The Economics of Health Demand and Human Aging: Health Capital vs. Health Deficits*

Carl-Johan Dalgaard† and Holger Strulik**

January 2015.

Abstract. This paper provides a rigorous comparative analysis of assumptions and predictions of alternative economic theories of health demand and human aging. The health-capital model, based on Grossman (1972) and the health-deficit model, based on Dalgaard and Strulik (2014). We show that both theories lead to fundamentally different predictions of health behavior and human life histories. We find that the health-deficit model provides a consistent approach to health behavior and aging whereas the health-capital model generates predictions that are hard to square with the stylized facts. We argue that the root of the disagreement of the theories is the following: The health-capital model postulates that of two people of the same age the healthier one looses more health in the next instant whereas the health-deficit model, based on insights from modern gerontology, proceeds from the opposite assumption.

Keywords: Aging, Longevity, Health.

JEL: D91, J17, J26, I12.

* We would like to thank Rob Alessie, Axel Boersch-Supan, Christian Dustman, Volker Grossmann, Ben Heijdra, Volker Meier, Petros Milionis, Timo Trimborn, and Matthias Weiss, as well as participants at seminars in Groningen, Nuremberg, Rotterdam, the Cesifo Institute, and the Munich Center for the Economics of Aging for discussion and helpful comments.

† Department of Economics, University of Copenhagen, Oester Farimagsgade 5, building 26, DK-1353 Copenhagen, Denmark; email: carl.johan.dalgaard@econ.ku.dk.

** University of Goettingen, Department of Economics, Platz der Goettinger Sieben 3, 37073 Goettingen, Germany; email: holger.strulik@wiwi.uni-goettingen.de.

1. Introduction

The health capital model of health demand, also known as the Grossman model, although the major tool in health economics for almost 40 years, is still not fully understood. In this paper we use phase diagram techniques to investigate in general the predictions of the health-capital model when health investment is motivated by the desire for a longer life and the utility- or productivity enhancing effects of good health. We compare the results with the predictions made by the alternative economic approach to health investments, the health-deficit model developed by Dalgaard and Strulik (2014). The main difference of both models lies in their distinct view on human aging. The health-capital model conceptualizes the human life course as accumulation of health capital while Dalgaard and Strulik conceptualize life as the accumulation of health deficits.

Applications of the health-deficit approach (Dalgaard and Strulik, 2012, 2014; Strulik, 2011) hitherto emphasized as their major advantage that the main health parameter – the force of aging – has a foundation in medical science where it has been estimated with great precision. This means that numerical calibrations of the model are straightforward and leave no degrees of freedom for the researcher. We have also stressed the model's formal elegance, which allows a straightforward interpretation of results, and the fact that health deficits are easily measurable, in contrast to health capital, which is a latent concept. Nevertheless, these features may not be regarded as sufficient to abandon the health-capital paradigm of health capital accumulation.

Advocates of the health-capital model could defend the approach based on Friedman's (1953) methodology of economic modeling: that it is not the scientific foundation of the assumptions but the predictive performance that constitutes the quality of a theory. In the words of a referee of one of our papers: "We don't usually look at engineers to measure the depreciation of physical capital in economics. So why should we ask doctors to measure the depreciation of health human capital?" In this paper we show that it is important to built on a scientific foundation of health depreciation in order to get the predictions for human aging and health demand over the life course right.

Specifically, we use phase diagram techniques in order to compare two simplified versions of the health-capital model and the health-deficit model. These models capture the basic underlying mechanism of the respective theories and are simple enough to be rigorously analyzed. Specifically we consider the impact of health investments on (i) longevity and utility, and (ii) longevity and productivity. With this distinction we follow the tradition in the mainstream health literature (e.g. Grossman, 2000). Variant (i) and (ii) are also known in the literature as "pure consumption model" and "pure investment model". These names have been coined because a large body of the health-capital literature focusses on either consumption or productivity effects of health investments and neglects longevity effects by imposing a fixed terminal time of death. Longevity effects were introduced into the health-capital context by Ehrlich and Chuma (1990).[1]

The original health-capital model was stated in discrete time, which has complicated the analysis and lead to confusing results concerning the endogenous determination of the age at death, which would be invariant to sufficiently small parametric changes (see the discussion in Ried, 1998; Grossman, 2000). This outcome, however, is solely an artefact of the discrete time concept. Here we thus follow the majority of more recent applications of the health-capital model and set it up in continuous time. Another speciality of Grossman's original model is that it distinguishes between income spent on health care and time as inputs in health investment. This certainly adds more realism but does not affect any of the qualitative implication made in the present study (as long as the commodity production functions are linear homogenous, as typically assumed). Like many other applications of the health-capital model we thus focus on goods inputs. A major simplification of the health-capital model is the assumption of constant returns in health investment. Since this assumption has been heavily criticized and debated in the literature (e.g. Ehrlich and Chuma, 1990; Grossman, 2000; Galama et al., 2012), we follow Grossman in the main text and assume constant returns but show in the Appendix that the main results are also obtained under decreasing returns. The main concerns with the health-capital model arise independently from the returns to scale assumption.

Several other extensions of the simple models are conceivable and needed in order apply them for the discussion of actual policy issues. These applications, due to increasing complexity, would rely on numerical solutions of one particular calibration of the specific type of model. As will become apparent below, a purely numerical analysis entails the danger to neglect general properties and implication of the model by focussing on a specific numerically predetermined

[1] An incomplete list of studies based on the health-capital paradigm includes Cropper (1977), Eisenring (1999), Forster (2001), Galama and Kapteyn (2011), Galama et al (2012), Jacobsen (2000), Kuhn et al. (2012), Laporte and Ferguson (2007), Liljas (1998), Meier (2000), Picone et al. (1998), and van Kippersluis and Galama (2014). Important empirical studies are Muurinen (1982) and Wagstaff (1986). Grossman (2000) provides an overview.

trajectory. Most importantly, requesting the arrival at a state of low terminal health may conceal the feasibility of an actually preferred life trajectory that converges towards a *permanently* better state of health. The simplified model versions presented here, in contrast, are ideal in order to elaborate the basic mechanisms in a general fashion by means of phase diagram analysis. These basic mechanisms are also at work in any extended version of the basic models.

The published variants of the health-capital model can also be distinguished by the imposed terminal condition (Forster, 2001). Here, we follow Ehrlich and Chuma (1990) and many others and consider a free terminal time problem. This means that the time of death is a free variable which can be influenced through health investments. Death is determined by a terminal condition on health capital or health deficits, respectively, i.e. survival is untenable below a minimum state of health. There exists, however, also a relatively large literature that considers health-capital accumulation for a fixed terminal time (e.g. Eisenring, 1999; Kuhn et al., 2012; van Kippersluis and Galama, 2014). According to this approach, the time of death is predetermined and death occurs irrespective of an individual's health status. Individuals "only" decide how healthy they want to be when they have to die (and, of course during the rest of their life). Here we do not follow this approach because we want to scrutinize the determinants of human aging and longevity, which is impossible under the fixed terminal time approach.

The predictive quality of a model of health demand could be, of course, evaluated in many dimensions. In keeping with Friedman's methodology, however, we consider it sufficient to show that a theory fails to predict some core features of the human life course. Specifically, we consider two features of a good theory: (i) It should be possible to find at least some initial conditions and parameters (income, health technology) for which the theory predicts that life is finite. (ii) The theory should be able to predict that health expenditure rises when health declines, at least in old age. Below we show that the health-capital model usually predicts eternal life. If this is ignored (by imposing a finite terminal time or by, perhaps inadvertently, disregarding this infinite horizon solution in a numerical approach), then the health-capital model predicts that health expenditure declines as people get unhealthier.

We show that the health-deficit model does not share these worrisome predictions. In contrast it predicts that eternal life does not exist when the state of medical technology is sufficiently low and that health expenditure may increase or decrease as individuals age, depending on the characteristics of the utility function.

These distinctive predictions of the two theories have an intuitive explanation. Consider two individuals of the same age and the same socio-economic background. Then the health-capital theory assumes that the healthier individual ages faster (loses more health in the next instant) while the health-deficit theory assumes that the unhealthier individual ages faster (loses more health in the next instant). To see this formally, consider health capital accumulation, $\dot{H} = f(I) - \delta H$, the core equation of the health-capital model. Here H is health capital, I is health investment, f is a positive function and δ is the depreciation rate. The change of health is given by $\dot{H} \equiv \partial H/\partial t$, in which t is age. Observe that health capital accumulation implies $\partial \dot{H}/\partial H = \delta > 0$. In words, the healthier individual, i.e. the one with the larger health capital stock, loses more health through health depreciation. Notice that this outcome is observed irrespective of whether the depreciation rate is constant or age-dependent. Below we show that this feature creates an equilibrating force, a movement towards a constant stock of health, which promotes immortality of the modeled agents.

Consider, in contrast health deficit accumulation, $\dot{D} = \mu D - f(I)$, in which D are health deficits, i.e. $\dot{D} \equiv \partial D/\partial t$ is the change of deficits that occurs with aging. Observe that $\partial \dot{D}/\partial D = \mu > 0$. In words, the unhealthier individual, i.e. the individual with more health deficits, is predicted to develop more deficits in the next instant. This creates a dis-equilibrating force, moving individuals away from a state of constant health and along a trajectory of increasing accumulation of health deficits towards death.

