Economy* 84(4): pp. 11-44.

Imbens, Guido, Donald Rubin, and Bruce Sacerdote (2001). "Estimating the Effect of Unearned Income on Labor Supply, Earnings, Savings and Consumption: Evidence from a Survey of Lottery Players", *American Economic Review* 91(4): pp. 778-794.

Jayachandran, S. and Lleras-Muney, A. (2009), "Life Expectancy and human capital investments: Evidence from Maternal Mortality Declines", *Quarterly Journal of Economics*, 124(1): pp. 349-397.

Juster, Thomas F., Joseph P. Lupton, James P. Smith, and Frank Stafford (2006), "The Decline in Household Saving and the Wealth Effect", *The Review of Economics and Statistics* 88(1): pp. 20-27.

Kuhn, Michael, Wrzaczek, Stefan, Prskawetz, Alexia, and Feichtinger, Gustav (2012), "Optimal choice of health and retirement in a life-cycle model", *ECON WPS - Vienna University of Technology Working Papers in Economic Theory and Policy*, 01/2012.

Lleras-Muney, A. (2005), "The Relationship Between Education and Adult Mortality in the United States", *Review of Economic Studies*, 72: 189-221.

Lleras-Muney, A. and F. Lichtenberg (2005), "The Effect of Education on Medical Technology Adoption: Are the More Educated More Likely to Use New Drugs", *Annales d'Economie et Statistique* in memory of Zvi Griliches, No 79/80.

Madsen, Jakob (2012), "Health, Human Capital Formation and Knowledge Production: Two Centuries of International Evidence", *NBER Working Paper* 18461.

Mazumder, B. (2008), "Does Education Improve Health? A Reexamination of the Evidence from Compulsory Schooling Laws", *Federal Reserve Bank of Chicago Economic Perspectives*, Q2: 2-16.

Mazumder, B. (2012), “The effects of education on health and mortality”,*Nordic Economic Policy Review* 1/2012: 261-301.

Meghir, C., M. Palme, and E. Simeonova (2012), “Education, Health and Mortality: Evidence from a Social Experiment”, NBER Working Paper No. 17932.

Mincer, J.A. (1974), *Schooling, Experience, and Earnings*, Columbia University Press.

Murphy, K. and Topel, R.H. (2006), “The Value of Health and Longevity”, *Journal of Political Economy* 114(5): 871-904.

Mushkin, S. (1962), “Health as an investment”, *Journal of Political Economy* 70(5): pp. 129-157.

Organization for Economic Cooperation and Development (2011), “Pensions at a Glance 2011: Retirement-Income Systems in OECD and G20 Countries”, OECD.

Oster, Emily, Ira Shoulson, and E. Dorsey (2013), “Limited Life Expectancy, Human Capital and Health Investments.” *The American Economic Review* 103(5): 1977-2002.

Pearson, T., Russell, J., Campbell, M. J., and Barker, M. E. (2005), “Do ’food deserts’ influence fruit and vegetable consumption?A cross-sectional study.” *Appetite*, 45(2), 195-197.

Perri, T.J. (1984), “Health Status and Schooling Decisions of Young Men”, *Economics of Education Review*, 3(3): 207-213.

Serdula, M. K., Gillespie, C., Kettel-Khan, L., Farris, R., Seymour, J., and Denny, C. (2004), “Trends in fruit and vegetable consumption among adults in the United States: behavioral risk factor surveillance system, 19942000. ” *American Journal of Public Health*, 94(6), 1014-1018.

Podor, M., and T.J. Halliday (2012), “Health Status and the Allocation of Time”, *Health Economics* 21, 514-527.

Seierstad, A. and K. Sydsaeter (1987), “Optimal Control Theory With Economic Applications”, In: Advanced Textbooks in Economics, Volume 24, Series Editors: C.J. Bliss and M.D. Intriligator, Elsevier, North Holland.

Schultz, T. (1961), “Investment in Human Capital,” *American Economic Review* 51: pp. 1-17.

Soares, Rodrigo R. (2005), "Mortality reductions, educational attainment, and fertility choice", *The American Economic Review* 95(3): 580-601.

Soares, Rodrigo R. (2006), " The Effect of Longevity on Schooling and Fertility: Evidence from the Brazilian Demographic and Health Survey", *Journal of Population Economics* 19(1), 71-97.

Strulik, H. (2013), "Health and Education: Understanding the Gradient", Center for European, Governance and Economic Development Research Discussion Papers 172, University of Goettingen, Department of Economics.

Van Kippersluis, H., O. O'Donnell, and E. van Doorslaer (2011), "Long Run Returns to Education: Does Schooling lead to an Extended Old Age?" *Journal of Human Resources* 96(4): 695-721.

Weil, D.N. (2007), "Accounting for the effect of health on economic growth", *Quarterly Journal of Economics* 122(3), pp. 1265-1306.

Willis, R. J. (1985), "Wage Determinants: A Survey and Reinterpretation of Human Capital Earnings Functions." In O. Ashenfelter and D. Card (eds.), Handbook of Labor Economics. North Holland, Amsterdam.

Zweifel, P., Felder, S., and Meiers, M. (1999), "Ageing of population and health care expenditures: a red herring? ", Health Economics, 8(6): pp. 485-496.

A Appendix

A.1 First-order (necessary) conditions: general framework

The first-order necessary conditions for the optimal control problem, consisting of maximizing the objective function (1) subject to the constraints (2) to (4) and begin and end conditions, follow from Pontryagin's maximum principle (e.g., Caputo, 2005). The Hamiltonian is given by (10). For the co-state variable $q_A(t)$ associated with assets we have

$$\begin{aligned}\frac{\partial q_A}{\partial t} &= -\frac{\partial \Im}{\partial A} = -q_A(t)r \Leftrightarrow \\ q_A(t) &= q_A(0)e^{-rt}.\end{aligned} \tag{37}$$

The co-state variable $q_E(t)$ associated with skill capital follows from

$$\begin{aligned}\frac{\partial q_E}{\partial t} &= -\frac{\partial \Im}{\partial E} \\ &= -q_A(t)\frac{\partial Y}{\partial E} - q_H(t)\frac{\partial f_H}{\partial E} + q_E(t)\left[d_E(t) - \frac{\partial f_E}{\partial E}\right]. \qquad S \leq t < R\end{aligned} \tag{38}$$

For non-working ages, income $Y(t)$ is fixed, and the evolution of the co-state variable $q_E(t)$ reduces to

$$\begin{aligned}\frac{\partial q_E}{\partial t} &= -\frac{\partial \Im}{\partial E} \\ &= -q_H(t)\frac{\partial f_H}{\partial E} + q_E(t)\left[d_E(t) - \frac{\partial f_E}{\partial E}\right]. \qquad 0 \leq t < S, R \leq t < T\end{aligned} \tag{39}$$

The co-state variable $q_H(t)$ associated with health capital follows from

$$\begin{aligned}\frac{\partial q_H}{\partial t} &= -\frac{\partial \Im}{\partial H} \\ &= -\frac{\partial U}{\partial H}e^{-\beta t} - q_A(t)\frac{\partial Y}{\partial H} \\ &\quad - q_E(t)\frac{\partial f_E}{\partial H} + q_H(t)\left[d_H(t) - \frac{\partial f_H}{\partial H}\right] - \lambda_{H_{\min}}(t). \qquad S \leq t < R\end{aligned} \tag{40}$$

For non-working ages, income $Y(t)$ is fixed, and the evolution of the co-state variable $q_H(t)$ is given by

$$\begin{aligned}\frac{\partial q_H}{\partial t} &= -\frac{\partial \Im}{\partial H} \\ &= -\frac{\partial U}{\partial H}e^{-\beta t} - q_E(t)\frac{\partial f_E}{\partial H} + q_H(t)\left[d_H(t) - \frac{\partial f_H}{\partial H}\right] \\ &\quad + \lambda_{\tau_w}(t)w[t, E(t)]\frac{\partial s}{\partial H} - \lambda_{H_{\min}}(t). \qquad 0 \leq t < S, R \leq t < T\end{aligned} \tag{41}$$

where the cost of sick time is now valued at $\lambda_{\tau_w}(t)$.

The first-order condition for investment in skill capital (18) follows from optimizing with respect to skill-capital investment goods and services $X_E(t)$ and time inputs $\tau_E(t)$:

$$\frac{\partial \Im}{\partial X_E} = 0 \Leftrightarrow \quad q_E(t)\frac{\partial f_E}{\partial I_E}\frac{\partial I_E}{\partial X_E} - q_A(t)p_E(t) = 0 \tag{42}$$

$$\frac{\partial \Im}{\partial \tau_E} = 0 \Leftrightarrow \quad q_E(t)\frac{\partial f_E}{\partial I_E}\frac{\partial I_E}{\partial \tau_E} - q_A(t)w[t, E(t)] = 0 \qquad S \leq t < R \tag{43}$$

$$\Leftrightarrow \quad q_E(t)\frac{\partial f_E}{\partial I_E}\frac{\partial I_E}{\partial \tau_E} - \lambda_{\tau_w}(t)w[t, E(t)] = 0. \qquad 0 \leq t < S, R \leq t < T \tag{44}$$

The first-order condition for investment in health capital (23) follows from optimizing with respect to health investment goods and services $X_H(t)$ and time inputs $\tau_H(t)$:

$$\frac{\partial \Im}{\partial X_H} = 0 \Leftrightarrow \quad q_H(t)\frac{\partial f_H}{\partial I_H}\frac{\partial I_H}{\partial X_H} - q_A(t)p_H(t) = 0 \tag{45}$$

$$\frac{\partial \Im}{\partial \tau_H} = 0 \Leftrightarrow \quad q_H(t)\frac{\partial f_H}{\partial I_H}\frac{\partial I_H}{\partial \tau_H} - q_A(t)w[t, E(t)] = 0 \qquad S \leq t < R \tag{46}$$

$$\Leftrightarrow \quad q_H(t)\frac{\partial f_H}{\partial I_H}\frac{\partial I_H}{\partial \tau_H} - \lambda_{\tau_w}(t)w[t, E(t)] = 0. \qquad 0 \leq t < S, R \leq t < T \tag{47}$$

The first-order condition for consumption (15) follows from optimizing with respect to consumption goods and services $X_C(t)$:

$$\frac{\partial \Im}{\partial X_C} = 0 \Leftrightarrow \quad \frac{\partial U}{\partial X_C}e^{-\beta t} - q_A(t)p_C(t) = 0. \tag{48}$$

The first-order condition for leisure time (16) follows directly from optimizing with respect to leisure time $L(t)$:

$$\frac{\partial \Im}{\partial L} = 0 \Leftrightarrow \quad \frac{\partial U}{\partial L}e^{-\beta t} - q_A(t)w[t, E(t)] = 0 \qquad S \leq t < R \tag{49}$$

$$\Leftrightarrow \quad \frac{\partial U}{\partial L}e^{-\beta t} - \lambda_{\tau_w}(t)w[t, E(t)] = 0. \qquad 0 \leq t < S, R \leq t < T \tag{50}$$

A.2 First-order (necessary) conditions: simpler model

The first-order conditions are obtained by taking the derivative of the Hamitonian

$$\Im = U\{X_C(t), L(t)[E(t)+H(t)]\}e^{-\beta t} + q_E(t)\frac{\partial E}{\partial t} + q_H(t)\frac{\partial H}{\partial t} + q_A(t)\frac{\partial A}{\partial t}, \tag{51}$$

with respect to the controls (not shown). Start with the first-order condition for the optimal expenditures on skill capital goods, $X_E(t)$, and for time inputs, $\tau_E(t)$, and divide the two resulting expressions by one another to obtain the relation

$$\tau_E(t)\,[E(t)+H(t)] = \frac{\alpha_E}{\beta_E} p_E(t) X_E(t). \tag{52}$$

Similarly for health investment one obtains the relation

$$\tau_H(t)\,[E(t)+H(t)] = \frac{\alpha_H}{\beta_H} p_H(t) X_H(t). \tag{53}$$

Now insert these relations back into the first-order condition for $X_E(t)$, $\tau_E(t)$, $X_H(t)$, and $\tau_H(t)$, to obtain the analytical solutions:

$$\begin{aligned}
X_E(t) &= \frac{\beta_E \mu_E(t)}{p_E(t)} q_{e/a}(t)^{\frac{1}{1-\gamma_E}}, & (54)\\
\tau_E(t)\,[E(t)+H(t)] &= \alpha_E \mu_E(t) q_{e/a}(t)^{\frac{1}{1-\gamma_E}}, & (55)\\
X_H(t) &= \frac{\beta_H \mu_H(t)}{p_H(t)} q_{h/a}(t)^{\frac{1}{1-\gamma_H}}, & (56)\\
\tau_H(t)\,[E(t)+H(t)] &= \alpha_H \mu_H(t) q_{h/a}(t)^{\frac{1}{1-\gamma_H}}, & (57)
\end{aligned}$$

where $\gamma_E = \alpha_E + \beta_E$, and $\gamma_H = \alpha_H + \beta_H$, and the functions $\mu_E(t)$ and $\mu_H(t)$ are defined in (34) and (35).

The co-state equations for $q_E(t)$ and $q_H(t)$ follow from the usual conditions $\partial q_E/\partial t = -\partial\Im/\partial E$ and $\partial q_H/\partial t = -\partial\Im/\partial H$, and using (54) to (57), we obtain

$$\begin{aligned}
\frac{\partial q_E}{\partial t} &= q_E(t) d_E(t) - q_A(t), & (58)\\
\frac{\partial q_H}{\partial t} &= q_H(t) d_H(t) - q_A(t). & (59)
\end{aligned}$$

Using the dynamic relation for skill- (2) and health-capital formation (3), the Ben-Porath production functions (32) and (33), and the solutions for the controls (54) to (57), one obtains analytical expressions for the relative marginal value of skill capital $q_{e/a}(t)$, skill capital $E(t)$, the relative marginal value of health capital $q_{h/a}(t)$, and health capital $H(t)$:

$$\begin{aligned}
q_{e/a}(t) &= \int_t^T e^{-\int_t^s [d_E(x)+r]dx} ds, & (60)\\
E(t) &= E_0 e^{-\int_0^t d_E(x)dx} + \int_0^t \mu_E(s) q_{e/a}(s)^{\frac{\gamma_E}{1-\gamma_E}} e^{-\int_s^t d_E(x)dx} ds, & (61)
\end{aligned}$$

$$
\begin{aligned}
q_{h/a}(t) &= q_{h/a}(0)e^{\int_0^t [d_H(x)+r]dx} - \int_0^t e^{\int_s^t [d_H(x)+r]dx} ds, && (62)\\
H(t) &= H_0 e^{-\int_0^t d_H(x)dx} + \int_0^t \mu_H(s) q_{h/a}(s)^{\frac{\gamma_H}{1-\gamma_H}} e^{-\int_s^t d_H(x)dx} ds, && (63)
\end{aligned}
$$

where we have used $q_{e/a}(T) = 0$, and the solution for the marginal value of assets $q_A(t) = q_A(0)e^{-rt}$, see (37).

Using the dynamic relation for assets (4), (31), and (54) to (67), we obtain

$$
\begin{aligned}
A(t)e^{-rt} &= A_0 + \int_0^t e^{-rs}\left[E(s) + H(s)\right] ds \\
&- \gamma_E \int_0^t \mu_E(s) q_{e/a}(s)^{\frac{1}{1-\gamma_E}} e^{-rs} ds \\
&- \gamma_H \int_0^t \mu_H(s) q_{h/a}(s)^{\frac{1}{1-\gamma_H}} e^{-rs} ds \\
&- q_A(0)^{-1/\rho} \Lambda \int_0^t p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta - r(1-\rho)r)}{\rho} s} ds. && (64)
\end{aligned}
$$

Finally, the analytical solutions for consumption $X_C(t)$ and leisure $L(t)\left[E(t)+H(t)\right]$ are obtained by dividing the two first-order conditions, leading to the relation

$$
L(t)\left[E(t)+H(t)\right] = \frac{(1-\zeta)}{\zeta} p_C(t) X_C(t). \tag{65}
$$

Inserting this relation back into the first-order conditions for consumption $X_C(t)$ and leisure $L(t)$, leads to the analytical solutions

$$
\begin{aligned}
X_C(t) &= \zeta \Lambda q_A(0)^{-1/\rho} p_C(t)^{-(1-\zeta+\zeta/\rho)} e^{-\frac{\beta - r}{\rho} t}, && (66)\\
L(t)[E(t)+H(t)] &= (1-\zeta)\Lambda q_A(0)^{-1/\rho} p_C(t)^{-\zeta(1-\rho)/\rho} e^{-\frac{\beta - r}{\rho} t}, && (67)
\end{aligned}
$$

where

$$
\Lambda \equiv \left[\zeta^{\zeta}(1-\zeta)^{1-\zeta}\right]^{\frac{1-\rho}{\rho}}. \tag{68}
$$

The analytical solutions for the controls, state variables, and co-state variables (54) to (67), are functions of the marginal value of initial wealth $q_A(0)$, the initial relative marginal value of skill-capital $q_{e/a}(0)$, and the initial relative marginal value of health-capital $q_{h/a}(0)$. These in turn are determined by initial, end, and transversality conditions.

