Strulik, Holger

**Working Paper**

# Health and Education: Understanding the Gradient

Diskussionsbeitrag, No. 487

**Provided in Cooperation with:**
School of Economics and Management, University of Hannover

*Suggested Citation:* Strulik, Holger (2011) : Health and Education: Understanding the Gradient, Diskussionsbeitrag, No. 487, Leibniz Universität Hannover, Wirtschaftswissenschaftliche Fakultät, Hannover

This Version is available at:
https://hdl.handle.net/10419/73105

# Health and Education: Understanding the Gradient

Holger Strulik*

Leibniz Universitat Hannover, Discussion Paper No. 487

ISSN 0949-9962

December 2011

**Abstract.** This study presents a novel view on education and health behavior of individuals constrained by aging bodies. The aging process, i.e. the accumulation of health deficits over time, is built on recent insights from gerontology. The loss of body functionality, which eventually leads to death, can be accelerated by unhealthy behavior and delayed through health expenditure. The proposed theory rationalizes why better educated people optimally choose a healthier lifestyle, that is why they spend more on health and indulge less in unhealthy behavior. The model is calibrated for the average male US citizen. In the benchmark case a difference of the return to education that motivates one year more of education motivates also about 8 percent less unhealthy behavior and 5 percent more health expenditure and thus explains half a year gain of longevity. Progress in medical technology explains why the education gradient gets larger over time.

*Keywords:* Schooling, Aging, Longevity, Health Expenditure, Unhealthy Behavior, Smoking.

*JEL:* D91, J17, J26, I12.

---

*University of Hannover, Wirtschaftswissenschaftliche Fakultät, Königsworther Platz 1, 30167 Hannover, Germany; email: strulik@vwl.uni-hannover.de. I would like to thank Carl-Johan Dalgaard, Gustav Feichtinger, Leonid and Natalia Gavrilov, Franz Hof, Wolfgang Lutz, Arnold Mitnitksi, Alexia Prskawetz, Patrick Puhani, Fidel Perez Sebastian, and Engelbert Theurl, for useful comments. This research was funded by the European Commission within the project "Long-Run Economic Perspectives of an Aging Society" (LEPAS) in the Seventh Framework Programme under the Socio-economic Sciences and Humanities theme (Grant Agreement: SSH7-CT-2009-217275)

## 1. Introduction

Better educated individuals are, on average, healthier and die later than less educated ones. The literature refers to this strong positive association between education and health as the education gradient or just "the gradient". According to one popular study, in the year 1990, US Americans aged 25 with any college education could expect to die 5.4 years later compared to those with only high school or less. By the year 2000 the gap increased to 7.0 more years for the better educated (Meara et al., 2008). In many other countries a similar association between education and health has been observed.[1]

One obvious explanation for the gradient is that the better educated care more about their health. They spend more on preventive care, smoke less, are less obese, and display " healthier behaviors along virtually every margin" (Cutler and Lleras-Muney, 2006). But then, of course, the question arises why do they do that? In particular the fact appears puzzling that the less educated, who are presumably less wealthy, spend more on costly unhealthy activities like smoking and eating a lot.

So far, the literature has suggested three different kinds of deeper explanations for the gradient: common third factors, productive efficiency and allocative efficiency. The "third factor" argument is based on the impact of general attitudes on behavior and becomes particularly intuitive if one thinks of time preference. More patient persons are presumably more willing to delay entry into the workforce for education as well as they are more willing to sacrifice pleasure from unhealthy consumption in exchange for a longer a life. The problem is that, empirically, general attitudes seem to play only a minor role for educational differences in health behavior. Cutler and Lleras-Muney (2010) estimate that attitudes like time preference account for about 10 percent of health behavior, similar to the contribution of health knowledge, whereas income (access to resources) and cognitive ability account for the greatest shares, each for about 30 percent.

The idea of productive efficiency is based on Becker's (1965) commodity theory. It postulates that less educated individuals "produce" less health out of any given inputs of time and medical care (see e.g. Grossman, 1972, 2000). Allocative efficiency, with contrast, puts the emphasis on

[1] An incomplete list of the literature on the gradient includes Elo and Preston (1996), Contoyannis and Jones (2004), Case and Deaton (2005), Lleras-Muney (2005), Mackenbach et al. (2008), Conti et al. (2010), and Cutler and Lleras-Muney (2010). See Grossman (2006), Cutler and Lleras-Muney (2006), and Cutler et al., (2011) for surveys of the by now large literature. See Glied and Lleras-Muney (2008) and Cutler et al. (2010) on the rising education gradient.

the inputs and suggests that less educated individuals use different inputs, presumably because they are less well informed about their "health technology" (see e.g. Kenkel, 1991). The common theme of both ideas is that less educated people behave less efficiently. If they had only access to the health technology and the knowledge of the better educated, they would care more about their health and live longer.

Acknowledging that the so far available theory certainly has a role in explaining the education gradient, the present paper offers an alternative, novel theory. Inspired by the empirical power of cognitive ability in accounting for health behavior it asks the following question. Assume that individuals share the same attitudes (preferences) and share the same allocative and productive efficiency, namely they are fully rational and perfectly foresighted. Assume that they face different returns to education. How much of the observable education gradient can then be explained by their individual-specific return to education and the implied optimal education and health behavior?

One obvious explanation for idiosyncratic differences of the return to education across a population is cognitive ability. Smarter people expect a higher payoff from further education and thus educate more (see e.g. Heckman and Vytlacil, 2001). But, of course, the return of education is potentially influenced by other factors as well, for example, by family background and school quality (e.g. Card, 1991). Since the theory cannot distinguish between these factors, the driver of the education gradient is "only" identified as idiosyncratic differences in the return to education, although the interpretation as cognitive ability is tempting in light of the above mentioned empirical evidence (see also Kaestner and Callison, 2011).

More specifically, we consider intertemporally optimizing individuals who decide about education and health behavior knowing precisely how education will affect their future salaries and how health behavior will affect their health and, eventually, the time of their death. Figuratively speaking individuals can decide about how healthy their *lifestyle* should be. At the extremes they could choose a rock'n'roll lifestyle – experience a lot of pleasure from unhealthy consumption and die early – or the lifestyle of a saint – abstain completely from unhealthy activities and die late. The average person, that is the "Reference US American" in the calibration of the model, of course, prefers an intermediate lifestyle.

Methodologically the present article is related to the literature on optimal health spending and longevity, to which Ehrlich and Chuma (1990) and Hall and Jones (2007) are presumably

the most popular contributions. With contrast to that literature, the present article considers education and unhealthy consumption as individual choice variables. More importantly the present article has the distinction of being solidly built upon recent research in gerontology. So far the literature has treated health like human capital, that is as a stock which can be accumulated by investment and which depreciates over time ($\dot{H} = I_H - \delta H$). Without further amendments this means that health depreciation is greater when the stock of health is large, that is when individuals are relatively young and healthy. Preserving health would thus require health expenditure to be high at young age and low at old age (see also Case and Deaton, 2005 for a critique of the conventional approach). In order to counteract this problem, the literature has assumed that the depreciation rate is increasing with age (Grossman, 1972). This however, entailed the next problem, namely to identify the functional form of the age-dependent depreciation rate.

The present article turns the problem upside down by stating that it is not health but *health deficits* $D$ which are accumulated with age, $\dot{D} = \mu D - E$. This means that the arrival of new health deficits increases with the number of deficits that a person already has. The law of deficit accumulation has a gerontological foundation. As suggested by McFadden (2005) it is built upon an application of reliability theory to the function of the human body (see Gavrilov and Gavrilova, 1991). In order to model deficit accumulation the present article employs the so called frailty index, established by Mitnitski and Rockwood and several coauthors in a series of articles (2002a, 2002b, 2005, 2006, 2007). The frailty index counts the proportion of the total potential deficits that an individual has, at a given age. The list of potential deficits ranges form mild ones like impaired vision to severe ones like dementia. Using the frailty index Mitnitski and Rockwoo d estimate with an $R^2$ around 95 percent the rate $\mu$ at which health deficits are accumulated. In developed countries the average adult individual accumulates 3-4% more deficits from one birthday to the next. The great precision of the gerontological estimates allows a reliable calibration of the present theory.[2]

The present study follows Dalgaard and Strulik (2010) by assuming that the factor $E$ in the law of deficit accumulation, which operates to slow down the aging process, is partly explained by health expenditure. Additionally we will consider that unhealthy behavior speeds up the aging process (by reducing $E$). Unhealthy behavior and the consideration of an endogenous schooling

[2]Recently, Chatterji et al. (2011) have shown a strong negative association between education and health deficits by computing the frailty index for 15 European countries and a population stratified by educational attainment.

decision are the main differences to Dalgaard and Strulik (2010). The present article can thus be seen as the logical follow up study. While Dalgaard and Strulik focus on the income gradient (the Preston curve) ignoring education, the present study considers the education gradient and its origin from health spending and unhealthy behavior.

The article is organized as follows. The next section sets up the model. The standard approach of human capital accumulation (Mincer, 1974) is modified to account for the aging body and mind and then integrated together with the opportunity of unhealthy consumption into a gerontologically founded life-cycle model. Section 3 presents analytical results. It shows that the optimal life-style is governed by conditions for (i) optimal expenditure profile (on health and unhealthy goods), (ii) optimal aging (the evolution of the expenditure profile with age), (iii) optimal schooling, (iv) optimal financial management, and (v) optimal death. Most of these optimality conditions are simple enough to allow for an intuitive interpretation.

In Section 4 the model is calibrated for a 16 year old male US American in the year 2000. Section 5 present the results. The most interesting experiment is, of course, to vary the return to education, $\theta$. It is shown that an increase of $\theta$ that motivates one year more of education implies half a year longer life. The reason is health behavior. A better educated person has more precious human capital to preserve which causes him or her to indulge less in unhealthy consumption and to spend more on health. The gradient is mildly non-linear and increasing with education. Eight years more education than the Reference American imply an eight years longer life.

The study continues with a series of numerical experiments whether there are alternatives to the return to education. It is shown which channels, perhaps surprisingly, do not work (income, time preference, rate of aging) and which channels, perhaps equally surprisingly, can motivate an education gradient as well (occupation-specific loss of human capital with age, productivity growth). From an aggregate viewpoint the results suggest a link between longevity and education running in the reverse direction of the causation so far discussed in the macro-literature. The so far established (and debated) link runs from longevity to education and ignores health behavior (Ben-Porath, 1967; Cervelatti and Sunde, 2005, 2010; Hazan, 2009). Here, a higher return on education triggers education, provides an incentive to protect precious human capital with healthy behavior, and leads to a longer life.

After corroborating the main result with a series of robustness checks the article finishes by comparing the education gradient for voluntary and compulsory education and by showing that medical technological progress widens the education gradient. The reason is that better educated persons spend more on health and benefit thus more from (human capital preserving) new technologies.