Of course these shortcomings of the health-capital model remained not unnoticed (see e.g. Wagstaff, 1986, Case and Deaton, 2005). We claim no originality here. We only rigorously prove the shortcomings. The previous criticism of the health-capital model, however, was non-constructive, without suggesting a (self-contained) alternative theory. Here we provide such a theory. We show that the predictive difficulties of the health-capital model do not carry over to the health-deficit model whose assumptions are based on a scientific foundation. That is, we demonstrate a positive correlation between the quality of the assumptions and the predictive quality for two alternative theories of health demand and human aging.

The paper is organized as follows. The next section provides an introduction to the biological foundation of the health deficit model. While the focus of this paper is on predictions, this section is nevertheless useful in order to understand how a scientific foundation of assumptions can improve the predictive quality of a theory. Section 3 solves the "pure consumption" model under the hypothesis of health capital accumulation and health deficit accumulation and compares the respective predictions for health demand over the life cycle, aging, and longevity. Section 4 repeats this analysis for the "pure investment" model. Section 5 concludes.

2. Inspiration and Foundation of the Theories of Health Capital Accumulation and Health Deficit Accumulation

It is perhaps fair to say that health capital theory has no foundation in the natural sciences. As evident from Grossman (1972) and Grossman (2000) the theory was heavily inspired from the accumulation of human capital in form of education (Becker, 1967; Ben-Porath, 1967). Yet, the seminal health capital papers provide no references to the biological, medical, or gerontological literature on human aging in order to justify the core assumption of health capital accumulation. Apparently the model was developed sui generis. One immediate problem arising from the fact that health capital is a latent variable and that its notion is exclusively shared within the economics profession concerned the empirical test of the theory. How should one proxi for the unobservable? A popular approach in empirical applications used the fact that health deficits are observable and tried to estimate health capital from a vector of (absent) health deficits (see Muurinen, 1982; Wagstaff, 1986). However, we argue that there is no need for such a 'detour'. A theory of health demand can be founded directly on observable health deficits without the construction of an unobservable health capital stock.

The theory of health deficit accumulation proposes to integrate into economics a notion of human aging founded in the natural sciences. Aging is defined as the intrinsic, cumulative, progressive, and deleterious loss of function that eventually culminates in death (Arking, 2006). In order to understand how humans (and other animals) age, modern gerontology has adopted insights and mechanisms from reliability theory in engineering (Gavrilov and Gavrilova, 1991). The theory is based on the notion of gradual depletion of redundancy in the human body.[2] In its simplest form, reliability theory regards a body (or any organism) as a whole as consisting of a fixed number of non-aging components, where non-aging is defined as exhibiting a constant failure rate.[3] The phenomenon of redundancy is captured by the assumption that components

[2] In economics, McFadden (2008) argued in favor of the desirability of a theory of health demand based on the concepts of reliability and redundancy.

[3] The perpetual youth model (In this sense, the models of Yaari (1965) and Blanchard (1982), by originating from a constant failure rate of the whole human body, conceptualize the whole person as non-aging, as aptly alluded to by the "perpetual youth" naming of the model (Blanchard and Fisher (1989).

are connected in parallel, and that the system as a whole is assumed to survive as long as there is one functioning component remaining. Aging is thus explained as a consequence of initial redundancy. Indeed, the functional capacity of organs in young adults is estimated to be tenfold higher than needed for mere survival (Fries, 1980). As a person gets older and redundancy in the system as a whole declines, the failure rate of the whole body increases. This way the simple model successfully reproduces a rising death rate with age. Many extensions of the basic model have been developed, enabling a comprehensive understanding of generalities and specific details of the human aging process (like Gompertz-Makeham law and the leveling off of aging for the oldest old, see e.g., Gavrilov and Gavrilova (1991).

As the redundancy of the body shrinks, a person becomes more fragile. The fact that humans, as they age, develop an increasing number of disorders or health deficits has been exploited empirically by the development of the frailty index by Mitnitski and Rockwood and various coauthors (e.g., Mitnitski et al, 2002a,b; 2005; Rockwood and Mitnitski, 2006). The frailty index computes from a sufficiently long list of potential health deficits the proportion of deficits that an individual has, at a given age. It constitutes a comprehensive measure of human health, aging, and mortality.[4]

Let $D(t)$ denote the number (percentage) of health deficits that an individual of age t has. Mitnitski et al. (2002a) show that aging is then well represented by a linear-exponential association of health deficits and age: $D(t) = E + Be^{\mu t}$. This "law of increasing frailty" explains around 95 percent of the variation in the data, and its parameters are estimated with great precision. While the parameter E turns out to be common for men and women, the parameters B and μ are gender specific. Most importantly the parameter μ, the rate of aging, has been estimated to be approximately of the same size across different populations (Rockwood and Mitnitski, 2007). On average adult individuals accumulate 3-4 percent more deficits from one birthday to the next.

In order to integrate the law of health deficit accumulation into economics we begin with restating it in flow form by differentiating with respect to age: $\dot{D}(t) = \mu (D(t) - E)$. Notice

[4]Methodologically, the frailty index exploits the fact that health deficits are largely interdependent. Approaches to estimate bio-markers (like health capital) using multivariate regression or principle component analysis are, according to the frailty literature, flawed because they assume statistical independence of the right hand side variables (Mitnitski et al., 2002).

that E works to slow down the speed of deficit accumulation by making health deficit accumulation less then exponential. In the natural science literature the parameter E is interpreted as capturing the impact of non-biological factors on deficit accumulation (Mitnitski et al., 2002a). Accordingly, we assume that E is amendable to change by way of deliberate health investments. Specifically, we propose the following parsimonious refinement of the process of health deficit accumulation: $\dot{D}(t) = \mu \left(D(t) - Af(I) - a\right)$, in which I is health investment and f is a positive, concave function. The parameter a captures environmental influences on aging beyond the control of the individual and the parameters A captures the state of the health care technology.

In principle, health care expenditure could also be thought of as having a bearing on the force of aging μ. So far, however, the idea that "standard" medical treatments have substantially modified the *rate* at which our bodies decay received little empirical support (Gavrilov and Gavrilova, 1991). Standard treatments like, for example a bypass operation or a liver transplant, are effectively delaying death (by removing one or several health deficits) but they are not manipulating the intrinsic rate of deficit accumulation. Death occurs when health deficits reach an upper boundary to deficit accumulation, $\bar{D}$. Rockwood and Mitnitski (2006) present evidence supporting the notion of a maximum number of viable health deficits. Notice that in the theory of health deficit accumulation chronological age does not play a role in itself. It takes into account that a aging has a molecular structure and it is not tied to a clock or calender. This is important since biologists emphasize that there should be no role for (calendar-) time in a successful model of human aging: "Only if we can substitute the operation of the actual physiological mechanism for time we have a firm idea of what we are talking about." (Arking, 2006, p. 10).

3. Health Consumption

3.1. **Health Capital Accumulation.** Suppose that health has two functions: it provides utility and it determines death. A long life is desirable because the utility function is concave in goods consumption C and health capital H but linear in longevity. Assume that instantaneous utility is strictly concave and iso-elastic such that

$$U(C,H) = \frac{\left(C^{\alpha} H^{1-\alpha}\right)^{1-\sigma} - 1}{1-\sigma} \tag{1}$$

for $0 < \alpha < 0$, $\sigma > 0$, and $\sigma \neq 1$. For $\sigma = 1$ the utility function reads $U = \alpha \log C + (1-\alpha) \log H$. The parameter σ reflects the inverse of the elasticity of intertemporal substitution and the parameter α reflects the relative weight of goods consumption in utility. Together they determine the sign of the derivatives:

$$U_{CC} < 0 \text{ for } 1-\alpha(1-\sigma) > 0 \tag{2a}$$

$$U_{HH} < 0 \text{ for } 1-(1-\alpha)(1-\sigma) > 0 \tag{2b}$$

$$U_{CH} > 0 \text{ for } (1-\alpha)(1-\sigma) > 0. \tag{2c}$$

Conditions (2a) and (2b) assure decreasing marginal utility. They are always fulfilled under the parameter restrictions made. The cross derivative (2c) is allowed to be positive (individuals prefer to consume a lot when they are healthy), negative, or zero (in the case of log utility).[5]

Individuals maximize life-time utility

$$V = \int_0^T U(C,H)\mathrm{e}^{-\rho t}\mathrm{d}t, \tag{3}$$

Here, t is age, ρ is the discount rate of future consumption, and T is the yet to be determined age of death. In contrast to the available literature, we do not impose a finite T. In principle, $T=\infty$. Of course, we expect from a plausible model that it is capable of generating a finite life, for example because the state of medical technological knowledge is not (yet) sufficiently advanced to life forever. In any case, however, mere logical consistency suggests the following assumption about the size-ordering of health capital stocks.