From (64), and the initial, $A(0) = A_0$, and end condition, $A(T) = A_T$, follows a condition for $q_A(0)$

$$
\begin{aligned}
A_T e^{-rT} &= A_0 + \int_0^T e^{-rs}\left[E(s)+H(s)\right] ds \\
&- \gamma_E \int_0^T \mu_E(s) q_{e/a}(s)^{\frac{1}{1-\gamma_E}} e^{-rs} ds \\
&- \gamma_H \int_0^T \mu_H(s) q_{h/a}(s)^{\frac{1}{1-\gamma_H}} e^{-rs} ds \\
&- q_A(0)^{-1/\rho} \Lambda \int_0^T p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta - r(1-\rho))}{\rho} s} ds. \qquad (69)
\end{aligned}
$$

From (63) and the initial, $H(0) = H_0$, and end condition, $H(T) = H_{\min}$, follows a condition for $q_{h/a}(0)$

$$
H_{\min} e^{\int_0^T d_H(x)dx} = H_0 + \int_0^T \mu_H(s) q_{h/a}(s)^{\frac{\gamma_H}{1-\gamma_H}} e^{\int_0^s d_H(x)dx} ds. \qquad (70)
$$

The condition for $q_{e/a}(0)$ follows from the transversality condition $q_E(T) = 0$ ($E(T)$ free) and is obtained from (60)

$$
q_{e/a}(0) = \int_0^T e^{-\int_0^s [d_E(x)+r]dx} ds. \qquad (71)
$$

The remaining endogenous parameters and functions in the above three conditions (69), (70), and (71), are T, which is determined by (14), $q_{e/a}(t)$, which is determined by (60), $E(t)$, which is determined by (61), $q_{h/a}(t)$, which is determined by (62), and $H(t)$, which is determined by (63).

A.3 Life cycle profiles of skill and health in the simpler model

The convenient choices made for the functional forms, referred to as "Ben-Porath neutrality" ensure that the relative marginal value of skill capital $q_{e/a}(t)$, and in our case also of health capital $q_{h/a}(t)$, are independent of the capital stocks (see 58 and 59). The system of equations for (the relative marginal value of) skill capital, and (the relative marginal value of) health capital reduces to the following system:

$$
\begin{aligned}
\frac{\partial q_{e/a}}{\partial t} &= q_{e/a}(t)\left[d_E(t) + r\right] - 1, \qquad (72) \\
\frac{\partial E}{\partial t} &= \mu_E(t) q_{e/a}(t)^{\frac{\gamma_E}{1-\gamma_E}} - d_E(t)E(t), \qquad (73) \\
\frac{\partial q_{h/a}}{\partial t} &= q_{h/a}(t)\left[d_H(t) + r\right] - 1, \qquad (74) \\
\frac{\partial H}{\partial t} &= \mu_H(t) q_{h/a}(t)^{\frac{\gamma_H}{1-\gamma_H}} - d_H(t)H(t). \qquad (75)
\end{aligned}
$$

Relations (72) and (73) depend endogenously only on the relative marginal value of skill $q_{e/a}(t)$ and the stock of skill capital $E(t)$. Similarly, relations (74) and (75) depend endogenously only on the relative marginal value of health $q_{h/a}(t)$ and the stock of health capital $H(t)$. Thus, we have obtained separate, quasi-independent systems for skill capital (72 and 73) and for health (74 and 75).[45] The phase diagrams for skill capital and for health capital are shown in Figure 2.

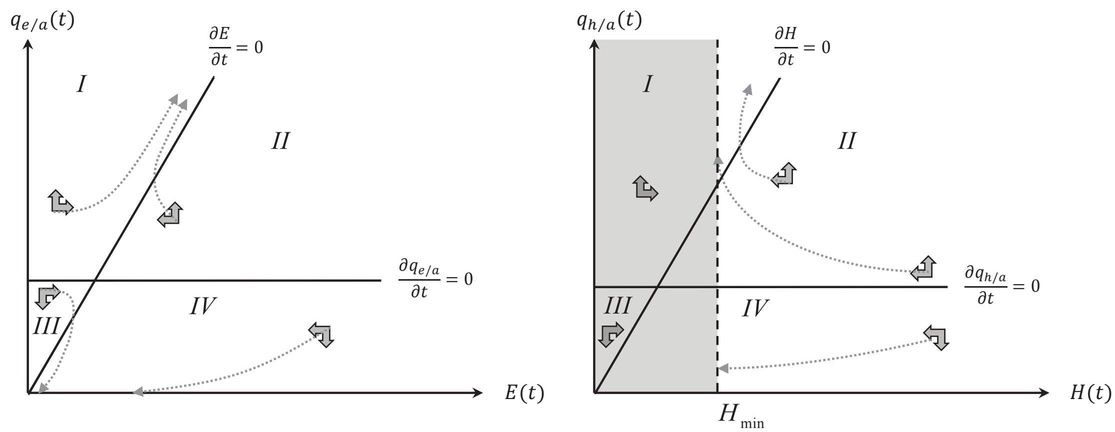

Figure 2: *Phase diagrams of the relative marginal value of skill $q_{e/a}(t)$ versus skill capital $E(t)$ (left) and of the relative marginal value of health $q_{h/a}(t)$ versus health capital $H(t)$ (right).*

The left-hand side of Figure 2 shows the direction of motion of the optimal solution of the system of first-order differential equations given by (72) and (73) as a function of the relative marginal value of skill $q_{e/a}(t)$ (vertical axis) versus the stock of skill $E(t)$ (horizontal axis). Regime switches occur when $\partial q_{e/a}/\partial t = 0$ and $\partial E/\partial t = 0$. These boundaries between regimes, so called null-clines, are shown by the thick lines in the figure and are obtained by setting the derivatives $\partial q_{e/a}/\partial t$ and $\partial E/\partial t$ to zero in (72) and (73), respectively. Since the null-cline for the relative marginal value of skill, $\partial q_{e/a}/\partial t = 0$, is independent of the level of skill capital (see 72) it runs horizontally.

For $\gamma_E < 0.5$, $\gamma_E = 0.5$, and $\gamma_E > 0.5$, the null-cline for skill capital, $\partial E/\partial t = 0$, defines, respectively, a convex, linear, and concave relation between $q_{e/a}(t)$ and $E(t)$ (see 73). Here, we show a linear relation ($\gamma_E = 0.5$). The two null-clines define four distinct dynamic regions: I, II, III and IV. For example, every point in region III is associated with an evolution toward lower relative value of skill capital $q_{e/a}(t)$ (i.e., toward lower levels of investment $\tau_E(t)$, $X_E(t)$) and higher skill capital $E(t)$. The left-up, right-up, left-down, and right-down block arrows indicate the direction of motion in the phase diagram and the

[45] The systems are not fully independent as they are still connected through the marginal value of wealth $q_A(t)$ (the budget constraint).

grey dotted lines provide example trajectories. Any trajectory in the phase diagram is an optimal solution for the relative marginal value of skill $q_{e/a}(t)$ and for skill capital $E(t)$. Its starting point is determined by the initial condition $E(0) = E_0$, and by the transversality condition $q_E(T) = 0$ (and hence $q_{e/a}(T) = 0$; see 18). This rules out trajectories that do not end on the horizontal axis.

The null-clines may shift with time (except for the autonomous system where $\mu_E(t)$, $\mu_H(t)$, $d_E(t)$, and $d_H(t)$ are constant), but the general properties of the phase diagram do not change. The fixed point (where the two null-clines cross), is of little interest as a potential solution for the system. First, it is saddle-point unstable. This is clear from visual inspection of the phase diagram: a small deviation (perturbation) from the fixed point will evolve away from the fixed point, except if the deviation landed on an infinitesimally narrow trajectory (the unique trajectory that eventually leads to the fixed point).[46] Second, if the trajectory starts at a point that is not a fixed point, it cannot reach a fixed point in a finite amount of time (Theorem 13.4, p. 350, Caputo 2005). Thus, the fixed point requires infinite length of life, exacting (i.e., highly unlikely) initial conditions and the absence of any perturbations (no matter how small).

The right-hand side of Figure 2 shows the phase diagram for health capital $H(t)$. Its interpretation is analogous to the diagram for skill capital, with two exceptions: there is no end condition for $q_H(T)$ and end of life occurs when health deteriorates to a minimum health level $H_T = H_{\min}$. A priori we do not know the location of $H_{\min}$ and it is shown to the right of the fixed point for illustration. The region to the left of $H_{\min}$ is not allowed as illustrated by the shaded area.

Solutions for skill capital are best described by region III: gradual building up of skill capital over the life-cycle, followed by declining skill capital, and decreasing levels of investment in skill capital (i.e. decreasing $q_{e/a}(t)$) as the returns to schooling diminish with shortening of the time horizon. Solutions in regions III and IV are feasible (they end at the horizontal axis) and solutions from region I and II have to be ruled out (except if, due to shifting of the null-clines, they eventually end up in regions III and IV). Solutions for health capital are best described by region II: declining health capital and increasing levels of investment in health with age (i.e. increasing $q_{h/a}(t)$).

In sum, skill capital is plausibly characterized by an investment process where investment in skill at young ages increases the level of skill initially, and the relative marginal value of skill decreases with age (investment decreases with shortening of the horizon). Health capital is plausibly characterized by an investment process where no matter how much one invests, health eventually declines, and the relative marginal value of health increases with age (investment in health increases with age till death).

[46] A formal proof of the instability of the fixed point can be straightforwardly obtained by calculating the Jacobian $J(q^*_{e/a}, E^*)$ of the linearized system at the fixed point $(q^*_{e/a}, E^*)$ and showing that the two eigenvalues are real, non-zero, and unequal ($\{\mathrm{tr}[J(q^*_{e/a}, E^*)]\}^2 > 4\det[J(q^*_{e/a}, E^*)]$). The same holds for health capital. See Theorem 13.6, p. 354 of Caputo (2005).

A.4 Comparison with the general theory

The simpler version of our model maintains the most important properties of the general model defined in section 3.1. There are some costs associated with the simplications associated with the assumption of "Ben-Porath neutrality" (see below), which we describe here, but the benefits arguably outweigh the costs. Most importantly, the assumption enables obtaining analytical results for the comparative dynamic analyses.

As for the general model, in the simpler model both skill and health contribute to earnings, the production processes of skill and health investment are increasing and concave in the investment inputs,[47] and they exhibit both self-reinforcing and cross-fertilizing self-productivity and dynamic complementarity. Not surprisingly, the dynamics of the simpler model are qualitatively similar to that of the general model (see Appendix A.3). Skill capital and health capital occupy different regions of the phase space. The relative marginal value of skill decreases with age (investment decreases with shortening of the horizon) and skill capital increases. The relative marginal value of health increases with age (investment in health increases with age till death) and health capital declines (conjectures 1 and 2).

Compared to the general theory, there are however a few differences. First, the assumed specific functional form for the utility, earnings, skill-production, and health-production functions, ensure that the marginal value of skill and of health are no longer functions of the stock of skill and health (compare 20 and 25 with 72 and 74). This is commonly known as "Ben-Porath neutrality", and as a result we can solve the model analytically.[48] In the general model, however, the relative marginal value of skill is likely to be decreasing in the stock of skill (due to decreasing returns to scale) and potentially increasing in the stock of health (due to complementarity between skill and health, in the generation of earnings, and in the production of skill and health investment $f_E(t)$ and $f_H(t)$, see 20). The opposite is true for the relative marginal value of health (see 25), which is likely decreasing in health and potentially increasing in skill.

Second, we assume that there are no separate periods exclusively devoted to schooling S and to retirement R. While the model no longer contains an explicit school-leaving age and retirement age, schooling and retirement phases do exist. Early in life individuals invest in skill capital as the stock of skill is low (and hence the marginal benefits high), the opportunity cost of time is low, and the horizon over which the benefits of skill-capital investment can be reaped is long. Individuals do not work much as low skill capital implies low earnings, such that this period of life corresponds to a schooling phase. As individuals develop skill capital they start investing less in skill due to gradually declining marginal benefits and shortening of the horizon, and work more (working phase). Later

[47]For $\alpha_E + \beta_E < 1$ and $\alpha_H + \beta_H < 1$, we have $\partial f_E/\partial X_E > 0$, $\partial f_E/\partial \tau_E > 0$, $\partial f_H/\partial X_H > 0$, $\partial f_H/\partial \tau_H > 0$, $\partial^2 f_E/\partial X_E^2 < 0$, $\partial^2 f_E/\partial \tau_E^2 < 0$, $\partial^2 f_H/\partial X_H^2 < 0$, $\partial^2 f_H/\partial \tau_H^2 < 0$, $\left(\partial^2 f_E/\partial X_E^2\right)\left(\partial^2 f_E/\partial \tau_E^2\right) > \left(\partial^2 f_E/\partial X_E \partial \tau_E\right)^2$ and $\left(\partial^2 f_H/\partial X_H^2\right)\left(\partial^2 f_H/\partial \tau_H^2\right) > \left(\partial^2 f_H/\partial X_H \partial \tau_H\right)^2$.

[48]To maintain Ben-Porath neutrality, skill also enters the utility function in the simpler version of the model, so that not only health, but also skill provides a consumption benefit.

in life individuals work less and invest more in health as a result of declining health, corresponding to a retirement phase. Thus the simpler model contains phases of schooling, work and retirement. The institutions of schooling and retirement, defined in the general model, only formalize and exacerbate this natural pattern.

Third, the simpler model assumes no sick time. Therefore health $H(t)$ does not protect time per period (let's say during a day), and the simpler model loses the characteristic of earnings being multiplicative in skill and health. Both health and skill still contribute to earnings, but do so in an additive way.[49] Since we find strong complementarity effects between skill and health even for the simpler model, the general model would only excacerbate these, but would not lead to a different conclusion.

A.5 Comparative dynamics: simpler model

Consider a generic control, state, or co-state function $g(t)$ and a generic variation δZ_0 in an initial condition or model parameter. The effect of the variation δZ_0 on the optimal path of $g(t)$ can be broken down into variation for fixed longevity T and variation due to the resulting change in the horizon T (see 36). In the below analyses (i) we first analyze the case for fixed T, from which we obtain $\partial g(t)/\partial Z_0|_T$ (see discussion below), (ii) we then determine $\partial T/\partial Z_0$, and (iii) last we obtain $\partial g(t)/\partial T|_{Z_0}$, so that we have determined the full comparative dynamic effect.

Comparative dynamics of length of life $\partial T/\partial Z_0$ For fixed length of life T we can take derivatives of the first-order conditions and state equations with respect to the initial condition or model parameter and study the optimal adjustment to the lifecycle path in response to variation in an initial endowment or other model parameter.

For free T, however, this is slightly more involved since the additional condition $\Im(T) = 0$ has to be satisfied. Varying the initial condition or model parameter Z_0, and taking into account $\Im(T) = 0$, we have

$$\left.\frac{\partial \Im(T)}{\partial Z_0}\right|_T \delta Z_0 + \left.\frac{\partial \Im(T)}{\partial T}\right|_{Z_0} \delta T = 0. \tag{76}$$

Using the expression for the Hamiltonian (51), taking the first derivative of the transversality condition $\Im(T) = 0$ with respect to the initial conditions or model parameter Z_0, and holding length of life T fixed, we obtain

[49] For simplicity we assume constant returns to scale of human capital $E(t) + H(t)$ in the production of wages $w[t, E(t), H(t)]$. Predictions are however not affected when imposing decreasing or increasing returns to scale $w[t, E(t), H(t)] = [E(t) + H(t)]^\sigma$ with $\sigma \neq 1$, as long as human capital affects the utility of leisure $U\{X_C(t), L(t)\,[E(t) + H(t)]^\sigma\}$, and the efficiency of time investments $\tau_E(t)\,[E(t) + H(t)]^\sigma$ and $\tau_H(t)\,[E(t) + H(t)]^\sigma$ in the production functions of skill capital and health capital in the same way.

$$
\begin{aligned}
\left.\frac{\partial \Im(T)}{\partial Z_0}\right|_T &= \left.\frac{\partial \Im}{\partial \xi}\frac{\partial \xi(T)}{\partial Z_0}\right|_T + \left.\frac{\partial \Im}{\partial E}\frac{\partial E(T)}{\partial Z_0}\right|_T + \left.\frac{\partial \Im}{\partial A}\frac{\partial A(T)}{\partial Z_0}\right|_T + \left.\frac{\partial \Im}{\partial H}\frac{\partial H(T)}{\partial Z_0}\right|_T \\
&\quad + \left.\frac{\partial \Im}{\partial q_E}\frac{\partial q_E(T)}{\partial Z_0}\right|_T + \left.\frac{\partial \Im}{\partial q_A}\frac{\partial q_A(T)}{\partial Z_0}\right|_T + \left.\frac{\partial \Im}{\partial q_H}\frac{\partial q_H(T)}{\partial Z_0}\right|_T \\
&= -\left.\frac{\partial q_E(t)}{\partial t}\right|_{T,t=T}\left.\frac{\partial E(T)}{\partial Z_0}\right|_T + \left.\frac{\partial q_A(T)}{\partial Z_0}\right|_T\left.\frac{\partial A(t)}{\partial t}\right|_{T,t=T} \\
&\quad + \left.\frac{\partial q_H(T)}{\partial Z_0}\right|_T\left.\frac{\partial H(t)}{\partial t}\right|_{T,t=T},
\end{aligned}
\tag{77}
$$

where $\xi(t)$ is the vector of control functions $X_C(t)$, $L(t)$, $X_E(t)$, $\tau_E(t)$, $X_H(t)$, and $\tau_H(t)$. The first-order conditions imply $\partial\Im(t)/\partial\xi(t) = 0$. Further, $\partial\Im(T)/\partial E = -\left.\partial q_E(t)/\partial t\right|_{t=T}$, $\partial A(T)/\partial Z_0 = \partial H(T)/\partial Z_0 = 0$ since $A(T)$ and $H(T)$ are fixed, and $\left.\partial q_E(T)/\partial Z_0\right|_T = 0$ since $q_E(T) = 0$.