## 2. Model Setup

2.1. **The objective function.** Consider a young adult at the end of the compulsory schooling period. Later on, in the numerically part, this will be a 16 year old with 9 years of education. For simplicity let the initial age at this stage be normalized to zero. The – yet to be determined – date of death is denoted by $T$. At each age $t$ the person experiences utility from consumption of health-neutral goods $c(t)$ and unhealthy goods $u(t)$. Without loss of generality (but more notational clutter) the health-neutral good good can be made healthy or a third, explicitly health enhancing good, could be introduced. Here we consider a minimum setup for the explanation of both health expenditure and unhealthy consumption. In a minimum setup it suffices to treat health expenditure as purely instrumental and not in itself utility-enhancing. Likewise, education is purely instrumental in achieving higher labor income and not for itself utility enhancing. Health deficits $D(t)$, however, are allowed to diminish utility. We follow Hall and Jones and let the health status enter utility additively to consumption. Let the instantaneous utility function be iso-elastic such that life-time utility is given by (1).

$$V = \int_0^T \mathrm{e}^{-\rho t} \left\{ v\left( \; c(t) + \beta u(t) \; \right) - \xi D(t)^{\nu} \right\} \; \mathrm{d}t \tag{1}$$

with $v(x) = (x^{1-\sigma} - 1)/(1-\sigma)$ for $\sigma \neq 1$ and $v(x) = \log(c)$ for $\sigma = 1$. The inverse $1/\sigma$ is the intertemporal elasticity of substitution. We measure consumption such that $x$ is always larger than one, implying that at each age utility is positive, a fact that makes a longer life, in principle, desirable.

The parameter $\beta$ measures how pleasurable consumption of the unhealthy good is. If $\beta > 1$, the person likes unhealthy consumption better than health-neutral consumption. If $\beta < 1$, the person prefers health-neutral consumption and unhealthy goods are consumed only because they are cheaper. Later on, the parameter $\beta$ is a useful device to model the price elasticity of demand for unhealthy goods. In the calibration, cigarettes will stand in for the unhealthy good.

2.2. **Budget constraint.** Let health expenditure be denoted by $h$. The price of health-neutral goods is normalized to unity, the price for health goods is denoted by $p$, and the price of unhealthy goods is denoted by $q$. Total expenditure is thus given by $e = c + ph + qu$. The – yet to be determined – length of the voluntary education period is denoted by $s$ and the predetermined age of retirement is at $R$. From $s$ to $R$ the individual receives a wage $w(t)$ per unit of human capital. The individual stock of human capital depends on education and age and is denoted by $H(s,t)$. For simplicity there are no restrictions on the capital market; the individual can borrow or lend at rate $r$. An individual that holds capital $k$ thus faces the budget constraint (2).

$$\dot{k}(t) = \chi w(t)H(s,t) + rk(t) - c(t) - ph(t) - qu(t) \tag{2}$$

with $k(0) = k_0$ and $k(T) = \bar{k}$. For simplicity we abstain from modeling a bequest motive such that $k_0$, and $\bar{k}$ as well as all prices are taken as given by the individual. In (2) $\chi$ is an indicator function, $\chi = 1$ if the individual is in work mode, that is for $t \in [s, R]$ and $\chi = 0$ otherwise. People are saving for consumption and health interventions after retirement. Although the arrival of health events is certainly stochastic, we follow the related literature (e.g. Ehrlich and Chuma, 1990, Hall and Jones, 2007) and treat, for simplicity, the problem deterministically. This means that the model neglects a precautionary savings motive and thus potentially underestimates the propensity to save for old age.[3]

2.3. **Health Deficit Accumulation.** Inspired by recent research in gerontology we model a physiologically founded aging process, according to which aging is understood as increasing loss of redundancy in the human body (Gavrilov and Gavrilova, 1991, Arking, 2006). When we are young the functional capacity of our organs is about tenfold higher than needed for mere survival (Fries, 1980). With preceding age and vanishing redundancy in our organism, we become more fragile. An empirical measure of human frailty has been developed by Mitnitski and Rockwood and various coauthors in a series of articles (Mitnitski et al, 2002a,b; 2005; Rockwood and Mitnitski, 2006). They propose to compute the frailty index as the proportion of the total potential health deficits that an individual has, at a given age. As suggested by aging theory,

[3]Strulik (2011) investigates a stochastic version of the simpler but structurally similar model of optimal aging by Dalgaard and Strulik (2010) and shows that the quantitative predications are relatively insensitive to the consideration of death as a stochastic event. While being more realistic the stochastic model is unattractive because many expressions can no longer be derived analytically and stated explicitly, a fact which delimits a rigorous economic understanding of the mechanisms at work.

Mitnitski et al. (2002a) confirm that the frailty index number (the number of health deficits), denoted by $D(t)$ increases exponentially with age $t$, $D\left(t\right) = E + b\mathrm{e}^{\mu t}$.

This "law of increasing frailty" explains around 95% of the variation in the data, and its parameters are estimated with great precision. Conceptually, the rate of aging $\mu$ is given to the adult individual. From a physiological viewpoint, however, it can be explained by applying reliability theory to human functioning (Gavrilov and Gavrilova, 1991). Mitnitski and Rockwood (2007) have obtained very similar estimates of $\mu$ for Australia, USA and Sweden (Rockwood and Mitnitski, 2007). In these four developed countries (in spite of differences in samples, the precise contents of the frailty index etc.) the average (male) individual accumulates about 3-4% more deficits from one birthday to the next.

In order to utilize the findings of Mitnitski and Rockwood for the present work we begin with differentiating the frailty law with respect to age, $\dot{D}\left(t\right) = \mu\left(D\left(t\right) - E\right)$. Integrating and using the initial condition $D(0) = D_0$ we get the solution $D(t) = \left(D_0 - E\right)\mathrm{e}^{\mu t} + E = D_0\mathrm{e}^{\mu t} - E(\mathrm{e}^{\mu t} - 1)$. From this expression it is obvious that a larger autonomous component $E$ implies less deficits for any given age $t$ and that the compound parameter $(D_0 - E)$ corresponds to Mitnitski et al.'s estimate of $b$. Following Dalgaard and Strulik (2010) we next introduce health expenditure by assuming that $E$ is amendable to change by way of deliberate health expenditure. Furthermore we assume that $E$ can be diminished by unhealthy behavior. Specifically, we propose the following parsimonious refinement of the process of deficit accumulation:

$$\dot{D}(t) = \mu\left[\ D(t) - a - Ah(t)^{\gamma} + Bu(t)^{\omega}\ \right], \qquad 0 < \gamma < 1, \qquad \omega > 1. \tag{3}$$

The parameter $a$ captures environmental influence on aging beyond the control of the individual (less pollution, say, implying a higher value for $a$), the parameters $A > 0$ and $0 < \gamma < 1$ reflect the state of the health technology, and $h$ is health investment. While $A$ refers to the general power of health expenditure in maintenance and repair of the human body, the parameter $\gamma$ specifies the degree of decreasing returns of health expenditure. The larger $\gamma$ the larger the relative productivity of cost-intensive high-technology medicine in maintaining and repairing highly deteriorated human bodies. Likewise, the parameter $B$ measures the general unhealthiness of the unhealthy good, the parameter $\omega$ measures the degree of "increasing returns" in terms of

deficits from excessive consumption, that is it helps to differentiate between health effects of another glass of red wine consumed by the occasional connoisseur and by the binge drinker.[4]

Initial frailty is given for the young adult, $D(0) = D_0$. Furthermore, following, Rockwood and Mitnitski (2006), we assume that a terminal frailty exists at which the individual expires, $D(T) = \bar{D}$. Problem (1) – (3) thus constitutes a free terminal time problem in which the terminal states $k(T)$ and $D(T)$ are known.

2.4. **Education.** The modeling of education introduces aging in a conventional Mincerian model of human capital accumulation (Mincer, 1974). Specifically we assume that human capital of an individual of age $t$ with $s$ years of schooling is given by (4).

$$H(s,t) = \bar{H}\mathrm{e}^{\theta s+\eta(t-s)-\alpha\mu t} - \delta D(t), \tag{4}$$

for $t > s$ and $H(t,s) = \bar{H} - \delta D(t)$ otherwise. The length of the schooling period $s$ is a choice variable for individuals. We can interpret the initial endowment $\bar{H}$ as predetermined skills and skills acquired through compulsory schooling. The parameter $\theta$ is the return to education and $\eta$ is the return to experience (learning on the job). For $\alpha = \delta = 0$ the schooling function boils down to the standard Mincer model (see e.g. Bils and Klenow, 2000).

The parameter $\alpha$ controls for the impact of the force of aging $\mu$ on cognitive skills and human capital. Most cognitive abilities start to decline between age 20 and 30 (Skirbekk, 2004). The modeling in (4) allows to discuss the impact of aging on cognitive skills and on (occupation-specific) human capital separately. It captures the fact that the rate of aging $\mu$ and the associated rate of cognitive skill loss appears to all individuals of a population alike and, in particular, independent from occupation whereas the *impact* of the skill loss on wages, measured by $\alpha$, is occupation specific (Skirbekk, 2004). For example, while aging economics professors probably experience the same loss of cognitive skills as aging race car drivers (and perhaps a similar decline of productivity, see Oster and Hamermesh, 1998), the skill loss has much less severe consequences on their salary. Since $\alpha$ is potentially job-specific it provides an explanation (beyond education)

[4]By way of contrast to $E$, the parameter $\mu$ – impressed by its empirical constancy across developed countries – is considered to be a physiological parameter. In the remaining we will refer to this physiological parameter as *the force of aging*, as it drives the inherent and inevitable process of human aging. In a science-fiction version of the model we could perhaps also address how health expenditure, for example through epigenetic regulation or hormone replacement therapy, affects $\mu$. In retrospect, however, there is so far very little evidence that "standard" medical treatments have substantially modified the *rate* at which our bodies decay (Gavrilov and Gavrilova, 1991).

why holders of certain occupations are healthier (Case and Deaton, 2005). Varying $\alpha$ constitutes an interesting experiment for the calibrated model.

The parameter $\delta$ controls the feedback of health deficits on human capital. It is also used to convert 'units' such that the outcomes from schooling and job experience can be summed up with physiological conditions to a unique human capital $H(s,t)$. The parameter $\delta$ is certainly job-specific as well. While all individuals age in the same way and – if they display the same health behavior – develop deficits in the same way, their health deficits may have different *impact* on their human capital. For individuals with jobs depending highly on "fluid abilities" like muscle function (a carpenter) we expect $\delta$ to be higher than for individuals with jobs depending highly on "crystallized abilities" like experience and wisdom (a priest).