Assumption 1. *The health capital stock at death is smaller than the health capital stock that would guarantee eternal life, $H_{min} < H^*$.*

According to the health-consumption model, individuals face a constant stream of income Y until they die. This assumption is inconsequential for the results because all decisions depend on discounted life-time income, which could alternatively result from a constant or non-constant income per age profile. The decisive assumption is that income is independent from health status. Income is spent on consumption and investment in health, denoted by I. The relative price of health goods is given by p and is constant over time. The budget constraint at any time is thus given by

$$Y = C + pI. \tag{4}$$

[5] In the case of $\sigma > 1$ we assume that consumption is scaled appropriately in order to avoid negative utility, which would lead to the degenerate outcome that life-time utility is decreasing in the length of life such that individuals would prefer immediate death; see Hall and Jones (2009) for an extensive discussion of this property.

As discussed in the Introduction, the health-capital model postulates that health is accumulated more or less in the same fashion as human capital in the form of education is accumulated in many economic models of human capital accumulation. Specifically health capital H evolves according to

$$\dot{H} = AI - \delta H, \tag{5}$$

in which δ is the rate of depreciation of health capital. The parameters $A > 0$ captures the power of medical technology. For simplicity, we focus on non-decreasing returns in health investment. Allowing for decreasing returns adds more realism at the expense of increasing analytical complexity. In the Appendix we show that all the main conclusions are robust against decreasing returns in health investment.

Individuals are endowed with an initial stock of health capital $H(0) = H_0$ and survival requires that the health stock exceeds $H_{min} \geq 0$. In other words, individuals die at age T when health deteriorates to the level $H(T) = H_{min}$. Following the literature, we treat the depreciation rate alternatively as constant or rising with age. We begin by considering the case of parametrically given δ and discuss age-dependent health depreciation later. Notice, however, that the health-capital model in any case implies that the loss of health at any age δH is greater when the stock of health is large, that is when individuals are relatively young. Formally, this can be seen from $\partial\dot{H}/\partial H < 0$. Ceteris paribus, individuals age at a high rate when they are young and healthy and at relatively slow rate when they are old. This feature appears to be not only unrealistic but it implies also that individuals converge towards a constant state of health, as shown below.

Individuals are assumed to chose optimal health expenditure over the life course by maximizing (3) subject to (4) and (5) given initial health H_0 and the boundary condition $H \geq H_{min}$. The associated current value Hamiltonian is

$$J = \frac{\left(C^{\alpha}H^{1-\alpha}\right)^{1-\sigma} - 1}{1-\sigma} + \lambda_H \left(AI - \delta H\right), \tag{6}$$

in which λ_H denotes the costate variable, i.e. the shadow price of health. The associated first order condition and costate equation are:

$$\frac{\partial J}{\partial I} = \frac{\alpha \left(C^{\alpha}H^{1-\alpha}\right)^{1-\sigma} - 1}{C}(-p) + \lambda_H A = 0 \tag{7}$$

$$\frac{\partial J}{\partial H} = \frac{(1-\alpha)\left(C^{\alpha}H^{1-\alpha}\right)^{1-\sigma} - 1}{H} - \lambda_D \delta = \lambda_H \rho - \dot{\lambda}_H. \tag{8}$$

In principle, there are infinitely many life-time trajectories fulfilling the first order conditions. The optimal solution moreover fulfills the transversality condition (see e.g. Acemoglu (2009, Theorem 7.1):

$$J(I(T), H(T), \lambda_H(T)) = 0 \text{ for finite } T \tag{9a}$$

$$\lim_{T\to\infty} J(I(T), H(T), \lambda_H(T))\mathrm{e}^{-\rho T} = 0 \text{ otherwise} \tag{9b}$$

If a fixed point exists such that $\lim_{T\to\infty} H(T) = H^*$, condition (9b) simplifies to

$$\lim_{T\to\infty} \lambda_H(T)H(T)\mathrm{e}^{-\rho T} = 0. \tag{9c}$$

It turns out to be most convenient to discuss the solution in the health-consumption space. Notice that optimal health investment is easily recovered as the residual $pI = Y - C$. Replacing I in (4) and differentiating (7) with respect to age and using the result to substitute λ_H and $\dot{\lambda}_H$ in (8) provides the equation of motion for goods consumption and health capital:

$$\frac{\dot{C}}{C} = \frac{1}{1-\alpha(1-\sigma)} \left\{ \frac{(1-\alpha)A}{\alpha p}\frac{C}{H} - (\delta+\rho) + (1-\alpha)(1-\sigma)\frac{\dot{H}}{H} \right\} \tag{10}$$

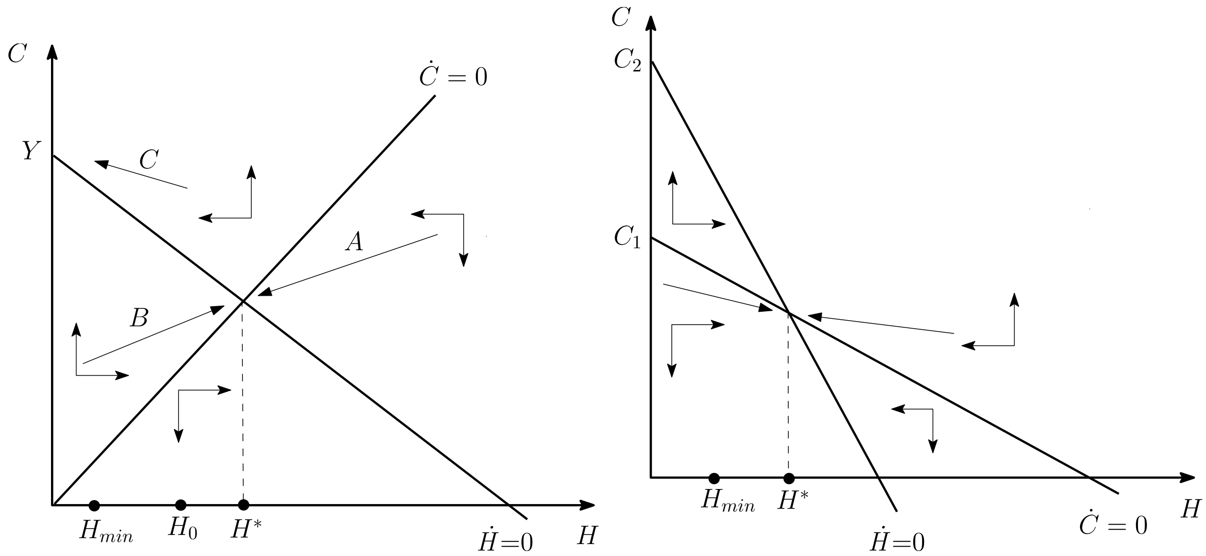
$$\dot{H} = \frac{A(Y-C)}{p} - \delta H. \tag{11}$$

Human aging as described by the dynamical system (10) and (11) can be conveniently analyzed using phase diagram techniques. Notice from (10) that $\partial\dot{C}/\partial C > 0$ when $\dot{H} = 0$, such that the arrows of motion point towards larger C. Most importantly, notice from (11) that $\partial\dot{H}/\partial H < 0$, implying that the arrows of motion point towards lower health when the health stock is above its equilibrium value and towards better health when the health stock is below its equilibrium value. This outcome is a consequence of the initial assumption about health capital accumulation and will be identified as the driving force behind the model's implausible predictions.

From (11) the $\dot{H} = 0$–isocline is given by $C = Y - \delta pH/A$. It is a falling straight line originating from Y. From (10) the $\dot{C} = 0$–isocline is obtained as

$$C = \frac{\alpha}{(1-\alpha)\left[1-\alpha(1-\sigma)\right]} \left[\left\{ \rho + \delta\left[1+(1-\alpha)(1-\sigma)\right]\frac{p}{A}H - (1-\alpha)(1-\sigma)Y \right\} \right]. \tag{12}$$

Figure 1: Health Capital Accumulation and Consumption



The slope of the isocline is generally ambiguous. A positive cross derivative $U_{CH} > 0$ is a sufficient, non-necessary condition for a positive slope. The isoclines intersect at the fixed point

$$H^* = \frac{(1-\alpha)A}{p(\delta+\alpha\rho)}, \qquad C^* = \frac{\alpha(\delta+\rho)}{\delta+\alpha\rho}. \tag{13}$$

Because health does not deteriorate at the fixed point, it characterizes a state of eternal life. Observe from (13) that eternal life always exists in the health-capital model, for any parameter constellation and any initial state of health.

The panel on the left hand side of Figure 1 shows the phase diagram for the case of a positive slope. Inspection of the phase diagram identifies eternal life as a saddlepoint: from each side of the fixed point a unique trajectory (stable manifold) leads towards the fixed point. This means that eternal life can always be reached, from any initial state of health. Because health capital is constant at the steady state, the transversality condition (9c) applies. Because consumption is constant at the steady state, the shadow price λ_H is constant as well, see (7), and thus the transversality condition is fulfilled. According to the first order conditions and the transversality condition, it is optimal to converge towards eternal life. The initial health endowment may be higher or lower than H^* but due to Assumption 1, the deadly health capital stock is necessarily situated to the right of H^*.