Note that we distinguish in notation between $\left.\partial f(t)/\partial t\right|_{t=T}$, which represents the derivative with respect to time t at time $t = T$, and $\left.\partial f(t)/\partial T\right|_{t=T}$, which represents variation with respect to the parameter T at time $t = T$.

From (76) and (77) we have

$$
\frac{\partial T}{\partial Z_0} = \frac{q_A(T)\left.\frac{\partial E(T)}{\partial Z_0}\right|_T + \left.\frac{\partial q_A(T)}{\partial Z_0}\right|_T\left.\frac{\partial A(t)}{\partial t}\right|_{T,t=T} + \left.\frac{\partial q_H(T)}{\partial Z_0}\right|_T\left.\frac{\partial H(t)}{\partial t}\right|_{T,t=T}}{-\left.\frac{\partial\Im(T)}{\partial T}\right|_{Z_0}},
\tag{78}
$$

where we have used $\left.\partial q_E(t)/\partial t\right|_{t=T} = q_E(T)d_E(T) - q_A(T) = -q_A(T)$ (see 58 and use $q_E(T) = 0$).

The denominator of (78) can be obtained from

$$
\begin{aligned}
\left.\frac{\partial \Im(T)}{\partial T}\right|_{Z_0} &= -\beta U[.]e^{-\beta T} \\
&\quad + q_A(T)\left.\frac{\partial E(T)}{\partial T}\right|_{Z_0} + \left.\frac{\partial q_A(T)}{\partial T}\right|_{Z_0}\left.\frac{\partial A(t)}{\partial t}\right|_{t=T} + \left.\frac{\partial q_H(T)}{\partial T}\right|_{Z_0}\left.\frac{\partial H(t)}{\partial t}\right|_{t=T},
\end{aligned}
\tag{79}
$$

which follows from differentiating (51) with respect to T and using the first-order conditions (54) to (67), the co-state equations (72) to (75), (37), and the transversality condition $q_{e/a}(T) = 0$.

Consistent with diminishing returns to life extension (Ehrlich and Chuma, 1990), we assume

$$
\left.\frac{\partial \Im(T)}{\partial T}\right|_{Z_0} < 0,
\tag{80}
$$

in which case we can identify the sign of the variation in life expectancy from

$$\text{sign}\left(\frac{\partial T}{\partial Z_0}\right) = \text{sign}\left(\left.\frac{\partial \Im(T)}{\partial Z_0}\right|_T\right), \tag{81}$$

where,

$$\left.\frac{\partial \Im(T)}{\partial Z_0}\right|_T = q_A(T)\left.\frac{\partial E(T)}{\partial Z_0}\right|_T + \left.\frac{\partial q_A(T)}{\partial Z_0}\right|_T \left.\frac{\partial A(t)}{\partial t}\right|_{T,t=T} + \left.\frac{\partial q_H(T)}{\partial Z_0}\right|_T \left.\frac{\partial H(t)}{\partial t}\right|_{T,t=T}. \tag{82}$$

As (81) shows, we can explore variation in initial conditions keeping length of life T initially fixed in order to investigate whether life would be extended as a result of such variation.

Comparative dynamics of variation in length of life: $\partial g(t)/\partial T|_{Z_0}$ The derivatives of the control functions, state function and co-state functions with respect to length of life T, holding constant Z_0, are identical for any initial condition or model parameter Z_0. We therefore first obtain their derivatives (using 60 to 71). The symbol $\gtreqless 0$ is used to indicate that the sign cannot unambiguously be determined.

$$\begin{aligned}
\left.\frac{\partial q_{e/a}(t)}{\partial T}\right|_{Z_0} &= e^{-\int_t^T [d_E(x)+r]dx} > 0,\\
\left.\frac{\partial q_{h/a}(t)}{\partial T}\right|_{Z_0} &= \frac{-\left.\frac{\partial H(t)}{\partial t}\right|_{t=T} e^{\int_0^t [2d_H(x)+r]dx} e^{\int_t^T d_H(x)dx}}{\frac{\gamma_H}{1-\gamma_H}\int_0^T \mu_H(s) q_{h/a}(s)^{\frac{2\gamma_H-1}{1-\gamma_H}} e^{\int_0^s [2d_H(x)+r]dx} ds} > 0,\\
\left.\frac{\partial E(t)}{\partial T}\right|_{Z_0} &= \frac{\gamma_E}{1-\gamma_E}\int_0^t \mu_E(s) q_{e/a}(s)^{\frac{2\gamma_E-1}{1-\gamma_E}} \left.\frac{\partial q_{e/a}(s)}{\partial T}\right|_{Z_0} e^{-\int_s^t d_E(x)dx} ds > 0,\\
\left.\frac{\partial H(t)}{\partial T}\right|_{Z_0} &= \frac{\gamma_H}{1-\gamma_H}\int_0^t \mu_H(s) q_{h/a}(s)^{\frac{2\gamma_H-1}{1-\gamma_H}} \left.\frac{\partial q_{h/a}(s)}{\partial T}\right|_{Z_0} e^{-\int_s^t d_H(x)dx} ds > 0,\\
\left.\frac{\partial X_E(t)}{\partial T}\right|_{Z_0} &= \frac{1}{1-\gamma_E}\frac{\beta_E}{p_E(t)}\mu_E(t) q_{e/a}(t)^{\frac{\gamma_E}{1-\gamma_E}} \left.\frac{\partial q_{e/a}(t)}{\partial T}\right|_{Z_0} > 0,\\
\left.\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial T}\right|_{Z_0} &= \frac{1}{1-\gamma_E}\alpha_E \mu_E(t) q_{e/a}(t)^{\frac{\gamma_E}{1-\gamma_E}} \left.\frac{\partial q_{e/a}(t)}{\partial T}\right|_{Z_0} > 0,\\
\left.\frac{\partial X_H(t)}{\partial T}\right|_{Z_0} &= \frac{1}{1-\gamma_H}\frac{\beta_H}{p_H(t)}\mu_H(t) q_{h/a}(t)^{\frac{\gamma_H}{1-\gamma_H}} \left.\frac{\partial q_{h/a}(t)}{\partial T}\right|_{Z_0} > 0,\\
\left.\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial T}\right|_{Z_0} &= \frac{1}{1-\gamma_H}\alpha_H \mu_H(t) q_{h/a}(t)^{\frac{\gamma_H}{1-\gamma_H}} \left.\frac{\partial q_{h/a}(t)}{\partial T}\right|_{Z_0} > 0,
\end{aligned}$$

$$\left.\frac{\partial q_A(0)}{\partial T}\right|_{Z_0} = \frac{-\left.\frac{\partial A(t)}{\partial t}\right|_{Z_0,t=T} e^{-rT} - \int_0^T \left.\frac{\partial \phi(s)}{\partial T}\right|_{Z_0} ds}{\frac{\Lambda}{\rho} q_A(0)^{-\left(\frac{1+\rho}{\rho}\right)} \int_0^T p_C(s)^{-\frac{\zeta(1-\rho)}{\rho}} e^{-\frac{(\beta-r(1-\rho))}{\rho}s} ds} \gtreqless 0,$$

$$\left.\frac{\partial A(t)}{\partial T}\right|_{Z_0} = e^{rt} \int_0^t \left.\frac{\partial \phi(s)}{\partial T}\right|_{Z_0} + \left[e^{rt} \frac{\Lambda}{\rho} q_A(0)^{-\frac{(1+\rho)}{\rho}} \int_0^t p_C(s)^{-\frac{\zeta(1-\rho)}{\rho}} e^{-\frac{(\beta-r(1-\rho))}{\rho}s} ds\right] \left.\frac{\partial q_A(0)}{\partial T}\right|_{Z_0} \gtreqless 0,$$

where

$$\begin{aligned}\left.\frac{\partial \phi(s)}{\partial T}\right|_{Z_0} \equiv\ & e^{-rs}\left[\left.\frac{\partial E(s)}{\partial T}\right|_{Z_0} + \left.\frac{\partial H(s)}{\partial T}\right|_{Z_0}\right] \\ & - \frac{\gamma_E}{1-\gamma_E} e^{-rs} \mu_E(s) q_{e/a}(s)^{\frac{\gamma_E}{1-\gamma_E}} \left.\frac{\partial q_{e/a}(s)}{\partial T}\right|_{Z_0} \\ & - \frac{\gamma_H}{1-\gamma_H} e^{-rs} \mu_H(s) q_{h/a}(s)^{\frac{\gamma_H}{1-\gamma_H}} \left.\frac{\partial q_{h/a}(s)}{\partial T}\right|_{Z_0},\end{aligned}$$

$$\begin{aligned}\left.\frac{\partial X_C(t)}{\partial T}\right|_{Z_0} &= -\frac{\zeta\Lambda}{\rho} q_A(0)^{-\left(\frac{1+\rho}{\rho}\right)} p_C(t)^{-\left(\frac{\rho+\zeta-\zeta\rho}{\rho}\right)} e^{-\left(\frac{\beta-r}{\rho}\right)t} \left.\frac{\partial q_A(0)}{\partial T}\right|_{Z_0} \gtreqless 0, \\ \left.\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial T}\right|_{Z_0} &= -\frac{(1-\zeta)\Lambda}{\rho} q_A(0)^{-\left(\frac{1+\rho}{\rho}\right)} p_C(t)^{-\left(\frac{\zeta-\zeta\rho}{\rho}\right)} e^{-\left(\frac{\beta-r}{\rho}\right)t} \left.\frac{\partial q_A(0)}{\partial T}\right|_{Z_0} \gtreqless 0. \end{aligned} \tag{83}$$

Comparative dynamics of initial wealth $\partial g(t)/\partial A_0$: First consider the case where T is fixed. Differentiating (70) with respect to A_0, using (62), and differentiating (71) with respect to A_0, one finds $\partial q_{e/a}(0)/\partial A_0|_T = 0$ and $\partial q_{h/a}(0)/\partial A_0|_T = 0$. Using (54) to (64), and (69), we obtain

$$\begin{aligned}
\left.\frac{\partial q_{e/a}(t)}{\partial A_0}\right|_T &= 0,\ \forall t & \left.\frac{\partial q_{h/a}(t)}{\partial A_0}\right|_T &= 0,\ \forall t \\
\left.\frac{\partial E(t)}{\partial A_0}\right|_T &= 0,\ \forall t & \left.\frac{\partial H(t)}{\partial A_0}\right|_T &= 0,\ \forall t \\
\left.\frac{\partial X_E(t)}{\partial A_0}\right|_T &= 0,\ \forall t & \left.\frac{\partial X_H(t)}{\partial A_0}\right|_T &= 0,\ \forall t \\
\left.\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial A_0}\right|_T &= 0,\ \forall t & \left.\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial A_0}\right|_T &= 0,\ \forall t
\end{aligned}$$

$$
\begin{aligned}
\left.\frac{\partial A(t)}{\partial A_0}\right|_T &= e^{rt}\left[1-\frac{\int_0^t p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta-r(1-\rho))}{\rho}s} ds}{\int_0^T p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta-r(1-\rho))}{\rho}s} ds}\right] \geq 0, \\
\left.\frac{\partial q_A(0)}{\partial A_0}\right|_T &= \frac{-1}{\frac{\Lambda q_A(0)^{-\frac{(1+\rho)}{\rho}}}{\rho} \int_0^T p_C(s)^{-\frac{\zeta(1-\rho)}{\rho}} e^{-\frac{(\beta-r(1-\rho))}{\rho}s} ds} < 0, \\
\left.\frac{\partial X_C(t)}{\partial A_0}\right|_T &= -\frac{1}{\rho}\zeta\Lambda q_A(0)^{-\frac{(1+\rho)}{\rho}} p_C(t)^{-(1-\zeta+\zeta/\rho)} e^{-\frac{\beta-r}{\rho}t} \left.\frac{\partial q_A(0)}{\partial A_0}\right|_T \\
&= \frac{\zeta p_C(t)^{-(1-\zeta+\zeta/\rho)} e^{-\frac{\beta-r}{\rho}t}}{\int_0^T p_C(s)^{-(1-\zeta+\zeta/\rho)} e^{-\frac{(\beta-r(1-\rho))}{\rho}s} ds} > 0, \\
\left.\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial A_0}\right|_T &= -\frac{1}{\rho}\left(1-\zeta\right)\Lambda q_A(0)^{-\frac{(1+\rho)}{\rho}} p_C(t)^{-\zeta(1-\rho)/\rho} e^{-\frac{\beta-r}{\rho}t} \left.\frac{\partial q_A(0)}{\partial A_0}\right|_T \\
&= \frac{\left(1-\zeta\right) p_C(t)^{-\zeta(1-\rho)/\rho} e^{-\frac{\beta-r}{\rho}t}}{\int_0^T p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta-r(1-\rho))}{\rho}s} ds} > 0. \qquad (84)
\end{aligned}
$$

Note that the relation for the variation in wealth has the desired properties $\partial A(0)/\partial A_0|_T = 1$, and $\partial A(T)/\partial A_0|_T = 0$.

Now allow length of life T to be optimally chosen. Using (78) we have

$$
\frac{\partial T}{\partial A_0} = \frac{\left.\frac{\partial q_A(0)}{\partial A_0}\right|_T e^{-rT}\left[\left.\frac{\partial A(t)}{\partial t}\right|_{t=T} + q_{h/a}(T)\left.\frac{\partial H(t)}{\partial t}\right|_{t=T}\right]}{-\partial\Im(T)/\partial T|_{A_0}} > 0, \qquad (85)
$$

where we have used $\partial E(T)/\partial A_0|_T = 0$ (see 61 and note that $\partial q_{e/a}(t)/\partial A_0\big|_T = 0,\ \forall t$), $\partial q_H(T)/\partial A_0|_T = q_{h/a}(T)\ \partial q_A(T)/\partial A_0|_T$ (since $\partial q_{h/a}(T)/\partial A_0\big|_T = 0$), $\partial H(t)/\partial t|_{t=T} < 0$ by definition as health approaches $H_{\min}$ from above, $\partial A(t)/\partial t|_{t=T} < 0$ as individuals draw from their savings in old age, and $-\partial\Im(T)/\partial T|_{A_0} > 0$ (see 80).