## 3. Solution

3.1. **Summary.** The problem of the individual is to maximize (1) subject to (2) – (4). In order to solve the problem conveniently it turns out to be helpful to define a measure of aggregate consumption $d \equiv c + \beta u$ and replace $c$ in (1) and (2). Details of the computation are delegated to the Appendix. From the first order conditions we obtain that the optimal solution is characterized by the following, nicely interpretable conditions.

$$u^{\omega-1} = \begin{cases} \frac{\gamma(\beta-q)A}{\omega p B} \cdot \frac{1}{h^{1-\gamma}} & \text{for } \beta > q \\ 0 & \text{for } \beta \leq q \end{cases} \tag{5}$$

$$g_c \equiv \dot{d}/d = \frac{r-\rho}{\sigma} \tag{6}$$

$$g_h \equiv \dot{h}/h = \frac{r-\mu}{1-\gamma} \tag{7}$$

$$g_u \equiv \dot{u}/u = \frac{r-\mu}{1-\omega} \tag{8}$$

$$\bar{H}\mathrm{e}^{(\theta-\alpha\mu)s} - \delta D(s) = \frac{\mathrm{e}^{(\theta-\alpha\mu)s}}{\eta + g_w - r - \alpha\mu}\left[e^{(g_w-r+\eta-\alpha\mu)(R-s)} - 1\right]. \tag{9}$$

3.2. **Optimal Consumption Profile.** Condition (5) constrains the optimal consumption expenditure. Optimal expenditure on health and on unhealthy goods are negatively correlated because there are increasing return of damage from unhealthy consumption ($\omega > 1$) and decreasing returns from health expenditure in repairing damage ($\gamma < 1$). A person who spends more on health is predicted to indulge less in health-damaging consumption. For unhealthy consumption to occur at all, $\beta > q$. Given that the price of health-neutral consumption has

been normalized to unity, the condition requires that utility derived from a unit of unhealthy consumption exceeds that from health-neutral consumption ($\beta > 1$), or that the unhealthy good is cheaper than the health-neutral good ($q < 1$), or both. For policy it is important to note that demand of unhealthy goods is lower at higher prices and that there exists a preemptive price, $q = \beta$, which deters unhealthy consumption.

If unhealthy consumption is optimal, then condition (5) furthermore predicts that its incidence is large if medical efficiency in repairing damage is large ($A$ large), if the resulting health damage is low ($B$ is low), or if the price of health goods $p$ is low.

3.3. **Optimal Aging.** The Euler equations (6)–(8) show how optimal expenditure evolve through life. Together they determine optimal aging of the person since the evolution of health deficits $D$ depends on health behavior ($h$ and $u$). Condition (6) is the familiar Euler equation for consumption, here stated for the aggregate measure of consumption $d$. It has the usual textbook interpretation. Equation (7) is the "Health-Euler" (Dalgaard and Strulik, 2010). Similar to the "Consumption- Euler" it suggests to postpone expenditure for health to later periods of life in favor of financial investment if return on investment $r$ is relatively high. If, on the other hand, the force of aging, $\mu$ is high, implying that health deficits accumulate very fast at the end of life, late-in-life health investments are a relatively ineffective way of prolonging life. It is then optimal to invest more heavily early in life. The dynamic expenditure profile for health growth is also influenced by $\gamma$, which captures the curvature of the health investment function: a larger $\gamma$ implies a higher growth rate of health expenditure. Intuitively, if $\gamma$ is small, diminishing returns set in rapidly, which makes it optimal to smooth health expenditure to make the deficit-reducing effect as large as possible.

Although theory does not exclude the reverse, it makes sense in light of the empirical background to assume that $r$ exceeds $\mu$ such that health expenditure rises with age. Then, recalling that $\omega > 1$, condition (8) prescribes that expenditure for unhealthy consumption should decrease with age. Intuitively, the damage done by, for example, binge-drinking is relatively harmless at young age when there is still a lot of redundancy in the body. At an advanced age, binge-drinking could be deadly and the model recommends to reduce drinking to an occasional glass of red wine. The expenditure profile requires, according to (5), that unhealthy consumption is negatively correlated with health expenditure. This implies that the expenditure profile is steep (in absolute value) if the interest rate is high and the rate of deficit accumulation is low.

3.4. **Optimal Schooling.** The optimal length of the education period $s$ requires that the marginal loss from postponing entry in the labor market, $\mathrm{e}^{-rs}w(s)H(s,s)$, equals the marginal gain from extending education, $\int_s^R \frac{\partial}{\partial s}\mathrm{e}^{-rt}w(t)H(s,t)\mathrm{d}t$. Inserting the respective values leads to condition (9), in which $g_w$ denotes the growth rate of the wage per unit of human capital, $w(t) = \bar{w}\exp(g_w t)$. It is assumed to be given by the rate of aggregate productivity growth and to be exogenous to the individual. The left hand side of (9) displays the marginal loss from postponing entry and the right hand side is the marginal gain from education.

For the special case in which health does not matter for human capital, that is for $\alpha = \delta = 0$ condition (9) has an explicit solution for $s$.

$$s = R - \frac{\log \Psi}{r - g_w - \eta}, \qquad \Psi \equiv \frac{\theta - \eta}{g_w + \theta - r}. \tag{10}$$

This solution coincides with Bils and Klenow's (2000) result. In the present context the most interesting observation is that for $\alpha = \delta = 0$ the schooling period is independent from the length of life $T$. The right hand side of (10) consists of constants given to the individual. For the general case, however, $\alpha$ and $\delta$ are positive and education and life-length are interdependent through the rate of health deficit accumulation. At this stage however it is impossible to infer in which particular way education and health interact. This requires the (numerical) solution of the full model and will be done below.

3.5. **Optimal Financial Management.** Integrating (2) and inserting initial and terminal values provides the life-time budget

$$k_0 + W(s,R) - \frac{d(0)}{g_c - r}\left(\mathrm{e}^{g_c - r)T} - 1\right) - \frac{ph(0)}{g_D}\left(\mathrm{e}^{g_D t} - 1\right) - \frac{(q-\beta)u(0)}{g_\omega}\left(\mathrm{e}^{g_\omega T} - 1\right) = \bar{k}\mathrm{e}^{-rT} \tag{11}$$

with $g_D \equiv (\gamma r - \mu)/(1-\mu)$ and $g_\omega \equiv (\omega r - \mu)/(1-\omega)$. The expression $W(s,R)$ denotes life-time labor income (human wealth) acquired between leaving school and retirement. After some algebraic transformations it is obtained as (12).

$$\begin{aligned} W(s,R) = {} & \frac{\bar{w}\mathrm{e}^{(\theta-\eta)s}}{\eta + g_A - r - \alpha\mu}\left[\mathrm{e}^{(\eta+g_A-r-\alpha\mu)R} - \mathrm{e}^{(\eta+g_A-r-\alpha\mu)s}\right] \\ & - \delta\bar{w}\left[\mathrm{e}^{(\mu+g_A-r)R} - \mathrm{e}^{(\mu+g_A-r)s}\right]\left[\frac{D_0 + a}{\mu + g_A - r} + \frac{\mu A h(0)^\gamma}{g_D(\mu + g_A - r)} - \frac{\mu B u(0)^\omega}{g_\omega(\mu + g_A - r)}\right] \\ & - \frac{\delta\bar{w}\mu A h(0)^\gamma}{g_D(\mu + g_A - r + g_D)}\left[\mathrm{e}^{\mu+g_A-r+g_D)R} - \mathrm{e}^{\mu+g_A-r+g_D)s}\right] \end{aligned} \tag{12}$$

$$+\frac{\delta\bar{w}\mu Bu(0)^{\omega}}{g_{\omega}(\mu+g_A-r+g_{\omega})}\left[\mathrm{e}^{\mu+g_A-r+g_{\omega})R}-\mathrm{e}^{\mu+g_A-r+g_{\omega})s}\right]+\frac{\delta\bar{w}}{g_A-r}\left[\mathrm{e}^{(g_A-r)R-\mathrm{e}^{(g_A-r)s}}\right].$$

3.6. **Optimal Death.** At the individually optimal time of expiry two conditions have to hold. The accumulated health deficits must have reached the terminal value $\bar{D}$ and the Lagrangian associated with Problem (1)– (4) must assume the value zero. In other words, it is not worthwhile to live any longer. Turning towards the first condition, integrating (3) provides $D(T)$ and thus (13).

$$\bar{D}=D(T)=D_0\mathrm{e}^{\mu T}-a\left(\mathrm{e}^{\mu T}-1\right)-\frac{\mu Ah(0)^{\gamma}\mathrm{e}^{\mu T}}{g_D}\left(\mathrm{e}^{g_D T}-1\right)+\frac{\mu Bu(0)^{\omega}\mathrm{e}^{\mu T}}{g_{\omega}}\left(\mathrm{e}^{g_{\omega}T}-1\right). \quad (13)$$

Evaluation of the Lagrangian at $T$ provides (14).

$$\begin{aligned}0=v(T)+\xi\bar{D}^{\nu}+d(T)^{-\sigma}\left\{r\bar{k}-d(T)-(q-\beta)u(T)-ph(T)\right\} \quad (14)\\ -\frac{d(T)^{-\sigma}ph(T)^{1-\gamma}}{\mu A\gamma}\left\{-\mu a-\mu Ah(T)^{\gamma}+\mu Bu(T)^{\omega}+\mu D(T)\right\}\end{aligned}$$

with $d(T)=d(0)\mathrm{e}^{g_c T}$, $h(T)=h(0)\mathrm{e}^{g_h T}$, $u(T)=u(0)\mathrm{e}^{g_u T}$, and $v(T)=(d(T)^{1-\sigma}-1)/(-\sigma)$ for $\sigma\neq 1$ and $v(T)=\log(d(T)$ otherwise. Together, (5), and (9) – (14) establish 5 equations in 5 unknowns: the initial values $d(0), h(0)$, and $u(0)$, optimal education $s$, and the optimal age of death $T$.

## 4. Calibration

In the following calibration study we consider an average 16 year old male US American in the year 2000. The initial age is set to 16 years, corresponding to model-age zero, because individuals below roughly the age of 16 are not subject to increasing morbidity (Arking, 2006) and are presumably not well described by the law of increasing frailty. Furthermore, in many states of the US as well as in many countries around the world schooling is compulsory up to an age of about 16. This means that there is not really an *individual decision* about education below this age. An implication is that the individual of model age zero has spent already 9 years on compulsory education. The effect of compulsory education is captured by the initial endowment $\bar{H}$.

In order to calibrate the model to US data we begin with employing the Health Euler (7). From the data in Keehan et al. (2004) we put the growth rate of health expenditure over the life cycle $g_h$ to 0.021. From Mitnitski et al. (2002a) we take the estimate of $\mu=0.043$ for

(Canadian) men, and finally we put $r = 0.06$ (e.g., Barro et al., 1995). This produces the estimate $\gamma = 1 - (r - \mu)/g_h = 0.19$, which squares well with the independent estimates obtained by Hall and Jones (2007).[5]

In the year 2000 the average life-expectancy of a 20 year old male US Americans was 75.6 years. From Mitnitski et al.'s (2002a) regression analysis we infer terminal health deficits $\bar{D} = D(75.6) = 0.1005$ and initial health deficits $D(0) = D(16) = 0.0261$. In order to get an estimate of $a$ we assume that before the onset of the 20th century the impact of medical technology on *adult* mortality was virtually zero. In the year 1900 the life expectancy of a 20 year old U.S. American was 62 years (Fries, 1980). Implying that a 16 year old expects to live for 46 more years. We set $a$ such that a person who abstains from unhealthy consumption and has no access to (effective) life prolonging technology expects $T = 46$. From this value we get the estimate $a = 0.01427$.

