Center for European, Governance and Economic Development Research
cege

Discussion Papers

Number 400 – August 2020

FETAL ORIGINS – A LIFE CYCLE MODEL OF HEALTH AND AGING FROM CONCEPTION TO DEATH

Carl-Johan Dalgaard
Casper Worm Hansen
Holger Strulik

GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

ISSN: 1439-2305

Fetal Origins – A Life Cycle Model of Health and Aging from Conception to Death

Carl-Johan Dalgaard*, Casper Worm Hansen† and Holger Strulik**

August 2020.

Abstract. The fetal origins hypothesis has received considerable empirical support, both within epidemiology and economics. The present study compares the ability of two rival theoretical frameworks in accounting for the kind of path dependence implied by the fetal origins hypothesis. We argue that while the conventional health capital model is irreconcilable with fetal origins of late-in-life health outcomes, the more recent health deficit model can generate shock amplification consistent with the hypothesis. In order to discuss human health over the life cycle from conception to death, we develop a theory of ontogenetic growth in utero and during childhood, unify it with the theory of adult aging, and discuss the transmission of early-life shocks to late-life health deficit accumulation.

Keywords: Fetal Origins; Health Capital; Health Deficits; Ontogenetic Growth; In Utero Development.

JEL: I10, J13, D91.

*Department of Economics, University of Copenhagen, CAGE and CEPR, London. Contact info: Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen, Denmark; email: carl.johan.dalgaard@econ.ku.dk.

† Department of Economics, University of Copenhagen, Øster Farimagsgade 5, Building 26, DK-1353 Copenhagen, Denmark; email: Casper.Worm.Hansen@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

Half a century ago epidemiologists would tend to view the fetal state as a protected one. Since then epidemiological evidence has been accumulating that this appears not to be the case, which has spawned the fetal origins hypothesis. The fetal origins hypothesis suggests that health deficits in utero may cause morbidities in old age though without being directly visible for most of the life course (e.g., Almond and Currie, 2011a). Within economics, research has considerably strengthened the case that *in utero* (or early-in-life) shocks indeed appear to impact on late-in-life health (e.g., Almond, 2006; Van den Berg et al., 2006; Almond and Mazumder, 2011; Lin and Lui, 2014; Bhalotra and Rawlings, 2011; Kesternich et al., 2015; Scholte et al., 2015). In addition, research has demonstrated effects beyond late-in-life health, including human capital and labor market outcomes (e.g. Bleakley, 2007; Almond et al., 2009; Nelson, 2010; Bhalotra and Venkataramani, 2016; Scholte et al., 2015); welfare dependence (Almond, 2006; Oreopoulos et al., 2008), and even investment behavior (Cronqvist et al., 2016). From a broader perspective, the fetal origins hypothesis thus seems to be a promising avenue through which to gain further insights into the causes and intergenerational transmission of inequality (e.g., Currie, 2011). For surveys from economists' perspective on the vast literature of fetal or developmental origins, see Almond and Currie (2011b) Almond et al. (2018), and Conti et al. (2019).

The first contribution of this paper is to provide a general mechanism, based on principles of human aging established in medical science and gerontology, that allows for a straightforward discussion of fetal origins in health economics. We argue a new mechanism is needed in order to explain the process by which shocks in utero are amplified during the course of life. In the existing paradigm, which involves health capital accumulation (Grossman, 1972), initial differences are depreciated away as individuals grow older. In contrast, in the health deficit model (Dalgaard and Strulik 2014), initial health deficits are conducive to faster development of new deficits. Consequently, initial health differences become larger as individuals grow older: small initial differences that are perhaps only visible at the cellular level are amplified to be visible as differences in biomarker quality in young adults (Belsky et al., 2005), and further amplified to be visible as differences in diagnosed diseases and frailties in old age (Mitnitski et al., 2002a,b, 2016).

While the underlying mechanism of aging applies at all ages, the health deficit model has been so far only available for adults. Accordingly, the second contribution of this paper is to extend

it by a childhood period and to provide a health economic model for the whole human life cycle, from conception to death. The unified health economic theory makes it possible to explain how late-in-life health challenges can have origins in initial health conditions. For that purpose we first develop a theory of in-utero and childhood development in terms of body growth and health and then integrate it with the health deficit model.[1]

The theory of health deficit accumulation is built on two basic features imported from gerontology: the frailty index and the reliability theory of aging. The frailty index measures the number of health deficits that a person has, relative to the number of potential deficits (Mitnitski et al., 2002a; Searle et al., 2008). Reliability theory explains aging by the loss of function through accumulating damage in the redundant building blocks and elements of the body, such as organs, tissue, bones, cells etc. (Gavrilov and Gavrilova, 1991). While (cellular) damage occurs throughout life, from the time when the first cells and tissue begin to form until death (Kirkwood, 2005), the in-utero and childhood period are distinct from adulthood because the body grows. We conceptualize ontogenetic growth as the build-up period of redundancy in body cells and biological aging as accumulated damage of redundant body cells (in organs, tissue, bones, etc.). In terms of the frailty index, growth in utero and childhood increase the denominator (the number of potentially damageable cells) while health deficit accumulation increases the numerator.

The unified model of human aging makes it possible to discuss how specific shocks of nutrition and health damage in utero and childhood affect adult health and longevity. Health deficits, if they remain unrepaired, have always a more severe impact on subsequent health the earlier in life they occur. The reason is the quasi-exponential nature of health deficit accumulation (Mitnitski et al., 2002a, 2016, Abeliansky and Strulik, 2018a). The exponential (or, more generally, convex) association of health deficits with age is a formal expression of the generality of biological aging understood as "intrinsic, cumulative, progressive, and deleterious loss of function that eventually culminates in death" (Arking, 2006). According to the terminology developed in Dragone and Vanin (2020), the accumulation of health deficits is a self-productive process: the presence of many health deficits is conducive to the faster development of new deficits. It is a natural outcome of theories of aging built on the redundancy and interdependence of health deficits such

[1]The basic model of health deficit accumulation has been adapted to study, among other things, the link between health and education (Strulik, 2018), years in retirement (Dalgaard and Strulik, 2017), the gender-gap in mortality (Schünemann et al., 2017b), and the health gap between married and unmarried individuals (Schünemann et al., 2020).

as reliability theory (Gavrilov and Gavrilova, 1991) and network theories of aging (Rutenberg et al., 2018). Here, we show that the feature of self-productivity in health deficit accumulation explains how invisibly small damages in utero or childhood are amplified over the human life cycle to be expressed in large (visible) health deficits in old age.

Our modeling of childhood development has also a foundation in human biology. While there are several purely statistical approaches to formally represent ontogenetic growth (for a review, see Karkach, 2006), West et al. (2001) provide a theory that derives the growth curve of humans (and other animals) from first principles in thermodynamics, i.e. the energy needs to create and maintain body cells. The integration of this theory of childhood growth into the model of human aging allows us to discuss nutritional shocks in utero as well as in childhood and their impact on child growth and adult health in a unified and scientifically founded way. We show that growth dynamics imply a natural separation of childhood into two distinct periods; an early period, in which initial differences in body size and frailty are amplified by child growth; and a later period, in which initial differences are dampened by child growth. This provides a micro-foundation of the frequent assumption in economic models of child development that such distinct periods exist (Heckman, 2007; Almond and Currie, 2011a,b). Health damages, in contrast, are amplified everywhere along the human life cycle, from conception to death.

The paper proceeds as follows. In the next section we compare basic versions of the health capital model and the health deficit model in their ability to account for the fetal origins hypothesis and elaborate on their testable implications. In Section 3, we provide a discussion of the impact of initial conditions on lifetime health outcomes within the health deficit model where investments are optimally determined. In Section 4, we set up the model of ontogenetic growth and derive its implication for the transmission of health and nutrition shocks during childhood. In Section 5, we unify the childhood period with the health deficit model for adults and use the new theory to discuss fetal- and early-life origins of late-life health. In Section 6, we conclude.

2. Basic Models

2.1. **Health Capital Accumulation.** The survey by Almond and Currie (2011a) provides an illustration of the inability of the health capital model (Grossman, 1972) to account for fetal origins. The illustration has the following law of motion for health capital as the main ingredient:

$$H_t = (1-\delta)H_{t-1} + I_t,\ H_0 \text{ given},\ H_t > \underline{H} \tag{1}$$

in which H_t is the stock of health capital at age t, δ is a constant rate of health capital depreciation, I represents health investment, H_0 is the initial health capital stock, and $\underline{H}$ is a hypothesized lower boundary for health capital beyond which individuals expire. Repeated substitution leaves us with the following expression for the stock of health capital at age t:

$$H_t = (1-\delta)^t H_0 + \sum_{i=0}^{t-1} (1-\delta)^i I_{t-(i+1)}.$$

The key observation to make is that shocks *in utero* that influence initial health, H_0, depreciate away with the passing of time. In general, events in the past are far less important to current health than recent events. This is an inevitable consequence of the basic assumption in the health capital model that health depreciates in proportion to the stock of health. In principle, the model therefore imposes that healthy individuals age faster than unhealthy (or elderly) individuals, ceteris paribus. Consequently, initial conditions will be of little consequence later in life.

The panel on the left hand side of Figure 1 provides a numerical illustration of this point, replicating Figure 1 of Almond and Currie (2011a). It shows how an initial shock, which creates a 25 percent deviation in initial health to a reference individual, depreciates with age for three different rates of health capital depreciation. At five percent depreciation the initial 25 percent deviation is melted down to about a five percent deviation at age 30. At 15 percent, initial differences are basically equalized at age 30.[2]

As it turns out, the health capital model actually holds a stronger prediction than what is indicated by the experiment conducted in Figure 1. Observe that the absolute difference in health capital between two individuals ($i = 1, 2$ respectively) with different initial conditions (i.e., different H_0), in the absence of health investments, is given by:

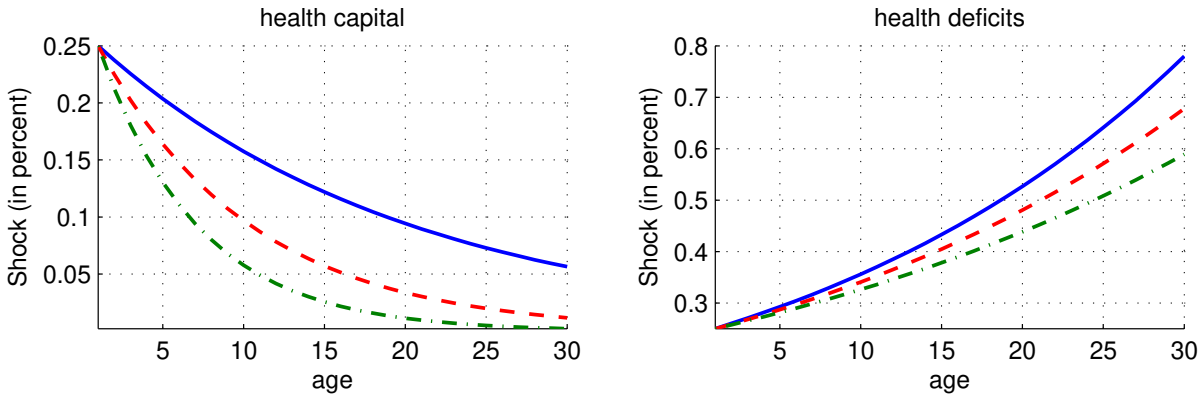
$$H_t^1 - H_t^2 = (1-\delta)^t \left(H_0^1 - H_0^2\right).$$

Hence, the health capital model implies a stronger version of non-persistence than convergence in relative health levels: namely, *absolute* convergence in health levels between individuals with different initial conditions, holding investments fixed.

[2]More formally, Figure 1 shows the impact on the long run relative level of health of two individuals (1 and 2, say) after one is hit by a shock at age zero:

$$d\left(H_t^1/H_t^2\right) = (1-\delta)^t \, dH_0^1$$

FIGURE 1: SHOCK PERSISTENCE BY AGE: HEALTH CAPITAL VS. HEALTH DEFICITS ACCUMULATION



The figure shows how persistent a 25 percent negative shock to the birth endowment would be given alternative annual depreciation rates. Left: blue (solid) line: 5 percent depreciation; red (dashed) line: 10% depreciation; green (dash-dotted): 15% depreciation. Right: blue (solid) line: $\mu = 0.04$; red (dashed): $\mu = 0.035$; green (dash-dotted): $\mu = 0.03$ ($E = 0.02$ and $D_0 = 0.02$).