Using (36), we obtain the following total responses to variation in wealth

$$
\begin{aligned}
\frac{\partial q_{e/a}(t)}{\partial A_0} &= \left.\frac{\partial q_{e/a}(t)}{\partial T}\right|_{A_0}\frac{\partial T}{\partial A_0} > 0, &\qquad \frac{\partial q_{h/a}(t)}{\partial A_0} &= \left.\frac{\partial q_{h/a}(t)}{\partial T}\right|_{A_0}\frac{\partial T}{\partial A_0} > 0, \\
\frac{\partial E(t)}{\partial A_0} &= \left.\frac{\partial E(t)}{\partial T}\right|_{A_0}\frac{\partial T}{\partial A_0} > 0, &\qquad \frac{\partial H(t)}{\partial A_0} &= \left.\frac{\partial H(t)}{\partial T}\right|_{A_0}\frac{\partial T}{\partial A_0} > 0, \\
\frac{\partial X_E(t)}{\partial A_0} &= \left.\frac{\partial X_E(t)}{\partial T}\right|_{A_0}\frac{\partial T}{\partial A_0} > 0, &\qquad \frac{\partial X_H(t)}{\partial A_0} &= \left.\frac{\partial X_H(t)}{\partial T}\right|_{A_0}\frac{\partial T}{\partial A_0} > 0,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial A_0} &= \left.\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial T}\right|_{A_0}\frac{\partial T}{\partial A_0} > 0, \\
\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial A_0} &= \left.\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial T}\right|_{A_0}\frac{\partial T}{\partial A_0} > 0,
\end{aligned}
$$

$$\frac{\partial A(t)}{\partial A_0} = \left.\frac{\partial A(t)}{\partial A_0}\right|_T + \left.\frac{\partial A(t)}{\partial T}\right|_{A_0} \frac{\partial T}{\partial A_0} \gtreqless 0,$$
$$\frac{\partial q_A(0)}{\partial A_0} = \left.\frac{\partial q_A(0)}{\partial A_0}\right|_T + \left.\frac{\partial q_A(0)}{\partial T}\right|_{A_0} \frac{\partial T}{\partial A_0} \gtreqless 0,$$

$$\begin{aligned}
\frac{\partial X_C(t)}{\partial A_0} &= -\left(\frac{1}{\rho}\zeta\Lambda q_A(0)^{-\frac{(1+\rho)}{\rho}} p_C(t)^{-(1-\zeta+\zeta/\rho)} e^{-\frac{\beta-r}{\rho}t}\right) \times \\
&\quad \left[\left.\frac{\partial q_A(0)}{\partial A_0}\right|_T + \left.\frac{\partial q_A(0)}{\partial T}\right|_{A_0} \frac{\partial T}{\partial A_0}\right] \gtreqless 0, \\
\frac{\partial L(t)\,[E(t)+H(t)]}{\partial A_0} &= -\left(\frac{1}{\rho}\,(1-\zeta)\,\Lambda q_A(0)^{-\frac{(1+\rho)}{\rho}} p_C(t)^{-\zeta(1-\rho)/\rho} e^{-\frac{\beta-r}{\rho}t}\right) \times \\
&\quad \left[\left.\frac{\partial q_A(0)}{\partial A_0}\right|_T + \left.\frac{\partial q_A(0)}{\partial T}\right|_{A_0} \frac{\partial T}{\partial A_0}\right] \gtreqless 0,
\end{aligned} \tag{86}$$

where we have used (83). Note that the total response of $q_A(0)$ with respect to initial wealth A_0 is ambiguous, since the additional wealth has to be spread over more time periods ($\partial T/\partial A_0 > 0$). But, a longer horizon also increases the returns to skill investment and to health investment, increasing the stocks, earnings and permanent income (lowering the marginal value of wealth $q_A(0)$). Hence, the effect of initial wealth on $q_A(0)$ and thereby on consumption and leisure is ambiguous for free T. Since wealthy individuals are generally found to consume more and retire earlier (e.g., Imbens, Rubin and Sacerdote, 2001; Juster et al. 2006; Brown, Coile, and Weisbenner, 2010), it is plausible that the wealth effect dominates $\partial q_A(0)/\partial A_0 < 0$, and consumption goods and services $X_C(t)$ and effective leisure $L(t)\,[E(t)+H(t)]$ are higher throughout life.

Comparative dynamics of initial skill $\partial g(t)/\partial E_0$: Again, first consider the case where T is fixed. Differentiating (71) with respect to E_0, one finds $\partial q_{e/a}(0)/\partial E_0\big|_T = 0$ and differentiating (70) with respect to E_0, using (62) we find $\partial q_{h/a}(0)/\partial E_0\big|_T = 0$. Using (54) to (64), and (69), we find

$$\left.\frac{\partial q_{e/a}(t)}{\partial E_0}\right|_T = 0,\ \forall t \qquad \left.\frac{\partial q_{h/a}(t)}{\partial E_0}\right|_T = 0,\ \forall t$$
$$\left.\frac{\partial E(t)}{\partial E_0}\right|_T = e^{-\int_0^t d_E(x)dx} > 0,\ \forall t \qquad \left.\frac{\partial H(t)}{\partial E_0}\right|_T = 0,\ \forall t$$
$$\left.\frac{\partial X_E(t)}{\partial E_0}\right|_T = 0,\ \forall t \qquad \left.\frac{\partial X_H(t)}{\partial E_0}\right|_T = 0,\ \forall t$$
$$\left.\frac{\partial \tau_E(t)\,[E(t)+H(t)]}{\partial E_0}\right|_T = 0,\ \forall t \qquad \left.\frac{\partial \tau_H(t)\,[E(t)+H(t)]}{\partial E_0}\right|_T = 0,\ \forall t$$

$$
\begin{aligned}
\left.\frac{\partial A(t)}{\partial E_0}\right|_T &= e^{rt}\left[\int_0^T e^{-\int_0^s [d_E(x)+r]dx} ds\right] \\
&\quad \times \left[\frac{\int_0^t e^{-\int_0^s [d_E(x)+r]dx} ds}{\int_0^T e^{-\int_0^s [d_E(x)+r]dx} ds} - \frac{\int_0^t p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta - r(1-\rho))}{\rho}s} ds}{\int_0^T p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta - r(1-\rho))}{\rho}s} ds}\right] \gtreqqless 0, \\
\left.\frac{\partial q_A(0)}{\partial E_0}\right|_T &= \frac{-\int_0^T e^{-\int_0^s [d_E(x)+r]dx} ds}{\frac{\Lambda}{\rho} q_A(0)^{-\left(\frac{1+\rho}{\rho}\right)} \int_0^T p_C(s)^{-\frac{\zeta(1-\rho)}{\rho}} e^{-\frac{(\beta - r(1-\rho))}{\rho}s} ds} < 0, \\
\left.\frac{\partial X_C(t)}{\partial E_0}\right|_T &= -\zeta\Lambda \frac{q_A(0)^{-\frac{(1+\rho)}{\rho}}}{\rho} p_C(t)^{-(1-\zeta+\zeta/\rho)} e^{-\frac{(\beta-r)}{\rho}t} \left.\frac{\partial q_A(0)}{\partial E_0}\right|_T \\
&= \frac{\zeta p_C(t)^{-(1-\zeta+\zeta/\rho)} e^{-\frac{(\beta-r)}{\rho}t} \int_0^T e^{-\int_0^s [d_E(x)+r]dx} ds}{\int_0^T p_C(s)^{-\frac{\zeta(1-\rho)}{\rho}} e^{-\frac{(\beta - r(1-\rho))}{\rho}s} ds} > 0, \\
\left.\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial E_0}\right|_T &= -(1-\zeta)\Lambda \frac{q_A(0)^{-\frac{(1+\rho)}{\rho}}}{\rho} p_C(t)^{-\frac{\zeta(1-\rho)}{\rho}} e^{-\frac{(\beta-r)}{\rho}t} \left.\frac{\partial q_A(0)}{\partial E_0}\right|_T \\
&= \frac{(1-\zeta) p_C(t)^{-\frac{\zeta(1-\rho)}{\rho}} e^{-\frac{(\beta-r)}{\rho}t} \int_0^T e^{-\int_0^t [d_E(x)+r]dx} dt}{\int_0^T p_C(s)^{-\frac{\zeta(1-\rho)}{\rho}} e^{-\frac{(\beta - r(1-\rho))}{\rho}s} ds} > 0. \qquad (87)
\end{aligned}
$$

Note that the relation for the variation in wealth has the desired properties $\partial A(0)/\partial E_0|_T = 0$, and $\partial A(T)/\partial E_0|_T = 0$. Further, the wealth effect of additional skill capital δE_0 is proportional to the effect we derived earlier of an additional amount of wealth δA_0,

$$
\begin{aligned}
\left.\frac{\partial q_A(0)}{\partial E_0}\right|_T &= \left\{\int_0^T e^{-\int_0^s [d_E(x)+r]dx} ds\right\} \left.\frac{\partial q_A(0)}{\partial A_0}\right|_T, \\
\left.\frac{\partial X_C(t)}{\partial E_0}\right|_T &= \left\{\int_0^T e^{-\int_0^s [d_E(x)+r]dx} ds\right\} \left.\frac{\partial X_C(t)}{\partial A_0}\right|_T, \\
\left.\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial E_0}\right|_T &= \left\{\int_0^T e^{-\int_0^s [d_E(x)+r]dx} ds\right\} \left.\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial A_0}\right|_T. \qquad (88)
\end{aligned}
$$

Note further, that

$$
\begin{aligned}
\left.\frac{\partial f_E[.]}{\partial E_0}\right|_T &= 0, \\
\left.\frac{\partial f_H[.]}{\partial E_0}\right|_T &= 0, \\
\left.\frac{\partial Y[.]}{\partial E_0}\right|_T &= \left.\frac{\partial E(t)}{\partial E_0}\right|_T. \qquad (89)
\end{aligned}
$$

Now allow length of life T to be optimally chosen. Using (78) we have

$$
\begin{aligned}
\frac{\partial T}{\partial E_0} &= \frac{q_A(0)e^{-\int_0^T [d_E(x)+r]dx} + \left.\frac{\partial q_A(0)}{\partial E_0}\right|_T e^{-rT}\left[\left.\frac{\partial A(t)}{\partial t}\right|_{t=T} + q_{h/a}(T)\left.\frac{\partial H(t)}{\partial t}\right|_{t=T}\right]}{-\left.\frac{\partial \Im(T)}{\partial T}\right|_T} \\
&= \frac{q_A(0)e^{-\int_0^T [d_E(x)+r]dx}}{-\left.\frac{\partial \Im(T)}{\partial T}\right|_T} + \left\{\int_0^T e^{-\int_0^s [d_E(x)+r]dx} ds\right\}\frac{\partial T}{\partial A_0} > 0. \qquad (90)
\end{aligned}
$$

Using (36), we obtain the following total responses to variation in skill capital

$$
\begin{aligned}
\frac{\partial q_{e/a}(t)}{\partial E_0} &= \left.\frac{\partial q_{e/a}(t)}{\partial T}\right|_{E_0}\frac{\partial T}{\partial E_0} > 0, &\qquad \frac{\partial q_{h/a}(t)}{\partial E_0} &= \left.\frac{\partial q_{h/a}(t)}{\partial T}\right|_{E_0}\frac{\partial T}{\partial E_0} > 0, \\
\frac{\partial E(t)}{\partial E_0} &= \left.\frac{\partial E(t)}{\partial E_0}\right|_T + \left.\frac{\partial E(t)}{\partial T}\right|_{E_0}\frac{\partial T}{\partial E_0} > 0, &\qquad \frac{\partial H(t)}{\partial E_0} &= \left.\frac{\partial H(t)}{\partial T}\right|_{E_0}\frac{\partial T}{\partial E_0} > 0, \\
\frac{\partial X_E(t)}{\partial E_0} &= \left.\frac{\partial X_E(t)}{\partial T}\right|_{E_0}\frac{\partial T}{\partial E_0} > 0, &\qquad \frac{\partial X_H(t)}{\partial E_0} &= \left.\frac{\partial X_H(t)}{\partial T}\right|_{E_0}\frac{\partial T}{\partial E_0} > 0,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial E_0} &= \left.\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial T}\right|_{E_0}\frac{\partial T}{\partial E_0} > 0, \\
\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial E_0} &= \left.\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial T}\right|_{E_0}\frac{\partial T}{\partial E_0} > 0,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial A(t)}{\partial E_0} &= \left.\frac{\partial A(t)}{\partial E_0}\right|_T + \left.\frac{\partial A(t)}{\partial T}\right|_{E_0}\frac{\partial T}{\partial E_0} \gtreqless 0, \\
\frac{\partial q_A(0)}{\partial E_0} &= \left.\frac{\partial q_A(0)}{\partial E_0}\right|_T + \left.\frac{\partial q_A(0)}{\partial T}\right|_{E_0}\frac{\partial T}{\partial E_0} \gtreqless 0,
\end{aligned}
$$

$$
\frac{\partial X_C(t)}{\partial E_0} = -\frac{\zeta}{\rho} q_A(0)^{-\left(\frac{1+\rho}{\rho}\right)} \Lambda p_C(t)^{-(1-\zeta+\zeta/\rho)} e^{-\left(\frac{\beta-r}{\rho}\right)t}\left[\left.\frac{\partial q_A(0)}{\partial E_0}\right|_T + \left.\frac{\partial q_A(0)}{\partial T}\right|_{E_0}\frac{\partial T}{\partial E_0}\right] \gtreqless 0,
$$

$$
\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial E_0} = \left.\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial E_0}\right|_T + \left.\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial T}\right|_{E_0}\frac{\partial T}{\partial E_0} \gtreqless 0, \qquad (91)
$$

where we have used (83).

Comparative dynamics of initial health $\partial g(t)/\partial H_0$: Again, first consider the case where T is fixed. Differentiating (70) with respect to H_0, using (62), and differentiating (71) with respect to H_0, one finds

$$
\begin{aligned}
\left.\frac{\partial q_{e/a}(0)}{\partial H_0}\right|_T &= 0, \\
\left.\frac{\partial q_{h/a}(0)}{\partial H_0}\right|_T &= \frac{-1}{\frac{\gamma_H}{1-\gamma_H}\int_0^T \mu_H(s) q_{h/a}(s)^{\frac{2\gamma_H-1}{1-\gamma_H}} e^{\int_0^s [2d_H(x)+r]dx} ds} < 0.
\end{aligned} \tag{92}
$$

Using (54) to (70), and (92), we obtain

$$
\begin{aligned}
\left.\frac{\partial q_{e/a}(t)}{\partial H_0}\right|_T &= 0, \ \forall t \\
\left.\frac{\partial q_{h/a}(t)}{\partial H_0}\right|_T &= \left.\frac{\partial q_{h/a}(0)}{\partial H_0}\right|_T e^{\int_0^t [d_H(x)+r]dx} < 0, \\
\left.\frac{\partial E(t)}{\partial H_0}\right|_T &= 0, \ \forall t \\
\left.\frac{\partial H(t)}{\partial H_0}\right|_T &= e^{-\int_0^t d_H(x)dx}\left[1 - \frac{\int_0^t \mu_H(s) q_{h/a}(s)^{\frac{2\gamma_H-1}{1-\gamma_H}} e^{\int_0^s [2d_H(x)+r]dx} ds}{\int_0^T \mu_H(s) q_{h/a}(s)^{\frac{2\gamma_H-1}{1-\gamma_H}} e^{\int_0^s [2d_H(x)+r]dx} ds}\right] \geq 0, \\
\left.\frac{\partial X_E(t)}{\partial H_0}\right|_T &= 0, \ \forall t \\
\left.\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial H_0}\right|_T &= 0, \ \forall t \\
\left.\frac{\partial X_H(t)}{\partial H_0}\right|_T &= \frac{X_H(t)}{1-\gamma_H}\frac{\left.\frac{\partial q_{h/a}(t)}{\partial H_0}\right|_T}{q_{h/a}(t)} < 0,
\end{aligned}
$$

$$
\begin{aligned}
\left.\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial H_0}\right|_T &= \frac{\tau_H(t)\left[E(t)+H(t)\right]}{1-\gamma_H}\frac{\left.\frac{\partial q_{h/a}(t)}{\partial H_0}\right|_T}{q_{h/a}(t)} < 0, \\
\left.\frac{\partial A(t)}{\partial H_0}\right|_T &= e^{rt}\left[\int_0^T \epsilon\left[H(s), q_{h/a}(s)\right] ds\right] \\
&\quad \times \left[\frac{\int_0^t \epsilon\left[H(s), q_{h/a}(s)\right] ds}{\int_0^T \epsilon\left[H(s), q_{h/a}(s)\right] ds} - \frac{\int_0^t p_C(s)^{-\zeta(1-\rho)/\zeta} e^{-\frac{(\beta-r(1-\rho))}{\rho}s} ds}{\int_0^T p_C(s)^{-\zeta(1-\rho)/\zeta} e^{-\frac{(\beta-r(1-\rho))}{\rho}s} ds}\right] \gtreqless 0, \\
\left.\frac{\partial q_A(0)}{\partial H_0}\right|_T &= \frac{-\int_0^T \epsilon\left[H(s), q_{h/a}(s)\right] ds}{q_A(0)^{\frac{-(1+\rho)}{\rho}} \frac{\Lambda}{\rho}\int_0^T p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta-r(1-\rho))}{\rho}s} ds} \\
&= \left\{\int_0^T \epsilon\left[H(s), q_{h/a}(s)\right] ds\right\} \left.\frac{\partial q_A(0)}{\partial A_0}\right|_T < 0,
\end{aligned} \tag{93}
$$

where

$$
\epsilon\left[H(s), q_{h/a}(s)\right] = \left.\frac{\partial H(s)}{\partial H_0}\right|_T - \frac{\gamma_H}{1-\gamma_H}\mu_H(s) q_{h/a}(s)^{\frac{\gamma_H}{1-\gamma_H}} \left.\frac{\partial q_{h/a}(s)}{\partial H_0}\right|_T e^{-rs} > 0, \ \forall s
$$

and we have used $\partial H(s)/\partial H_0|_T > 0$ and $\partial q_{h/a}(s)/\partial H_0\big|_T < 0$ (see 93).