The parameters entering the equation for optimal education (9) are potentially individual-specific and varying their size is our most interesting numerical experiment. In the following we try to fix parameter values for the average US American. For the return to education there exists a variety of estimates, depending on method and sample, but a consensus value in recent estimates for the average return to education in the US seems to be 0.1 (Psacharopoulos and Patrinos, 2004). We thus put $\theta = 0.1$ and, following Bils and Klenow (2000), $\eta = 0.05$. We set the growth rate of the wage per unit of human capital $g_w$, that is aggregate productivity, to an annual rate of one percent based on US TFP growth in 1995-2000 (Jorgenson et al., 2008). Finally we adjust $\alpha$ and $\delta$ such that (i) labor income peaks at age 55 (as observed by French, 2005) and (ii) the representative individual optimally chooses $s = 4.5$ years of extra education, implying in total 13.5 years of education, which corresponds with the US average in the year 2000 (Turner et al., 2007). This leads to the estimates $\alpha = 1.058$ and $\delta = 11.2$.

We normalize $\bar{H} = 1$, set $R = 48 = 64 - 16$ corresponding to the average US retirement age, and adjust the initial unit wage $\bar{w}$ such that total labor income across all working ages equals \$35,320, that is the average annual pay for workers in the year 2000 (BLS, 2011). This implies $\bar{w} = 29170$. In the basic run we put $k_0 = \bar{k} = 0$. Later on we investigate the impact of

[5]As explained in Section 2: the force of aging within the US and Canada are similar (Rockwood and Mitnitski, 2007). Thus, using the estimate from the Canadian sample should be a good approximation. While Rockwood and Mitnitski (2007) stress the similarity of their results for US and Canadian populations they do not report the detailed results for their US analysis, for which reason we are forced to rely on the results from the Canadian sample.

bequests and inheritances on education and health behavior as numerical experiments. We will also investigate later on the impact of a rudimentary public health system. In the basic run, however, the government plays no role.

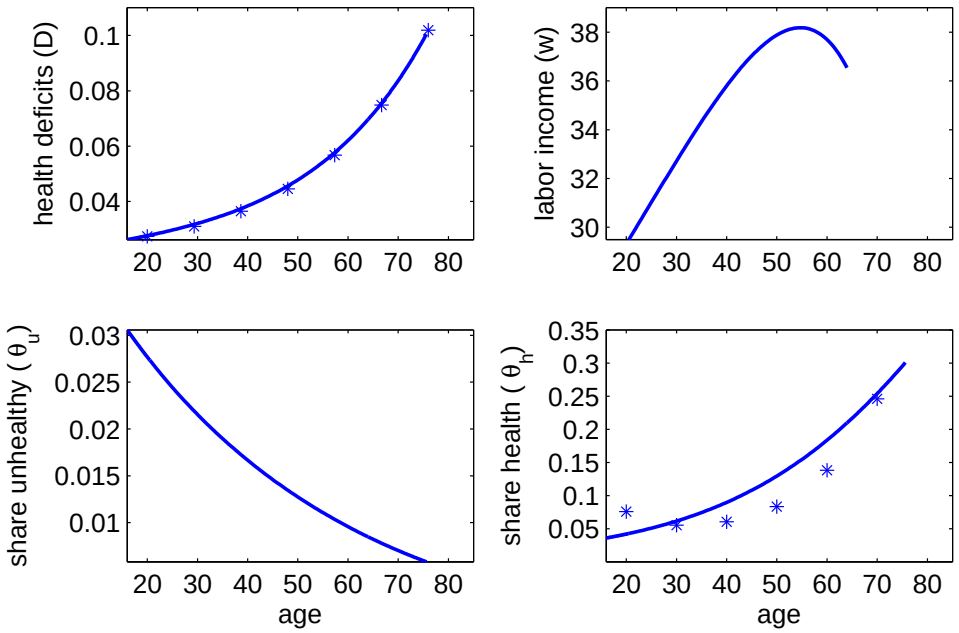
Following Hall and Jones (2007) the benchmark calibration treats health expenditure purely instrumental in prolonging life, that is $\xi = 0$. Later on, it will be shown that results are relatively insensitive to the consideration of health status in the utility function. We normalize the price of health $p$ to unity and experiment with alternative values of the price of the unhealthy good. For that purpose we exploit the fact that $q$ enters equations (5), (9), and (11)-(14) only in form of the compound $(\beta - q)$. We thus proceed in the following way. We first calibrate $(\beta - q)$ and the remaining parameters, $A$, $B$, $\rho$, $\sigma$ and, $\omega$ and then adjust $q$ in order to capture a particular price elasticity of demand for the unhealthy good. We adjust the remaining parameters such that (i) the model predicts the actual accumulation of health deficits over life (as estimated by Mitnitski and Rockwood, 2002), (ii) such that death occurs at the moment when $\bar{D}$ health deficits have been accumulated at an age of 75.6 years, (iii) such that the health share of total expenditure approximates average age specific expenditure shares of American adults, and (iv) such that consumption of the unhealthy good costs 2.5 years of longevity.

These 2.5 lost years are explained as follows. Most of the available empirical literature on the consumption of unhealthy goods is about cigarettes and tobacco. It thus seems reasonable to capture the characteristics of cigarette consumption in a benchmark case and then proceed with sensitivity analysis. Preston et al. (2010) estimate that smoking takes away 2.5 years of life-expectancy of 50 year old US males. To let the model produce this particular unhealthiness of consumption we proceed iteratively. After running a particular specification of the model we compute how damaging the unhealthy behavior actually was by integrating its effect on deficit accumulation. From that we can compute the counter-factual, that is how many extra years the individual would lived without the unhealthy behavior. Parameters are adjusted until the model predicts the loss of 2.5 years. This leads altogether to the following estimates: $(\beta - q) = 3.3$, $B = 1.7 \cdot 10^{-7}$, $\omega = 1.4$, $A = 0.001645$, $\rho = 0.085$ and $\sigma = 1.18$. The estimate of $\sigma$ fits nicely with recent empirically studies suggesting that the intertemporal elasticity of substitution is around unity (e.g. Chetty, 2006).

Finally we adjust $q$ (and thus $\beta$). From Chaloupka and Warner (2000) we know that the price elasticity of demand for cigarettes is most likely between $-0.5$ and $-0.3$ and we alternatively

match both values. In our first scenario we consider a price elasticity $\epsilon_u$ of $-0.3$. This leads to the estimate $q = 1.08$ and thus $\beta = 4.3$.

FIGURE 1: OPTIMAL SCHOOLING AND AGING: BASIC RUN




Solid lines: basic run. Parameters: $a = 0.01427$, $A = 0.00165$, $\alpha = 1.058$, $\delta = 11.2$, $\eta = 0.05$, $\theta = 0.1$, $\bar{H} = 1$, $\bar{w} = 29170$, $g_A = 0.01$, $\mu = 0.043$, $r = 0.06$, $\rho = 0.088$, $\sigma = 1.18$, $D_0 = 0.0274$, $\bar{D} = 0.10$, $p = 1$, $\beta = 4.38$, $q = 1.08$, $B = 1.7 \cdot 10^{-7}$, $\omega = 1.40$, $k_0 = \bar{k} = 0$. Stars: data, red dot: US mean male age and age-structure-weighted expenditure share. Wages in thousands, wealth (capital) in hundred thousands.

Figure 1 shows the implied trajectories over the life cycle of the Reference American. Stars in the health deficit panel indicate the actual estimates of Mitnitski and Rockwood (2002). The model explains the actual accumulation of health deficits quite well. Stars in the lower right panel show the actual age-specific share of health expenditure. The health expenditure data is taken from Meara et al. (2004) and the data for total expenditure is taken from BLS (2002). Both are from the year 2000. The health data, however, is per person and the original expenditure data is per household. I thus converted the household data into age-specific consumption per adult by crudely following Deaton (1997) and computing equivalence scales. Specifically I assigned household members under 18 a weight of 0.5 of adult consumption (The BLS data does not differentiate between children of different ages in the household). The model mildly underestimates health expenditure shares at young ages and overestimates it at middle ages. The main reason for this is that the model predicts somewhat too little curvature for labor income over age. Altogether, however, the predicted health fits the data reasonably well.

The lower left panel of Figure 1 shows the expenditure share of unhealthy consumption $\theta_u$ for the Reference American. When young he is predicted to spend about 3 percent on unhealthy goods. The expenditure share is declining until death. On average the reference American spends $ 403 per year on unhealthy consumption, a figure that squares reasonably well with the $ 319 that Americans spent on average for cigarettes in the year 2000 (BLS, 2002).

The upper right panel displays the calibrated invertedly-u-shaped trajectory of labor income across ages. The model predicts that a 20 year old receives about 77 percent of the peak wage and a 64 year old about 96 percent. These figures match the OECD average about right. In the US there is actually a bit more curvature of the wage for age curve (a 20 year old gets about 60 percent of the peak wage (OECD, 1998). Gourinchas and Parker (2002) estimate that peak wages – there obtained at about age 50 – are 1.35 times wages at age 25 whereas the presents model predicts a factor of 1.25.

Finally we calculate the model's predictiong for the value of life of the Reference American. Following the established methodology we divide instantaneous utility by the unit value of an "util", $\partial v/\partial c$, before summing up to get the value of life-long utility in monetary terms:

$$\begin{aligned}\tilde{V} &= \int_0^T \frac{v(c,u)}{\partial v/\partial c} \mathrm{e}^{-\rho t} \\ &= \frac{[c(0)+\beta u(0)]^{\sigma}}{1-\sigma} \left\{ \frac{[c(0)+\beta u(0)]^{1-\sigma}}{\rho - g_c} \left(1 - \mathrm{e}^{-(\rho - g_c)T}\right) - \frac{1}{\rho - \sigma g_c} \left(1 - \mathrm{e}^{-(\rho - \sigma g_c)T}\right) \right\}.\end{aligned}$$

The benchmark calibration predicts a value of life of $ 5.7 million for a 20 year old, a value that corresponds well with Murphy and Topel's (2006) estimate of $ 6.3 million for the value of a statistical life.

## 5. Results

5.1. **The Education Gradient.** The major interest of this paper is to identify the education gradient. The first experiment thus considers a person endowed with a higher return to education. The experiment is inspired by research in labor economics which acknowledges that the "return to education is not a single parameter in the population, but rather a random variable that may vary with other characteristics of individuals" (Card, 1999). Since our experiment holds constant preferences (attitudes) and time (calender year), the variation of the return to

education could perhaps best be rationalized as originating from a variation in cognitive ability, implying that individuals who are more or less able than the average take up more or less education (Heckman and Vytlacil, 2001). Given this interpretation, the results presented below will fit nicely with the empirical evidence on cognitive ability and health behavior (Cutler and Lleras-Muney, 2010). But actually the model takes no stand in this regard. The variation in the return of education could equally well originate from family background or school quality, two alternatives suggested in the labor literature (Card, 1999).

Specifically, we increase $\theta$ such that the person is motivated to one more year of education. We then observe the implied optimal changes in behavior predicted by the model and their impact on life-length. Results are shown in the first row of Table 1. Ceteris paribus, the person educates a year longer when $\theta$ rises from 0.100 to 0.104. This motivates the person to reduce unhealthy consumption by 8.8 percent and increase health expenditure by 4.4 percent compared to the benchmark citizen. These values are calculated on the basis of average expenditure on the respective good over the life-cycle. As a consequence of the behavioral changes, the better educated person lives half a year longer.

The result accords well with the empirical observation that education is positively associated with health through behavior as well as through income (sometimes called resources or access to health in the literature) and that both channels are about equally important (Cuter and Lleras-Muney, 2010). The model helps to identify causality, a problem, which has tormented the related empirical literature. The mechanism goes as follows. Higher cognitive skills make education more worthwhile. Better educated persons are endowed with more "precious" human capital, which they care more to protect by indulging less in unhealthy consumption and by spending more on health. Consequently they live longer.