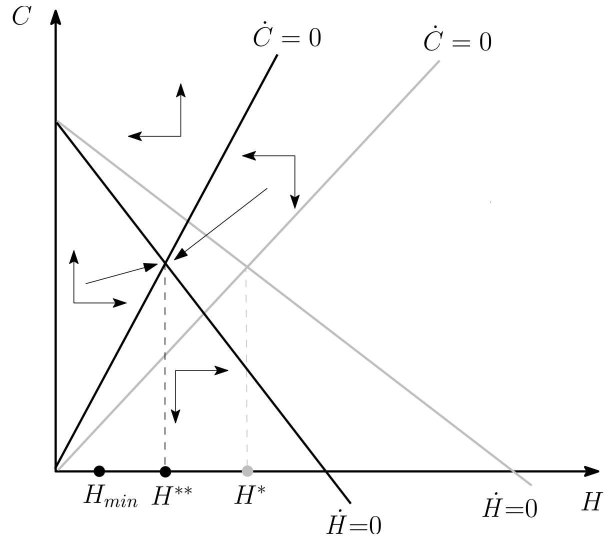
Many applications of the health-capital model do not identify eternal life as the optimal solution because they assume a fixed terminal T or because they simply ignore from the outset the possibility that a variable T could be infinite. The latter could easily happen in a purely numerical application that requires arrival at the boundary value H_{min} without caring about the potential existence of H^*. If, however, arrival at H_{min} is enforced, the health-capital model implies that consumption rises as health capital declines. Trajectory $\mathcal{C}$ shows an example solution. Notice that H_0 cannot be reached by any trajectory in which consumption is declining. This means that the health-capital model – when it ignores eternal life – predicts that health expenditure declines as the state of health erodes. This paper is, of course not the first one to arrive at this conclusion (see e.g. Wagstaff, 1986).

The panel on the right hand side of Figure 1 shows the case of a negative slope of the $\dot{C}/C = 0$–isocline. Notice from (12) that the $\dot{C}/C = 0$–isocline cuts the C-axis at $C_1 \equiv \alpha(\sigma-1)/\left[1+\alpha(\sigma-1)\right]Y$. It thus follows from (2a) that $C_1 < Y$. The $\dot{H} = 0$–isocline cuts the C-axis at $C_2 = Y$ and thus $C_2 > C_1$. This means that the $\dot{H} = 0$–isocline cuts the $\dot{C} = 0$–isocline from above. The arrows of motion point upwards above the $\dot{C} = 0$-isocline and from the right of the $\dot{H} = 0$-isocline to the left and vice versa. The fixed point is thus again identified as a saddlepoint. Eternal life is the preferred solution. However, if we ignore this outcome, it would be impossible to reach any terminal H_{min} smaller than the initial health H_0 as long as the initial health capital stock is smaller than the one guaranteing eternal life. The arrows of motion are always pointing away from H_{min}, towards eternal life.

The fact that eternal life is not only possible but inescapable questions the plausibility of theoretical and empirical studies of the health-capital model. An ad hoc remedy of the problem, suggested by the literature, is to assume that health depreciation δ is increasing with age. An undesirable side-effect of age-dependent health depreciation is that the comparative statics can no longer be assessed qualitatively. This conclusion becomes obvious by writing age-dependent depreciation as $\dot{\delta}(t) = f(\delta(t))$ such that the depreciation rate becomes another state variable. This means that health dynamics are now governed by a three-dimensional dynamic system, a fact that renders qualitative phase diagram analysis basically impossible and prevents the application of Oniki's (1973) method of comparative statics. Consequently, the available qualitative discussion of the health-capital model focussed on models with constant δ (Eisenring, 1999; Meier, 2000; Forster, 2001).[6]

[6] Ehrlich and Chuma (1990) did not mention that they made this simplifying assumption in order to derive the comparative statics of their model (Table 3). But Oniki's method, which they apparently apply, requires the reduction to a two-dimensional system; see also Eisenring (1999).

Figure 2: Eternal Life is Inescapable (Grossman)



The introduction of age-dependent health depreciation, $\delta(t)$ entails also methodological problems. The question arises how exactly the function is to be specified. Insofar as one does not get this choice right the resulting model will represent an inaccurate description of the aging process. More importantly, the implied notion that chronological age matters to aging per se is seen as deeply problematic from the biological perspective on aging. Biologists emphasize the distinction between biological and age and calender age, i.e. that aging has a molecular structure and it is not tied to a clock or calender. The main research goal is to get an explanation of aging independent from calender time (Arking, 2006).

Most disconcerting is perhaps that the introduction of age-dependent health depreciation only seemingly solves the problem of inescapable eternal life. In order to see this conveniently, it is helpful to imagine the increase of δ in discrete steps (say, yearly deterioration of the depreciation rate). Diagrammatically, an increasing δ can then be captured by an increasing slope of both isoclines. This means that as the individual ages the fixed point moves to the left. Figure 3 returns to the case of an upward sloping $\dot{C} = 0$-isocline and draws the movement to the left for one incremental increase of δ. Grey curves show the initial situation. Notice that the equilibrium of eternal life does not disappear. It just occurs at a lower state of health, H^{**}. It is tempting to argue that eternal life eventually disappears when the fixed point falls below $H_{\min}$. However, this line of reasoning violates Assumption 1: the health capital stock at death would be higher than the health capital stock that enables eternal life.

It can be argued that the trouble with the health-capital model originates from the very notion of "health capital" itself. The health-capital model inadvertently assumes that of two individuals of the same age and different health, the healthier one loses more health through health depreciation in the next period (or instant). Notice that this assumption is also maintained with age-dependent health depreciation, i.e. for any given age t: $\delta(t)H_1(t) > \delta(t)H_2(t)$ when $H_1(t) > H_2(t)$. Diagrammatically, the assumption causes the arrows of motion for health to point to lower health when the state of health is good (to the right of the $\dot{H} = 0$-isocline) and to better health at a state of bad health (to the left of the $\dot{H} = 0$-isocline). The arrows of motion thus always point towards the fixed point of eternal life. These worrying results are a consequence of a notion of health that is exclusively held within the profession of economists and that finds no support in gerontology. Gerontology, in contrast, finds a strong positive path dependence of health *deficits*, i.e. in Grossman's terminology it finds a strong negative association between the state of health and the loss of health (Mitnitski et al, 2002, 2005, 2006).

3.2. **Health Deficit Accumulation.** As explained in Section 2 the health-deficit model conceptualizes human life histories as a sequence of accumulated health deficits. Neglecting an environmental constant and focussing, as for the health-capital model, on constant returns of health investment, health deficits evolve according to (14).[7]

$$\dot{D} = \mu(D - AI). \tag{14}$$

As explained above, μ is the force of aging as estimated in gerontology. The parameters A controls the state of the health technology, as in the health-capital model. Investment in health care delays the "natural" accumulation of deficits through maintenance and repair. The individual life (after puberty) starts out with $D(0) = D_0$ deficits and ends when $D_{max} \geq D_0$ health deficits have been accumulated.

In order to get health into the utility function in analogous way to the health-capital model we define the state of best health $\bar{H}$ as the absence of any health deficits and assume that utility declines with increasing appearance of health deficits. Instantaneous utility is thus given by

$$U(C, D) = \frac{\left[C^{\alpha}(\bar{H} - D)^{1-\alpha}\right]^{1-\sigma} - 1}{1 - \sigma} \tag{15}$$

[7]The environmental constant is important for fitting the model to historical data but neglecting it is inconsequential for the mechanics of the model and alleviates the comparison with the health-capital model.

for $D < \bar{H}$ and $U(C, D) = 0$ otherwise. The normalization of zero utility when health deficits fall below $\bar{H}$ and individuals have "zero health" is arbitrary and could be altogether avoided by assuming $\bar{H} > D_{max}$.[8] Analogously to the health-capital case we make the following assumption.

Assumption 2. *Health deficits at death are larger than the stock of health deficits which would guarantee eternal life, $D_{max} > D^*$.*

Individuals maximize lifetime utility $\max \int_0^T U(C, D)e^{-\rho t}\mathrm{d}t$ by the appropriate choice of health investment, given (14), the boundary conditions and the budget constraint (4). The associated Hamiltonian is $J = \frac{\left[C^\alpha(\bar{H}-D)^{1-\alpha}\right]^{1-\sigma}-1}{1-\sigma} + \lambda_D\left(\mu D - \mu A I\right)$, in which λ_D denote the value of health deficits. The first order conditions for a maximum are:

$$\frac{\partial J}{\partial I} = \frac{\alpha\left[C^\alpha(\bar{H}-D)^{1-\alpha}\right]^{1-\sigma}-1}{C}(-p) - \lambda_D \mu A = 0 \tag{16}$$

$$\frac{\partial J}{\partial H} = \frac{(1-\alpha)\left[C^\alpha(\bar{H}-D)^{1-\alpha}\right]^{1-\sigma}-1}{H} \cdot (-1) + \lambda_D \mu = \lambda_D \rho - \dot{\lambda}_D. \tag{17}$$

The optimal solution furthermore fulfils the transversality condition

$$J(I(T), D(T), \lambda_D(T)) = 0 \text{ for finite } T$$

$$\lim_{T\to\infty} J(I(T), D(T), \lambda_D(T))\mathrm{e}^{-\rho T} = 0 \qquad \text{or } \lim_{T\to\infty} \lambda_D(T)D(T)\mathrm{e}^{-\rho T} = 0.$$

if a fixed point exists.