So far we have assumed a constant rate of health depreciation, δ. Naturally, in the health capital model the depreciation rate is not constant, as it theoretically would enable individuals to "live forever" contingent on sufficient health capital investments (see e.g. Grossman, 1972, section III). Instead the depreciation rate is assumed to increase with the passing of time as the individual ages. Obviously, this only serves to strengthen the prediction that initial health shocks lose significance with the passing of time. In this case, depreciation of initial health differences become faster than geometric.

Notice that $\delta > 0$ is sufficient to generate convergence. This implies that convergence is also obtained in refinements of the health capital model. For example, we could allow for individual-specific and/or stochastic depreciation rates. In Appendix B, we discuss a variant of the health capital model where early-life shocks imply a greater depreciation rate of health capital. This leads to the prediction that the health impact of early-life shocks is amplified at young ages but preserves the feature of convergence of initial health differences in old age.

The exercise of Figure 1 and the subsequent discussion set health investments to zero. This may seem to open the door to a simple way of reconciling the health capital model with fetal origins, namely through investments. However, as pointed out by Almond and Currie (2011a, p. 158) in the context of the illustration depicted in Figure 1:

> If investments in all periods subsequent to the shock are affected by the shock, then prenatal exposures could be important for adult health in the Grossman

(1972) framework. However, the fetal origins literature posits an important and persistent biological effect of the prenatal period – that is, holding investments fixed.

It is important to appreciate that fetal origins, from the point of view of the medical literature, involve a specific mechanism. The early literature argued that environmental shocks would "program" the fetus with a predisposition towards various diseases, like coronary heart disease (e.g., Barker, 1995). Today a widespread view is that early-in-life shocks affect late-in-life health outcomes due to epigenetic changes. That is, changes in hereditary traits brought on by environmental influence (e.g., Gluckman and Hanson, 2004; Wu et al., 2004; Hilakivi-Clark and De Assis, 2006; Dolinoy el al. 2007; Waterland and Michels, 2007; Sinclair et al, 2007; Thompson and Einstein, 2010). Animal trials have been instrumental in providing proof of the principle of the fetal origin's hypothesis (McMullen and Mostyn, 2009, for a review). Accordingly, in order to fully account for the fetal origins hypothesis a theory would have to allow for an influence from initial conditions on long run health outcomes *conditional* on investments, for purely biological reasons. The health capital model does not allow for such a line of influence, as seen above.

Finally, it should be observed that while the standard health capital model does not suggest that initial conditions influence subsequent investments, the more recent work by Heckman (2007) does. The theory of human capability formation creates dynamic complementarities by assuming that health investments happen at two (or more) distinct periods in life such that health outcomes are produced with the distinct health investments as inputs. Since negative early-in-life shocks, or low initial investments, reduce the productivity of future investments, early-in-life events can have very persistent effects. The end result is not a given, of course. At present, the evidence in favor of dynamic complementarities seems to be largely descriptive in nature. Moreover, while some studies find that parental investments reinforce shocks, implying persistence in early-in-life shocks through investments, other studies find that parents act in a compensatory fashion.[3]

When dynamic complementarities are introduced into the health capital model the resulting framework can generate persistence, which is broadly consistent with the fetal origins hypothesis. As should be clear, however, initial conditions only influences eventual outcomes *via* investments.

[3]See Currie and Almond (2011b), Almond and Mazumder (2013), and Almond et al. (2018) for a detailed discussion of the Heckmann (2007) model and reviews of the parental investment literature. In Section 4 we propose a model of child development that endogenously generates a division of childhood in two distinct periods with shock amplification in early childhood and shock dampening in adolescence.

As a consequence, the "Grossman-Heckman" framework cannot account for an impact of initial conditions on late-in-life outcomes holding investments fixed, and by extension it cannot account for the fetal origins hypothesis as it is conventionally understood in the medical literature. Combining the "Heckman mechanism" with the health deficit model would not have this drawback, as will be clear from the discussion to follow.

2.2. **Health Deficit Accumulation.** The health deficit model of Dalgaard and Strulik (2014), in its simplest form, can be written as:

$$D_t - D_{t-1} = \mu(D_{t-1} - E), \text{ for } D < \bar{D}, \ D_0 \text{ given}, \tag{2}$$

where D is the frailty index. It measures the relative number of health deficits that a person has out of a long list of potential deficits. Accordingly, the index is defined on a 0 to 1 scale, and aging (declining health status) occurs as the index gradually traverses towards one. In general, individuals with a higher frailty index are to be considered more frail, and thus physiologically older. In practise the process of deficit accumulation continues until an upper boundary for deficits, $\bar{D}$, is reached at which point the individual expires. The parameter μ is the "natural" rate of aging, and E is an "environmental constant". Equation (2) derives from the literature on gerontology and the underlying parameters have been estimated with great precision (Mitnitski et al., 2002; Searle et al., 2008). By extension, it is worth noting that in contrast to the health capital model where the object of interest – health capital – is an unobserved variable, the object of interest in the present model – health deficits – is empirically observable.

Repeated substitution of (2) leads to

$$D_t = (1+\mu)^t D_0 - \sum_{i=0}^{t-1} \mu^{1+i} (1+\mu)^i E_{t-(i+1)}.$$

Since $(1+\mu)^t$ grows with increasing age t, an inherent feature of the health deficit model is that early-in-life shocks that influence the initial relative number of deficits, D_0, are amplified over time. This creates a force of divergence: initially unhealthier individuals accumulate health deficits faster than initially healthy individuals. The reason is that health deficit accumulation is a self-productive process (Dragone and Vanin, 2020): the presence of many health deficits is conducive to a faster development of new health deficits. Empirically this features is confirmed by the exponential (or, more generally, convex) association of health deficits with age (Mitnitski

et al., 2002; Searle at al., 2008; Shi et al., 2011; Harttgen et al., 2013; Mitnitski and Rockwood, 2013, 2016; Abeliansky and Strulik, 2018a,b, Abeliansky et al., 2020).

The feature of self-productivity of health deficits explains shock amplification and can explain why small health shocks early in life have large consequences in late life. The panel on the right hand side of Figure 1 provides a numerical illustration of this feature. As in the previous section we study the impact a health shock that creates a 25% deviation in initial deficits relative to a reference individual. The deviation from benchmark increases over time. For $\mu = 0.04$, the initial 25% deviation has reached 80% percent at the age of 30.[4]

As noted above, the interpretation of E in the natural science literature is that of "environmental" influences. While some such influence can be external to individuals (such as pollution), E may also be influenced by deliberate health investments. By increasing E such investments will serve to slow down the process of deficit accumulation and thus provide the prospect of a longer life. In the illustration in Figure 1, the level of E is ignored so as to provide a clean comparison with the properties of the health capital model in the absence of health investments. Nevertheless it is also of interest to understand the consequences of allowing for optimal health investments in the presence of shocks to initial deficits within the deficit model. The next section therefore studies the impact from initial deficits in the baseline health deficit model.

As a final remark on the properties of the basic deficit model its worth observing that it also holds radically different implications from the health capital model in terms of the evolution of *absolute* health differences. Comparing the absolute difference in health deficits between two individuals ($i = 1, 2$ respectively) with different initial conditions (i.e., different D_0), in the absence of health investments ($E = 0$), is given by:

$$D_t^1 - D_t^2 = (1+\mu)^t \left(D_0^1 - D_0^2\right).$$

Hence, initial differences in health deficits are amplified and the model thus predicts absolute divergence in health holding investments fixed.

[4]More formally, Figure 1 shows the impact on the long run relative level of health deficits of two individuals (1 and 2, say) after one is hit by a shock in utero (time zero):

$$d\left(D_t^1/D_t^2\right) = (1+\mu)^t \, dD_0^1$$

3. Shock Persistence in the Health Deficit Model

In this section we review the health deficit model of Dalgaard and Strulik (2014) with a special focus on the persistence and amplification of early-life health shocks, which were not addressed in the original contribution. We begin by rewriting equation (2) for a continuous notion of age and separating E into the impact of health investment on health deficit accumulation and a residual ϵ, capturing "real" environmental conditions:

$$\dot{D}(t) = \mu \left(D(t) - Ah(t)^{\gamma} + \epsilon \right). \tag{3}$$

Here, 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 γ specifies the degree of decreasing returns of health expenditure. The larger γ the larger the relative productivity of cost-intensive high-technology medicine in maintaining and repairing deteriorated human bodies. Bad health promotes death such that individuals die when $\bar{D}$ health deficits have been accumulated. Formally, this defines a free terminal time problem: the conditions for death are given but the length of life is variable and death occurs as an endogenous event, depending on life cycle choices.

Individuals are interested only in maximizing their lifetime utility from consumption:

$$\int_{\tau}^{T} \mathrm{e}^{-\rho(t-\tau)} u(c(t)) \ \mathrm{d}t, \tag{4}$$

with $u(c) = (c^{1-\sigma} - 1)/(1-\sigma)$ for $\sigma \neq 1$ and $u(c) = \log(c) + b$ for $\sigma = 1$. The parameter σ is the inverse of the elasticity of intertemporal substitution and ρ is the rate of time preference. Allowing for death to be a stochastic event and considering health as an element in the utility function leads to some further interesting results but does not change the basic insight on the accumulation of health deficits (see Strulik, 2015, Schünemann et al., 2017a). We thus focus on the simpler model here.

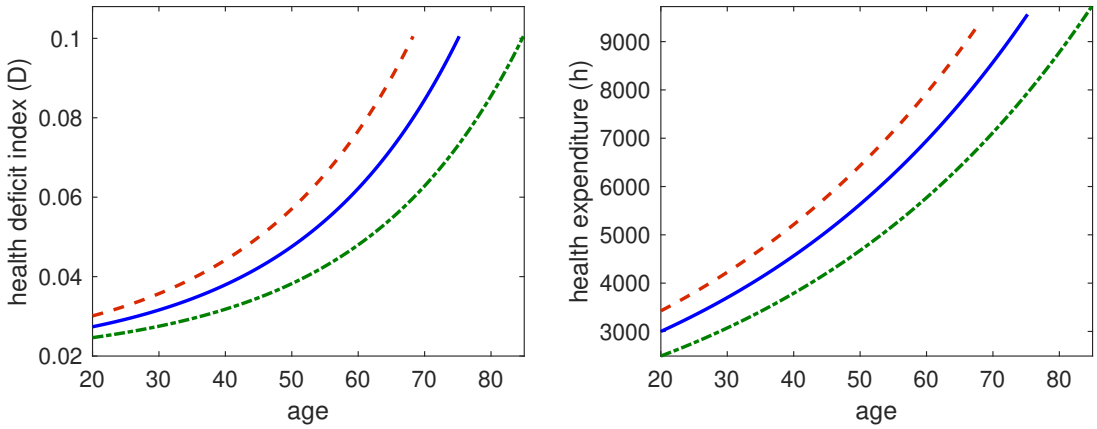
Besides spending income on final goods, individuals may save or borrow at a net interest rate r. Individuals take all prices as exogenously given. The law of motion for individual wealth k is thus given by (5):

$$\dot{k}(t) = w + rk(t) - c(t) - ph(t), \tag{5}$$

in which w is the (annual) wage, r is the interest rate, and p is the price of health goods. The problem is to maximize (4) subject to the accumulation equations (3) and (5), the initial conditions $D(\tau) = D_\tau$, $k(\tau) = k_\tau$, and the terminal conditions $k(T) = \bar{k}$, $D(T) = \bar{D}$. At the very basic level the problem is to trade off the benefits and costs of health investments over the life cycle. The benefits consists in, by slowing down the process of aging, a longer life which allows for more consumption along the extensive margin. However, by increasing health investments, individuals forego consumption in the current period. In Appendix A, we provide the details on the analytical solution of this free terminal time problem. Here, we present the numerical solution of a calibrated version.