The effect of education on life-length is non-linear. If $\theta$ is reduced from 0.1 to 0.096, as shown in the second row of Table 1, the person is motivated to attain school for one year less and as a consequence of the entailed behavior he or she lives 0.40 years shorter. Again, the explanation is that less "precious" human capital has been accumulated, which makes health expenditure less essential and pleasure from unhealthy behavior less costly in terms of human capital.

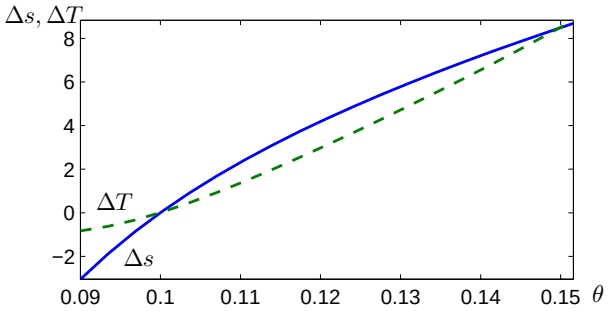
Figure 2 investigates this non-linearity more closely. It shows the desired extra years of schooling and the associated length of life motivated by alternative returns on education. Whereas the response of schooling is, naturally, concave, the response of life-length is mildly convex.

Table 1: The Education Gradient: Alternative Mechanisms

| $\Delta s$ | $\Delta T$ | $\Delta u/u$ | $\Delta h/h$ | par. change |
|---|---|---|---|---|
| $+1$ | $+0.50$ | $-8.8$ | $+4.4$ | $\theta = 0.104$ |
| $-1$ | $-0.40$ | $+7.9$ | $-3.2$ | $\theta = 0.096$ |
| $+1$ | $+1.9$ | $-26$ | $+17$ | $\alpha = 0.930$ |
| $-1$ | $-1.9$ | $+31$ | $-12$ | $\alpha = 1.185$ |
| $+1$ | $+4.2$ | $+57$ | $-1.8$ | $\mu = 0.0038$ |
| $-1$ | $-4.4$ | $-30$ | $+1.7$ | $\mu = 0.0048$ |
| $+1$ | $+1.2$ | $-18$ | $+10$ | $g_w = 1.47\%$ |
| $-1$ | $-1.1$ | $+19$ | $-8.1$ | $g_w = 0.51\%$ |

$\Delta T$ measured in years, $\Delta u/u$ and $\Delta h/h$ measured in percent.

Figure 2: Response of Schooling and Health to the Return to Education

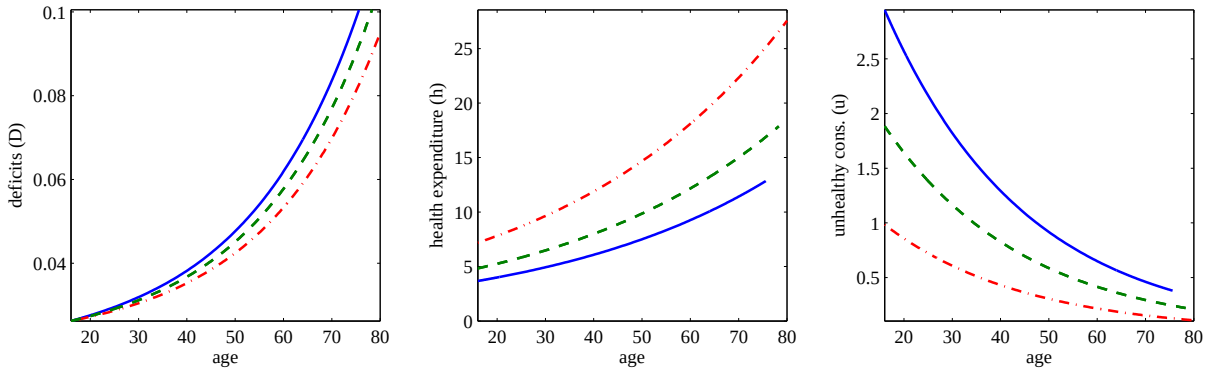



Schooling and lifespan relative to basic run for alternative rate of returns on education.

This implies an increasing education gradient. The more people educate the larger is the return in terms of longevity. Eight years of education on top of the 13.5 from the basic run, that is basically a PhD degree, is predicted to result in a about eight more years of life. This result fits nicely with Cutler and Meara's (2010) observation that in 1990 a 25 year college graduate could expect 8 years longer than a high school dropout of the same age (referring to Richards and Barry, 1998).

The associated evolution of health and health behavior is presented in Figure 3. It displays optimal age-trajectories for the Reference American (blue lines), another person endowed with $\theta = 0.12$ who takes up four more years of education (green lines), and a third person endowed with $\theta = 0.14$ and seven years more education (red lines). The better educated persons display at any given age a better health status (less deficits). The health differences are explained by health behavior. The better educated persons spend more on health and less on unhealthy consumption at any given age.

Figure 3: Education, Health, and Health Behavior




Health expenditure and unhealthy expenditure in thousands. Blue (solid) lines: basic run (Figure 1). Green (dashed) lines: $\theta = 0.12$ (four more years education). Red lines: $\theta = 0.14$ (seven years more education).

On the aggregate level the result suggests that the secular increase of the average return to education over the last decades (Katz and Autor, 1999) may have had a causal impact on the simultaneously observed increase in life-expectancy (Oeppen and Vaubel, 2010). This means there exists a mechanism that links adult longevity and education that reverses the causality of the channel so far proposed in the literature (Ben Porath, 1967, Cervelatti and Sunde, 2005). The empirical fact that the return to education increased more for persons with high cognitive skills and for occupations requiring a lot of cognitive skills (Murnane et al., 1995) may have contributed to the disproportionate increase of longevity for the well educated.

Before we begin to investigate other potential drivers of the education gradient, it is worthwhile to note that two seemingly natural candidates are already excluded by theory, namely the time preference rate, $\rho$, and income income for given education, that is $\bar{w}$. The reason is that both parameters – while having a strong impact on life-length – leave education unaffected. To see this, reconsider equation (9) which determines optimal $s$ and conclude that it is independent from $\bar{w}$ and $\rho$. The independence of the schooling decision from time-preference and the level of wages is not a particularity of the current approach but a standard result from the literature on optimal education (see e.g. Bils and Klenow, 2000, Card, 1999).[6]

The next couple of rows in Table 1 investigates the impact of the rate of "aging of human capital", $\alpha$. Since the parameter measures how important the decline of (cognitive) ability is

[6]For the time preference rate this statement is actually not exactly true because time preference affects the desired length of life and thus health expenditure. Through this channel it affects the number of bodily deficits accumulated when the persons leaves school $D(s)$. The size of this indirect effect, however, is insignificant. This suggests that if there is an observable association between education and time preference it is driven by reversed causation, that is more patience acquired through longer schooling .

for labor income, it can be probably best be thought of as identifying occupation, varying from pilots and race car drivers to professors and artists. The parameter is a very powerful predictor of life-length but to a somewhat lesser degree of education. Inverting the result from the Table, it means that one additional year of life is associated with $0.6 = 1/1.67$ more years of education. Diagrammatically, the gradient is "too steep", which means it leaves a lot of the variation in education unexplained.

This observation is even more true for the experiment, which investigates the impact of the rate of aging $\mu$. The rate of aging is a very powerful determinants of longevity. But it is of little help in explaining the education gradient. According to the aging channel there is too much longevity associated with an extra year of education. More importantly, the aging channel gets health behavior wrong. A generally more healthy individual (with lower $\mu$) uses these "health reserves" to indulge more in unhealthy consumption and spend less on health. More education and longevity are – counter-factually – associated with *more* unhealthy less health expenditure. Similar results are obtained for idiosyncratic differences of initial health deficits $D(0)$. These results are very intuitive. Generally healthier persons allow themselves more pleasure from unhealthy consumption with less serious consequences on longevity. Physiologically frail persons, on the other hand, stay away from deliberately damaging their bodies and spend more on health. For the identification of the education gradient, however, these results mean that the physiological parameters are useless in motivating the education gradient through health behavior.

The final rows in Table 1 considers an interesting *economic* channel. Returning to equation (9) we see that, while the level of income is irrelevant for the schooling decision, the growth rate of income (of productivity), $g_w$, matters. Economic growth devalues the costs of not working today and increases the benefit of working tomorrow. At higher rate of economic growth delaying entry into the work-life in favor for additional education becomes less costly and the person educates more. Inspecting (9) we see also that, once individuals are working, economic growth operates formally like experience on the job. It makes human capital more worthwhile. Consequently individuals spend more on health and indulge less in unhealthy behavior. An increase of productivity growth from 1.0 to 1.47 percent triggers one year more of education and behavioral changes that enable the person to live about one year longer.

On the aggregate level the result is interesting since it motivates a causal effect from economic growth to education and longevity while the so far established literature has focussed on causality in the oppositive direction (e.g. Cervelatti and Sunde, 2005). The induced behavioral changes are also interesting in light of the discussion of the moral consequences of economic growth (Friedman, 2005). The problem is, however, that aggregate productivity growth varies within a limited range, which makes it impossible to motivate a large and further increasing education gradient. In order to exploit productivity growth to rationalize the education gradient it seems more reasonable to consider occupation-specific growth of productivity. In this context, the model predicts that people are more motivated to educate longer in order to get an occupation in a high growth sector of the economy and then to behave in order to protect their human capital, which in turn raises their longevity.

5.2. **Robustness of Results.** Table 2 displays some robustness checks. It focusses on the return to education (cognitive ability), which appears to be the most convincing driver of the education gradient. The first two rows, case 1 and 2, document that the education gradient does not operate trough (education unrelated) income. Results are shown when, ceteris paribus, the annual wage per unit of human capital is assumed to be 50 percent higher or lower. These income variations have very strong effects on longevity, which rises by 4 years or, respectively, falls by 7 years compared to the basic run. The gradient, however, that is the gain in longevity that is associated with one more year of education, is not much affected. The gradient is generally a bit higher for poorer individuals.