We eliminate λ_D and $\dot{\lambda}_D$ by time-differentiating (16) and summarize the two equations in one equation of motion for consumption, (18). Inserting the budget constraint (4) into (14) expresses the equation of motion for health deficits in terms of goods consumption, (19).

$$\frac{\dot{C}}{C} = \frac{1}{1-\alpha(1-\sigma)}\left\{\frac{(1-\alpha)\mu A}{\alpha p}\frac{C}{(\bar{H}-D)} - (\rho-\mu) - (1-\alpha)(1-\sigma)\frac{\dot{D}}{(\bar{H}-D)}\right\} \tag{18}$$

$$\dot{D} = \mu D - \mu\frac{A(Y-C)}{p}. \tag{19}$$

The optimal life cycle trajectory is described by the dynamic system (18)–(19) and the transversality condition. Observe from (18) that $\partial\dot{C}/\partial C > 0$ (when $\dot{D} = 0$) such that the arrows of motion point towards higher consumption above the $\dot{C} = 0$-isocline, as in the health-capital

[8] It is easily imaginable that utility actually gets negative when health deficits get large enough, a situation under which the individual would prefer euthanasia. For simplification, and to keep symmetry with the health-capital model, this interesting case is neglecting subsequently.

model. Observe from (19) that $\partial\dot{D}/\partial D > 0$. The arrows of motion are thus pointing *away* from the $\dot{D} = 0$-isocline. This feature follows directly from the law of health deficit accumulation (14) and provides the most distinguishing feature of the health-deficit model. Individuals with many health deficits develop even more health deficits in the next instant than healthy individuals.

From (18) and (19) the isoclines for constant consumption and health deficits are obtained as:

$$C = \frac{\alpha p}{(1-\alpha)\mu A\left[1-\alpha(1-\sigma)\right]}\left[(\rho-\mu)\bar{H} - \left\{\rho - \mu\left[1+(1-\alpha)(1-\sigma)\right]\right\}D - (1-\alpha)(1-\sigma)\mu\frac{AY}{p}\right] \quad (20)$$

$$C = Y - \frac{p}{A}D. \quad (21)$$

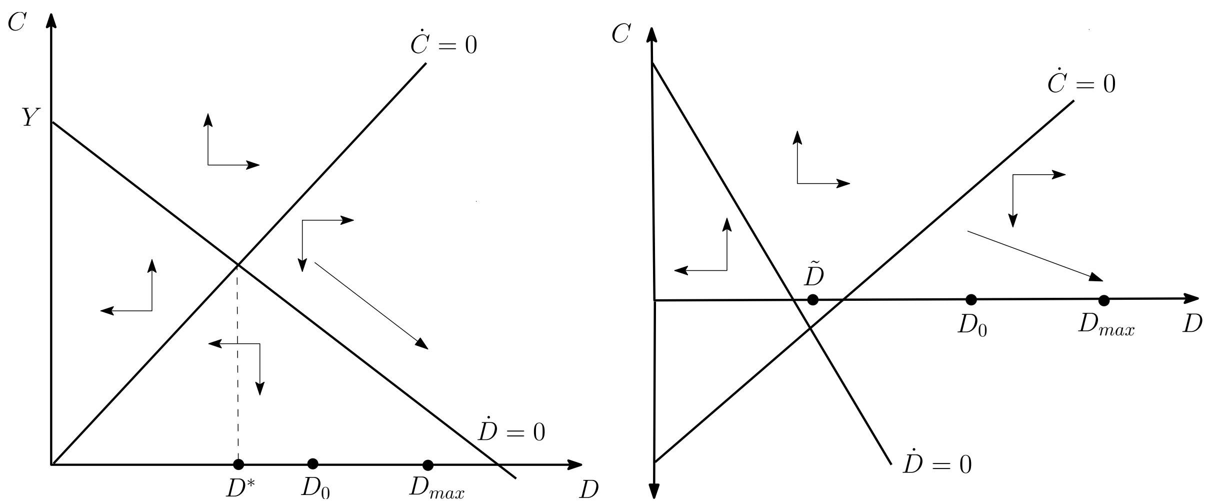
From (20) we see that at the steady state, $(1-\alpha)\mu AC/(\bar{H}-D) = (\rho-\mu)\alpha p$. The left hand side of the equation is strictly positive, implying that $\rho > \mu$ is a necessary, not sufficient condition for a steady state of eternal life to exist. The condition is not sufficient because consumption at the intersection of the isoclines could be negative such that no positive steady state exists. Specifically, the intersection of the isoclines is obtained from (20) and (21) as

$$C^* = \frac{\alpha\rho - \mu}{\alpha(\rho-\mu)}\cdot\frac{A}{AY - p\bar{H}}, \qquad D^* = \bar{H} - \frac{(1-\alpha)\mu A}{\alpha p(\rho-\mu)}C^*. \quad (22)$$

Notice that for $\alpha\rho > \mu$ consumption is positive at the intersection of the isoclines only if $AY > p\bar{H}$. This means that one can always find an A or Y low enough such that no positive intersection exists. In other words, eternal life does not exist if medical technology or income are assumed to be sufficiently low.

Altogether we have to distinguish three cases: (i) the isoclines intersect in the positive quadrant (a steady state of eternal life exists), (ii) the isoclines intersect at a negative value for consumption, (iii) no intersection. Figure 3 shows the first two cases, obtained for $\rho > \mu$. The panel on the left side shows the case in which a fixed point of eternal life, characterized by constant health deficits D^*, exists. As indicated by the arrows of motion, the fixed point is globally unstable. In contrast to the health-capital model, there exists no life trajectory leading to eternal life. If people start out with more health deficits than tolerable for eternal life, consumption is monotonously declining as more health deficits are accumulated, as shown by the path from D_0 to D_{max} in the diagram. Otherwise, there might be an initial increasing branch of consumption. In any case, consumption is declining late in life, meaning that – in contrast to the health-capital model – health expenditure is increasing with declining health.

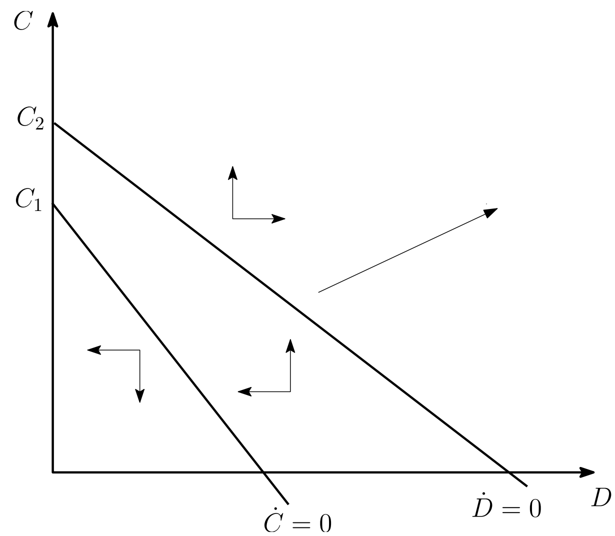
Figure 3: Health Deficit Accumulation and Consumption ($\rho > \mu$)



The panel on the right hand side of Figure 3 shows the case in which there exists no positive steady state because income is too low or health technology is too weak to allow for an eternal life. The isoclines intersect at $\tilde{D}$ in the negative quadrant. For $D_0 > \tilde{D}$, health deficits are increasingly accumulated with age and consumption declines, implying that health expenditure increases with declining health. Death in any case is the constraint optimal solution because income and technology do not (yet) allow for an eternal life. These features are a natural consequence of the modeling of health deficits as suggested by gerontologists. It is not necessary to introduce a third state variable that moves the $\dot{D}$–isocline with age. Age as such is not a determinant of health.

Figure 5 shows the phase diagram when $\rho < \mu$, i.e. when the isoclines do not intersect. The $\dot{C} = 0$–isocline hits the ordinate at $C_1 < C_2$ and thus lies everywhere below the $\dot{D} = 0$-isocline. Figure 4 thus shows the unique way to draw the case of $\rho < \mu$. Interestingly, increasing health deficit accumulation is now associated with increasing consumption, i.e. declining health expenditure. The result, though perhaps unexpected, is intuitive. When $\rho < \mu$, the time preference is smaller than the rate at which the body deteriorates. Consequently, it is optimal to make most health investments early in life and consume more at a later stage. Depending on parameter choice the model is thus capable to explain both increasing and declining health expenditure patterns.

Figure 4: Health Deficit Accumulation and Consumption ($\rho < \mu$)



In a more general version of the model (Dalgaard and Strulik, 2014) we integrate access to the capital market. Then, ρ is replaced by rate of return on capital investment r and individuals optimally chose a declining health investment trajectory when $r < \mu$. Again, the result is intuitive. If the rate of return on savings is smaller than the rate of bodily deterioration, it is optimal to invest in health early in life. On the other hand, for $r > \mu$, it is optimal to delay most health investments to later stages of life and save at young age for the large health expenditures expected in old age. Empirically, the force of aging is precisely estimated to be between 3 and 4 percent (Mitnitski et al., 2002). A long-run interest rate of around 6 percent as suggested by the historical record (Siegel, 1992) would then explain the more frequently observed case of increasing health expenditure with age.