We take the calibration of the model for an average 20 years old male U.S. American in the year 2000 from Dalgaard and Strulik (2014). This means that we set the rate of aging μ to 0.043, which is the rate of health deficit accumulation estimated for Canadian men by Mitnitski et al. (2002a). Rockwood and Mitnitski (2007) stress the similarity of their results for U.S. and Canadian populations but they do not report the detailed results for the U.S. analysis. Recently, Abeliansky et al. (2020, Table 6) found that US. American men living in the North East age at the same rate as Canadians in the Mitnitski et al. (2002a) study. We set the interest rate to 6 percent, following Barro et al. (1995), and we set $\gamma = 0.19$ to capture the growth of health spending at an annual rate of two percent over the life cycle (Keehan et al., 2004). From the estimates of Mitnitski et al. (2002a) we set $D(0) = 0.0274$ as the relevant initial value at age 20 and $\bar{D} = 0.1005$, i.e. 55.2 years later since the life-expectancy of a 20 year old U.S. American in the year 2000 was 55.2 years. We set $\epsilon = -0.013$ such that the model predicts a life-expectancy at age 20 of 42 years for $A = 0$ (corresponding to the life expectancy in the late 19th century when adult life expectancy was only modestly affected by medical technology). We set $\rho = r$ such that the age-consumption profile is constant over the life as obtained by Browning and Ejrnæs (2009) for childless households. We take GDP per worker in the U.S. in the year 2000 (PPP$ 77,003) and assume a capital share of 1/3, which implies an annual labor income (in international dollars) of $ 51,335. We normalize $p = 1$ and set $\sigma = 1$ in order to obtain a value of life at age 20 consistent with the estimate of Murphy and Topel (2006). Finally, we estimate $A = 0.00139$ such that the individual dies with deficits $\bar{D}$ at age 75.2, according to the life-expectancy of 20 years old U.S. Americans in the year 2000.

FIGURE 2: INITIAL HEALTH AND HEALTH DEFICIT ACCUMULATION



Blue (solid) lines replicate results for the Reference American in Dalgaard and Strulik (2014). Green (dashed) lines: individual with 10 percent less initial health deficits at age 20. Red (dash-dotted) lines: individual with 10 percent more initial health deficits at age 20.

In Figure 2 we replicate the benchmark run of Dalgaard and Strulik (2014), represented by blue (solid lines). We then look at an individual that is initially 10 percent less healthy than the Reference American, represented by red (dash-dotted) lines, and an individual that is initially 10 percent healthier than the Reference American. These differences in initial health deficits at age 20 can be thought of as resulting from negative health shocks earlier in life (and perhaps in utero). The model predicts that unhealthier individuals spend more on health, in line with observations in de Nardi et al. (2017). But the higher health expenditure is not not powerful enough to equalize initial health differences. In fact, initial health differences get amplified over time: as individuals age, the vertical distance between the individuals' deficit trajectories becomes larger, see the panel on the left-hand side of Figure 2. The underlying reason for this pattern is that initial deficits influence the effectiveness of health investments: the greater the health deficits the smaller the impact of a given amount of health investments in prolonging life. In this sense the model involves dynamic complementarities akin to those found in the human capability theory (Heckman, 2007).

4. ONTOGENETIC DEVELOPMENT AND CHILDHOOD HEALTH AND FRAILTY

4.1. **Basic Principles: Accumulation and Depletion of Redundancy.** The baseline health deficit model tracks the evolution of health deficits, health expenditure, and consumption over the life cycle of adults, starting from about the age of 20. Hence, in a strict sense, the analysis does not fully capture fetal origins. Fetal origins are "only" represented as initial values of health

deficits in adulthood. In this section we extend the theory to include a childhood period and an embryonic period.

There are at least two main differences between health development of embryos and children compared to adults. (i) There is no rational health investment from the side of the fetus or child. Instead the fetus receives nutrients and perhaps negative health shocks from the mother, and children receive nutrition and health investments from their parents. (ii) There is body growth. Child development can be understood as the accumulation of body cells such that an initially nonviable fetus gets healthier over time. Here we propose a model of ontogenetic growth that is naturally unified with the model of human aging through the notion of redundancy. Growth in utero and in childhood can be understood as the build-up of body-cell redundancy and organ reserve. Aging, in gerontology, is understood as depletion of redundancy in (functioning) body cells, i.e. as accelerated loss of organ reserve such that individuals become increasingly frail. It has been estimated that initially, as a young adult, the functional capacity of human organs is tenfold higher than needed for survival (Fries, 1980). Gavrilov and Gavrilova (1991) provide a micro-foundation of human aging from reliability theory, understood as the gradual loss of functionality of basic elements (such as body cells), which causes the loss of organ functionality and, eventually, death. By combining the periods of childhood (build-up of redundancy) and adulthood (depletion of redundancy) we arrive at a novel unified theory of human development from conception to death.

The frailty index has been developed to measure visible health deficits in elderly persons. However, in line with basic principles of reliability theory, damage and depletion of functioning body cells occurs at any age (Kirkwood, 2005). In animal studies it has been shown that loss of redundancy at the cellular level is positively associated with the frailty index (Rockwood et al., 2015). A study by Belsky et al. (2015) extends the measurement of aging to young adults, aged 26 to 38. Since in this cohort only 1% of members had been diagnosed with an age-related chronic disease, the study used biomarkers to measure physiological deterioration of multiple organ systems (pulmonary, periodontal, cardiovascular, renal, hepatic, and immune function). Biological age at chronological age 38 was computed from biomarker function and found to be normally distributed ranging from age 28 to 61. In line with the health deficit model it was found that biologically older individuals aged at a faster rate. On average, each year increase of biological age was associated with a 5 percent larger deterioration of biomarkers, implying

that deficits in organ systems grow exponentially at a rate of about 5 percent, similar to the growth of health deficits in elderly person, which were found to grow between 3 and 5 percent (Mitnitski et al., 2002, 2016; Abeliansky et al., 2020).

In order to apply the health deficit model to individuals of all ages, we thus interpret the frailty index as a measure of the relative number of health deficits at the organ or cellular level, i.e. as the number of damaged body cells (or organs) divided by the number of potentially damageable cells (or organs). We associate the beginning of adulthood with the end of ontogenetic growth (at about age 20). For adults, the denominator of the frailty index is given. Fetal and childhood development, in contrast, are characterized as the period of body cell accumulation, in which the denominator of the frailty index increases through the build-up of organ reserve and redundancy through ontogenetic growth. Formally, the frailty index is defines as $D = D_A/P$, in which P are potential deficits (at the cellular level) and D_A is the number of actual deficits. Using the new notation we rewrite equation (3) as:

$$\frac{\dot{D}_A}{P} = \mu \left(\frac{D_A}{P} - Ah^{\gamma} + \epsilon \right). \tag{6}$$

It shows the evolution of the frailty index for a given size of the denominator, i.e. it applies in adulthood, after ontogenetic growth. In utero and in childhood the denominator of D_A/P also grows because ontogenetic growth accumulates redundancy in cells and organs. The frailty index for children thus evolves as:

$$\left(\frac{\dot{D_A}}{P} \right) = \frac{\dot{D}_A}{P} - \frac{D_A}{P}\frac{\dot{P}}{P}. \tag{7}$$

The first terms on the right hand side is shared by children and adults. The second term is unique to children. It shows how the build-up of redundancy in the body reduces the frailty index. The missing link between childhood and adulthood is thus the accumulation of cell and organ reserve through ontogenetic growth.

4.2. **A Simple Theory of Ontogenetic Growth.** A theory of ontogenetic growth has been proposed by West et al. (2001). It is based on simple thermodynamic regularities and produces empirical observable age-growth patterns for humans and other animals. The growth equation originates from an energy balance stating that the energy consumed is equal to the energy used for the creation and maintenance of body cells such that $B = bm + e_c\dot{m}$, in which B is the energy consumed, m is the number of body cells, b is the energy needed to maintain a cell and

e_c is the energy needed to create a cell. Assuming that a cell has unit weight and that $e_c = 1$, the number of cells and thus the weight of the individual evolves as:

$$\dot{m} = B - bm. \tag{8}$$

Energy flow per unit of time is given by:

$$B = \min\left\{am^{\beta},\ \ \bar{B}\right\}, \tag{9}$$

in which $\bar{B}$ is an energy supply constraint. If $B = am^{\beta}$, energy supply is unconstrained and equation (9) states that the metabolic rate scales with the body mass of the child. This allometric relationship between energy consumption B and body mass m is known as Kleiber's Law (Kleiber, 1932). The scaling parameter β is estimated with high precision as 3/4 for mammals and almost all terrestrial animals, yielding the famous "mouse-to-elephant curve" (Brody, 1945). West et al. (1997) provide a microfoundation of the scaling law by showing that organisms, viewed as energy transporting networks that minimize energy dissipation, fulfil Kleiber's law with $\beta = 3/4$. The fact that $\beta < 1$ implies that larger bodies are more energy-efficient in the sense that they need less energy to maintain a body cell.

Inserting $B = am^{\beta}$ into (8) and solving the resulting Bernoulli differential equation provides the von Bertalanffy equation for body size (von Bertalanffy, 1957):

$$m(t) = \left\{a/b - \left[a/b - m(\tau)^{1-\beta}\right] \mathrm{e}^{-(1-\beta)b(t-\tau)}\right\}^{\frac{1}{1-\beta}}, \tag{10}$$

in which τ is the initial time and $m(\tau)$ is the initial body mass. Empirically, the von Bertalanffy equation is a good approximation of body growth if food supply is abundant and it is frequently used to describe ontogenetic growth of humans and other animals in biology and related natural sciences (Kooijman, 2000; Karkach, 2006). Graphically, the von Bertalanffy equation has a sigmoid shape. If growth would never stop, body size would converge towards $m = (a/b)^{1-\beta}$, as can be read off from (10) for $t \to \infty$. However, human growth is determinate, which means that it stops at a certain age $t = \tilde{T}$. While $\tilde{T}$ is idiosyncratic it is typically reached a few years after sexual maturity.

The sigmoid shape implies that there exists an inflection point (at age t_I) at which the shape of the growth curve changes from convex to concave. In order to develop this feature in more detail, consider the differential equation $\dot{m} = am^{\beta} - bm$ that follows from (8) and (9) in the

case of no energy supply restrictions. Taking the second derivative we obtain $\ddot{m} = \beta a m^{\beta-1} - b$. Noticing that the age-growth pattern is convex for $\ddot{m} > 0$, we conclude a convex growth for:

$$m(t) < m_I \equiv \left(\frac{\beta a}{b}\right)^{\frac{1}{1-\beta}}, \tag{11}$$

in which m_I is body size at the inflection point. Inserting m_I for $m(t)$ in (10) and solving for age, we obtain age at the inflection point (t_I) and conclude that human growth follows a convex pattern at ages:

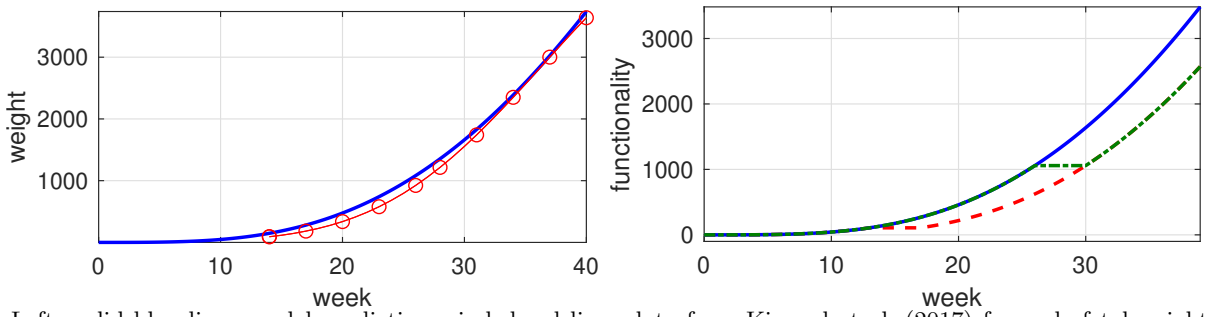
$$t < t_I \equiv \frac{1}{(1-\beta)b} \log\left[\left(\frac{a/b - m(\tau)^{1-\beta}}{a/b}\right)\frac{1}{1-\beta}\right] + \tau, \tag{12}$$

in which $m(\tau)$ is initial body size at age $\tau < t_I$. Since $\ddot{m} > 0$ for $t < t_I$, initial differences in body weight are amplified by body growth at ages smaller than t_I, whereas initial differences at ages larger than t_I are dampened. Notice that there exists at most one inflection point. If growth stops before m_I is reached, the growth curve is always convex and if initial size is larger than m_I, the growth curve is always concave. See Appendix C for a detailed formal discussion of shock amplification.