During the education period the Reference American voluntarily acquires 4.5 years of education and accumulates debt of about 150 k. This relatively high accumulation of debt in young ages is an unreasonable artefact originating from the assumption that the reference American does not benefit from supporting parents. In order to accommodate this criticism, case 3 endows the person with an inheritance of $5\bar{w}$, which is about four times the average annual labor income, a sum which can be regarded as sufficient to finance 4 or 5 years of voluntary education. Interestingly, the inheritance reduces the gradient by about one third to 0.33. The reason is that financial wealth, ceteris paribus, reduces the incentive to educate. The wealthy person prefers to finance larger parts of consumption and health expenditure in old age by returns on capital rather than human capital and savings from labor income. Case 4 requires, additionally, that

Table 2: The Education Gradient: Robustness Checks

| Case | $\Delta s$ | $\Delta T$ | $\Delta u/u$ | $\Delta h/h$ | comment |
|---|---|---|---|---|---|
| 1) $\Delta\bar{w} = +50\%$ | +1 | +0.44 | −10 | +5.0 | a richer individual |
| 2) $\Delta\bar{w} = -50\%$ | +1 | +0.51 | −6.1 | +2.8 | a poorer one |
| 3) $k_0 = 5\bar{w}$ | +1 | +0.33 | −7.8 | +3.7 | inheritance (parental support) |
| 4) $k_0 = 5\bar{w}$, $\bar{k} = 6\bar{w}$ | +1 | +0.32 | −7.7 | +3.7 | inheritance and bequest |
| 5) $\epsilon_q = -0.5$ | +1 | +0.50 | −8.8 | +4.4 | higher price elasticity of unhealthy good |
| 6) $B = 10^{-6}$, $\omega = 1.16$ | +1 | +0.65 | −17 | +3.6 | unit consumption more unhealthy |
| 7) $B = 10^{-8}$, $\omega = 1.77$ | +1 | +0.43 | −4.9 | +4.6 | unit consumption less unhealthy |
| 8) $\xi = 10$, $\nu = 0.77$ | +1 | +0.44 | −7.9 | +3.8 | health in utility (decreasing damage) |
| 9) $\xi = 1000$, $\nu = 2.70$ | +1 | +0.40 | −7.5 | +3.5 | health in utility (increasing damage) |
| 10) $\tau = 0.1$, $\psi = 0.67$ | +1 | +0.57 | −7.9 | +3.8 | public health system |
| 11) $\tau = 0.15$, $\psi = 0.77$ | +1 | +0.60 | −7.4 | +3.6 | more generous public health system |

In all cases the experiment increases $\theta$ from 0.1 to 0.104. In order to match the data, re-calibration for case 5: $q = 0.715$, $\beta = 3.95$. For case 8: $\sigma = 1.28$. For case 9: $\sigma = 1.40$. For case 10: $A = 0.00142$, $\alpha = 1.059$. For Case 11: $A = 0.00139$, $\alpha = 1.061$. All other parameters from benchmark case (Figure 1). See text for details.

the person leaves a bequest of $6\bar{w}$. It demonstrates that the mechanism runs mainly through the inheritance received rather then through the bequest left.

Case 5 in Table 2 adjusts prices and preferences such that the price elasticity of demand for the unhealthy good equals -0.5, an elasticity observed at the upper end of estimates from cigarette demand. The experiment keeps the difference $(\beta - q)$ from the basic model such that initial education and longevity are preserved. The result verifies the claim that it is indeed the preference-price differential rather than absolute values, which are driving the education gradient and health behavior.

With case 6 and 7 we investigate the unhealthy character of good $u$. Case 5 assumes that the good is much more unhealthy than cigarettes. $B$ is raised by about factor 10. The scale parameter $\omega$ is adjusted to 1.16 such that the calibration continues to produce education, expenditure, and longevity from the basic run. The greater unhealthiness of consumption of small quantities of the good implies a higher incentive for the better educated to stay away from this good. The experiment consequently predicts a larger education gradient.

Inverting the result from above, the model produces a smaller gradient if $u$ is generally less unhealthy. This is confirmed by case 7 in which $B$ has been reduced by about factor 10 and the scale parameter has been increased from 1.4 to 1.77 in order to match the data. The new parameter value reflects a good for which consumption of small quantities entails relatively small

effects on health whereas large excess consumption has severe consequences. It could perhaps be thought of as alcohol. The predicted education gradient is 0.43 and thus a bit smaller than for the basic run (representing cigarettes). The reason is that better educated individuals have less incentive to stay away from consuming the good.

Case 8 and 9 document that the education gradient is relatively robust against considering health deficits in the utility function. In order to calibrate the utility parameters $\xi$ and $\nu$ we employ an idea of Hall and Jones (2007) and associate age-specific flow utility with quality adjusted life years (QALYs). Specifically, we use the estimate from Cutler and Richardson (1997) that compared to a newborn the QALY of 20 year old is 94 percent and the QALY of a 65 year old is 73 percent. We then assume a particular level parameter $\xi$ and estimate $\nu$ such that utility $v\,(\ c(t)+\beta u(t)\ )-\xi D(t)^{\nu}$ evaluated at age 65 equals 73/94 times the utility at age 20. Furthermore $\sigma$ is adjusted in order to match the empirical data on deficits accumulation and health spending. For $\xi = 10$ this implies the estimate $\nu = 0.77$ and thus decreasing marginal loss of utility with increasing health deficits. For $\xi = 1000$ the corresponding estimate is 2.7, indicating increasing marginal loss of utility. In order to understand this results it is helpful to recall that health deficits, measured by the frailty index, are a number between zero and one. Irrespective of the shape of the utility function the education gradient is relatively unaffected and estimated to be a bit smaller than obtained for the basic run. The reason is the entailed larger estimate for $\sigma$, indicating more incentive to smooth consumption over life and thus a higher desire for a longer life for both well and less well educated persons.

Finally we integrate a rudimentary public health system into the analysis. For that purpose we assume that a government taxes labor income at rate $\tau$ and uses the revenue to finance a subsidy $\psi$ on the price of health. The budget constraint of the individual (2) is thus modified to (15).

$$\dot{k}(t) = (1-\tau)\chi w(t)H(s,t) + rk(t) - c(t) - (1-\psi)ph(t) - qu(t). \tag{15}$$

A balanced budget of the government implies that the present value of taxes collected over the lifetime equals the present value of health subsidies granted, that is

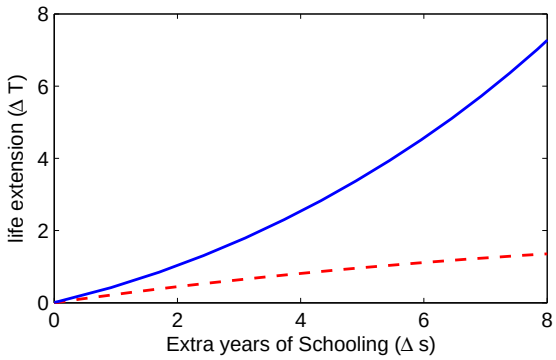
$$\tau w(s,R) = \frac{\psi p h(0)}{g_h}\left[\exp(g_h T) - 1\right] \tag{16}$$

with $w(s, R)$ denoting the present value of life time labor earnings of a person with $s$ years education, as established in (12).

The public health system distorts the price of health and labor income. The well educated suffer more from income lost to taxation but benefit more from a lower price of health because of their generally higher propensity to spend on health. The total effect of the public health system is thus not a priori obvious. Case 10 and 11 in Table 2 document that the price effect dominates for the calibrated model. The public health system over-proportionately encourages health expenditure of the well educated. It makes the health gradient larger. An income tax of 10 percent suffices to finance 2/3 of all health outlays occurring through a lifetime. A 15 percent tax finances 77 percent of all health expenditure. In both cases the education gradient increases to around 0.6.

5.3. **Voluntary vs. Compulsory Education.** Empirical studies trying to identify causality have looked at the effect of (the introduction or change of) compulsory education on health and life-expectancy. These natural experiments have been very helpful at low levels of education but their scope is, of course, limited by the length of the compulsory education period. They cannot be used to identify the education gradient for higher education. With the model at hand we can run the counter-factual experiment and introduce compulsory education up to the PhD level.

Figure 4: Longevity Gain: Voluntary vs. Enforced Education




The blue (solid) line reiterates $\Delta T$ from Figure 4 (life-time gain when $\theta \in [0.1,\ 0.15]$ and education is adjusted optimally. The red (dashed) line shows the life-time gain for fixed $\theta = 0.1$ and constraint optimization for given years schooling ($s$).

We discuss compulsory education by solving the individual problem as a constrained maximization problem, imposing that the person has to experience $s$ years of education on top of the 9 years initial endowment and on top of the 4.5 additional years that the Reference American

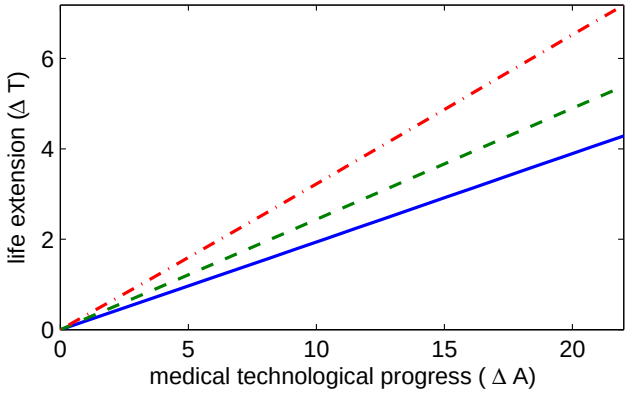
anyway wants to achieve. Compulsory education thus becomes *binding* after 13.5 years of education when it disables the optimality condition (9) but leaves otherwise all individual calculus intact. The blue (solid) line in Figure 4 re-iterates the education gradient for voluntary schooling from Figure 2, that is when more education is driven by a higher (individual-specific) return to education. A value of $\Delta s = 8$ corresponds with 21.5 years of education, that is, basically, a PhD degree.

The red (dashed) line shows the result when $\theta$ is kept at 0.1, and higher education is imposed. There is a gradient observable under compulsory education. But the perhaps more surprising result is that the gradient is comparatively flat. While the difference between the voluntary and compulsory gradient is not large for one or two years of extra education it widens as the compulsory education period becomes longer. The reason is that compulsory education leads to relatively less precious human capital than voluntary education. At the extreme, forcing a person who would prefer to enter the workforce after high school to do a PhD leads to relatively less payoff in terms of future wages. With human capital being less precious, the individual is less averse against unhealthy consumption, spends less on health, and lives less long compared to a person who decided for the long education on his or her own accord because of a high return on education.

5.4. **Medical Technological Progress and the Health Gradient.** In our last experiment we investigate the impact of medical technological progress on life-time extension and in particular on the education gradient. It has been hypothesized that the observable secular increase of the education gradient may have its origin in technological progress because better educated persons have better access (more resources) to utilize technological advances to their benefit (Cutler et al., 2010). The analysis presented in this paper has moreover suggested that better educated person, ceteris paribus, demand more health services in order to protect or repair their human capital. There exists thus a twofold motive for a widening education gradient.

The results presented in Figure 5 confirm these hypotheses. The Figure shows the longevity gains resulting from alternative increases of $A$, that is the power of health expenditure to reduce health deficits; $\Delta A$ is measured in percent of the benchmark run. If medical technological advances at an annual rate of 1 percent (3 percent) the level of $A$ is 20 percent higher after about 18 years (6 years). The solid line shows the predicted longevity for the Reference American (endowed with a return to education of $\theta = 0.1$). The dashed line shows the prediction for a

Figure 5: Medical Technological Progress and the Health Gradient




The Figure shows the gain in longevity for alternative progress of medical technology ($\Delta A$). Blue (solid): benchmark run (13.5 years of education, $\theta = 0.1$). Green (dashed): 4 years more of education ($\theta = 0.12$). Red (dash-dotted): 7.2 years more of education ($\theta = 0.14$) The longevity gain is measured relative to the own initial life-span for both types.

person with $\theta = 0.12$, which educates for 4 years longer and the dash-dotted line reflects longevity of a person with $\theta = 0.14$ and 7.2 more years of education. Although everybody experiences an increase in longevity, the predicted gain of the better educated persons is higher. When $A$ advances by 20 percent the longevity gap between a high school graduate (solid) and a college graduate (dashed) has widened by about 2 more years.