4. Health and Productivity

4.1. **Health-Capital Model.** We next turn to the "pure investment model", i.e. we assume that, aside from its impact on longevity, health expenditure exerts a positive effect on productivity. According to Grossman's original version , productivity is a function of an individual's production of healthy time and sick time, which are functions of health capital. For simplicity, we consider here a "reduced form" approach in which productivity and thus income is a strictly concave function of an individual's health status. This does not change the basic mechanism of the model and has the convenient side effect that the model becomes structurally equivalent to the neoclassical growth model. Since this textbook model is basically known by every economist, the transitional dynamics are particularly easy to convey. Specifically we assume that income is given by

$$Y = \theta H^{\alpha}, \tag{23}$$

in which α controls the return to health in terms of productivity, $0 < \alpha < 1$.

Following the tradition in the literature we now ignore consumption aspects of health such that life time utility is given by

$$\int_0^T U(C)e^{-\rho t}\mathrm{d}t \tag{24}$$

with $U(C) = (C^{1-\sigma} - 1)/(1-\sigma)$ for $\sigma \neq 1$ and $U(C) = \log(C)$ otherwise. Individuals maximize (24) subject to (23) and (4) and (5), obeying the boundary conditions $H(0) = H_0$ and death happens at $t = T$ when $H(T) \leq H_{min}$, as before. The first order conditions for a maximum are

$$C^{-\sigma} = \lambda_H A/p \tag{25}$$

$$\lambda_D \left(\frac{A}{p}\theta\alpha H^{\alpha-1} - \delta\right) = \lambda_H \rho - \dot{\lambda}_H, \tag{26}$$

in which λ_H is the shadow price of health. Additionally the optimal solution fulfils the transversality condition (9). Equation (25) equates the return from health investment on the right hand side with the opportunity cost in terms of forgone utility from consumption. Equation (26) shows that the productivity effects of health lead to a less steeply optimal increase of shadow price of health. Proceeding as in Section 3.1 we obtain from the first order conditions an equation of motion for consumption (27), which together with the equation of motion for health capital (28) constitutes the dynamic system:

$$\sigma\frac{\dot{C}}{C} = \frac{A}{p}\theta\alpha H^{\alpha-1} - (\delta + \rho) \tag{27}$$

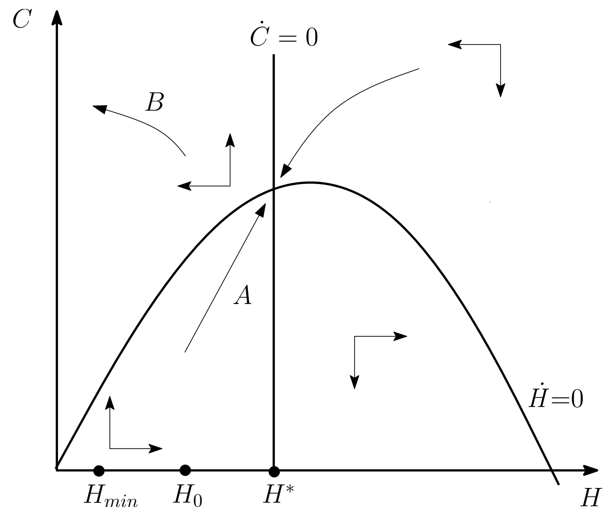
$$\dot{H} = \frac{A(\theta H^{\alpha} - C)}{p} - \delta H. \tag{28}$$

Setting for a moment $H = K$ and $A = p = 1$, it becomes evident that the model is structurally identical with the neoclassical growth model (e.g. Barro and Sala-i-Martin, 2005, Chapter 3). Since transitional dynamics for this type of model are well-known, the discussion can be brief.

From (27) and (28) we observe a unique steady state of eternal life at

$$H^* = \left[\frac{\alpha\theta A}{p(\delta + \alpha\rho)}\right]^{\frac{1}{1-\alpha}}, \qquad C^* = \theta(H^*)^{\alpha} - \frac{\alpha p}{A}H^*. \tag{29}$$

Figure 5: Investment Model (Health Capital Accumulation)



Observe from (27) that the $\dot{C} = 0$-isocline is a horizontal line through H^* and that $\partial \dot{C}/\partial C < 0$ for $H > H^*$ such that the arrows of motion point towards the isocline. From (28) the $\dot{H} = 0$-isocline is given by the concave curve $C = f(H) = \theta H^\alpha - \alpha p A H/A$ and $\partial \dot{H}/\partial C < 0$. The arrows of motion point towards lower health above the isocline and towards better health below the isocline.

Figure 5 shows the famous phase diagram, which is known (albeit with capital K at the axis) from the textbooks on economic growth. The health isocline is intersected by the consumption isocline before it reaches its maximum.[9] The intersection identifies the unique steady state of eternal life. The steady state fulfils the transversality condition (9c) since health and consumption and thus the shadow price of health are constant. This means it is optimal from any initial endowment of health capital to converge towards eternal life. However, if a researcher ignores this result and enforces convergence towards a terminal state of health H_{min}, then consumption is increasing, as shown by path B, implying that the health expenditure share declines as health deteriorates.

As for the "pure consumption" model, the introduction of an age-dependent depreciation would only seemingly solve these problems. Diagrammatically, it would shift the consumption isocline and thus H^* to the left. But the steady state of eternal life continues to exist and it remains optimal to converge to it. Only if H^* falls below H_{min}, death occurs. But this situation

[9]In order to see this, compute from $f'(H) = 0$ the maximum of the isocline at $H_{max} = [\alpha\theta A/(\alpha p)]^{1/(1-\alpha}$ and observe $H_{max} < H^*$.

violates Assumption 1. Death occurs at a health capital stock that is larger than the one needed for eternal life. In summary, the investment-variant of the health-capital model displays all the troubling features of the health-consumption variant.

4.2. **Health Deficit Accumulation.** In order to maintain the comparability with the health-capital model we again define a state of best health $\bar{H}$, at which human productivity is largest. It is achieved in the absence of any health deficits. Analogously to the health-capital model, individual income is given by

$$Y = \theta(\bar{H} - D)^{\alpha} \tag{30}$$

for $D < \bar{H}$ and $Y = 0$ otherwise. As for the health-capital case, individuals maximize lifetime utility from consumption (24). They face the equation of motion for health deficits (14), the budget constraint (4), productivity according to (30) and the initial health deficits $D(0) = D_0$. They die at time T when D exceeds D_{max}. In the following we focuss on the case where $\bar{H} > D$ such that health actually matters for productivity.[10] The first order condition for maximizing the associated Hamiltonian are:

$$C^{-\sigma} + \lambda_D A/p = 0 \tag{31}$$

$$\lambda_D \left[\mu + \mu \frac{A}{p} \alpha\theta(\bar{H} - D)^{\alpha-1} \right] = \lambda_D \rho - \dot{\lambda}_D \tag{32}$$

As before, the first order condition can be reduced to an equation of motion for optimal consumption (33). Inserting income (30) into (14) provides the equation of motion for health deficits (34) such that the optimal life trajectory fulfils (33)–(34) and the transversality condition.

$$\sigma \frac{\dot{C}}{C} = \mu - \rho + \mu \frac{\alpha\theta A}{p} (\bar{H} - D)^{\alpha-1} \tag{33}$$

$$\dot{D} = \mu D - \mu \frac{A}{p} \left[\theta(\bar{H} - D)^{\alpha} - C \right]. \tag{34}$$

Setting the left hand side of (33) to zero we obtain the consumption isocline as horizontal line through D_1 with

$$\bar{H} - D_1 = \left[\frac{\mu\alpha\theta A}{(\rho - \mu)} \right]^{\frac{1}{1-\alpha}}. \tag{35}$$

Since $\bar{H} - D > 0$ for health to matter and for income to be positive, $\rho > \mu$ is a necessary assumption for the isocline to exist. From (33) we see that $\partial\dot{C}/\partial D > 0$ such that the arrows of

[10]Otherwise we would obtain $\dot{\sigma}C/C = \mu - \rho$ independently from the state of health and $\dot{D} = \mu(D - AC/p)$.

motion point towards higher consumption at the right hand side of the isocline. From (34) we obtain the $\dot{D} = 0$-isocline as

$$C = \theta(\bar{H} - D)^{\alpha} - \frac{p}{A} D. \tag{36}$$

From (34) we have again that $\partial \dot{D}/\partial D > 0$ such that the arrows of motion point away from the isocline. The isocline cuts the D-axis at $D2$, which is implicitly given by