Further using (66) and (67) it follows that

$$
\begin{aligned}
\left.\frac{\partial X_C(t)}{\partial H_0}\right|_T &= -\frac{1}{\rho}\zeta\Lambda q_A(0)^{-\frac{(1+\rho)}{\rho}} p_C(t)^{-(1-\zeta+\zeta/\rho)} e^{-\frac{\beta-r}{\rho}t} \left.\frac{\partial q_A(0)}{\partial H_0}\right|_T > 0,\\
\left.\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial H_0}\right|_T &= -\frac{1}{\rho}(1-\zeta)\Lambda q_A(0)^{-\frac{(1+\rho)}{\rho}} p_C(t)^{-\zeta(1-\rho)/\rho} e^{-\frac{\beta-r}{\rho}t} \left.\frac{\partial q_A(0)}{\partial H_0}\right|_T > 0. \qquad (94)
\end{aligned}
$$

Note that the relation for the variation in the health stock has the desired properties $\partial H(0)/\partial H_0|_T = 1$, and $\partial H(T)/\partial H_0|_T = 0$, and the relation for the variation in wealth has the desired properties $\partial A(0)/\partial H_0|_T = 0$, and $\partial A(T)/\partial H_0|_T = 0$. Also note that

$$
\begin{aligned}
\left.\frac{\partial f_E[.]}{\partial H_0}\right|_T &= 0,\\
\left.\frac{\partial f_H[.]}{\partial H_0}\right|_T &= \frac{\gamma_H}{1-\gamma_H}\frac{f_H[.]}{q_{h/a}(t)}\left.\frac{\partial q_{h/a}(t)}{\partial H_0}\right|_T < 0, \qquad (95)
\end{aligned}
$$

so that the additional productivity $f_E[.]$ from greater health, $\partial H(t)/\partial H_0|_T > 0$, is exactly offset by the reduction in time inputs, $\partial \tau_E(t)/\partial H_0|_T < 0$, and, the additional productivity $f_H[.]$ from greater health, $\partial H(t)/\partial H_0|_T > 0$, is more than offset, $\partial f_H[.]/\partial E_0|_T < 0$, in order to ensure that length of life remains of the same duration (we assumed fixed T).

Now allow length of life T to be optimally chosen. Using (78) we have

$$
\begin{aligned}
\frac{\partial T}{\partial H_0} &= \frac{\left.\frac{\partial q_A(0)}{\partial H_0}\right|_T e^{-rT}\left.\frac{\partial A(t)}{\partial t}\right|_{t=T} + \left[q_{h/a}(T)\left.\frac{\partial q_A(0)}{\partial H_0}\right|_T e^{-rT} + q_A(T)\left.\frac{\partial q_{h/a}(T)}{\partial H_0}\right|_T\right]\left.\frac{\partial H(t)}{\partial t}\right|_{t=T}}{-\left.\frac{\partial \Im(T)}{\partial T}\right|_{H_0}}\\
&= \frac{q_A(T)\left.\frac{\partial q_{h/a}(T)}{\partial H_0}\right|_T\left.\frac{\partial H(t)}{\partial t}\right|_{t=T}}{-\left.\frac{\partial \Im(T)}{\partial T}\right|_{H_0}} + \left\{\int_0^T \epsilon\left[H(s), q_{h/a}(s)\right]ds\right\}\frac{\partial T}{\partial A_0} > 0. \qquad (96)
\end{aligned}
$$

Using (36), we obtain the following total responses to variation in skill capital

$$
\begin{aligned}
\frac{\partial q_{e/a}(t)}{\partial H_0} &= \left.\frac{\partial q_{e/a}(t)}{\partial T}\right|_{H_0}\frac{\partial T}{\partial H_0} > 0, &
\frac{\partial q_{h/a}(t)}{\partial H_0} &= \left.\frac{\partial q_{h/a}(t)}{\partial H_0}\right|_T + \left.\frac{\partial q_{h/a}(t)}{\partial T}\right|_{H_0}\frac{\partial T}{\partial H_0} \gtreqless 0,\\
\frac{\partial E(t)}{\partial H_0} &= \left.\frac{\partial E(t)}{\partial T}\right|_{H_0}\frac{\partial T}{\partial H_0} > 0, &
\frac{\partial H(t)}{\partial H_0} &= \left.\frac{\partial H(t)}{\partial H_0}\right|_T + \left.\frac{\partial H(t)}{\partial T}\right|_{H_0}\frac{\partial T}{\partial H_0} > 0,\\
\frac{\partial X_E(t)}{\partial H_0} &= \left.\frac{\partial X_E(t)}{\partial T}\right|_{H_0}\frac{\partial T}{\partial H_0} > 0, &
\frac{\partial X_H(t)}{\partial H_0} &= \left.\frac{\partial X_H(t)}{\partial H_0}\right|_T + \left.\frac{\partial X_H(t)}{\partial T}\right|_{H_0}\frac{\partial T}{\partial H_0} \gtreqless 0,
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial H_0} &= \left.\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial T}\right|_{H_0}\frac{\partial T}{\partial H_0} > 0,\\
\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial H_0} &= \left.\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial H_0}\right|_T + \left.\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial T}\right|_{H_0}\frac{\partial T}{\partial H_0} \gtreqless 0,
\end{aligned}
$$

$$\frac{\partial A(t)}{\partial H_0} = \left.\frac{\partial A(t)}{\partial H_0}\right|_T + \left.\frac{\partial A(t)}{\partial T}\right|_{H_0} \frac{\partial T}{\partial H_0} \gtreqless 0,$$

$$\frac{\partial q_A(0)}{\partial H_0} = \left.\frac{\partial q_A(0)}{\partial H_0}\right|_T + \left.\frac{\partial q_A(0)}{\partial T}\right|_{H_0} \frac{\partial T}{\partial H_0} \gtreqless 0,$$

$$\frac{\partial X_C(t)}{\partial H_0} = -\frac{\zeta}{\rho} q_A(0)^{-\left(\frac{1+\rho}{\rho}\right)} \Lambda p_C(t)^{-(1-\zeta+\zeta/\rho)} e^{-\left(\frac{\beta-r}{\rho}\right)t} \left[\left.\frac{\partial q_A(0)}{\partial H_0}\right|_T + \left.\frac{\partial q_A(0)}{\partial T}\right|_{H_0} \frac{\partial T}{\partial H_0} \right] \gtreqless 0,$$

$$\frac{\partial L(t)\,[E(t)+H(t)]}{\partial H_0} = \left.\frac{\partial L(t)\,[E(t)+H(t)]}{\partial H_0}\right|_T + \left.\frac{\partial L(t)\,[E(t)+H(t)]}{\partial T}\right|_{H_0} \frac{\partial T}{\partial H_0} \gtreqless 0, \tag{97}$$

where we have used (83).

Skill productivity The comparative dynamics for the skill productivity factor $\mu_E(x)$ and the generalized health productivity factor $\mu_H(x)$ are summarized in Table 2.

Table 2: Comparative dynamic effects of the generalized skill productivity factor $\mu_E(t)$ and the generalized health productivity factor $\mu_H(t)$ on the state and co-state functions, control functions and the parameter T.

	$\mu_E(t)$		$\mu_H(t)$	
Function	T fixed	T free	T fixed	T free
$E(t)$	> 0	> 0	0	> 0
$q_{e/a}(t)$	0	> 0	0	> 0
$X_E(t)$	> 0	> 0	0	> 0
$\tau_E(t)\,[E(t)+H(t)]$	> 0	> 0	0	> 0
$H(t)$	0	> 0	≥ 0	> 0
$q_{h/a}(t)$	0	> 0	< 0	+/-
$X_H(t)$	0	> 0	+/-	+/-
$\tau_H(t)\,[E(t)+H(t)]$	0	> 0	+/-	+/-
$A(t)$	+/-	+/-	+/-	+/-
$q_A(0)$	< 0	$< 0^\dagger$	< 0	$< 0^\dagger$
$X_C(t)$	> 0	$> 0^\dagger$	> 0	$> 0^\dagger$
$L(t)\,[E(t)+H(t)]$	> 0	$> 0^\dagger$	> 0	$> 0^\dagger$
T	n/a	> 0	n/a	> 0

Notes: 0 is used to denote 'not affected', $+/-$ is used to denote that the sign is 'undetermined', n/a stands for 'not applicable', and † is used to denote that the 'sign holds under the plausible assumption that the wealth effect dominates the effect of life extension'. This is consistent with the empirical finding (Imbens, Rubin and Sacerdote 2001; Juster et al. 2006; Brown, Coile, and Weisbenner, 2010) that additional wealth leads to higher consumption, even though the horizon over which consumption takes place is extended.

Consider the case where T is fixed. Differentiating (70) with respect to $\mu_E(x)$, using (62), and differentiating (71) with respect to $\mu_E(x)$, one finds

$$\begin{aligned}
\left.\frac{\partial q_{e/a}(0)}{\partial \mu_E(x)}\right|_T &= 0, \\
\left.\frac{\partial q_{h/a}(0)}{\partial \mu_E(x)}\right|_T &= 0. \qquad (98)
\end{aligned}$$

Using (54) to (63) we obtain

$$\begin{aligned}
\left.\frac{\partial q_{e/a}(t)}{\partial \mu_E(x)}\right|_T &= 0, \ \forall t \\
\left.\frac{\partial q_{h/a}(t)}{\partial \mu_E(x)}\right|_T &= 0, \ \forall t \\
\left.\frac{\partial E(t)}{\partial \mu_E(x)}\right|_T &= q_{e/a}(x)^{\frac{\gamma_E}{1-\gamma_E}} e^{-\int_x^t d_E(u)du} > 0, \ \text{for } t \geq x, \\
\left.\frac{\partial H(t)}{\partial \mu_E(x)}\right|_T &= 0, \ \forall t \\
\left.\frac{\partial X_E(t)}{\partial \mu_E(x)}\right|_T &= \frac{\beta_E}{p_E(t)} q_{e/a}(t)^{\frac{1}{1-\gamma_E}} \delta(t-x) > 0, \ \forall t \\
\left.\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial \mu_E(x)}\right|_T &= \alpha_E q_{e/a}(t)^{\frac{1}{1-\gamma_E}} \delta(t-x) > 0, \ \forall t \\
\left.\frac{\partial X_H(t)}{\partial \mu_E(x)}\right|_T &= 0, \ \forall t \\
\left.\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial \mu_E(x)}\right|_T &= 0, \ \forall t \qquad (99)
\end{aligned}$$

where $\delta(t-x)$ is the Dirac Delta function, which is zero everywhere except for $t = x$ and has a total area of 1 (it is the continuous-time equivalent of the discrete Kronecker delta function).

Differentiating (69) with respect to $\mu_E(x)$, we have (for $t \geq x$)

$$\begin{aligned}
\left.\frac{\partial A(t)}{\partial \mu_E(x)}\right|_T &= e^{rt} q_{e/a}(x)^{\frac{\gamma_E}{1-\gamma_E}} \left\{ \left[\int_0^t e^{-rs} e^{-\int_x^s d_E(u)du} ds - \gamma_E q_{e/a}(x) e^{-rx} \right] - \right. \\
& \left. \left[\int_0^T e^{-rs} e^{-\int_x^s d_E(u)du} ds - \gamma_E q_{e/a}(x) e^{-rx} \right] \frac{\int_0^t p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta - r(1-\rho))}{\rho} s} ds}{\int_0^T p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta - r(1-\rho))}{\rho} s} ds} \right\} \gtreqless 0, \qquad (100) \\
\left.\frac{\partial q_A(0)}{\partial \mu_E(x)}\right|_T &= \frac{\gamma_E q_{e/a}(x)^{\frac{1}{1-\gamma_E}} e^{-rx} - q_{e/a}(x)^{\frac{\gamma_E}{1-\gamma_E}} \int_0^T e^{-rs} e^{\int_x^s d_E(u)du} ds}{q_A(0)^{\frac{-(1+\rho)}{\rho}} \frac{\Lambda}{\rho} \int_0^T p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta - r(1-\rho))}{\rho} s} ds} < 0, \qquad (101)
\end{aligned}$$

where, in signing the term, we have assumed that the transient effect (first term in the numerator) is dominated by the permanent effect (second term in the numerator).

Further using (66) and (67) it follows that

$$\left.\frac{\partial X_C(t)}{\partial \mu_E(x)}\right|_T = -\frac{1}{\rho}\zeta\Lambda q_A(0)^{-\frac{(1+\rho)}{\rho}} p_C(t)^{-(1-\zeta+\zeta/\rho)} e^{-\frac{\beta-r}{\rho}t} \left.\frac{\partial q_A(0)}{\partial \mu_E(x)}\right|_T > 0, \tag{102}$$
$$\left.\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial \mu_E(x)}\right|_T = -\frac{1}{\rho}(1-\zeta)\Lambda q_A(0)^{-\frac{(1+\rho)}{\rho}} p_C(t)^{-\zeta(1-\rho)/\rho} e^{-\frac{\beta-r}{\rho}t} \left.\frac{\partial q_A(0)}{\partial \mu_E(x)}\right|_T > 0. \tag{103}$$

Now allow length of life T to be optimally chosen. Using (78) we have

$$\frac{\partial T}{\partial \mu_E(x)} = \frac{q_A(T) \left.\frac{\partial E(T)}{\partial \mu_E(x)}\right|_T + \left.\frac{\partial q_A(0)}{\partial \mu_E(x)}\right|_T e^{-rT}\left[\left.\frac{\partial A(t)}{\partial t}\right|_{t=T} + q_{h/a}(T)\left.\frac{\partial H(t)}{\partial t}\right|_{t=T}\right]}{-\left.\frac{\partial \Im(T)}{\partial T}\right|_{\mu_E(x)}} > 0. \tag{104}$$

Using (36), we obtain the following total responses to variation in the generalized productivity of skill investment, $\mu_E(x)$:

$$\frac{\partial q_{e/a}(t)}{\partial \mu_E(x)} = \left.\frac{\partial q_{e/a}(t)}{\partial T}\right|_{\mu_E(x)} \frac{\partial T}{\partial \mu_E(x)} > 0, \quad \frac{\partial q_{h/a}(t)}{\partial \mu_E(x)} = \left.\frac{\partial q_{h/a}(t)}{\partial T}\right|_{\mu_E(x)} \frac{\partial T}{\partial \mu_E(x)} > 0,$$
$$\frac{\partial E(t)}{\partial \mu_E(x)} = \left.\frac{\partial E(t)}{\partial \mu_E(x)}\right|_T + \left.\frac{\partial E(t)}{\partial T}\right|_{\mu_E(x)} \frac{\partial T}{\partial \mu_E(x)} > 0, \quad \frac{\partial H(t)}{\partial \mu_E(x)} = \left.\frac{\partial H(t)}{\partial T}\right|_{\mu_E(x)} \frac{\partial T}{\partial \mu_E(x)} > 0,$$
$$\frac{\partial X_E(t)}{\partial \mu_E(x)} = \left.\frac{\partial X_E(t)}{\partial \mu_E(x)}\right|_T + \left.\frac{\partial X_E(t)}{\partial T}\right|_{\mu_E(x)} \frac{\partial T}{\partial \mu_E(x)} > 0, \quad \frac{\partial X_H(t)}{\partial \mu_E(x)} = \left.\frac{\partial X_H(t)}{\partial T}\right|_{\mu_E(x)} \frac{\partial T}{\partial \mu_E(x)} > 0,$$

$$\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial \mu_E(x)} = \left.\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial \mu_E(x)}\right|_T + \left.\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial T}\right|_{\mu_E(x)} \frac{\partial T}{\partial \mu_E(x)} > 0,$$
$$\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial \mu_E(x)} = \left.\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial T}\right|_{\mu_E(x)} \frac{\partial T}{\partial \mu_E(x)} > 0,$$

$$\frac{\partial A(t)}{\partial \mu_E(x)} = \left.\frac{\partial A(t)}{\partial \mu_E(x)}\right|_T + \left.\frac{\partial A(t)}{\partial T}\right|_{\mu_E(x)} \frac{\partial T}{\partial \mu_E(x)} \gtreqless 0,$$
$$\frac{\partial q_A(0)}{\partial \mu_E(x)} = \left.\frac{\partial q_A(0)}{\partial \mu_E(x)}\right|_T + \left.\frac{\partial q_A(0)}{\partial T}\right|_{\mu_E(x)} \frac{\partial T}{\partial \mu_E(x)} \gtreqless 0,$$

$$\frac{\partial X_C(t)}{\partial \mu_E(x)} = \left.\frac{\partial X_C(t)}{\partial \mu_E(x)}\right|_T + \left.\frac{\partial X_C(t)}{\partial T}\right|_{\mu_E(x)} \frac{\partial T}{\partial \mu_E(x)} \gtreqless 0,$$

$$\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial \mu_E(x)} = \left.\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial \mu_E(x)}\right|_T + \left.\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial T}\right|_{\mu_E(x)} \frac{\partial T}{\partial \mu_E(x)} \gtreqless 0, \tag{105}$$

where we have used (83).