## 6. Conclusion

This study has proposed a new view on the education gradient. It has assumed away any explanation based on attitudes, non-cognitive skills, and allocative or productive inefficiency of the uneducated. Instead it has asked how large a gradient can be motivated by optimal decisions on education and health behavior of individuals who know precisely how their behavior affect their future health status and the time of death. The theory has been firmly built on insights from modern gerontology which allowed a robust calibration for a "Reference American". It predicts that a person whose return to education (cognitive skills) motivate one year more of education, spends more on health and less on unhealthy behavior such that he or she lives about half a year longer.

It has been shown that the result is robust against various modifications and extension of the basic model and that almost all other potential channels that the model allows to investigate can been excluded as an explanation of the education gradient. However, there is at least one other serious channel, which provided similar but somewhat less strong results, namely the

(occupation-specific) loss of human capital through aging. If human capital deteriorates slowly it is more worthwhile to educate in order to achieve it and it is more worthwhile to protect it with healthy behavior.

The theory suggests furthermore that the education gradient is mildly non-linear, increasing with the length of education (eight years more education imply eight years more longevity), and that medical technological progress can explain why the gradient gets larger over time. The suggested reason is that well educated persons demand relatively more health services in order to protect their precious human capital and thus benefit to a larger degree from health innovations.

The study has focussed on the human life cycle from young adulthood onwards. At this age, taking cognitive skills as approximately given is presumably a fair enough simplification. But since cognitive skills seem to be malleable at younger ages (see e.g. Heckman, 2006), the present study also highlights the importance of childhood development for later life. Equipped with a high return on education, individuals are not only predicted to educate longer and earn more labor income but also to lead a healthier life and to live longer. On the other hand, for a given low return on education the model supports also the choice of an unhealthy lifestyle as a rational best response. With little human capital to protect it makes sense to experience a lot of pleasure from unhealthy behavior in exchange for a shorter life.

## References


Acemoglu, D. and Johnson, S., 2007, Disease and development: the effect of life expectancy on economic growth, *Journal of Political Economy* 115, 925-985.

Altonji, J.G. and Dunn, T.A., 1996, The effects of family characteristics on the return to education, *Review of Economics and Statistics* 78, 692–704.

Barro, R., G. Mankiw and X. Sala-i-Martin, 1995. Capital mobility in neoclassical models of growth. *American Economic Review* 85, 103–115

Ben-Porath, Y., 1967, The production of human capital and the life cycle of earnings, *Journal of Political Economy* 75, 352–365.

Bils, M. and Klenow, P.J., 2000, Does schooling cause growth?,*American Economic Review* 90, 1160–1183.

BLS, 2002, Consumer Expenditures in 2000, United States Department of Labor, Bureau of Labor Statistics, Report 958.

BLS, 2011, State average annual pay for 2000 and 2001, United States Department of Labor, Bureau of Labor Statistics (`http://www.bls.gov/news.release/annpay.t01.htm`).

Brunello, G., Fort, M., Schneeweis, N., and Winter-Ebmer, R., 2001, The causal effect of education on health: What is the role of health behaviors, mimeo.

Card, D., 1999, The causal effect of education on earnings, *Handbook of Labor Economics* Vol. 3, 1801–1863, Elsevier, Amsterdam.

Case, A. and Deaton, A., 2005, Broken down by work and sex: how our health declines, in: David Wise (ed.), Analyses in the Economics of Aging, University of Chicago Press, Chicago.

Cervellati, M. and U. Sunde, 2005. Human capital formation, life expectancy, and the process of development, *American Economic Review* 95, 1653–1672.

Cervellati, M. and Sunde, U., 2010, Longevity and lifetime labor supply: evidence and implications revisited, Discussion Paper.

Chaloupka, F.J. and Warner, K.E., 2000, The economics of smoking, *Handbook of Health Economics* Vol. 1, Elsevier, Amsterdam, 1539–1627.

Chatterji, S., Harttgen, K., Kowal, P., Strulik, H., Vollmer, S., 2011, Patterns in disability and frailty in older adults: Results from the Survey of Health, Ageing and Retirement in Europe (SHARE), mimeo, Harvard Medical School.

Chetty, R., 2006, A new method of estimating risk aversion, *American Economic Review* 96, 1821-1834.

Conti, G. and Heckman, J. and Urzua, S., 2010, The education-health gradient, *American Economic Review* 100, 234–38.

Contoyannis, P. and Jones, A.M., 2004, Socio-economic status, health and lifestyle, *Journal of Health Economics* 23, 965–995.

Cutler, D.M., and Richardson, E., 1997, Measuring the health of the US population, *Brookings papers on economic activity. Microeconomics* (1997), 217–282.

Cutler, D.M. and Lleras-Muney, A., 2006, Education and health: Evaluating theories and evidence, NBER Working Paper 12352.

Cutler, D.M., Lleras-Muney, A. and Vogl, T., 2008, Socioeconomic status and health: dimensions and mechanisms, NBER Working Paper 14333.

Cutler, D.M., Lleras-Muney, A. and Vogl, T., 2011, Socioeconomic status and health: dimensions and mechanisms, in: Glied, S. and Smith, P.C. (eds), *The Oxford Handbook of Health Economics*, Oxford University Press, Oxford, pp. 124-163.

Cutler, D.M. and Lleras-Muney, A., 2010, Understanding differences in health behaviors by education, *Journal of Health Economics* 29, 1–28.

Cutler, D.M., Lange, F., Meara, E., Richards, S., and Ruhm, C.J., 2010, Explaining the rise in educational gradients in mortality, NBER Working Paper 15678.

Dalgaard, C.J. and Strulik, H., 2010, Optimal aging and death, PGDA Working Paper No. 58, Harvard University.

Deaton, A., 1997, *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*, Johns Hopkins University Press.

Ehrlich, I. and Chuma, H., 1990, A Model of the Demand for Longevity and the Value of Life Extension, *Journal of Political Economy* 98, 761-82

Elo, I.T. and Preston, S.H., 1996, Educational differentials in mortality: United States, 1979-1985, *Social Science & Medicine* 42, 47–57.

Foster, M., 2001. The meaning of death: some simulations of a model of healthy and unhealthy consumption. *Journal of Health Economics* 20, 613-638.

French, E., 2005, The effects of health, wealth, and wages on labour supply and retirement behaviour, *Review of Economic Studies*72, 395-427.

Friedman, B.M. , 2005, *The Moral Consequences of Economic Growth*, First Vintage Books.

Fries, J.F., 1980. Aging, natural death, and the compression of morbidity. *New England Journal of Medicine* 303, 130-135.

Gavrilov, L.A. and Gavrilova, N.S., 1991, *The Biology of Human Life Span: A Quantitative Approach*, Harwood Academic Publishers, London.

Glied, S. and Lleras-Muney, A., 2008, Technological innovation and inequality in health, *Demography* 45, 741–761.

Grossman, M., 1972, On the concept of health capital and the demand for health, *Journal of Political Economy* 80, 223-255.

Grossman, M., 2006, Education and nonmarket outcomes, in: *Handbook of the Economics of Education* Vol. 1, 577–633, Elsevier, Amsterdam.

Grossman, M., 2000, The human capital model, in: *Handbook of Health Economics* Vol. 1, 347–408, Elsevier, Amsterdam.

Gourinchas, P.-O., and Parker, J.A., 2002, Consumption over the life cycle, *Econometrica* 70, 47-89.

Hall, R.E. and Jones, C.I., 2007. The Value of Life and the Rise in Health Spending. *The Quarterly Journal of Economics* 122, 39–72

Hazan, M., 2009, Longevity and lifetime labor supply: Evidence and implications, *Econometrica* 77, 1829–1863.

Heckman, J.J. and Vytlacil, E., 2001, Identifying the role of cognitive ability in explaining the level of and change in the return of schooling, *Review of Economics and Statistics* 83, 1–12.

Heckman, J.J., 2006, Skill formation and the economics of investing in disadvantaged children, *Science* 312, 1900–1902.

Jorgenson, D.W., Ho, M.S. and Stiroh, K.J., 2008, A retrospective look at the US productivity growth resurgence, *Journal of Economic Perspectives* 22 (1), 3–24.

, Kaestner, R. and Callison, K., 2011,Adolescent Cognitive and Noncognitive Correlates of Adult Health, *Journal of Human Capital* 5, 29–69.

Katz, L.F. and David, H., 1999, Changes in the wage structure and earnings inequality, *Handbook of Labor Economics* Vol. 3, 1463–1555.

Keehan, S. P, H. Lazenby, M. Zezza, and A.C. Catlin, 2004. Age Estimates in the National Health Accounts. *Health Care Financing Review* 1, 1–16.

Kenkel, D.S., 1991, Health behavior, health knowledge, and schooling, *Journal of Political Economy* 99, 287–305.

Mackenbach, J.P., Stirbu, I., Roskam, A.J.R., Schaap, M.M., Menvielle, G., Leinsalu, M., and Kunst, A.E., 2008, Socioeconomic inequalities in health in 22 European countries, *New England Journal of Medicine* 358, 2468–248.

McFadden, D., 2005, Comment, in: David Wise (ed.), Analyses in the Economics of Aging, University of Chicago Press, Chicago.

Meara, E., White, C., and Cutler, D.M., 2004, Trends in medical spending by age, 1963–2000, *Health Affairs* 23, 176–183.

Meara, E.R., Richards, S., and Cutler, D.M., 2008, The gap gets bigger: changes in mortality and life expectancy, by education, 1981–2000, *Health Affairs* 27, 350–360.

Mincer, J., 1974 *Schooling, Experience, and Earnings*, Columbia University Press, New York.

Mitnitski, A.B. and Mogilner, A.J. and MacKnight, C. and Rockwood, K., 2002a, The accumulation of deficits with age and possible invariants of aging, *Scientific World* 2, 18161822.

Mitnitski, A.B. and Mogilner, A.J. and MacKnight, C. and Rockwood, K., 2002b, The mortality rate as a function of accumulated deficits in a frailty index, *Mechanisms of Ageing and Development* 123, 1457-1460.

Mitnitski, A. and Song, X. and Skoog, I. and Broe, GA and Cox, J.L. and Grunfeld, E. and Rockwood, K., 2005, Relative fitness and frailty of elderly men and women in developed countries and their relationship with mortality, *Journal of the American Geriatrics Society* 53, 2184–2189.

Mitnitski, A. and Bao, L. and Rockwood, K., 2006, Going from bad to worse: a stochastic model of transitions in deficit accumulation, in relation to mortality, *Mechanisms of Ageing and Development* 127, 490–493.

Murnane, R.J. and Willett, J.B. and Levy, F., 1995, The growing importance of cognitive skills in wage determination, *Review of Economics and Statistics* 77, 251–266.

Murphy, K.M. and Topel, R.H., 2006, The value of health and longevity, *Journal of Political Economy* 114, 871–904.

OECD, 1998, Work force aging in OECD countries, *OECD Employment Outlook*, Chapter 4.

Oster, S.M. and Hamermesh, D.S., 1998, Aging and productivity among economists, *Review of Economics and Statistics* 80, 154–156.