The literature usually analyzes in-utero growth and childhood growth separately, captured by two distinct growth curves. This distinction is intuitively plausible since the embryo is connected to the metabolism of the mother in an ambient temperature of about 37 degrees Celsius. These conditions change drastically with birth when the child depends on its own metabolism, loses energy through heat dissipation etc. Consequently in utero growth requires less energy for cell maintenance and creation (b and e_c are smaller in the simple model of ontogenetic growth). In other words, in utero growth is more energy efficient and faster. We thus consider both periods separately.

4.3. **In Utero Development.** In Figure 3, the panel on the left-hand side shows a calibration of the model for in-utero development when energy supply is unconstrained and there are no shocks. Parameters are $a = 1.08$, $b = 0.07$, $\beta = 3/4$ and $m(0) = 0.0001$. Body weight is measured in grams. The blue solid line shows the model prediction for body weight by week of gestation, the red dotted line shows the data for American boys from Kiserud et al. (2017). The age–growth curve is clearly convex during in-utero growth, implying that initial differences in body weight (and fetus frailty) will be amplified as the fetus develops.

FIGURE 3: ONTOGENETIC HUMAN GROWTH IN UTERO



Left: solid blue line: model prediction; circled red line: data from Kiserud et al. (2017) for male fetal weight during gestation. Right: nutritional constraint binding at $\bar{B} = bm(\tau)$ for 4 weeks (Ramadan) with onset at $\tau = 13$ (red dashed lines) and $\tau = 26$ (green dash-dotted lines).

Next consider energy constraints. If food is not abundant, fetal growth is constrained by maternal supply of energy (or, more generally, nutrients), $B = \bar{B}$. Naturally, if the fetus is energy-constrained, a larger share of energy is used for maintenance and less is available for cell creation such that growth is retarded in these periods. Suppose energy supply in utero is proportional to the metabolic rate of the mother and that the mother's metabolic rate scales with her size according to Kleiber's law such that $\bar{B} \propto M^{3/4}$ where M is the size of the mother and $M^{3/4}$ is the metabolic rate of the mother (Kleiber, 1932; West et al., 1997). We then conclude that, ceteris paribus, fetuses with larger mothers are less likely (or less frequently) energy-constrained and are thus larger at birth. More generally, we conclude from equations (8)–(10) that (i) fetuses who are less frequently energy-constrained in utero are bigger at any time in utero and thus also at birth, (ii) individuals who are born too early, i.e. at low t, are smaller at birth.

As an example application, consider a women who is pregnant during the Ramadan (Almond and Mazumder, 2011). We model this in a drastic sense such that for the four weeks during Ramadan there is no child growth and consider an onset of Ramadan at the beginning of the second or third trimester. Formally, we normalize the nutritional shock such that the nutritional constraint binds at $\bar{B} = bm(\tau)$ for 4 weeks with $\tau = 13$ and $\tau = 26$. The panel on the right hand side of Figure 3 shows the results. The solid line reiterates again the growth trajectory for unconstrained energy supply. Dashed lines show results for Ramadan onset at $\tau = 13$ and dash-dotted lines for $\tau = 26$. Since all body cells are treated equally, there is no impact of the timing of Ramadan during pregnancy. In order to explain a greater impact of early-life nutrition shocks one could differentiate between the importance of body cells and introduce a mechanism

why energy constraints have more severe impact early in the gestation period or one could argue that inferior nutrition does not only retard body growth but also induces health damages, a feature that we discuss in Section 4.5.[5]

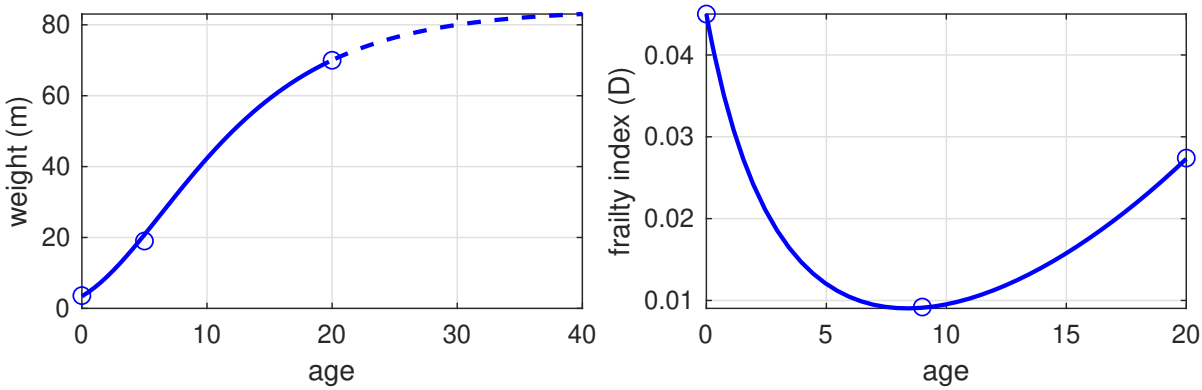
4.4. **Childhood Development.** From birth onwards, humans rely on their own metabolism and the metabolic parameters e_c and b change. Simple thermodynamics suggest that ontogenetic growth is still captured by the differential equation (8) and the nutrition constraint (9). Although the supply of energy (nutrients) in childhood is no longer determined by the metabolism of the mother, it is still determined by the provision of food by the parents. In Dalgaard and Strulik (2015; 2016) we developed a theory of human growth where nutritional investments were conceptualized as a choice variable of parents. Here, for simplicity, we take the position of the developing child for whom nutrition is exogenous. In case of sufficient food supply, $\bar{B} > am^b$, and child growth is still captured by the von Bertalanffy equation (10) with $\beta = 3/4$. However, the size of the parameters a and b changes, compared to in utero growth, due to the new metabolic conditions. Moreover, for child growth, it is more convenient to measure age in units of years and body mass in kilograms.

The panel on the left-hand side of Figure 4 shows a calibration of the growth curve that starts at average birth weight ($m(0) = 3.4$ kg) and a and b are targeted to match weight at age 5 and 20. This leads to the estimates $a = 1.52$ and $b = 0.50$. Dots show the actual average weight of US boys in the year 2000 according to CDC (2000). We assume that ontogenetic growth stops at age 20, as indicated by the end of the solid line. The dashed line indicates how growth would proceed if growth would be indeterminate and would never stop. The growth trajectory is mildly convex-concave with an inflection point at about age 7. This means that there is small amplification initial differences in body weight (and implied child frailty) in early childhood until about age 7 followed by a period of small dampening of initial differences in body weight until the onset of adulthood.

The final step is to map body weight m into organ reduncancy P. Here we venture into uncharted terrain since the frailty index has not yet been applied to children. As discussed above, in symptomless young adults, biomarkers measuring functionality of multiple organ systems

[5]Alternatively, one could argue that the energy constraint (12) is more likely to be binding if the fetus is already large. This feature would imply that the last trimester is more important for body growth and nutritional constraints. Such an outcome would be consistent with the finding of Doblhammer and Vaupel (2001) and Abeliansky and Strulik (2020) that individuals born in autumn (for whom fresh fruits have been abundant in summer) age slower than individuals born in spring.

FIGURE 4: ONTOGENETIC GROWTH IN CHILDHOOD AND THE FRAILTY INDEX



Left-hand side: solid line: calibrated model; dots: weight for age data for boys from CDC (2000); dashed line: hypothetical growth if growth were indeterminate. Right hand side: predicted frailty index; dots: constructed frailty index from mortality data.

decline at about the same rate at which measurable health deficits are accumulated among the elderly. The reliability theory of aging suggests that damages are accumulated at the same rate irrespective of age. We thus assume that the parameter of deficit accumulation are the same for children and adults. The numerical values are those from Section 3.

We assume that there are decreasing returns of redundancy. While this feature is obvious for some organs and tissues whose size is clearly limited by functionality (e.g. eyes and ears) it is perhaps less obvious for bone mass and muscle size. The mapping from size to functionality is thus coarse-grained and does not capture organ specifics. We assume that $P = \kappa m^{\alpha}$ and calibrate κ and α such the implied path of D approximates the evolution of a constructed frailty index for childhood.

In order to construct the frailty index, we exploit the feature that, for adults, it has been shown that the frailty index is a very good predictor of the mortality rate. The prediction of mortality can be so accurate that chronological age adds insignificant explanatory power when added to the regression (Rockwood and Mitnitski, 2007). At the population level, Mitnitski and Rockwood (2002b) obtain an R^2 of 0.99 in a simple log-log regression of the frailty index and the mortality rate. We thus hypothesize that a similar association exists between the frailty index of infants and children and their mortality rate. We use US mortality by age in the year 1999 for the calibration (cf. Figure 2 in Goldsmith, 2014).

For adults, mortality increases exponentially with age, a regularity known as Gompertz law. For children, however, mortality evolves non-monotonically. At birth it is as high as at age 50 (in the U.S.). For infants and small children, mortality declines quickly with age and starts

rising again at about age 10. A similar age-pattern is observed for the burden of chronic disease, measured by disability-adjusted life years: at birth the DALY is about as high as at age 50, it then declines quickly and starts rising again at about age 10 (see Belsky et al., 2015, Figure 1). We thus calibrate the remaining parameters κ and α in the following way. We take the parameters μ, A, ϵ, and γ as calibrated by Dalgaard and Strulik for adults. We feed into (6) average expenditure on child health care per year in the year 2000 (of \$1939) and then determine initial deficits and α and κ such that the solution of system (6)–(10) originates at a frailty index of 0.0450 at birth (equalling the calibrated frailty index at age 50), declines until age 10 and then rises again to a value of 0.0272 at age 20 (matching the actual value at age 20, stemming from the calibration of adults). This leads to the estimates $\alpha = 0.30$ and $\kappa = 0.18$.

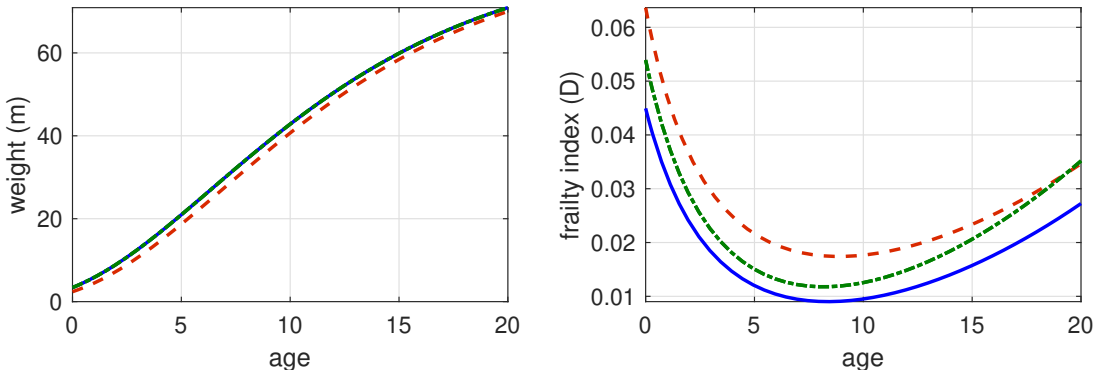
The evolution of the constructed frailty index is shown in the panel on the right-hand side of Figure 4. Although health damages at the cellular level start immediately when the first somatic cells are formed in utero (Kirkwood, 2005), ontogenetic growth and the formation of new body cells dominates during early childhood such that the frailty index declines and the young child becomes healthier and more robust as it grows. This process ends around puberty (at about age 11), the period, in which physiological functions gradually start to decline (Kirkwood and Mathers, 2009). In this period, body growth slows down and new damage accumulates faster than new cells and tissue are formed.