Health productivity Consider the case where T is fixed. Differentiating (70) with respect to $\mu_H(x)$, using (62), and differentiating (71) with respect to $\mu_H(x)$, one finds

$$\begin{aligned}
\left.\frac{\partial q_{e/a}(0)}{\partial \mu_H(x)}\right|_T &= 0, \\
\left.\frac{\partial q_{h/a}(0)}{\partial \mu_H(x)}\right|_T &= \frac{-q_{h/a}(x)^{\frac{\gamma_H}{1-\gamma_H}} e^{\int_0^x d_H(u)du}}{\frac{\gamma_H}{1-\gamma_H}\int_0^T \mu_H(s) q_{h/a}(s)^{\frac{2\gamma_H-1}{1-\gamma_H}} e^{\int_0^s [2d_H(u)+r]du} ds} < 0.
\end{aligned} \tag{106}$$

Using (54) to (63), and (106), we obtain

$$\begin{aligned}
\left.\frac{\partial q_{e/a}(t)}{\partial \mu_H(x)}\right|_T &= 0, \ \forall t \\
\left.\frac{\partial q_{h/a}(t)}{\partial \mu_H(x)}\right|_T &= \left.\frac{\partial q_{h/a}(0)}{\partial \mu_H(x)}\right|_T e^{\int_0^t [d_H(u)+r]du} < 0, \ \forall t \\
\left.\frac{\partial E(t)}{\partial \mu_H(x)}\right|_T &= 0, \ \forall t \\
\left.\frac{\partial H(t)}{\partial \mu_H(x)}\right|_T &= q_{h/a}(x)^{\frac{\gamma_H}{1-\gamma_H}} e^{-\int_x^t d_H(u)du} \times \\
&\qquad \left[1 - \frac{\int_0^t \mu_H(s) q_{h/a}(s)^{\frac{2\gamma_H-1}{1-\gamma_H}} e^{\int_0^s [2d_H(u)+r]du} ds}{\int_0^T \mu_H(s) q_{h/a}(s)^{\frac{2\gamma_H-1}{1-\gamma_H}} e^{\int_0^s [2d_H(u)+r]du} ds}\right] \geq 0, \ \text{for } t \geq x, \\
\left.\frac{\partial X_E(t)}{\partial \mu_H(x)}\right|_T &= 0, \ \forall t \\
\left.\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial \mu_H(x)}\right|_T &= 0, \ \forall t \\
\left.\frac{\partial X_H(t)}{\partial \mu_H(x)}\right|_T &= \frac{\beta_H}{p_H(t)} q_{h/a}(t)^{\frac{1}{1-\gamma_H}} \delta(x-t) + \\
&\qquad \frac{\beta_H \mu_H(t)}{p_H(t)} \frac{1}{1-\gamma_H} q_{h/a}(t)^{\frac{\gamma_H}{1-\gamma_H}} \left.\frac{\partial q_{h/a}(t)}{\partial \mu_H(x)}\right|_T < 0 \text{ for } t \neq x, \\
\left.\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial \mu_H(x)}\right|_T &= \alpha_H q_{h/a}(t)^{\frac{1}{1-\gamma_H}} \delta(x-t) + \\
&\qquad \alpha_H \mu_H(t) \frac{1}{1-\gamma_H} q_{h/a}(t)^{\frac{\gamma_H}{1-\gamma_H}} \left.\frac{\partial q_{h/a}(t)}{\partial \mu_H(x)}\right|_T < 0 \text{ for } t \neq x.
\end{aligned} \tag{107}$$

Differentiating (69) with respect to $\mu_H(x)$, we have

$$\left.\frac{\partial A(t)}{\partial \mu_H(x)}\right|_T = e^{rt}\left\{\chi(t,x) - \chi(T,x)\frac{\int_0^t p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta-r(1-\rho))}{\rho}s} ds}{\int_0^T p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta-r(1-\rho))}{\rho}s} ds}\right\} \gtreqless 0, \tag{108}$$

$$\left.\frac{\partial q_A(0)}{\partial \mu_H(x)}\right|_T = \frac{-\chi(T,x)}{q_A(0)^{\frac{-(1+\rho)}{\rho}} \frac{\Lambda}{\rho} \int_0^T p_C(s)^{-\zeta(1-\rho)/\rho} e^{-\frac{(\beta-r(1-\rho))}{\rho}s} ds} < 0, \tag{109}$$

where

$$\chi(t,x) = -e^{-rx}\gamma_H q_{h/a}(x)^{\frac{1}{1-\gamma_H}} + \int_0^t \left[\frac{\partial H(s)}{\partial \mu_H(x)}\bigg|_T - \frac{\gamma_H}{1-\gamma_H}\mu_H(s) q_{h/a}(s)^{\frac{\gamma_H}{1-\gamma_H}} \frac{\partial q_{h/a}(s)}{\partial \mu_H(x)}\bigg|_T \right] e^{-rs} ds > 0, \tag{110}$$

and, in signing the terms, we have assumed once more that permanent effects dominate transient effects.

Further using (66) and (67) it follows that

$$\frac{\partial X_C(t)}{\partial \mu_H(x)}\bigg|_T = -\frac{1}{\rho}\zeta\Lambda q_A(0)^{-\frac{(1+\rho)}{\rho}} p_C(t)^{-(1-\zeta+\zeta/\rho)} e^{-\frac{\beta-r}{\rho}t} \frac{\partial q_A(0)}{\partial \mu_H(x)}\bigg|_T > 0, \tag{111}$$

$$\frac{\partial L(t)\left[E(t)+H(t)\right]}{\partial \mu_H(x)}\bigg|_T = -\frac{1}{\rho}(1-\zeta)\Lambda q_A(0)^{-\frac{(1+\rho)}{\rho}} p_C(t)^{-\zeta(1-\rho)/\rho} e^{-\frac{\beta-r}{\rho}t} \frac{\partial q_A(0)}{\partial \mu_H(x)}\bigg|_T > 0. \tag{112}$$

Now allow length of life T to be optimally chosen. Using (78) we have

$$\frac{\partial T}{\partial \mu_H(x)} = \frac{\frac{\partial q_A(T)}{\partial \mu_H(x)}\Big|_T \frac{\partial A(t)}{\partial t}\Big|_{t=T} + \left[q_{h/a}(T)\frac{\partial q_A(T)}{\partial \mu_H(x)}\Big|_T + q_A(T)\frac{\partial q_{h/a}(T)}{\partial \mu_H(x)}\Big|_T \right] \frac{\partial H(t)}{\partial t}\Big|_{t=T}}{-\frac{\partial \Im(T)}{\partial T}\Big|_{\mu_H(x)}} > 0. \tag{113}$$

Using (36), we obtain the following total responses to variation in the generalized productivity of health investment, $\mu_H(x)$:

$$\frac{\partial q_{e/a}(t)}{\partial \mu_H(x)} = \frac{\partial q_{e/a}(t)}{\partial T}\bigg|_{\mu_H(x)} \frac{\partial T}{\partial \mu_H(x)} > 0, \quad \frac{\partial q_{h/a}(t)}{\partial \mu_H(x)} = \frac{\partial q_{h/a}(t)}{\partial \mu_H(x)}\bigg|_T + \frac{\partial q_{h/a}(t)}{\partial T}\bigg|_{\mu_H(x)} \frac{\partial T}{\partial \mu_H(x)} \gtreqless 0,$$

$$\frac{\partial E(t)}{\partial \mu_H(x)} = \frac{\partial E(t)}{\partial T}\bigg|_{\mu_H(x)} \frac{\partial T}{\partial \mu_H(x)} > 0, \quad \frac{\partial H(t)}{\partial \mu_H(x)} = \frac{\partial H(t)}{\partial \mu_H(x)}\bigg|_T + \frac{\partial H(t)}{\partial T}\bigg|_{\mu_H(x)} \frac{\partial T}{\partial \mu_H(x)} > 0,$$

$$\frac{\partial X_E(t)}{\partial \mu_H(x)} = \frac{\partial X_E(t)}{\partial T}\bigg|_{\mu_H(x)} \frac{\partial T}{\partial \mu_H(x)} > 0, \quad \frac{\partial X_H(t)}{\partial \mu_H(x)} = \frac{\partial X_H(t)}{\partial \mu_H(x)}\bigg|_T + \frac{\partial X_H(t)}{\partial T}\bigg|_{\mu_H(x)} \frac{\partial T}{\partial \mu_H(x)} > 0,$$

$$\frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial \mu_H(x)} = \frac{\partial \tau_E(t)\left[E(t)+H(t)\right]}{\partial T}\bigg|_{\mu_H(x)} \frac{\partial T}{\partial \mu_H(x)} > 0,$$

$$\frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial \mu_H(x)} = \frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial \mu_H(x)}\bigg|_T + \frac{\partial \tau_H(t)\left[E(t)+H(t)\right]}{\partial T}\bigg|_{\mu_H(x)} \frac{\partial T}{\partial \mu_H(x)} > 0,$$

$$\frac{\partial A(t)}{\partial \mu_H(x)} = \frac{\partial A(t)}{\partial \mu_H(x)}\bigg|_T + \frac{\partial A(t)}{\partial T}\bigg|_{\mu_H(x)} \frac{\partial T}{\partial \mu_H(x)} \gtreqless 0,$$

$$\frac{\partial q_A(0)}{\partial \mu_H(x)} = \frac{\partial q_A(0)}{\partial \mu_H(x)}\bigg|_T + \frac{\partial q_A(0)}{\partial T}\bigg|_{\mu_H(x)} \frac{\partial T}{\partial \mu_H(x)} \gtreqless 0,$$

$$\frac{\partial X_C(t)}{\partial \mu_H(x)} = \left.\frac{\partial X_C(t)}{\partial \mu_H(x)}\right|_T + \left.\frac{\partial X_C(t)}{\partial T}\right|_{\mu_H(x)} \frac{\partial T}{\partial \mu_H(x)} \gtreqless 0,$$

$$\frac{\partial L(t)\,[E(t)+H(t)]}{\partial \mu_H(x)} = \left.\frac{\partial L(t)\,[E(t)+H(t)]}{\partial \mu_H(x)}\right|_T + \left.\frac{\partial L(t)\,[E(t)+H(t)]}{\partial T}\right|_{\mu_H(x)} \frac{\partial T}{\partial \mu_H(x)} \gtreqless 0, \quad (114)$$

where we have used (83).

A.6 Comparative dynamics for the general model

Prediction 4: Absent ability to extend life T, associations between wealth, skill and health are absent or small.

This prediction is true also for the general model. In both the general and the simpler model, the end condition applies that end of life occurs at $t = T$ at the minimum health level $H(T) = H_{\min}$. Hence, even though in the general model additional resources in the form of wealth, skill, or health, may lead to an initial increase in the relative marginal value of health $q_{h/a}(t)$ (see 25) and therefore greater health investment and greater health, for fixed length of life T this needs to be compensated by eventual lower health investment in order for health to reach $H_{\min}$ at $t = T$. The response to additional resources of health investment and thereby health is therefore muted.

While skill may be more responsive to additional resources, as its terminal level $E(T)$ is allowed to be free, the response to wealth of skill investment and skill is also muted due to strong complementarity between skill and health: the initial benefits derived from higher levels of health (earnings, self-productivity, and dynamic complementarity) are offset by subsequent lower benefits from reduced health. Moreover, one of the key drivers of skill-capital investment is the horizon (longevity). This important pathway is shut down when forcing length of life T to be fixed. As a result, there are no strong associations between wealth, skill, and health for fixed T, and $\partial g(t)/\partial Z_0|_T$, the first term on the RHS of (36), is generally small for variation δZ_0 in any model parameter of interest.

Prediction 5: Wealthy, skilled, and healthy individuals live longer.

Individuals optimally choose longevity T such that the marginal value of life extension is zero at this age, $\Im(T) = 0$ (see 14),

$$\Im(T) = U(T)e^{-\beta T} + q_H(T) \left.\frac{\partial H}{\partial t}\right|_{t=T} + q_A(T) \left.\frac{\partial A}{\partial t}\right|_{t=T} = 0, \quad (115)$$

where we have used the transversality condition $q_E(T) = 0$. As the expression shows, the marginal benefit of extending life consists of the additional utility from consumption and effective leisure, and the marginal costs consist of the increasingly binding wealth and health constraints, due to declining wealth and declining health near the end of life.[50] In particular, health is increasingly constraining relative to wealth as the marginal value of wealth $q_A(t) = q_A(0)e^{-rt}$ declines with age while the relative marginal value of health increases with age $q_{h/a}(t) = q_H(t)/q_A(t)$ (i.e. even if $q_H(t)$ declines [but more likely, it increases] it does so less rapidly than does $q_A(t)$). In addition, declining health reduces utility $U(t)$ and thereby the marginal benefit of life extension.

The conditions (115), (81) and (82) for optimal length of life do not depend on the characteristics of the simpler model. They also apply to the general model. We have argued in section 4.2.1 that the lifecycle trajectories of $A(t)$ and $H(t)$ are similar in the general and simpler model – in particular in both models health and assets decline towards the end of life. Thus, in order to establish proof, using (82), we need only establish that $\partial E(T)/\partial Z_0|_T > 0$, $\partial q_A(T)/\partial Z_0|_T < 0$, and $\partial q_H(T)/\partial Z_0|_T < 0$, for $Z_0 = \{A_0, E_0, H_0\}$. From prediction 6 follows $\partial E(T)/\partial Z_0|_T > 0$, for $Z_0 = \{A_0, E_0, H_0\}$. By the assumption of diminishing returns to wealth, we have $\partial q_A(T)/\partial A(0)|_T = e^{-rT}\, \partial q_A(0)/\partial A(0)|_T < 0$ (wealth increases life time utility but at a diminishing rate).

For the remainder of the proof we follow the reasoning of prediction 4. For the simple model we have $\partial q_A(T)/\partial E(0)|_T < 0$ and $\partial q_A(T)/\partial H(0)|_T < 0$, which for the general model plausibly holds as well, as follows. For fixed T any additional investment in health early in life, as a result of the additional resources δE_0 or δH_0, needs to be compensated by reduced investment later in life for health to reach $H_{\min}$ at $t = T$. Hence, we expect $\partial q_{h/a}(t)/\partial Z_0\big|_T$ to be positive up to some $t = t^\dagger$, and negative afterwards (see also Galama and Van Kippersluis, 2010). In particular, $\partial q_{h/a}(T)/\partial Z_0\big|_T < 0$. This also affects decisions regarding investment in skill and the response in terms of skill and health investment is muted for fixed T (see prediction 4). Since in aggregate not much additional investment is made (positive and negative variations in investment balance out), the additional resources can only be spend on consumption. This would be associated (see 15) with a reduced marginal value of wealth at any age. Therefore, $\partial q_A(T)/\partial E(0)|_T < 0$ and $\partial q_A(T)/\partial H(0)|_T < 0$. Last, we need to establish that $\partial q_H(T)/\partial Z_0|_T < 0$, for $Z_0 = \{A_0, E_0, H_0\}$. Now, $\partial q_H(T)/\partial Z_0|_T = q_A(T)\, \partial q_{h/a}(T)/\partial Z_0\big|_T + q_A(T)^{-1}\, \partial q_A(T)/\partial Z_0|_T$. Both terms on the RHS are negative as discussed above. Q.E.D.

Thus we have established that prediction 5 also plausibly holds in the general model.

Prediction 6: Wealthy and healthy individuals value skill more, invest more in skill, and are more skilled at every age. Individuals with more endowed skill are more skilled at every age, but potentially value skill less.

In the general model wealthy and healthy individuals also value skill more. Investment in skill is one margin of adjustment individuals can choose with several benefits: skill capital increases earnings $\partial Y/\partial E > 0$, the efficiency of skill production $\partial f_E/\partial E > 0$, and

[50] Both $\partial H(t)/\partial t|_{t=T}$ and $\partial A(t)/\partial t|_{t=T}$ are negative since health declines near the end of life as it approaches $H_{\min}$ from above, and assets decline near the end of life in absence of a very strong bequest motive.

the efficiency of health production $\partial f_H/\partial E > 0$, and skill capital extends the horizon (prediction 5), thereby increasing the return on skill-capital investment (see 19). Wealth provides additional resources that can be devoted to skill investment and so does health, but health also raises the various benefits of skill capital as $\partial^2 Y/\partial E \partial H > 0$ (skill raises wages and health increases time devoted to work), $\partial^2 f_E/\partial E \partial H > 0$ (both skill and health raise the productivity of skill formation), and plausibly $\partial^2 f_H/\partial E \partial H > 0$. Thus, both wealth and health increase skill investment and skill. In the general model the effect of health on skill is plausibly even larger than in the simpler model, given the strong complementarities of health and skill in earnings, and in the production of skill and health.

Whether additional skill increases skill investment is less clear. Since in the general model $\partial^2 Y/\partial E^2 < 0$, $\partial^2 f_E/\partial E^2 < 0$, and $\partial^2 f_H/\partial E^2 < 0$, the various benefits of skill capital are decreasing in endowed skill, providing incentives to reduce skill-capital investment. Nonetheless, starting out with higher skill, under standard economic assumptions regarding the functional forms of the utility and production functions and our assumed complementarity between skill and health (Y, f_E and f_H), skill investment will not be reduced to such an extent that skill is eventually lower for individuals who started out with a greater endowment of skill.

In sum, the only difference with the simpler model is that skilled individuals could potentially value skill less $\partial q_{e/a}(t)/\partial E_0 < 0$, and therefore invest less, but still have greater skill at every age.

Prediction 7: Wealthy and skilled individuals value health more, invest more in health, and are healthier at every age. Individuals with more endowed health are healthier at every age, but potentially value health less.