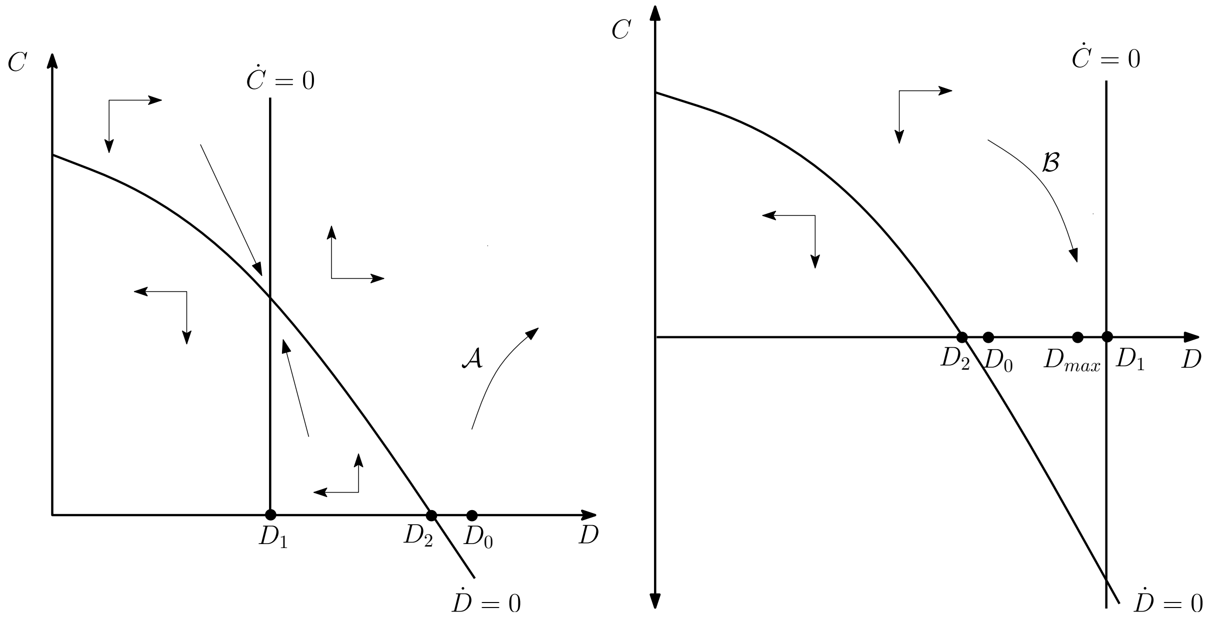
$$\bar{H} - D_2 = \left(\frac{pD_2}{A\theta}\right)^{\alpha}. \tag{37}$$

The intersection is thus at some D lower than $\bar{H}$, i.e. in the relevant range where health matters for productivity.

The intersection of the isoclines identifies a fixed point. Similar to the health-consumption case, the fixed point may be in the positive quadrant, such that a steady state of eternal life exists, or in the negative quadrant such that no steady state exists. Figure 6 shows both cases. The panel on the left hand side shows the case where the fixed point is located in the positive quadrant. It is identified as a saddlepoint. However, in contrast to the health-capital model, eternal life cannot be approached from everywhere. If the initial endowment of health deficits D_0 is larger than D_2, eternal life is not an option. Increasing health deficit accumulation is associated with more consumption and thus less health expenditure as, for example, along the trajectory $\mathcal{A}$. This outcome however could be avoided by assuming a sufficiently low state of the health technology.

The panel on the right hand side of Figure 6 shows the case when no steady state of eternal life exists, which is observed when $D_2 < D_1$. In this case, one can always find a life time trajectory along which individuals consume less as their health deteriorates as, for example, along trajectory $\mathcal{B}$. This does not necessarily imply that health expenditure in absolute terms rises since income, due to deteriorating health, declines. But it unambiguously implies that the health expenditure share rises with declining health. Most importantly, we can always find a health technology weak enough such that eternal life is impossible. To see this, compute $\partial D_1/\partial A < 0$ from (35) and $\partial D_2/\partial A > 0$ from (36). This means that, as the state of medical technology declines, D_1 shifts to the right and D_2 shifts to the left. Currently it is safe to assume that eternal life cannot (yet) be approached. The situations is thus characterized by sufficiently low medical technology such that the panel on the right hand side of Figure 6 is the relevant one.

Figure 6: Health and Productivity (Health Deficit Accumulation)



5. Conclusion

This paper has demonstrated that assumptions matter in the economics of health demand and longevity. It makes sense to listen to doctors and natural scientists in order to get health depreciation right. The health-capital theory and the health-deficit theory lead to fundamentally different predictions concerning longevity and the optimal age-health expenditure profile. The health-capital theory predicts that humans are immortal, irrespective of their initial conditions. Ignoring this prediction leads to the prediction that health expenditure declines as health deteriorates. The health-deficit model, in contrast suggest that one can always find a health technology low enough and an income level low enough such that immortality does not exists. The theory predicts that health expenditure usually increases as health deficits decline but allows as well, given certain parameters of the utility function, for a declining health expenditure pattern.

The crucial difference is that dynamics of the health-capital model by construction lead towards lower health H when the health stock exceeds its equilibrium value and towards better health when the health stock is below equilibrium. Self-equilibriating forces lead towards the fixed point of eternal life. Given the genesis of the model this outcome should not be surprising. The health-capital model was inspired by the human-capital model, which was inspired by the (neo-classical) growth model of physical capital accumulation. In neoclassical growth theory, inherent stability of the steady state was a desirable feature. But, as discussed by McFadden (2008), different stocks of capital follow different depreciation patterns and a reasonable modeling for physical capital constitutes not automatically a reasonable modeling for human health.

In contrast, dynamics of the health-deficit model lead away from a state of constant health. The reason is that the health-deficit model captures a fundamental insight of modern gerontology, namely that the speed of health deficit accumulation is a positive function of the health deficits that an individual already has.

This paper has compared simple versions of the two theories in order to allow for general inferences from phase diagram analysis. It is conceivable that amendments of the health-capital model exist that help avoiding its worrisome predictions. Yet, by introducing sufficiently many degrees of freedom any model can be perfectly fitted to the data. The goal for (health) economists should be to explain a lot by imposing only a little, another desirable feature of a good theory according to Friedman's (1953) methodology. We believe that the health-deficit theory fulfils this criterion. The main driving force of the theory is given by one single parameter, μ, the force of aging. The natural sciences provide precise estimates of the size of this parameter for men and women of different populations. For the health economists there are thus few degrees of freedom to 'manipulate' the theory. Instead, non-counterfactual predictions follow from simple principles of human gerontology.

Appendix: Robustness with Respect to Declining Returns of Health Investment

The main text has for simplicity focussed on linear returns to health investment. Taking into account that returns are potentially decreasing adds more realism to the model but it does not affect the central mechanism. The distinct adjustment dynamics of the two health models are driven by the assumption that according to Grossman healthier paper lose more health while according to Dalgaard and Strulik unhealthier people accumulate more health deficits. The core dynamics are independent from the specific health investment function. In order to verify this claim we focus on the health-consumption variant and furthermore on the case of $\sigma = 1$, i.e. the instantaneous utility function is $U(C, H) = \alpha \log C + (1 - \alpha) \log H$.

5.1. **Decreasing Returns in the Health-Capital Model.** For the health-capital model the equation of health capital accumulation (5) is rewritten as

$$\dot{H} = AI^{\gamma} - \delta H, \tag{A.1}$$

with $0 < \gamma < 1$. The individual maximizes life time consumption (3) given (A.1), the budget constraint (4) and the boundary conditions. The associated Hamiltonian reads:

$$J = \alpha \log C + (1 - \alpha) \log H + \lambda_H \left[AI^{\gamma} - \delta H \right].$$

From the first order condition and costate equation we obtain the equation of motion for optimal consumption (A.2) and after inserting the budget constraint the equation of motion for health capital is obtained as (A.3).

$$\frac{\dot{C}}{C}\left(1 + (1 - \gamma)\frac{C}{Y - C}\right) = \frac{(1 - \alpha)\gamma AC}{\alpha p^{\gamma}(Y - C)^{1-\gamma} H} - (\delta + \rho) \tag{A.2}$$

$$\dot{H} = A\left(\frac{Y - C}{p}\right)^{\gamma} - \delta H. \tag{A.3}$$

Observe from (A.2) that the sign of $\partial \dot{C}/\partial C$ equals the sign of $C(Y - C)^{\gamma - 1}|\frac{\partial}{\partial C}$ and thus $\partial \dot{C}/\partial C > 0$. The arrows of motion for consumptions point away from the $\dot{C} = 0$-isocline. From (A.3) we observe that $\partial \dot{H}/\partial H < 0$. The arrows of motion for health capital point towards the $\dot{H} = 0$-isocline. As obtained for the linear model, this central feature of the health-capital model implies again saddlepoint stability of the fixed point of eternal life.

Setting $\dot{C}/C = 0$ in (A.2) we cannot explicitly solve for the isocline $C(H)$. But we can solve H to get the inverse function $H(C)$:

$$H = \frac{(1-\alpha)\gamma AC}{\alpha p^{\gamma}(Y-C)^{1-\gamma}(\delta+\rho)}. \tag{A.4}$$

It is a strictly concave curve through the origin. The $\dot{H} = 0$-isocline is obtained from (A.3) as

$$C = Y - p\left(\frac{\delta H}{A}\right)^{\frac{1}{\gamma}}. \tag{A.5}$$

It is a monotonously falling curve originating from $C = Y$. Thus there exists a unique intersection of the isoclines as shown in Figure A.1. The arrows of motion identify the fixed point H^* as saddlepoint. Notice that the fixed point is approached from everywhere (for any initial condition) and for any set of parameter values (income, health technology, etc). Analogously to the analysis of the linear investment model it is verified that the fixed point fulfils the transversality condition (9c). It is optimal to live forever. However, if one would ignore this result, the phase diagram also reveals that any trajectory of declining health and terminal health lower than H^* (any trajectory fulfilling Assumption 1) implies that consumption rises, i.e. health expenditure declines, as individuals are getting unhealthier. The model replicates the main shortcomings of the linear investment model from the main text.