Although we argued that deficits are measured at the cellular level and thus their impact on diagnosed health impairments is only expressed when sufficient damage has been accumulated, it is interesting to note that infants and small children are also frail in the original meaning of the index. Several items that are frequently included in the frailty index applied to the elderly, apply to infants and small children as as well. For instance, infants and small children have "difficulties walking", "difficulties lifting weight", and have low grip strength. However, in contrast to the elderly they have a period of ontogenetic growth ahead, during which these difficulties are eliminated by the accumulation of redundant muscle and bone mass or, more generally, the accumulation of more healthy body cells.

4.5. **Two Types of Shocks and Childhood Frailty.** In utero and childhood shocks affect the frailty index in two distinct ways. Nutritional shocks affect the denominator of the frailty index through retarded ontogenetic growth and slower build-up of healthy body cells. Health shocks (infections and accidents) affect the numerator of the index by reducing the number of

functioning body cells. We could also imagine a combination of both shocks, e.g. if fighting infections needs energy and thus reduces the energy available for body growth. This could be conceptualized as an increase in the energy needed to maintain body cells, i.e. a larger b during the infection period.

FIGURE 5: FETAL ORIGINS: IN UTERO SHOCKS AND CHILDHOOD FRAILTY INDEX



Blue (solid) lines: reiteration of benchmark case (Figure 4). Red (dashed) lines: 1 kg lower birth weight. Green (dash-dotted) lines: 20 percent more health deficits at birth.

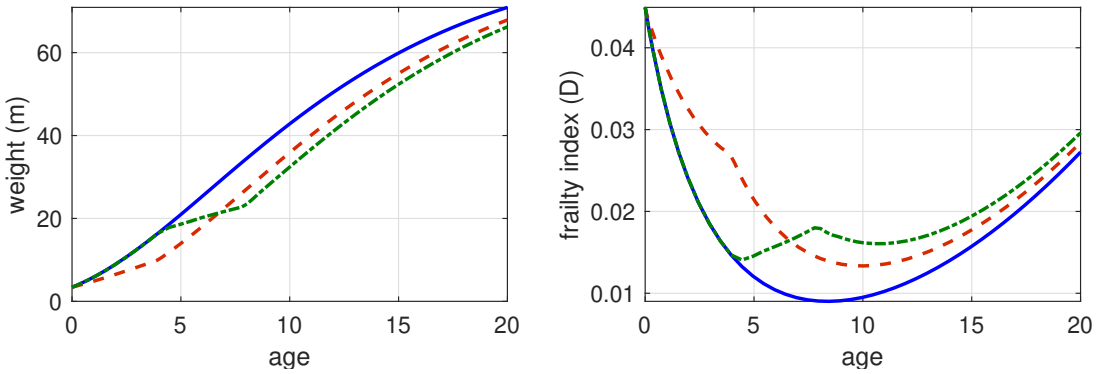
Figures 5 illustrates the two types of shocks. Blue (solid) lines reiterate benchmark development from Figure 4. Red (dashed) lines show the evolution of body growth and frailty when constrained nutrition in utero lowered birth weight by 1kg (to 2.4kg). Green (dash-dotted) lines show the development when the child is born with normal weight (3.4kg) but in utero health shocks increased health deficits at birth by 20 percent.

In case of low birth weight, the frailty index mildly diverges from the benchmark case during the convex growth phase of the body (which, as explained above, amplifies initial differences) and converges mildly towards benchmark in the concave growth phase (where initial differences are dampended). This pattern is distinct from the one obtained for birth with additional health deficits. Health deficits accumulate in a quasi-exponential way at any age and thus there is no convergence phase. Throughout life, the distance to benchmark gets larger, a feature that becomes salient in the second phase of childhood. For the qualitative result that early life shocks matter for adult health, however, the type shock as well as the assumed functional form is irrelevant. The only feature needed for 'fetal origins' is that shocks are at least partly preserved during childhood, i.e., diagrammatically, that the red and green curves do not converge fully back to the unshocked blue curve.

We next discuss shocks in early and middle childhood. We first consider a numerical experiment of prolonged nutritional restriction as, for example, experienced in the times of war

(Kesternich et al., 2015; Abeliansky and Strulik, 2018b). In Figure 6, the blue lines reiterate again the benchmark case (Figure 4). Red (dashed) lines reflect the case of food restriction in early childhood: from ages 0 to 4, children receive only 80 percent of the optimal energy provision for childhood growth. Green (dash-dotted) lines consider the same nutritional restriction experienced from age 4 to 8. Both cases lead to an initial divergence from healthy development with subsequent convergence in late childhood (when the body growth curve is concave). The later nutritional shock has a greater impact on childhood development and health at entry into adulthood. The reason is intuitive: Older children are bigger and need more energy for body cell maintenance. This means that the accumulation of new body cells is reduced more in older children when both age groups experience a nutritional shock of the same relative size. Most importantly, however, both childhood shocks are partly preserved by stunted body growth and increased frailty until entry into adulthood.

FIGURE 6: NUTRITION SHOCKS AT CHILDHOOD AGES 0 TO 4 AND 4 TO 8

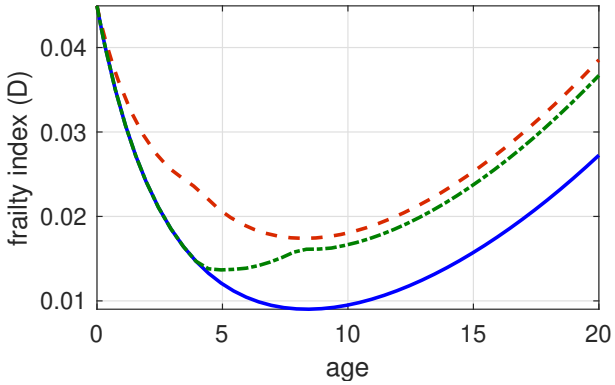


Blue (solid) lines: reiteration of benchmark case (Figure 4). Red (dashed) lines: energy supply 80% of optimal supply, $\bar{B} = 0.8am^b$, from age 0 to 4. Green (dash-dotted lines): energy supply 80% of optimal provision from age 4 to 8.

At first sight these considerations suggest that shocks during early childhood are worse than shocks in utero. But this is not generally the case. In particular, shocks that produce health damages in utero are of course worse than health damages that occur during early childhood, because they are self-amplified over a longer period by quasi-exponential growth of health deficits. With the same logic, health damages experienced in early child have worse consequences in old age than those experienced in late childhood. The next numerical experiment illustrates this feature. It considers a period of additional health deficit accumulation resulting, for example, from a prolonged period of infection or prolonged exposure to an unhealthy environment. In Figure 7, the blue (solid) line reiterates the benchmark case (from Figure 4). The red (dashed)

curve shows the consequences on the frailty index of a health shock ($\Delta\epsilon$) of 2 percent experienced from ages 0 to 4. The green (dash-dotted) curve shows the outcome when the same shock is experienced from ages 4 to 8. The body growth curve is omitted since, ontogenetic growth is, by assumption, the same in all cases. The experiment thus controls for the potential impact on body growth that an infection or detrimental environmental exposure may have. For both shocks there is no subsequent period of convergence to benchmark and the early shock has a stronger effect on the frailty index at entry into adulthood because damages have been accumulated in quasi-exponential fashion for a longer period of time in childhood.

FIGURE 7: HEALTH SHOCKS AT CHILDHOOD AGES 0 TO 4 AND 4 TO 8



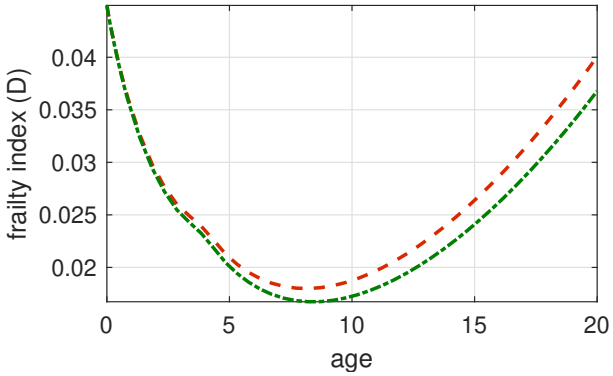
Blue (solid) lines: reiteration of benchmark case (Figure 4). Red (dashed) lines: health shock $\Delta\epsilon = 2$ percent, from age 0 to 4. Green (dash-dotted lines): same shock from age 4 to 8.

On the computer we can conduct controlled numerical experiments that disentangle nutritional shocks in terms of energy supply (affecting the denominator of the frailty index) and health damages (affecting the numerator). Actual shocks or positive health- or nutritional interventions from the parents, however, are likely to affect both childhood growth and health. Heckman (2007) and Almond and Currie (2011a,b) propose a model of childhood development inspired from production theory in economics. There, childhood is divided into two periods and child investments in the two periods combine via a CES production function to produce health or human capital of the child. For specific assumptions on the parameters, the result can be achieved that shocks matter more in the first period. Here, we consider child development based on biological foundations of ontogenetic growth and health deficit accumulation, without arbitrary separation into periods and without imposing an economic production function. We also observe shock dampening and amplification but the intricate interaction between child

growth and health suggests that it can be hard to make inferences about the timing of childhood shocks on future child development.

4.6. **The Widening Socioeconomic Gradient.** An apparent regularity of childhood development is that the socioeconomic gradient of childhood health widens with increasing age of children. The empirical literature has argued that the phenomenon is explained by income because richer parents are better able to purchase health care and provide safe environments (Case et al., 2002), or by the feature that children in families with low socio-economic status (SES) experience more or more severe health shocks (Currie and Stabile, 2003), or both (Condliff and Link, 2008). Our model of childhood development can motivate both channels due to the self-productive nature of health deficit accumulation. To see the second channel, it suffices to reconsider Figure 7 and assume that the low-shock outcome represented by the solid (blue) curve is associated with high SES background while the high-shock outcome represented by the dashed (red) or dashed-dotted lines is associated with low SES background. The SES-gradient, measured as the distance to the solid (blue) line, widens with child age, in particular for young and middle aged children. This numerical experience controls for income since $\bar{h}$ has been assumed to be the same for all three outcomes. This scenario is likely to be observed in an environment with a generous social safety net such that child care and child health care do not depend much on the individual income of parents.

FIGURE 8: HEALTH SHOCKS AND CHILD HEALTH EXPENDITURE



Health shock $\Delta\epsilon = 2$ percent, from age 0 to 4. Red (dashed) lines: half of benchmark child health expenditure ($\bar{h} = 968$). Green (dash-dotted lines): double of benchmark child health expenditure ($\bar{h} = 3874$).

To highlight the first channel, we consider in the next experiment two children exposed to the same health shock ($\Delta\epsilon = 0.02$ from age 0 to 4). In Figure 8, the child represented by red (dashed)

lines receives half of benchmark health expenditure ($\bar{h} = 968$) while the child represented by green (dash-dotted) lines receives double of benchmark health expenditure ($\bar{h} = 3874$). The social gradient (distance between the two lines) widens as the children grow older. This can be interpreted as increasing difficulty of low SES-parents to cope with a chronic condition developed in early childhood. The reason is the self-productive nature of health deficit accumulation.