In the general model, this prediction plausibly applies too. Wealthy and skilled individuals value health more for its many benefits: health provides utility $\partial U/\partial H > 0$, increases earnings $\partial Y/\partial H > 0$, the efficiency of skill production $\partial f_E/\partial H > 0$, and potentially the efficiency of health production $\partial f_H/\partial H > 0$, and health capital extends the horizon (prediction 5), thereby increasing the return on health-capital investment (see 24). Wealth provides additional resources that can be devoted to health investment and so does skill, but skill also raises the various benefits of health capital as $\partial^2 Y/\partial E \partial H > 0$ (skill raises wages and health increases time devoted to work), $\partial^2 f_E/\partial E \partial H > 0$ (both skill and health raise the productivity of skill formation), and plausibly $\partial^2 f_H/\partial E \partial H > 0$. These effects are plausibly larger in the general model than in the simpler model due to strong complementarity between skill and health in earnings and in the production of skill and health. Thus, both wealth and skill increase health investment and health.[51]

[51] If these additional resources can only moderately extend life (see prediction 7) then any initial higher levels of health investment may have to be somewhat offset by subsequent lower investment for health to reach $H_{\min}$ (as this case more closely resembles that of fixed T). Thus, for the wealthy and skilled, health is higher and in aggregate health investment is higher, but later in the lifecycle investment may be reduced.

Similar to the discussion for prediction 6, for greater endowed health, the demand for health investment is reduced since in the general model the various benefits of health capital are decreasing in endowed health. Nonetheless, starting out with higher health, and for the same reasons provided in prediction 6 for skill, health investment will not be reduced to such an extent that health is eventually lower when starting out with a greater endowment. Thus endowed health also leads to greater health at every age. Further, health extends the horizon, thereby increasing the return on health-capital investment, so that health investment may be higher at every age, in particular if endowed health enables substantial life extension.

In sum, analogous to the case for skill (prediction 6), healthy individuals could potentially value health less $\partial q_{h/a}(t)/\partial H_0 < 0$ in both the simpler and the general model. This scenario seems plausible for health since health is more constrained than skill (the terminal value of health is fixed at $H_{\min}$ and, unlike skill, health potentially lowers the efficiency of the health-production process $\partial f_H/\partial H < 0$).

Prediction 8: Gains in life expectancy reinforce the associations between wealth, skill, health, and technology.

The discussion for prediction 8 relied on the simpler model only in establishing that $\partial E(t)/\partial T|_{Z_0} > 0$ and $\partial H(t)/\partial T|_{Z_0} > 0$, for $Z_0 = \{A_0, E_0, H_0, \mu_E(t), \mu_H(t)\}$. If we can show these also hold for the general model, the proof is completed. While we cannot provide a formal proof, we can invoke a simple heuristic argument based on (36). From prediction 4 we know that $\partial E(t)/\partial Z_0|_T$ and $\partial H(t)/\partial Z_0|_T$ are small. So that $\partial E(t)/\partial Z_0 \approx \partial E(t)/\partial T|_{Z_0} (\partial T/\partial Z_0)$ and $\partial H(t)/\partial Z_0 \approx \partial H(t)/\partial T|_{Z_0} (\partial T/\partial Z_0)$. From prediction 5 we have $\partial T/\partial Z_0 > 0$, for $Z_0 = \{A_0, E_0, H_0, \mu_E(t), \mu_H(t)\}$ and from predictions 6 and 7 we have $\partial E(t)/\partial Z_0 > 0$ and $\partial H(t)/\partial Z_0 > 0$, for $Z_0 = \{A_0, E_0, H_0, \mu_E(t), \mu_H(t)\}$. Thus, if predictions 4, 5, 6, and 7 hold, we find $\partial E(t)/\partial T|_{Z_0} > 0$ and $\partial H(t)/\partial T|_{Z_0} > 0$, for $Z_0 = \{A_0, E_0, H_0, \mu_E(t), \mu_H(t)\}$, and $\forall t$. Q.E.D.

Prediction 9: Complementarity effects, operating through longevity, reinforce the associations between wealth, skill, health, and technology.

In many cases, variations in two (or more) parameters that affect longevity, reinforce each other. To see this, differentiate (36) with respect to an additional generic variation δW_0 in an initial condition or model parameter, to obtain:

$$\frac{\partial^2 g(t)}{\partial Z_0 \partial W_0} = \left.\frac{\partial^2 g(t)}{\partial Z_0 \partial W_0}\right|_T + \left.\frac{\partial^2 g(t)}{\partial T \partial W_0}\right|_{Z_0} \frac{\partial T}{\partial Z_0} + \left.\frac{\partial g(t)}{\partial T}\right|_{Z_0} \frac{\partial^2 T}{\partial Z_0 \partial W_0}. \tag{116}$$

The first term is small (for fixed T) for the same reasons discussed in prediction 4. The second term increases with the extent of life extension $\partial T/\partial Z_0$. If $\partial^2 g(t)/\partial T \partial W_0|_{Z_0} > 0$, then there is complementarity, and variation in Z_0 is reinforced by variation in

W_0. For example, $\partial E(t)/\partial T|_{Z_0}$ increases in the technology of skill production $\mu_E(x)$ ($\partial^2 E(t)/\partial T \partial \mu_E(x) > 0$), leading to prediction 9b. More generally, the effect of life expectancy on skill $\partial E(t)/\partial T|_{Z_0}$ increases in any factor that increases the marginal value of skill $q_{e/a}(t)$ (see the expression for $\partial E(t)/\partial T|_{Z_0}$ in 83), such as initial wealth A_0 (see 86), initial skill E_0 (see 91), initial health H_0 (see 97), skill productivity $\mu_E(t)$ (see 105), and health productivity $\mu_H(t)$ (see 114).

Another type of complementarity between Z_0 and W_0 could arise from the third term in (116). This term increases in $\partial g(t)/\partial T|_{Z_0}$, which is positive for skill $E(t)$ and health $H(t)$ (see prediction 8). It is cumbersome to mathematically establish that $\partial^2 T/\partial Z_0 \partial W_0$ is positive. First, there are many such possible combinations and second these higher-order expressions are substantially more complicated to analyze (see, e.g., 78 and 79). But intuitively, two factors that both increase longevity could operate together, acting as complements. For example, it is plausible that the effect of initial assets on life expectancy $\partial T/\partial A_0$ is increasing in health.

A.7 Comparative dynamics of schooling and retirement

Conjecture 3a: Wealthy individuals have shorter working lives, staying in school longer and retiring earlier.

Conjecture 3b: The relationship between skill and years of employment follows an inverse U-shape, with low- and high-skilled individuals having shorter working lives, and medium-skilled individuals longer working lives.

Conjecture 3c: Healthy individuals stay in school longer and retire later.

We first provide a brief summary of results followed by detailed analyses. There are various effects that influence the optimal schooling decision S (118) and the optimal retirement decision R (122). The sign of each effect and the conditions under which they hold are summarized in Table 3.

First, there is a "wealth" effect: those with more resources, in terms of endowed wealth, endowed skill, and endowed health, can afford the additional utility from leisure during schooling years and during retirement. Second, wealthy and healthy individuals value skill more (prediction 6). Thus they value the additional investment in skill during schooling ("value of skill" effect).[52] Likewise, wealthy and skilled individuals value health more (prediction 7). Thus they value the additional investment in health during retirement ("value of health" effect).[53] Third, individuals endowed with more wealth, skill, and health, are more skilled and healthier at all ages (predictions 6 and 7). Since higher earnings encourage labor-force participation while the benefits of enhanced utility, skill and health production favor not working, the signs of these "skill" and "health" effects

[52]Theoretically, however, it is possible that skilled individuals value skill less (prediction 6).

[53]However, healthy individuals may value health less (prediction 7).

are undetermined, except for wealthy individuals where the skill effect plausibly provides incentive to prolong schooling and the health effect plausibly enourages early retirement. Finally, in the case of retirement, there is a "longevity" effect: those with more resources, in terms of endowed wealth, endowed skill, and endowed health, live longer (see prediction 5), delaying retirement (if delaying retirement increases benefits, $\partial b/\partial R > 0$) since they need to finance consumption over a longer lifespan.

Table 3: Comparative dynamic effects of initial wealth A_0, initial skill E_0, and initial health H_0, on the optimal schooling age S, and the optimal retirement age R.

	δA_0	δE_0	δH_0
Schooling age S			
Wealth effect	$+$	$+$	$+$
Value of skill effect	$+$	$+/-^{*}$	$+$
Skill effect	$+$	$+/-^{\dagger}$	$+/-^{\dagger}$
Retirement age R			
Longevity effect	$+^{\P}$	$+^{\P}$	$+^{\P}$
Wealth effect	$-$	$-$	$-$
Value of health effect	$-$	$-$	$+^{\ddagger}$
Health effect	$-$	$+/-^{\S}$	$+/-^{\S}$

Notes: A $+$ denotes an increase, and a $-$ denotes a decrease in the optimal schooling age or retirement age. $*$ Sign is positive under enhancement through life extension, self-productivity, and / or dynamic complementarity. † Sign is positive if the benefits of skill in enhancing skill (and health) production outweigh the benefit of skill in enhancing earnings and negative otherwise. ¶ Sign holds if benefits increase with retirement age $\partial b(R)/\partial R > 0$. ‡ Sign holds if a higher health stock leads to a lower relative marginal value of health $\partial q_{h/a}(t)/\partial H_0 < 0$. § Sign is negative if the benefits of health in enhancing utility, health, and skill production outweigh the benefit of health in enhancing earnings, and positive if otherwise. See Appendix section A.7 for a detailed explanation of the conditions for which the signs hold.

Schooling The comparative dynamic result for the optimal schooling age S follows from differentiating the transversality condition (28) with respect to a variation δZ_0 in an initial condition or model parameter, and noting that the condition also contains the term

$$\frac{w(S)}{q_A(S)}\left[\lambda_{\tau_w}(S_-)\tau_w(S_-) - \lambda_{\tau_w}(S_+)\tau_w(S_+)\right], \tag{117}$$

which is zero in (28) because $\tau_w(S_-) = 0$ and $\lambda_{\tau_w}(S_+) = 0$. Since variation of this term is not necessarily zero, it needs to be included. This leads to the following expression

$$\begin{aligned}
&\left\{ \frac{\partial w}{\partial S}\tau_w(S_+) - \frac{\partial q_{e/a}(S)}{\partial S}\left[f_E(S_-) - f_E(S_+)\right] - \frac{\partial q_{h/a}(S)}{\partial S}\left[f_H(S_-) - f_H(S_+)\right] \right.\\
&\left. + \frac{(\beta - r)}{q_A(S)}\left[U(S_-) - U(S_+)\right]e^{-\beta S} \right\} \times \frac{\partial S}{\partial Z_0}\\
= &-\frac{1}{q_A(S)^2}\left[U(S_-) - U(S_+)\right]e^{-\beta S} \times \frac{\partial q_A(S)}{\partial Z_0}\\
&+ \left[f_E(S_-) - f_E(S_+)\right] \times \frac{\partial q_{e/a}(S)}{\partial Z_0} + \left[f_H(S_-) - f_H(S_+)\right] \times \frac{\partial q_{h/a}(S)}{\partial Z_0}\\
&+ \left[\left.\frac{\partial q_{e/a}}{\partial t}\right|_{t=S_+} - \left.\frac{\partial q_{e/a}}{\partial t}\right|_{t=S_-} \right] \times \frac{\partial E(S)}{\partial Z_0} + \left[\left.\frac{\partial q_{h/a}}{\partial t}\right|_{t=S_+} - \left.\frac{\partial q_{h/a}}{\partial t}\right|_{t=S_-} \right] \times \frac{\partial H(S)}{\partial Z_0},
\end{aligned} \tag{118}$$

where, we have used the first-order conditions (15, 16, 18, 23), $\tau_w(S_-) = 0$, and $\lambda_{\tau_w}(S_+) = 0$, and, for simplicity, we neglect terms due to variation in the price of consumption goods, skill investment goods, and health investment goods with respect to time $\partial p_C/\partial S$, $\partial p_E/\partial S$, and $\partial p_H/\partial S$, since these are arguably random (unrelated to joining the labor force).[54]

The last two terms can also be expressed as

$$\begin{aligned}
&\left[\left.\frac{\partial q_{e/a}}{\partial t}\right|_{t=S_+} - \left.\frac{\partial q_{e/a}}{\partial t}\right|_{t=S_-} \right] \times \frac{\partial E(S)}{\partial Z_0}\\
= &\left\{ q_{e/a}(S)\left[\frac{\partial f_E(S_-)}{\partial E(S)} - \frac{\partial f_E(S_+)}{\partial E(S)}\right] \right.\\
&\left. + q_{h/a}(S)\left[\frac{\partial f_H(S_-)}{\partial E(S)} - \frac{\partial f_H(S_+)}{\partial E(S)}\right] - \frac{\partial w(S)}{\partial E(S)}\tau_w(S_+) \right\} \times \frac{\partial E(S)}{\partial Z_0},
\end{aligned} \tag{119}$$

and

$$\begin{aligned}
&\left[\left.\frac{\partial q_{h/a}}{\partial t}\right|_{t=S_+} - \left.\frac{\partial q_{h/a}}{\partial t}\right|_{t=S_-} \right] \times \frac{\partial H(S)}{\partial Z_0}\\
= &\left\{ \frac{1}{q_A(S)}\left[\frac{\partial U(S_-)}{\partial H(S)} - \frac{\partial U(S_+)}{\partial H(S)}\right]e^{-\beta S} + q_{e/a}(S)\left[\frac{\partial f_E(S_-)}{\partial H(S)} - \frac{\partial f_E(S_+)}{\partial H(S)}\right] \right.\\
&\left. + q_{h/a}(S)\left[\frac{\partial f_H(S_-)}{\partial H(S)} - \frac{\partial f_H(S_+)}{\partial H(S)}\right] + \frac{\partial s(S)}{\partial H(S)}w(S)\left[1 - \frac{\lambda_{\tau_w}(S_-)}{q_A(S)}\right] \right\} \times \frac{\partial H(S)}{\partial Z_0}.
\end{aligned} \tag{120}$$

The comparative dynamic expression (118) for the optimal schooling age S can be understood as follows. After optimization, by employing the first-order conditions for the

[54]However, the rate of change of the price of education $\partial p_E/\partial S$ may change when individuals start working, as schooling is no longer subsidied or subsidies may be different for those who are working.

optimal controls $X_C(t)$, $L(t)$, $I_E(t)$, $I_H(t)$, and by employing the transversality conditions for R and T, the model is no longer a function of the control functions $X_C(t)$, $L(t)$, $I_E(t)$, $I_H(t)$, and optimized parameters R, and T, but is instead a function of the optimal schooling age S, state functions $A(t)$, $E(t)$, $H(t)$, and co-state functions $q_A(t)$, $q_{e/a}(t)$, $q_{h/a}(t)$ (as well as a function of initial conditions and other exogenous model parameters and functions, such as prices; e.g., Caputo 2005). The comparative dynamic result therefore consist of terms in the variation of S, $E(t)$, $H(t)$, $q_A(t)$, $q_{e/a}(t)$, and $q_{h/a}(t)$ (there is no term in $A(t)$ because it is not explicitly contained in the transversality condition for S) with respect to Z_0.

The large term in brackets on the LHS of (118) represents the value of staying in school. It is the difference between the various benefits from work and from schooling at S as a result of variation in the optimal schooling age $\partial S/\partial Z_0$. The term is plausibly positive.[55] The decision to continue schooling $\partial S/\partial Z_0$ is then determined by five terms (see RHS of 118). These five terms are essentially the difference between the various benefits from work and from schooling at S as a result of, respectively, variation in the marginal value of wealth $\partial q_A(S)/\partial Z_0$, the relative marginal value of skill $\partial q_{e/a}(S)/\partial Z_0$, the relative marginal value of health $\partial q_{h/a}(S)/\partial Z_0$, skill $\partial E(S)/\partial Z_0$, and health $\partial E(S)/\partial Z_0$. The first three terms represent variations in the marginal value of 'existing' differences between school S_- and work S_+. The last two terms represent variations in these differences.

Early in life the marginal value of skill is high and investment in skill is high ($f_E(t)q_{e/a}(t)$ large), while the marginal value of health is low and investment in health is low ($f_H(t)q_{h/a}(t)$ small), so that factors affecting skill formation (second and fourth term on the RHS of 118) are plausibly more important than factors regarding health formation (third and fifth term on the RHS of 118) in affecting the schooling decision. The opposite holds for retirement. For this reason, we focus in this section on the first, second, and fourth terms in discussing the comparative dynamic effect of schooling (and in the retirement section we focus on the first, second, fourth and sixth terms).