Figure A.1: Declining Returns to Health Investment (Health Capital Accumulation)

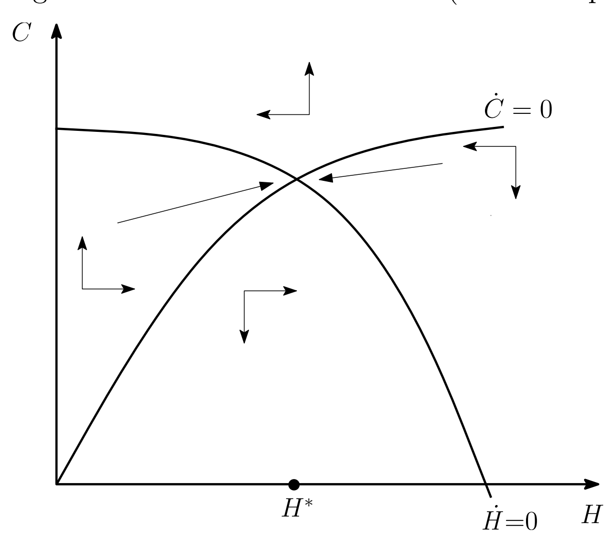
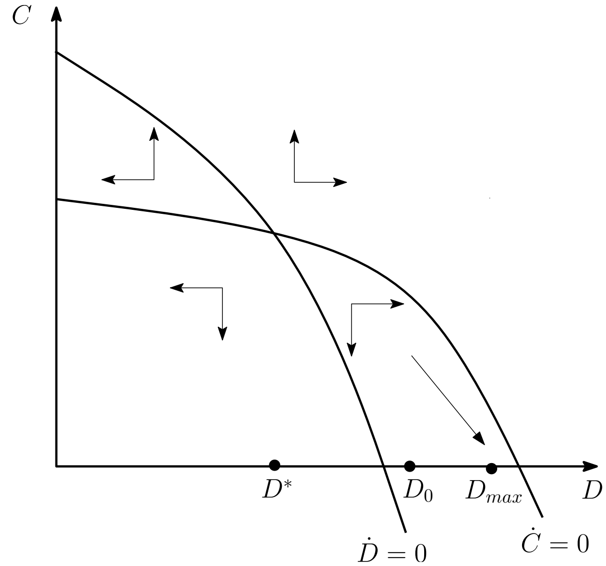


Figure A.2: Declining Returns to Health Investment (Health Deficit Accumulation)



5.2. **Decreasing Returns in the Health-Deficit Model.** The associated Hamiltonian for the health-deficit model reads

$$J = \alpha \log C + (1-\alpha)\log(\bar{H} - D) + \lambda D \left[\mu D - \mu A I^{\gamma}\right].$$

From the first order condition and costate equation we obtain the equation of motion for optimal consumption (A.6) and after inserting the budget constraint the equation of motion for health deficits is obtained as (A.7).

$$\frac{\dot{C}}{C}\left(1 + (1-\gamma)\frac{C}{Y-C}\right) = \frac{(1-\alpha)\gamma A C}{\alpha p^{\gamma}(Y-C)^{1-\gamma}(\bar{H}-D)} + (\mu - \rho) \tag{A.6}$$

$$\dot{D} = \mu D - \mu A \left(\frac{Y-C}{p}\right)^{\gamma}. \tag{A.7}$$

Notice that in (A.6) the term on the left hand side and the first term on the right hand side are both positive. Thus there exists no steady state of eternal life for $\mu > \rho$. As above, the sign of $\partial \dot{C}/\partial C$ equals the sign of $C(Y-C)^{\gamma-1}|\frac{\partial}{\partial C}$ and thus $\partial \dot{C}/\partial C > 0$. The arrows of motion for consumptions point away from the $\dot{C} = 0$-isocline. From (A.7) we observe that $\partial \dot{D}/\partial D > 0$. The arrows of motion for health deficits point away from the $\dot{D} = 0$-isocline.

Again, we cannot explicitly solve for the consumption isocline $C(D)$. But setting $\dot{C}/C = 0$ in (A.6) we can solve for D to get the inverse function $D(C)$:

$$D = \bar{H} - \frac{(1-\alpha)\mu\gamma A C}{\alpha p^{\gamma}(Y-C)^{1-\gamma}(\rho - \mu)}. \tag{A.8}$$

Consider the case $\rho > \mu$, i.e. the necessary condition for a steady state to exist. Then the isocline has negative slope with $\partial D/\partial C \to -\infty$ for $C \to 0$. The $\dot{H} = 0$-isocline is obtained from (A.7) as

$$C = Y - p\left(\frac{D}{A}\right)^{\frac{1}{\gamma}}. \tag{A.9}$$

It is a monotonously falling curve originating from $C = Y$. Thus there exists a unique intersection of the isoclines as shown in Figure A.2. The arrows of motion identify the fixed point D^* as globally unstable. No life trajectory leads to immortality. A plausible trajectory is shown in Figure A.2. Deteriorating health (increasing health deficits) are observed together with falling consumption, i.e. increasing health care expenditure.

References

Arking, R. (2006). *The Biology of Aging: Observations and Principles*, Oxford University Press, Oxford.

Becker, G.S. (1967). *Human Capital and the Personal Distribution of Income: an Analytical Approach*, Woytinsky Lecture No. 1, University of Michigan. Ann Arbor.

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

Blanchard, O.J. (1985). Debt, deficits, and finite horizons. *Journal of Political Economy* 93, 223-247.

Blanchard, O. J. and Fischer, S. (1989). *Lectures on Macroeconomics*, MIT press, Cambridge MA.

Case, A. and Deaton, A. S. (2005). Broken down by work and sex: How our health declines. In: Analyses in the Economics of Aging (pp. 185-212). University of Chicago Press.

Cropper, M. L. (1977). Health, investment in health, and occupational choice. *Journal of Political Economy* 85, 1273-1294.

Dalgaard, C-J. and Strulik, H. (2012). The genesis of the golden age: Accounting for the rise in health and leisure. Working Paper, University of Copenhagen.

Dalgaard, C-J. and Strulik, H. (2014). Optimal aging and death: Understanding the Preston Curve, *Journal of the European Economic Association* 12, 672-701.

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.

Eisenring, C. (1999). Comparative dynamics in a health investment model. *Journal of Health Economics* 18, 655-660.

Forster, M. (2001). The Meaning of death: Some simulations of a model of healthy and unhealthy consumption. *Journal of Health Economics*, 20, 613-638.

Galama, T., and Kapteyn, A. (2011). Grossman's missing health threshold. *Journal of Health Economics* 30, 1044-1056.

Galama, T. J., Hullegie, P., Meijer, E., and Outcault, S. (2012). Is there empirical evidence for decreasing returns to scale in a health capital model? *Health Economics* 21, 1080-1100.

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

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

Grossman, M. (2000). The human capital model. *Handbook of Health Economics*, Vol. 1, 347-408.

Hall, R.E. and Jones, C.I. (2007). The value of life and the rise in health spending. *Quarterly Journal of Economics* 122, 39–72.

Jacobson, L. (2000). The family as producer of healthan extended Grossman model. *Journal of Health Economics* 19, 611-637.

Kuhn, M., Wrzaczek, S., Prskawetz, A., and Feichtinger, G. (2012). Optimal choice of health and retirement in a life-cycle model, Vienna University of Technology, Working Papers in Economic Theory and Policy No. 01/2012.

Laporte, A. and Ferguson, B. S. (2007). Investment in health when health is stochastic, *Journal of Population Economics* 20, 423-444.

Liljas, B. (1998). The demand for health with uncertainty and insurance. *Journal of Health Economics* 17, 153-170.

McFadden, D. (2008). Human capital accumulation and depreciation. Applied Economic Perspectives and Policy 30, 379-385.

Meier, V. (2000). On the demand for preventive care. *OR-Spektrum* 22, 381-402.

Mitnitski, A., 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.

Mitnitski, A.B., Mogilner, A.J., MacKnight, C., and Rockwood, K. (2002a). The accumulation of deficits with age and possible invariants of aging. *Scientific World* 2, 1816-1822.

Mitnitski, A.B., Mogilner, A.J., 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., Song, X., Skoog, I., Broe, G.A., Cox, J.L., 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., Lang, I.A., Rockwood, K. (2011). Measuring aging in individuals and populations: A deficit accumulation approach, Lepas Working Papers

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

Muurinen, J. M. (1982). Demand for health: a generalised Grossman model. *Journal of Health Economics* 1(1), 5-28.

Oniki, H. (1973). Comparative dynamics (sensitivity analysis) in optimal control theory. *Journal of Economic Theory* 6, 265-283.

Picone, G., Uribe, M., and Mark Wilson, R. (1998). The effect of uncertainty on the demand for medical care, health capital and wealth. *Journal of Health Economics* 17, 171-185.

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

Strulik, H. (2011). Health and education: Understanding the gradient. Discussion Paper, University of Hannover.

van Kippersluis, H. and Galama, T.J. (2014). Wealth and health behavior: Testing the concept of a health cost. *European Economic Review*, forthcoming.

Wagstaff, A. (1986). The demand for health: some new empirical evidence. *Journal of Health Economics* 5 195-233.

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