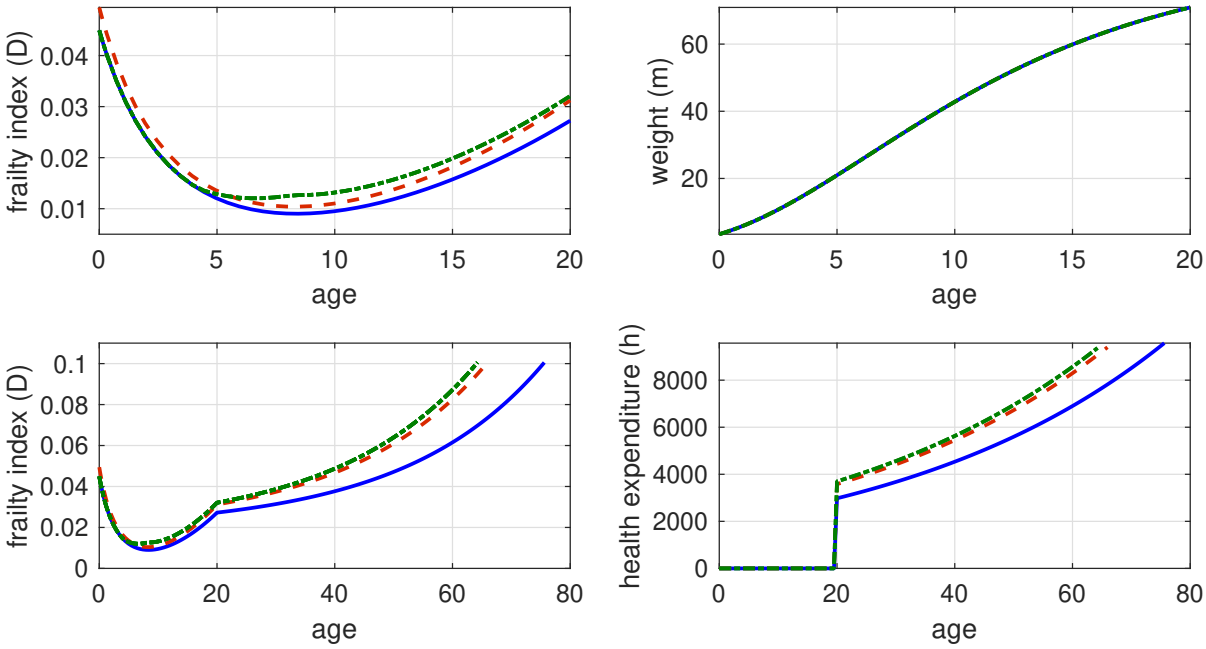
5. From Birth to Death: Growth, Aging, and Late-Life Health

Finally, we fit the pieces together and propose a unified model of child development and health at all ages from birth to death. Including an in-utero period would be straightforward but it would require a change of parameters because of the changed metabolic conditions when the child leaves the womb. We thus, represent in utero development in terms of initial values of birth weight $m(0)$ and frailty at birth $D(0)$ and represent the whole human life cycle (outside the womb) by a unique set of parameters. We take all parameter values as calibrated above, i.e. we take the values from Dalgaard and Strulik (2014) and Section 3 and add the childhood parameter values as calibrated in Section 4. The model is solved with the unique set of parameters in one go. However, with entry into adulthood (at age 20) several things happen: ontogenetic growth ceases, the person receives an income and he starts consuming, saving, and spending on health care in order to maximize lifetime utility.

Results for the benchmark run are shown by blue (solid) lines in Figure 9. Life cycle health and aging unifies the two periods that were previously explained by two separate models for adults (Dalgaard and Strulik, 2014, and Section 3) and for children (Section 4). The upper left panel is a zoom into the lower left panel, showing the first 20 years of development. The frailty index fits the actual data for adult men (from Mitnitski et al. (2002) and the curvature follows closely the curvature of the mortality rate (Goldsmith, 2014, Figure 2). It increases in a quasi-exponential way (following Gompertz law) in adulthood and is u-shaped in childhood. Only the mildly overshooting mortality in late adolescence is not traced by the frailty curve. This indicates that overshooting mortality is likely an outcome of non-aging related health damages explained by adolescent behavior, such as excessive drug consumption and reckless driving.

Red (dashed) lines in Figure 9 show the development of an identical individual that enters life with 10 percent more health deficits at birth than the benchmark individual. Due to the self-productive nature of deficit development the trajectory diverges from benchmark already in

FIGURE 9: CHILD DEVELOPMENT AND ADULT AGING



Blue (solid) lines: unshocked (optimal) growth in utero and childhood (death at 75.2). Red (dashed) lines: in utero health damage $\Delta D(0) = 10\%$. Green (dash-dotted) lines: Health shock in middle childhood: $\Delta\epsilon = 1\%$ from age 4 to 8. Parameters as for Figures 2 and 4-7.

childhood, although this is hardly visible from lifetime perspective (of the lower left panel). At age 20, health deficits exceed the benchmark by 20 percent. As the adult individual grows older, the divergence of the frailty index becomes more visible. In young adulthood, we can imagine these developments as differences in organ and tissues redundancy and biomarker quality as observed in the study of Belsky et al. (2015). In elderly individuals, health differences become visible as diagnosed health deficits (such as cardiovascular diseases or arthritis). Eventually, sufficiently many health deficits have been accumulated and death occurs, about 10 years earlier than in the benchmark case. Since adults behave fully rational, are endowed with the same preferences, receive the same income, and received the same nutrition and health expenditure in childhood, we identify fetal origins as the cause of adult differences in health and longevity.

Green (dash-dotted) lines in Figure 9 reflect development of an individual who experienced a period of exposure to unhealthy environment in childhood such that ϵ is 1 percent higher, compared to benchmark, from age 4 to 8. As result, health deficits are about 20 percent higher than benchmark at entry into adulthood. Subsequent development is similar as for the in-utero shock.

The panel on the right hand side of the figure shows the associated optimal health expenditure. Fetal- or childhood origins of inferior health are one possibility to motivate a negative association between health and health care expenditure. The standard health deficit model (as well as the health capital model) would predict the oppositive (with reversed causality), i.e. individuals that spend more on health are healthier. Here, we see that, for given age, less healthier individuals spend more on health in order to reduce their faster aging. The higher expenditure, however, does not fully remediate the initial disadvantage. In other words, the health difference between individuals diverges *despite* optimal countermeasures. Without the additional health investments of the shocked individuals, divergence would be greater.

A remaining question is whether individuals are able to observe slight deteriorations of health. Alternatively, these damages may go unnoticed in young and middle adult age and are only diagnosed in old age when the associated health deficits become sufficiently visible. Given such a visibility threshold, health deficit become earlier diagnosed in old age if they were initially larger. These perceptibility issues may explain why aging related health problems are diagnosed in old age although they are de facto always present and originate from sub-optimal development in young age and perhaps in utero. While health deficits are slowly accumulating (from e.g. mild hypertension to difficulties running or lifting weight to more severe cardiovascular problems), the empirical association between early-life health shocks and late life-life health outcomes is in many studies only observed when health problems became sufficiently severe. In this sense, the health deficit model, fills the "invisible" gap from early-life health shocks to health outcomes in old age.

6. Discussion and Conclusion

An influential strand of literature has provided convincing evidence in favor of the fetal origins hypothesis: *in utero* shocks have the ability to influence late-in-life outcomes. Relevant outcomes involve both health issues as well as a range of socio-economic outcomes. In this study we have argued that the current workhorse model of health economics, the Grossman (1972) model, is incapable of accounting for such effects. Indeed, since the notion of health – health capital– is analogous to physical capital, the model posits that health status depreciates more when the health status of individuals is high and less when the health status is low. These features imply, as demonstrated above, that the health capital model generates the prediction that individuals

with different initial conditions, prompted by *in utero* shocks, converge in health status during life, holding investments fixed. Convergence in health status in the aftermath of early-in-life shocks occurs both in a relative and in an absolute sense. This prediction is strengthened if one allows the health depreciation rate to grow over time, as it is required for the health capital model to be reconcilable with the fact of mortality. It is possible to generate important persistence in health outcomes through investments; for example, by introducing the human capability theory of Heckman (2007). But since the fetal origins hypothesis asserts an impact from *in utero* influences *conditional* on investments, the health capital model remains irreconcilable with the hypothesis.

The health deficit model offers radically different predictions. At its core the model conceptualizes aging as a continual process of loss of function – increasing frailty – that culminates in death. The notion of reduced functionality is captured by way of the frailty index: as humans age (health declines) the relative fraction of potential age-related health conditions climbs steadily upward. This underlying process, which can be slowed down by health investments, is exponential in nature. By implication, small differences in initial conditions at young age are amplified during life. The exponential nature of increasing deficits during life has been confirmed repeatedly by empirical work within gerontology. Overall, the deficit model seems well positioned to account for the type of dynamics implied by the fetal origins hypothesis.

Our study clarifies how an empirical researcher can discriminate between the health capital model and the health deficit model. Indirectly, any evidence of late-life health repercussions that can be causally related to early-life health shocks rejects the health capital model. This is so because the inherent mechanism of the health capital model is self-depleting: the loss of health is large when the state of health is good. Shock amplification, however, needs a self-productive process. The health deficit model is built on a self-productive process: the loss of health is large when health is bad. The health deficit model is thus not rejected by the observation that early-life health shocks matter for late-life health. Direct evidence in favor of the health deficit model, however, would employ a measure of health that can be tracked over time for individuals and ideally it would be based on the same metric to measure health deficits as the theoretical model, i.e. the frailty index.

A discriminating test between the two models is to check if the health effects of an environmental shock in utero (or during childhood) are increasing during adulthood. This is, of course,

somewhat data demanding in that it requires that the affected and non-affected individuals are observed more than once in adulthood (i.e., a panel of individuals). Yet first evidence along these lines is available. Abeliansky and Strulik (2018b) show that individuals exposed to hunger episodes during childhood diverge year-by-year during adulthood, in terms of health deficits, from comparable non-exposed individuals. Abeliansky and Strulik (2020) investigate a very mild health shock, namely the season of birth. They show that individuals born in autumn develop less health deficits than those born in spring and that the difference in health deficits gets larger with advancing age. This corresponds with the finding of Doblhammer and Vaubel (2001) that individuals born in autumn live longer than individuals born in spring.

In order to investigate the transition from early-life health to late-life health we have extended the health deficit mode by a childhood period. The unified model combines a period of aging and health deficit accumulation (originating from a loss in redundancy in organ reserve) with a period of ontogenetic growth (conceptualized as the build-up of cell redundancy and organ reserve). The model of ontogenetic growth model has been developed in a collaboration between biologists and physicists (West et al., 2001) and is based on energy needs for cell creation and maintenance that fulfil a simple law of energy conservation and Kleiber's scaling law for human metabolism. The model of ontogenetic growth naturally divides childhood in two distinct periods, a period of convex growth in utero and early childhood, in which initial differences in body weight and frailty are amplified by ontogenetic growth and a later period where initial differences are dampened. Health damage, in contrast, is always amplified in the course of human development. We have shown how early-life shocks are transmitted during childhood towards the onset of health deficit accumulation in adulthood and how they affect aging and health in old age. Taken together, we have thus proposed a new model of human development from conception to death, which motivates the fetal origins hypothesis of late-life health deficits.

References

Abeliansky, A. and Strulik, H. (2018a). How we fall apart: Similarities of human aging in 10 European countries. *Demography* 55, 341-359.

Abeliansky, A. and Strulik, H. (2018b). Hungry children age faster. *Economics and Human Biology* 29, 211-22.

Abeliansky, A. and Strulik, H. (2020). Season of birth, health, and aging. *Economics and Human Biology* 36, 100812.

Abeliansky, A., Erel, D., and Strulik, H. (2020). Aging in the USA: Similarities and Disparities Across Time and Space. *Nature Scientific Reports*, forthcoming.

Almond, D. (2006). Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population. *Journal of Political Economy* , 114(4), 672-712.

Almond, D., and Currie, J. (2011a). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives* 25(3), 153-72.

Almond, D., and Currie, J. D. (2011b). Human capital development before age five. In: *Handbook of Labor Economics*, Vol. 4, pp. 1315-1486. Elsevier, Amsterdam.

Almond, D., and Mazumder, B. (2011). Health Capital and the Prenatal Environment: The Effect of Ramadan Observance during Pregnancy. *American Economic Journal: Applied Economics*, 3(4), 56-85.

Almond, D., and Mazumder, B. (2013). Fetal origins and parental responses. *Annual Review of Economics* 5(1), 37-56.

Almond, D., Edlund, L. and Palme, M. (2009). Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden. *Quarterly Journal of Economics,* 124(4): 1729–72.

Almond, D., Currie, J. and Duque, V. (2018). Childhood circumstances and adult outcomes: Act II. *Journal of Economic Literature* 56(4), 1360-1446.

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

Barker, D.J. (1995). Fetal origins of coronary heart disease. *British Medical Journal* 311(6998), 171-174.