The first term on the RHS is the change at S in the value of utility derived from additional leisure during schooling, valued at (the inverse of) the marginal value of wealth. This term represents a "wealth" effect. Wealthy individuals value wealth less $\partial q_A(0)/\partial A_0 < 0$, which raises the value of the additional utility $U(S_-) - U(S_+)$ derived

[55] The value of staying in school consists of several terms. The first term represents the benefit of higher wages from continued schooling. It is positive as wages increase with age early in life. The second term represents the change in the value of greater skill investment during schooling $q_{e/a}(S)[f_E(S_-) - f_E(S_+)]$ as a result of variation in the marginal value of skill resulting from continued schooling. It is also positive, since the marginal value of skill declines with age, hence $\partial q_{e/a}(S)/\partial S < 0$. Likewise, the third term represents the change in the value of greater health investment during schooling $q_{h/a}(S)[f_H(S_-) - f_H(S_+)]$ as a result of variation in the marginal value of health resulting from continued schooling. Early in life the marginal value of skill is high and investment in skill is high, while the marginal value of health is low and investment in health is low. Thus the third term, likely negative as the marginal value of health increases with age $\partial q_{h/a}(S)/\partial S > 0$, is plausibly smaller than the second term. The fourth term represents the value of postponing entry into the labor market in terms of the utility from additional leisure $U(S_-) - U(S_+) > 0$. It is positive if the subjective discount rate β exceeds the rate of return on capital r.

from additional leisure during the schooling period (essentially wealthy individuals can "afford" leisure time). Skill and health are also forms of wealth and therefore skilled and healthy people also plausibly value leisure more, encouring them also to stay in school longer.

The second term is the change at S in the value of additional skill investment $[f_E(S_-) - f_E(S_+)]$ valued at the marginal value of skill. Wealthy and healthy individuals value skill more $\partial q_{e/a}(S)/\partial Z_0 > 0$ (prediction 6) so that the second term is positive, encouraging individuals to stay in school. It is theoretically possible, however, that the skilled value skill less $\partial q_{e/a}(S)/\partial E_0 < 0$ (prediction 6).

The fourth term (see 119) is the difference between the schooling and working phases of life in the rate of change of the relative marginal value of skill. Wealthy, skilled, and healthy individuals are more skilled at every age $\partial E(S)/\partial Z_0 > 0$ (prediction 6) so that the term is positive if the rate of change in the relative marginal value of skill increases after S. The fourth term is also (see 119) the effect of variation in skill $\partial E(S)/\partial Z_0$ on the difference between the various benefits from work and from schooling at S. Skill capital affects the difference between S_- and S_+ in skill production (first term on the RHS of 119), health production (second term on the RHS of 119), and earnings (last term on the RHS of 119). Thus the fourth term is positive, i.e. the rate of change in the relative marginal value of skill increases after S, if the difference in the marginal benefit of earnings before and after the school-leaving age S (negative sign) is smaller than the difference in the sum of the other two marginal benefits: the effect of greater endowed skill on the difference between S_- and S_+ in skill production and in health production (both of which have a positive sign).[56] And, it is negative otherwise.[57]

While we cannot unambiguously sign the fourth term, we can assess under which conditions it is more likely to be positive. For the usual assumptions made in this paper of diminishing returns to scale in investment, in skill, and in health, of the skill-production function $f_E(t)$, the first term in (119) $q_e/a(S)[\partial f_E(S_-)/\partial E - f_E(S_+)/\partial E]$ is positive, increasing in investment $I_E(t)$ and in health $H(t)$, and decreasing in skill $E(t)$.[58] Since the relative marginal value of skill increases in wealth and health, the term is larger for the wealthy and the healthy. Since the relative marginal value of skill is high and the relative marginal value of health is low early in life the second term in (119) $q_{h/a}(S)[\partial f_H(S_-)/\partial E - f_H(S_+)/\partial E]$ is likely small. The third term in (119) is negative,

[56]Skill investment and health investment are higher during schooling than during work because of the lower opportunity cost of time. If skill capital is complementary to skill investment and health investment, i.e. $\partial^2 f_E/\partial E \partial I_E > 0$ and $\partial^2 f_H/\partial E \partial I_H > 0$, then the terms $\partial f_E(S_-)/\partial E - f_E(S_+)/\partial E$ and $\partial f_H(S_-)/\partial E - f_H(S_+)/\partial E$ are positive.

[57]The top-left panel of Figure 1 shows such a pattern for illustrative purposes, where skill investment $I_E(t)$ decreases more rapidly during working ages than during the schooling phase.

[58]The direct effect of skill on the derivative of the production function with respect to skill $\partial f_E(S_-)/\partial E$ is negative (decreasing returns to skill), but the derivative increases in the level of skill investment, and skill investment in turn may increase with skill as a result of life extension, self-productivity, and / or dynamic complementarity effects (see main text). Further, the marginal value of skill $q_{e/a}(S)$ may increase with skill for these same reasons. Still, direct effects are usually stronger than indirect effects.

decreasing in skill (through diminishing returns) and decreasing in wealth (as demand for leisure reduces time that can be devoted to work), and increasing in health (as reduced sick time increases time that can be devoted to work). Hence, given that the first term is positive and increasing in wealth, and the last term is negative, and decreasing in wealth the total term is more likely to be positive for wealthy individuals.

In sum, the "wealth" effect encourages individuals with more endowed wealth, skill, and health to stay in school. The "value of skill" effect encourages wealthy and healthy individuals (who value skill more) to continue schooling, but the skilled potentially value skill less. The "skill" effect is ambiguous for skilled and healthy individuals but likely positive for wealthy individuals. Wealthy individuals thus stay in school longer. Casual observation, however, suggests skilled individuals also value continued schooling, suggesting that potentially self-productivity and dynamic complementarity are operating, and / or that endowed skill increases longevity, raising the relative value of skill $q_{e/a}(t)$ ("value of skill" effect).

Retirement Likewise, the comparative dynamic result for the optimal retirement age R follows from differentiating the transversality condition (29) with respect to a variation δZ_0 in an initial condition or model parameter, and noting that the condition also contains the term

$$\frac{w(R)}{q_A(R)}\left[\lambda_{\tau_w}(R_+)\tau_w(R_+) - \lambda_{\tau_w}(R_-)\tau_w(R_-)\right], \tag{121}$$

which is zero in (29) because $\tau_w(R_+) = 0$ and $\lambda_{\tau_w}(R_-) = 0$. Since variation of this term is not necessarily zero, it needs to be included. This leads to the following expression

$$\begin{aligned}
&\left\{\frac{\partial b}{\partial R}\left[1+e^{-r(T-R)}\right] - \frac{\partial^2 b}{\partial R^2}\frac{1}{r}\left[1-e^{-r(T-R)}\right] - \frac{(\beta-r)}{q_A(R)}\left[U(R_+)-U(R_-)\right]e^{-\beta R}\right.\\
&\left.+\frac{\partial q_{e/a}(R)}{\partial R}\left[f_E(R_+)-f_E(R_-)\right] + \frac{\partial q_{h/a}(R)}{\partial R}\left[f_H(R_+)-f_H(R_-)\right] - \frac{\partial w}{\partial R}\tau_w(R_-)\right\} \times \frac{\partial R}{\partial Z_0}\\
=&\ \frac{\partial b}{\partial R}e^{-r(T-R)} \times \frac{\partial T}{\partial Z_0} + \frac{1}{q_A(R)^2}\left[U(R_+)-U(R_-)\right]e^{-\beta R} \times \frac{\partial q_A(R)}{\partial Z_0}\\
&-\left[f_E(R_+)-f_E(R_-)\right] \times \frac{\partial q_{e/a}(R)}{\partial Z_0} - \left[f_H(R_+)-f_H(R_-)\right] \times \frac{\partial q_{h/a}(R)}{\partial Z_0}\\
&-\left[\left.\frac{\partial q_{e/a}}{\partial t}\right|_{t=R_-} - \left.\frac{\partial q_{e/a}}{\partial t}\right|_{t=R_+}\right] \times \frac{\partial E(R)}{\partial Z_0} - \left[\left.\frac{\partial q_{h/a}}{\partial t}\right|_{t=R_-} - \left.\frac{\partial q_{h/a}}{\partial t}\right|_{t=R_+}\right] \times \frac{\partial H(R)}{\partial Z_0},
\end{aligned} \tag{122}$$

where, we have used the first-order conditions (15, 16, 18, 23), $\tau_w(R_+) = 0$, and $\lambda_{\tau_w}(R_-) = 0$, and, for simplicity, we neglect terms due to variation in the price of consumption goods, skill investment goods, and health investment goods with respect to time $\partial p_C/\partial R$, $\partial p_E/\partial R$, and $\partial p_H/\partial R$, since these are arguably random (unrelated to retirement).[59]

The last two terms can also be expressed as

$$\left[\frac{\partial q_{e/a}}{\partial t}\bigg|_{t=R_-} - \frac{\partial q_{e/a}}{\partial t}\bigg|_{t=R_+} \right] \times \frac{\partial E(R)}{\partial Z_0}$$
$$= \Bigg\{ q_{e/a}(R) \left[\frac{\partial f_E(R_+)}{\partial E(R)} - \frac{\partial f_E(R_-)}{\partial E(R)} \right]$$
$$+ q_{h/a}(R) \left[\frac{\partial f_H(R_+)}{\partial E(R)} - \frac{\partial f_H(R_-)}{\partial E(R)} \right] - \frac{\partial w(R)}{\partial E(R)} \tau_w(R_-) \Bigg\} \times \frac{\partial E(R)}{\partial Z_0}, \tag{123}$$

and

$$\left[\frac{\partial q_{h/a}}{\partial t}\bigg|_{t=R_-} - \frac{\partial q_{h/a}}{\partial t}\bigg|_{t=R_+} \right] \times \frac{\partial H(R)}{\partial Z_0}$$
$$= \Bigg\{ \frac{1}{q_A(R)} \left[\frac{\partial U(R_+)}{\partial H(R)} - \frac{\partial U(R_-)}{\partial H(R)} \right] e^{-\beta R} + q_{e/a}(R) \left[\frac{\partial f_E(R_+)}{\partial H(R)} - \frac{\partial f_E(R_-)}{\partial H(R)} \right]$$
$$+ q_{h/a}(R) \left[\frac{\partial f_H(R_+)}{\partial H(R)} - \frac{\partial f_H(R_-)}{\partial H(R)} \right] + \frac{\partial s(R)}{\partial H(R)} w(R) \left[1 - \frac{\lambda_{\tau_w}(R_+)}{q_A(R)} \right] \Bigg\} \times \frac{\partial H(R)}{\partial Z_0}. \tag{124}$$

The large term in brackets on the LHS of (122) represents the value of retiring. It is the difference between the various benefits from work and from retirement at R as a result of variation in the optimal retirement age $\partial R/\partial Z_0$. The term is plausibly positive.[60] The decision to enter retirement $\partial R/\partial Z_0$ is then determined by six terms (see RHS of 122).

Analogous to the discussion for the schooling phase, since the marginal value of skill is initially high but decreases with age, while the marginal value of health is initially low

[59] Retirement in the United States often coincides with the age of Medicare eligibility, in which case the price of health investment may drop, suggesting the term due to variation in the price of health investment goods may need to be included.

[60] The value of entering retirement consists of several terms. The first term represents the benefit of higher pension benefits when retiring later, and the second term relates to possible diminishing returns to postponing retirement. Plausibly pension benefits increase with the retirement age, but at a diminishing rate such that the first two terms are positive. The third term represents the value of entering retirement in terms of the utility from additional leisure $U(R_+) - U(R_-) > 0$. It is negative if the subjective discount rate β exceeds the rate of return on capital r, and positive otherwise. The fourth term represents the change in the value of greater skill investment during retirement $q_{e/a}(R)[f_E(R_+) - f_E(R_-)]$ as a result of variation in the marginal value of skill resulting from later retirement. Likewise, the fifth term represents the change in the value of greater health investment during retirement $q_{h/a}(R)[f_H(R_+) - f_H(R_-)]$ as a result of variation in the marginal value of health resulting from later retirement. This term is likely positive, since the marginal value of health increases with age, hence $\partial q_{h/a}(R)/\partial R > 0$. Later in life the marginal value of skill is low and investment in skill is low, while the marginal value of health is high and investment in health is high. Thus the fourth term, likely negative as the marginal value of skill decreases with age, $\partial q_{e/a}(R)/\partial R < 0$, is plausibly smaller than the fifth term. The sixth term is the effect of delaying retirement on wages while working, and the sign of this term depends on whether wages still increase, remain flat, or decrease near retirement.

but increases with age, for the retirement decision we focus on the first, second, fourth, and sixth terms in discussing the comparative dynamic effect of retirement.

The first term is the "longevity" effect. Individuals with more endowed wealth, skill, and health live longer (see prediction 5), and this encourages individuals to work longer, as they require more resources to finance consumption over a longer lifespan.

The second term represents a "wealth" effect. Wealthy individuals value wealth less $\partial q_A(0)/\partial A_0 < 0$, which raises the value of the additional utility $U(R_+) - U(R_-)$ from additional leisure during the retirement phase (essentially wealthy individuals can "afford" leisure time). Skill and health are also forms of wealth and therefore skilled and healthy people also plausibly value leisure more, encouring them also to retire earlier.

The fourth term is the change at R in the value of additional health investment $[f_H(R_+) - f_H(R_-)]$, valued at the marginal value of health $q_{h/a}(R)$. Wealthy and skilled individuals value health more $\partial q_{h/a}(R)/\partial Z_0 > 0$ (prediction 7) so that the fourth term is negative, encouraging individuals to retire earlier. It is theoretically possible, however, that the healthy value health less $\partial q_{h/a}(R)/\partial H_0 < 0$ (prediction 7), suggesting that healthy individuals potentially care less about the additional time for health investment that the retirement phase offers.

The sixth term (see 119) is the difference between the working and retirement phases of life in the rate of change of the relative marginal value of health. Note that wealthy, skilled, and healthy individuals are healthier at every age $\partial H(S)/\partial Z_0 > 0$ (prediction 7), so that the term is negative (encouraging early retirement) if the rate of change in the relative marginal value of health decreases after R. The sixth term is also (see 124) the effect of variation in health $\partial H(R)/\partial Z_0$ on the difference between the various benefits from work and from retirement at R. Both before and after retirement, health provides utility, is used in the production of skill and health, and reduces sick time.[61] The sixth term is negative, i.e. the rate of change in the relative marginal value of health decreases after R, if the difference in the marginal benefit of reduced sicktime before and after the retirement age R (negative sign) is smaller than the difference in the sum of the other three marginal benefits: the effect of greater endowed health on the difference between R_- and R_+ in utility, skill production and in health production (all of which have a positive sign).[62] And, it is positive otherwise.[63]

While we cannot unambiguously sign the sixth term, we can assess under which conditions it is more likely to be positive. For the usual assumptions made in this paper of diminishing returns to scale of utility in health, the first term in (124) $(1/q_A(R))[\partial U(R_+)/\partial H - \partial U(R_-)/\partial H]e^{-\beta R}$ increases in endowed wealth and endowed

[61] The reduction in sick time is valued at the wage rate $w(R)$ during working life and at a reduced value $w(R)\lambda_{\tau_w}(R_+)/q_A(R)$ during the retirement phase.

[62] Leisure time, skill investment and health investment are higher during retirement than during work because of the lower opportunity cost of time. If health capital is complementary to leisure time, skill investment and health investment, i.e. $\partial^2 U/\partial H \partial L > 0$, $\partial^2 f_E/\partial H \partial I_E > 0$ and $\partial^2 f_H/\partial H \partial I_H > 0$, then the terms $\partial f_E(R_+)/\partial H - f_E(R_-)/\partial H$ and $\partial f_H(R_+)/\partial H - f_H(R_-)/\partial H$ are positive.

[63] The center-left panel of Figure 1 shows such a pattern for illustrative purposes, where health investment $I_H(t)$ increases less rapidly during working ages than during the retirement phase.

skill (a form of wealth / permanent income) and decreases in endowed health. Since the relative marginal value of skill is low and the relative marginal value of health is high later in life, the second term in (124) $q_{e/a}(R)[\partial f_E(R_+)/\partial H - \partial f_E(R_-)/\partial H]$ is likely small. For the usual assumptions made in this paper of diminishing returns to scale in investment, in skill, and in health, of the health-production function $f_H(t)$, the third term in (124) $q_{h/a}(R)[\partial f_H(R_+)/\partial H - \partial f_H(R_-)/\partial H]$ is positive, increasing in investment $I_H(t)$ and in skill $E(t)$, and decreasing in health $H(t)$. Since investment in health increases in wealth and skill, the term is larger for the wealthy and the skilled, and smaller for the healthy. The fourth term in (124) $(\partial s(R)/\partial H)w(R)[1(\lambda_{\tau_w}(R_+)/q_A(R)]$ is negative, increasing in skill (through the wage rate), and decreasing in health (through diminishing returns from reduced sick time). Thus, combining these, the sixth term is more likely to be positive for wealthy individuals, encouraging early retirement (but we cannot establish the overall sign of the sixth term for endowed skill and endowed health).

In sum, the "longevity" effect encourages individuals with more endowed wealth, skill, and health to retire later, while the "wealth" effect encourages them to retire early. The "value of health" effect encourages wealthy and skilled individuals (who value health more) to retire early, but the healthy potentially value health less, encouraging them to retire later. Last, the "health" effect encourages the wealthy to retire later but is ambiguous for skill and health endowments. While the combined result is thus ambiguous, each of these effects represents a potential pathway that may be empirically evaluated to explain retirement behavior.