Preston, S.H., D.A. Glei, and J.R. Wilmoth, 2010, Contribution of smoking to international differences in life expectancy, in: E.M. Crimmins, S.H. Preston, and B. Cohen (eds.), *International Differences in Mortality at Older Ages: Dimensions and Sources*, Washington, D.C.: National Academies Press, pp. 105-131.

Psacharopoulos, G. and Patrinos, H.A., 2004, Returns to investment in education: a further update, *Education Economics* 12, 111–134.

Richards, H., Barry, R., 1998. U.S. life tables for 1990 by sex, race, and education, *Journal of Forensic Economics* 11, 9–26.

Ried, W., 1998, Comparative dynamic analysis of the Grossman model. *Journal Health Economics* 17, 383-425.

Rockwood, K. and Mitnitski, A.B., 2007, Frailty in relation to the accumulation of deficits, *Journals of Gerontology Series A: Biological and Medical Sciences* 62, 722

Skirbekk, V., 2004, Age and individual productivity: A literature survey, *Vienna Yearbook of Population Research*, 133–154.

Turner, C., Tamura, R., Muholland, S., and Baier, S., 2007, Education and income of the states of the United States: 1840-2000, *Journal of Economic Growth* 12, 101-158.

## Mathematical Appendix

6.1. **Setup of the Problem.** Integrating (3) provides the following solution.

$$D(t) = D(0)\exp(\mu t) - \int_0^t \mu a \exp(\mu(t-v))\mathrm{d}v - \mu A \int_0^t h(v)^\gamma \exp(\mu(t-v))\mathrm{d}v + \mu B \int_0^t u(v)^\omega \exp(\mu(t-v))\mathrm{d}v. \tag{A.1}$$

Integrating (2) and using $d \equiv c + \beta u$ we get (A.2)

$$k(t) = k(0)\exp(rt) - \int_0^R \exp(r(t-v))w(v)H(s,v)\mathrm{d}v - \int_0^t \exp(r(t-v))d(v)\mathrm{d}v - \int_0^t \exp(r(t-v))(q-\beta)u(v)\mathrm{d}v - \int_0^t \exp(r(t-v))ph(v)\mathrm{d}v. \tag{A.2}$$

Using (A.1) and (A.2), the initial conditions $D(0) = D_0$, $k(0) = k_0$, and the terminal conditions $D(T) = \bar{D}$, $k(T) = \bar{k}$, the Lagrangian associated with problem (1)- (4) is given by

$$\begin{aligned}\max_{d,h,s,T} L = &\int_0^T \mathrm{e}^{-\rho t}\frac{d^{1-\sigma}}{1-\sigma}\ \mathrm{d}t - \xi\int_0^T \mathrm{e}^{-\rho t}D^\nu \mathrm{d}t \\ &+ \phi\left\{k_0 + \int_s^R \mathrm{e}^{-rt}w(t)H(s,t)\mathrm{d}t - \int_0^T \mathrm{e}^{-rt}d(t)\mathrm{d}t - \int_0^T \mathrm{e}^{-rt}ph(t)\mathrm{d}t - \int_0^T \mathrm{e}^{-rt}(q-\beta)u(t)\mathrm{d}t - \bar{k}\mathrm{e}^{-rT}\right\} \\ &+ \lambda\left\{D_0 - \mu a\int_0^T \mathrm{e}^{-\mu t}\mathrm{d}t - \mu A\int_0^T h(t)^\gamma \mathrm{e}^{-\mu t}\mathrm{d}t + \mu B\int_0^T u(t)^\omega \mathrm{e}^{-\mu t}\mathrm{d}t - \bar{D}\mathrm{e}^{-\mu T}\right\}.\end{aligned} \tag{A.3}$$

Using (4), $w(t)H(s,t)$ in (A.3)is determined as (A.4).

$$\begin{aligned}w(t)H(t,s) = &\ w(0)\exp(g_w t)\left[\bar{H}\exp(\theta s + \eta(t-s) - \alpha\mu t\right] \\ &- \delta\left[D_0\exp(\mu t) - \int_0^t \mu a\exp(\mu(t-v))\mathrm{d}v - \mu A\int_0^t h(v)^\gamma \exp(\mu(t-v))\mathrm{d}v\right] \\ &+ \delta\mu B\int_0^t u(v)^\omega \exp(\mu(t-v)).\end{aligned} \tag{A.4}$$

**Solution.** We begin with the Euler equations. The first order conditions for consumption, health expenditure, and unhealthy consumption are:

$$0 = \mathrm{e}^{-\rho t}d^{-\sigma} - \phi\mathrm{e}^{-rt} \tag{A.5}$$

$$0 = -\phi\mathrm{e}^{-rt}p - \lambda\mu A\gamma h^{\gamma-1}e^{-\mu t}. \tag{A.6}$$

$$0 = -\phi\mathrm{e}^{-rt}(q-\beta) + \lambda\mu B\omega u(t)^{\omega-1}\mathrm{e}^{-\mu t}. \tag{A.7}$$

Differentiating (A.5) with respect to time we get (6), differentiating (A.6) with respect to time we get (7), and differentiating (A.7) with respect to time we get (8). Next, solving (A.6) for $\lambda$ and using the result to substitute $\lambda$ in (A.7) provides (5) in the text.

The first order condition for optimal is schooling is $\partial L/\partial s = 0$, that is

$$0 = \int_0^R \frac{\partial}{\partial s}\mathrm{e}^{-rt}w(t)H(s,t)\mathrm{d}t - \mathrm{e}^{-rt}w(s)H(s,s),$$

requiring that the gain from a marginal extension of education, the first term on the right hand side, equals the income lose from a marginal extension of education, the second term. Inserting $\partial H(s,t)/\partial s = (\theta-\eta)\bar{H}\exp(\theta s+\eta(t-s)-\alpha\mu t)$, $w(t) = w(s)\exp(g_w(t-s))$, and $H(s,s) = \bar{H}\exp(\theta s-\alpha\mu s)-\delta D(s)$, the optimal schooling condition becomes

$$(\theta-\eta)\bar{H}\exp((\theta-\eta-g_w+r)s)\int_s^R \exp((g_w-r+\eta-\alpha\mu)t)\mathrm{d}t = \bar{H}\exp((\theta-\alpha\mu)s)-\delta D(s).$$

Solving the integral provides (9) in the text.

Two conditions have to be fulfilled at the optimal time to die $T$. The first one is that $D(T) = \bar{D}$. Evaluating (A.1) at $T$ and employing the fact of constant growth rates of $h$ and $u$ according to (7) and (8) this can be expressed as:

$$\begin{aligned}\bar{D} = D_0\exp(\mu T) - \mu a\int_0^T \exp(\mu(T-t)\mathrm{d}t - \mu A\int_0^T h(0)^\gamma\exp(\gamma g_h t)\exp(\mu(T-t))\mathrm{d}t\\ + \mu B\int_0^T u(0)^\omega\exp(\omega g_u t)\exp(\mu(T-t))\mathrm{d}t.\end{aligned}$$

Solving the integrals provides (13) in the text.

The second condition for optimal death is that the Lagrangian evaluated at $T$ assumes the value of zero, that is, using (A.3) and the Euler equations (5)-(7):

$$\begin{aligned}0 = \left(\frac{d(T)^{1-\sigma}-1}{1-\sigma}\right)\exp(\rho T) - \xi\bar{D}^\nu\exp(\rho T)\\ + \phi\left[-\exp(-rT)d(T) - (q-\beta)\exp(-rT)u(T) - p\exp(-rT)h(T) + r\exp(-rT)\bar{k}\right]\\ + \lambda\left[-\mu\exp(-\mu T) - \mu A h(T)^\gamma\exp(-\mu T) + \mu B u(T)^\omega\exp(-\mu T) + \mu\bar{D}\exp(-\mu T)\right]\end{aligned}$$

Inserting from (A.5)–(A.7) that $\phi\exp(-rT) = d(T)^{-\sigma}\exp(-\rho T)$, that $\lambda\exp(-\mu T) = -\phi\exp(-rT)$ $\cdot p h(t)^{1-\gamma}/(\mu A\gamma)$, and that $\lambda\exp(-\mu T) = \phi\exp(-rT)(q-\beta)u(t)^{1-\omega}/(\mu B\omega)$ provides (14) in the text.

Using the Euler conditions (A.5)–(A.7) the budget constraint (A.2) can be written as:

$$\begin{aligned}0 = k(0) + w(s,R) - \int_0^T d(0)\exp((g_d-r)t) - p\int_0^T h(0)\exp((g_h-r)t)\\ - (q-\beta)\int_0^T u(0)\exp((g_u-r)t) - \bar{k}\exp(-rT).\end{aligned}$$

Solving the integrals provides (11) in the text. Finally human wealth $w(s,R)$ is obtained as

$$\begin{aligned}w(s,R) &= \int_s^R \exp(-rt)w(t)H(s,t)\mathrm{d}t\\ &= \int_s^R \exp(-rt)\bar{w}\exp(g_w t)\left[\bar{H}\exp(\theta s+\eta(t-s)-\alpha\mu t)-\delta D\right]\mathrm{d}t = \bar{w}(x_1+x_2)\end{aligned}$$

with

$$\begin{aligned}x_1 &= \int_s^R \exp((-r+g_w)t)\bar{H}\exp(\theta s+\eta(t-s)-\alpha\mu t)\mathrm{d}t\\ &= \frac{\bar{H}\exp((\theta-\eta)s)}{\eta+g_w-r-\alpha\mu}\left[\exp(\eta+g_w-r-\alpha\mu)R - \exp(\eta+g_w-r-\alpha\mu)s\right]\end{aligned}$$

The second component $x_2$ is obtained as follows.

$$
\begin{aligned}
x_2 &= \delta \int_s^R \exp((-r+g_w)t) \left\{ D_0 \exp(\mu t) + \frac{\mu A h(0)^\gamma \exp(\mu t)}{g_D} \left[1 - \exp(g_D t)\right] \right. \\
&\quad \left. - \frac{\mu B u(0)^\omega \exp(\mu t)}{g_\omega} \left[1 - \exp(g_\omega t)\right] - a \left[\exp(\mu t) - 1\right] \right\} \mathrm{d}t \\
&= \delta D_0 \int_s^R \exp((\mu + g_w - r)t) \mathrm{d}t + \frac{\delta \mu A h(0)^\gamma \exp(\mu t)}{g_D} \int_s^R \exp((\mu + g_w - r)t) \mathrm{d}t \\
&\quad - \frac{\delta \mu A h(0)^\gamma \exp(\mu t)}{g_D} \int_s^R \exp((\mu + g_w - r + g_D)t) \mathrm{d}t \\
&\quad - \frac{\delta \mu B u(0)^\omega \exp(\mu t)}{g_\omega} \int_s^R \exp((\mu + g_w - r)t) \mathrm{d}t + \frac{\delta \mu B u(0)^\omega \exp(\mu t)}{g_\omega} \int_s^R \exp((\mu + g_w - r + g_\omega)t) \mathrm{d}t \\
&\quad - \delta a \int_s^R \exp(\mu + g_w - r) \mathrm{d}t + \delta a \int_s^R \exp((g_w - r)t) \mathrm{d}t.
\end{aligned}
$$

Solving the integrals provides the final building block for (12) in the text.