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

Belsky, D. W., Caspi, A., Houts, R., Cohen, H. J., Corcoran, D. L., Danese, A., ... and Sugden, K. (2015). Quantification of biological aging in young adults. *Proceedings of the National Academy of Sciences* 112(30), E4104-E4110.

Bhalotra, S., and Rawlings, S.B. (2011). Intergenerational persistence in health in developing countries: The penalty of gender inequality?. *Journal of Public Economics* 95(3), 286-299.

Bhalotra, S., and Venkataramani, A. (2016). Shadows of the Captain of the Men of Death: Early Life Health, Human Capital Investment and Institutions. *Working paper.*

Brody, S., 1945, *Bioenergetics and Growth*, Van Nostrand-Reinhold, New York.

Browning, M. and Ejrnæs, M. (2009). Consumption and children. *Review of Economics and Statistics* 91, 93–111.

Case, A., Lubotsky, D., and Paxson, C. (2002). Economic status and health in childhood: The origins of the gradient. *American Economic Review* 92(5), 1308-1334.

CDC (2000). Growth Charts: United States, Boys 2 – 20 years. National Center for Health Statistics in collaboration with the National Center for Chronic Disease Prevention and Health Promotion. `https://www.cdc.gov/growthcharts/`

Condliffe, S., and Link, C.R. (2008). The relationship between economic status and child health: evidence from the United States. *American Economic Review* 98(4), 1605-18.

Conti, G., Mason, G., and Poupakis, S. (2019). Developmental origins of health inequality, IZA Discussion Paper 12448.

Cronqvist, H., Previtero, A., Siegel, S., & White, R. E. (2016). The fetal origins hypothesis in finance: Prenatal environment, the gender gap, and investor behavior. *Review of Financial Studies* 29(3), 739-786.

Currie, J. (2011). Inequality at Birth: Some Causes and Consequences. *American Economic Review* 101(3), 1-22.

Currie, J., and Stabile, M. (2003). Socioeconomic status and child health: why is the relationship stronger for older children? *American Economic Review* 93(5), 1813-1823.

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

Dalgaard, C.J., and Strulik, H. (2015). The physiological foundations of the wealth of nations. *Journal of Economic Growth* 20(1), 37-73.

Dalgaard, C.J., and Strulik, H. (2016). Physiology and development: Why the West is taller than the rest. *Economic Journal* 126, 2292–2323.

Dalgaard, C.-J. and Strulik, H. (2017). The genesis of the golden age: Accounting for the rise in health and leisure. *Review of Economic Dynamics*, 24, 132-151.

De Nardi, M., Pashchenko, S., and Porapakkarm, P. (2017). The lifetime costs of bad health (No. w23963). National Bureau of Economic Research.

Doblhammer, G., and Vaupel, J. W. (2001). Lifespan depends on month of birth. *Proceedings of the National Academy of Sciences* 98(5), 2934-2939.

Dolinoy, D.C., Das, R., Weidman, J.R., and Jirtle, R.L. (2007). Metastable epialleles, imprinting, and the fetal origins of adult diseases. *Pediatric Research* 61, 30R-37R.

Dragone, D. and Vanin, P. (2020). Substitution effects in intertemporal problems. Working Paper DSE 1147, University of Bologna.

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

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

Goldsmith, T.C. (2014). *The Evolution of Aging* (3rd ed.) Azinet Press, Crownsville, MD.

Gluckman, P.D. and Hanson, M.A. (2004). Developmental origins of disease paradigm: a mechanistic and evolutionary perspective. *Pediatric Research*, 56(3), 311-317.

Grossman, M. (1972a). On the concept of health capital and the demand for health. *Journal of Political Economy* 80(2), 223-255.

Harttgen, K., Kowal, P., Strulik, H., Chatterji, S., and Vollmer, S. (2013). Patterns of frailty in older adults: comparing results from higher and lower income countries using the Survey of Health, Ageing and Retirement in Europe (SHARE) and the Study on Global AGEing and Adult Health (SAGE). *PLOS One* 8(10), e75847.

Hilakivi-Clarke, L., & De Assis, S. (2006). Fetal origins of breast cancer. Trends in Endocrinology & Metabolism, 17(9), 340-348.

HCCI (2012). 2007-2010 Children's Health Care Spending Report, Health Care Cost Institute (`http://www.healthcostinstitute.org/childrensreport`, accessed January, 7th, 2016.)

Heckman, J.J. (2007). The economics, technology, and neuroscience of human capability formation. *Proceedings of the national Academy of Sciences* 104(33), 13250-13255.

Karkach, A.S. (2006). Trajectories and models of individual growth. *Demographic Research* 15, 347-400.

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.

Kesternich, I., Siflinger, B., Smith, J.P., and Winter, J.K. (2015). Individual behaviour as a pathway between early-life shocks and adult health: Evidence from hunger episodes in post-war Germany. *Economic Journal* 588(125), F372-F393.

Kirkwood, T.B.L. (2005). Understanding the odd science of aging. *Cell* 120(4), 437-447.

Kirkwood, T.B.L., and Mathers, J.C. (2009). The basic biology of aging. In: Stanner S., Thompson R., Buttriss J. (eds.) *Healthy Aging – The Role of Nutrition and Lifestyle*, NY: Wiley-Blackwell.

Kiserud, T., Piaggio, G., Carroli, G., Widmer, M., Carvalho, J., Jensen, L. N., ... and Benachi, A. (2017). The World Health Organization fetal growth charts: a multinational longitudinal

study of ultrasound biometric measurements and estimated fetal weight. *PLoS medicine* 14(1), e1002220.

Kleiber, M., 1932, Body size and metabolism, *Hilgardia* 6, 315-353.

Kooijman, S.A.L.M. (2000). *Dynamic energy and mass budgets in biological systems*. Cambridge University Press.

Lin, M.J. and Liu, E.M. (2014). Does in utero exposure to illness matter? The 1918 influenza epidemic in Taiwan as a natural experiment. *Journal of Health Economics* 37: 152-163.

McMullen, S., and Mostyn, A. (2009). Animal models for the study of the developmental origins of health and disease. Proceedings of the Nutrition Society, 68(03), 306-320.

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.B, Song, X., and Rockwood, K. (2013). Assessing biological aging: the origin of deficit accumulation. *Biogerontology* 14(6), 709-717.

Mitnitski, A.B, and K. Rockwood. (2016). The rate of aging: the rate of deficit accumulation does not change over the adult life span. *Biogerontology*, 17(1), 199–204.

Murphy, K.M., and Topel, R.H. (2006). The value of health and longevity. *Journal of Political Economy* 114(5), 871-904.

Nelson, R.E. (2010). Testing the fetal origins hypothesis in a developing country: evidence from the 1918 influenza pandemic. *Health Economics*, 19(10), 1181-1192.

Oreopoulos, P., Stabile, M., Walld, R. and Roos, L. (2008). "Short-, Medium-, and Long-Term Consequences of Poor Infant Health: An Analysis Using Siblings and Twins." *Journal of Human Resources*, 43(1): 88–138.

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, 722-727.

Rockwood, K., Mitnitski, A., and Howlett, S. E. (2015). Frailty: Scaling from Cellular Deficit Accumulation?, in: Theou, O. and Rockwood, K. (eds.) *Frailty in Aging.*, Basel: Karger, 1-14.

Rutenberg, A.D., Mitnitski, A.B., Farrell, S.G., and Rockwood, K. (2018). Unifying aging and frailty through complex dynamical networks. *Experimental Gerontology* 107, 126-129.

Scholte, R.S., Van Den Berg, G.J., and Lindeboom, M. (2015). Long-run effects of gestation during the Dutch Hunger Winter famine on labor market and hospitalization outcomes. *Journal of Health Economics* 39, 17-30.

Searle, S.D., Mitnitski, A.B., Gahbauer, E.A., Gill, T.M., and Rockwood, K. (2008). A standard procedure for creating a frailty index. *BMC Geriatrics* 8(1), 24.

Shi, J., Song, X., Yu, P., Tang, Z., Mitnitski, A., Fang, X., and Rockwood, K. Analysis of frailty and survival from late middle age in the Beijing Longitudinal Study of Aging, *BMC Geriatrics* 11(1), 17.

Sinclair, K.D., Lea, R.G., Rees, W.D., and Young, L.E. (2007). The developmental origins of health and disease: current theories and epigenetic mechanisms. Society of Reproduction and Fertility supplement, 64, 425.

Schüenemann, J., Strulik, H., and Trimborn, T. (2017a).Going from bad to worse: Adaptation to poor health health spending, longevity, and the value of life. *Journal of Economic Behavior & Organization* 140, 130-146.

Schüenemann, J., Strulik, H., and Trimborn, T. (2017b). The gender gap in mortality: How much is explained by behavior? *Journal of Health Economics* 54, 79-90.

Schüenemann, J., Strulik, H., and Trimborn, T. (2020). The marriage gap: Optimal aging and death in partnerships. *Review of Economic Dynamics* 36, 158-176.

Strulik, H. (2015). Frailty, mortality, and the demand for medical care, *Journal of the Economics of Ageing* 6, 5-12 .

Strulik, H. (2018). The return to education in terms of wealth and health, *Journal of the Economics of Ageing* 12, 1-14.

Thompson, R.F., and Einstein, F.H. (2010). Epigenetic basis for fetal origins of age-related disease. *Journal of Women's Health* 19(3), 581-587.

Van Den Berg, G.J., Lindeboom, M. and Portrait, F. (2006). Economic conditions early in life and individual mortality. *American Economic Review* 96(1), 290-302.

von Bertalanffy, L. (1957). Quantitative laws in metabolism and growth. *Quarterly Review of Biology* 32(3), 217-231.

Waterland, R.A., and Michels, K.B. (2007). Epigenetic epidemiology of the developmental origins hypothesis. Annual Review of Nutrition 27, 363-388.

West G.B., J. H. Brown and B.J. Enquist, 1997, A general model of the origin of allometric scaling laws in biology, *Science* 276, 122-26.

West, G.B., Brown, J.H., and Enquist, B. J. (2001). A general model for ontogenetic growth. *Nature* 413(6856), 628-31.

Wu, G., Bazer, F.W., Cudd, T.A., Meininger, C.J., and Spencer, T.E. (2004). Maternal nutrition and fetal development. *Journal of Nutrition* 134(9), 2169-2172.

Appendix A: Solution of the Health Deficit Model

The maximization problem given in equations (3)–(5) together with the initial conditions $k(\tau) = 0$, $D(\tau) = D_\tau$ and the terminal conditions $k(T) = \bar{k}$ and $D(T) = \bar{D}$ constitutes a free terminal time problem of optimal control. The unknown terminal time is the age at death. The Hamiltonian associated with this problem reads

$$H = \frac{c^{1-\sigma} - 1}{1-\sigma} + \lambda\mu \left(D - Ah^\gamma + \epsilon\right) + \phi \left(rk + w - c - ph\right).$$

For $\sigma = 1$ the first term is replaced by $\log(c)$. The first order conditions wrt. c and h and the co-state equations are

$$c^{-\sigma} - \phi = 0 \qquad \Rightarrow \qquad c^{-\sigma} = \phi \qquad \Rightarrow \qquad \sigma\dot{c}/c = -\dot{\phi}/\phi \tag{A.1}$$

$$-\lambda\mu A\gamma h^{\gamma-1} - p\phi = 0 \tag{A.2}$$

$$\lambda\mu = \lambda\rho - \dot{\lambda} \qquad \Rightarrow \qquad \mu - \rho = -\dot{\lambda}/\lambda \tag{A.3}$$

$$\phi r = \phi\rho - \dot{\phi} \qquad \Rightarrow \qquad r - \rho = -\dot{\phi}/\phi. \tag{A.4}$$

Equation (A.4) is the well known Euler equation requiring that the shadow price of consumption (ϕ) grows at the rate of the interest rate less the time preference rate. Analogously, the Euler equation (A.3) requires that the shadow price of health grows at the rate of health deterioration (μ) less the time preference rate. At time of death, the end conditions $D(T) = \bar{D}$ and $k(T) = \bar{k}$ apply. Additionally, the condition for optimal T requires that the Hamiltonian assumes the value of zero at T. Otherwise, it would have been optimal to live longer or die earlier. Thus the Hamiltonian at age T fulfils

$$0 = H(T) = u(c(T)) + \lambda(T)\mu \left[\bar{D} - Ah(T)^\gamma + \epsilon\right] + \phi(T) \left[r\bar{k} + w - c(T) - ph(T)\right]. \tag{A.5}$$

To solve the model, we log-differentiate (A.2) wrt. time and insert (A.3) and (A.4) to obtain optimal growth of health expenditure:

$$\frac{\dot{\lambda}}{\lambda} - \frac{\dot{\phi}}{\phi} = (1-\gamma)\frac{\dot{h}}{h} \qquad \Rightarrow \qquad -\mu + \rho + r - \rho = (1-\gamma)\frac{\dot{h}}{h}.$$

Solving for the growth rate of health expenditure we obtain the "Health Euler equation"

$$g_h \equiv \frac{\dot{h}}{h} = \frac{r-\mu}{1-\gamma}. \tag{A.6}$$

By inserting (A.3) into (A.1) we obtain the Ramsey rule

$$g_c \equiv \frac{\dot{c}}{c} = \frac{r-\rho}{\sigma}. \tag{A.7}$$

The dynamic model can be solved without numerical integration since the differential equations can be solved analytically. We set (wolog) $\tau = 0$ and begin with noting that, because g_h is optimally constant according to (A.6), the differential equation (3) can be rewritten as $\dot{D} = \mu(D - Ah(0)^\gamma \exp(\gamma g_h t) + \epsilon)$. Given $D(0) = D_0$ the solution at time T is:

$$D(T) = D_0 \exp(\mu T) - \mu A h(0)^\gamma \exp(\mu T) \int_0^T \exp(\gamma g_h - \mu)\mathrm{d}t - \mu\epsilon \exp(\mu T) \int_0^T \exp(-\mu t)\mathrm{d}t.$$

At the time of expiry the boundary condition requires $D(T) = \bar{D}$. Solving the integrals in the above equation we get

$$\bar{D} = D_0 \exp(\mu T) - \frac{\mu A h(0)^{\gamma} \exp(\mu T)}{g_D} \left[\exp(g_D T) - 1\right] + \epsilon \left[\exp(\mu T) - 1\right], \tag{A.8}$$

where $g_D \equiv (\gamma r - \mu)/(1-\gamma)$. Next, we integrate (3) and insert $k(0) = k_0$ and $k(T) = \bar{k}$ to obtain

$$\begin{aligned}\bar{k} = {} & k_0 \exp(rT) + w \exp(rT) \int_0^T \exp(-rt) \mathrm{d}t \\ & - c(0) \exp(rT) \int_0^T \exp[(g_c - r)t] \mathrm{d}t - ph(0) \exp(rT) \int_0^T \exp[(g_h - r)t] \mathrm{d}t.\end{aligned}$$

Divide by $\exp(rT)$. Note that $g_h - r = (\gamma r - \mu)/(1-\gamma) \equiv g_D$ and solve the integrals to obtain

$$\bar{k} \exp(-rT) = k_0 - \frac{w}{r} \left[\exp((r)T) - 1\right] - \frac{c(0)}{g_c - r} \left[\exp((g_c - r)T) - 1\right] - \frac{ph(0)}{g_D} \left[\exp(g_D T) - 1\right]. \tag{A.9}$$

Finally, insert $\lambda(T)$ and $\phi(T)$ from (A.1) and (A.2) into (A.5) to obtain

$$0 = u(c(T)) - \frac{p}{c(T)^{\sigma}} \left[\frac{(\bar{D} + \epsilon) h(T)^{1-\gamma}}{\gamma A} - \frac{h(T)}{\gamma} - \frac{w + r\bar{k}}{p} + c(T) + h(T)\right]$$

where $u_T \equiv \log c(T)$ in the case of log-utility and $u_T \equiv [c(T) - 1]^{1-\sigma}/(1-\sigma)$ otherwise. Noting that $c(T) = c(0) \exp(g_c T)$ and $h(T) = h(0) \exp(g_h T)$ this provides

$$\begin{aligned}0 = {} & u_T - \frac{\exp(-\sigma g_c T)}{c(0)^{\sigma}} \times \\ & \left\{\frac{(\bar{D} + \epsilon)}{\gamma A} p h(0)^{1-\gamma} \exp((1-\gamma) g_h T) - \frac{1-\gamma}{\gamma} p h(0) \exp(g_h T) - w - r\bar{k} + c(0) \exp(g_c T)\right\},\end{aligned} \tag{A.10}$$

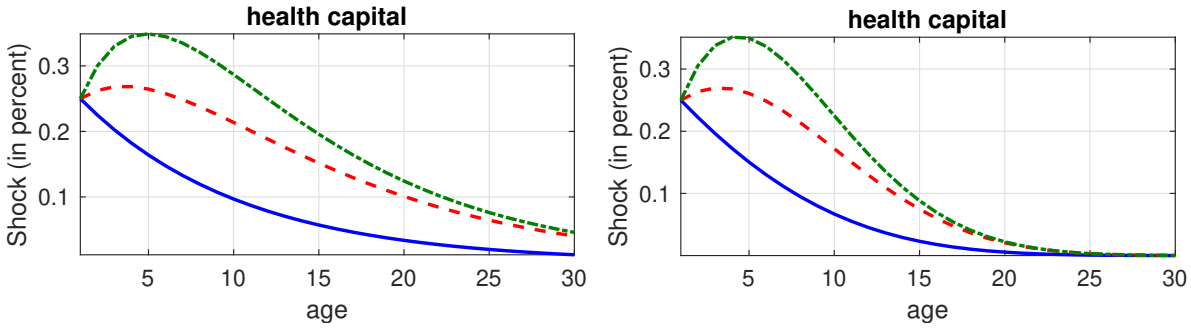
with $u_T \equiv \log(c(0)) + g_c T$ in the case of log-utility and $u_T \equiv [c(0) \exp(g_c T) - 1]^{1-\sigma}/(1-\sigma)$ otherwise. The three equations (A.8) – (A.10) can be solved for the three unknowns: $c(0)$, $h(0)$, and T. Having found the optimal initial values and the optimal terminal time T, the dynamic system consisting of (3), (5), and (A.5) – (A.7) is fully specified and it can be solved for the optimal life-cycle trajectories of c, h, k and D.

Appendix B: Health Shocks Increase Health Capital Depreciation

Here, we consider an interesting refinement of the health capital model and show that it does not resolve the problem of convergence of initial health differences in old age. Specifically, we implement the assumption that the rate of depreciation of the stock of health capital is a positive function of the initial health shock. To see the implications by way of example, let us reconsider the 25% health shock from the computational experiment in the main text (and from Almond and Currie, 2011). The blue line in the panel on the left-hand side in Figure A.1. reiterates the result from Figure 1 in the main text. It shows the difference of health capital between an unshocked individual and an individual who experienced 25% health reduction at time 0. The initial health difference is depreciated away as the individuals grow older. Suppose now that health depreciates faster for individuals who experienced a negative initial health shock. Red

(dashed) lines show the case where the health shock implies an increase of the depreciation rate to 15%. Green (dash-dotted) lines repeat the experiment when the shock increases the depreciation rate to 20%. We see that this leads to an increase in the health difference between individuals in young adulthood, implying divergence and, temporarily, shock amplification. As individuals grow older, however, convergence sets in and health differences are depreciated away. This variant of the model thus predicts that early-life health shocks lead to large health impediments at young age but only to insignificant health impediments in old age. It fails to motivate fetal origins of late-life health deficits.

FIGURE A.1: HEALTH CAPITAL DEPRECIATION INCREASES WITH EXPERIENCE OF IN-UTERO SHOCK



The figure shows how a 25 percent negative shock at time 0 develops given alternative annual depreciation rates. Left: constant rate of depreciation; blue (solid) line: 10 percent depreciation; red (dashed) line: initial shock associated with increase of depreciation rate to 15%; green (dash-dotted) line: initial shock associated with increase of depreciation rate to 20%. Right: Same exercise with the additional assumption that health depreciation increases at a rate of 5 percent per year as individuals age.

The result can also be derived more formally, following the argument made in the main text (subsequent to equation (1)). The difference in health capital between two individuals (1 and 2) is given by

$$H_t^1 - H_t^2 = (1-\delta^1)^t H_0^1 - (1-\delta^2)^t H_0^2.$$

Let the individual who experienced the shock be indexed by 2, $H_0^2 < H_0^1$ and assume that the shock is associated with a higher rate of health capital depreciation, $\delta^1 < \delta^2$. Then, the difference $H_t^1 - H_t^2$ becomes larger for small positive t (compared to the case of shock-independent δ). For large t, however, the difference $H_t^1 - H_t^2$ converges to zero, irrespective of the difference in the δ's. The reason is that the sufficient condition for convergence to zero is $0 < \delta < 1$, which is always fulfilled in the Grossman model, irrespective of any additional assumptions on the size of δ. It is an inherent feature of the Grossman model. This can be illustrated further by assuming that δ is increasing with age, an assumption frequently made in applications of the Grossman model.

In the panel on the left hand side of Figure A.1, we assume that the depreciation rate increases exponentially in age, $\delta^i(t) = \delta^i(0) \cdot e^{\nu t}$, $\nu > 1$ and $i = 1, 2$. We set $\nu = 0.05$ and assume that $\delta^1(0) = 0.1$ for the unshocked individual while for the individual who experienced the in-utero shock $\delta^2(0) = 0.15$ (red dashed lines) and $\delta^2(0) = 0.2$ (green dash-dotted lines). We see that, as explained in the main text, the positive age-dependence leads to faster convergence of initial health differences.

Appendix C: Proof that Shocks are Amplified iff $t < t_I$

In order to see the condition for shock amplification more clearly, consider a shock that changes body size at time τ and compute how it translates into a change in body size at age t, $t > \tau$:

$$\frac{\partial m(t)}{\partial m(\tau)} = \left\{a/b - \left[a/b - m(\tau)^{1-\beta}\right] \mathrm{e}^{-(1-\beta)b(t-\tau)}\right\}^{\frac{\beta}{1-\beta}} \mathrm{e}^{-(1-\beta)b(t-\tau)} m(\tau)^{-\beta}.$$

Next, compute how this effect changes with advancing age of the individual:

$$\begin{aligned}\frac{\partial^2 m(t)}{\partial m(\tau)\partial t} = \frac{\beta}{1-\beta} x^{\frac{\beta}{1-\beta}-1} \left[a/b - m(\tau)^{1-\beta}\right] (1-\beta) b \mathrm{e}^{-(1-\beta)b(t-\tau)} \mathrm{e}^{-(1-\beta)b(t-\tau)} m(\tau)^{-\beta} \\ - x^{\frac{\beta}{1-\beta}} (1-\beta) b \mathrm{e}^{-(1-\beta)b(t-\tau)} m(\tau)^{-\beta}, \qquad \text{(A.11)}\end{aligned}$$

with $x \equiv a/b - \left[a/b - m(\tau)^{1-\beta}\right] \mathrm{e}^{-(1-\beta)b(t-\tau)}$. In order to find the threshold for shock amplification, we solve (A.11) for $\frac{\partial^2 m(t)}{\partial m(\tau)\partial t} = 0$, i.e. we solve

$$\begin{aligned}&\left[\frac{a}{b} - m(\tau)^{1-\beta}\right] \frac{\beta}{1-\beta} \mathrm{e}^{-(1-\beta)b(t-\tau)} - a/b + \left[a/b - m(\tau)^{1-\beta}\right] \mathrm{e}^{-(1-\beta)b(t-\tau)} = 0 \\ \Rightarrow \quad &\left[\frac{a}{b} - m(\tau)^{1-\beta}\right] \left(\frac{\beta}{1+\beta} + 1\right) \mathrm{e}^{-(1-\beta)b(t-\tau)} = \frac{a}{b} \\ \Rightarrow \quad &\left[\frac{\frac{a}{b} - m(\tau)^{1-\beta}}{a/b}\right] \frac{1}{1+\beta} = \mathrm{e}^{(1-\beta)b(t-\tau)}.\end{aligned}$$

Taking logs and solving for t provides t_I from (12) in the main text.