A Unified Law of Mortality: Implications for Economic Analysis

Adriana Lleras-Muney and Flavien Moreau*

March 31, 2017

PRELIMINARY AND INCOMPLETE–PLEASE DO NOT CITE.

Abstract

How do social and economic conditions affect the evolution of health and mortality rates over the lifetime? To answer this question, we build a simple model that combines an age-dependent process with a random health hazard. A key insight of the model is that population mortality rates place constraints on the evolution of the underlying distribution of (unobserved) health, and can be used to infer how health has evolved over time and across countries. Using cohort life tables provided in the Human Mortality Database, we estimate our model and trace out the evolution of structural parameters since 1850. We use the model to understand how unexpected shocks, like wars, affect the age-profile of health and mortality; and investigate implications for SES gradients and optimal health care expenditures.

JEL: I10, J11

Keywords: Mortality, health, evolution, life course.

*Corresponding author: alleras@econ.ucla.edu. We are very grateful to Andy Atkeson, David Cutler, Jeff Ely, and Bo Honoré for their advice and to seminar participants at University of Connecticut, USC, and UCLA. This work used computational and storage services associated with the Hoffman2 Shared Cluster provided by UCLA Institute for Digital Research and Education's Research Technology Group. All errors are our own.

Many key questions in economics (and in other areas) require knowledge of how health and mortality evolve with age. A large literature in medicine, demography, sociology and economics documents for example that circumstances early in life affect health and mortality throughout the lifetime (for recent summaries see Almond and Currie, 2011 and Almond et al., 2017). Understanding of the full consequences of early circusmtance, as well as of programs aimed to address them, requires a full accounting of their long term effects. When and how to distribute resources over the lifetime depends crucially on both short and long term impacts. Yet there is no known parametric model of the evolution of health over time that can be used to trace how early insults affect the evolution of health and mortality. The key intuition in this paper is that, if mortality depends on health, then the age-profile of mortality imposes strong restrictions for models of the evolution of the health stock over the lifetime. And cohort mortality rates exhibit a remarkably consistent pattern, high in infancy and old age, and low but variable during reproductive ages. The stability and consistency of this shape across humans and primates, suggests there exists an underlying "law of mortality" and health (Gompertz, 1871, Carnes et al. 1996, Bronikowski et al., 2011). Yet a simple model of mortality for all ages has proved elusive.[1]

In this paper we provide a unified law of mortality that tracks health and mortality from birth to death. It is a simple dynamic model of the evolution of the health stock that accounts for the basic features of mortality and can be characterized by only five parameters. In the spirit of the classic demographic work by Vaupel et al. (1979), populations are born with an initial distribution of health (or frailty), and individuals with low levels of health die. This distribution of health is dynamic over the lifetime: the environment can increase (or decrease it) it by providing (health) resources, though these are not equally distributed in the population. But health also deteriorates at a deterministic and increasing rate with age. Finally individuals can also die from external causes, unrelated to "biological" processes and health status–this last force is not necessary to explain the basic age-profile of mortality. But these external shocks play an important role in explaining observed deaths during the reproductive period, while biological processes are most visible in childhood and old age.

This basic framework can be used to investigate how temporary shocks at a given age, or changes in permanent circumstances, affect survival at subsequent ages. There are no easily computable analytical solutions for the parameters of the model, or for the mortality rate at a point in time. But we can characterize the behavior of mortality at all ages as a function of the parameters, and we can estimate these parameters using the method of simulated moments.

We use high-quality data from cohort life tables provided in the Human Mor-

[1]Most notably Gompertz (1820), Gompertz (1825), Gompertz (1862), Gompertz (1871) noted log mortality is linear after 45. Models that succesfully predict mortality from birth to death typically model the hazard rates (or some function of the rates, like survival or probabilities of dying in a given interval) using complex mathematical models. Carriere (1992)for instance shows a mixture of gompertz, Weibull, inverse Gompertz and Inverse Weibull can fit the data nicely. We provide an more in depth discussion of how our model compares to others in the literature later in the paper.

tality Database (HMD) to estimate these parameters. The estimation recovers five (or more) parameters from each cohort table and results in relatively accurate prediction of life expectancy. We independently estimate the parameters of all cohorts born between 1860 and 1945 in France, and report their evolution. The results suggest that the main sources of mortality gains for these cohorts were driven by increases in the level of health at birth, increases in the level of resources and decreases in the variance of these resources. We also estimate parameters for primates using data from Kohler et al. (2006). Relative to other primates, humans have substantially highe initial health, but very similar aging rates.

We use our model to investigate the effects of temporary and permanent shocks on lifetime mortality. We estimate the effects of WWII and find it had long lasting effects on mortality, consistent with lower investments in war times possibly due to lower GDP, lower availability of food, and a worsening of sanitary conditions during the 1939-1945 years. We are also able to rationalize mortality during reproductive ages, which can best be characterized as an increase in the accident rate in the population. We also use of model to replicate and intepret findings in the literature, such as the recent those by Case and Deaton (2015).

Finally we derive optimal investment profiles by age. If we assume that resources are independent of health, we find that health investments are u-shaped: they are highest at birth, fall with age and rise again with age. We consider extensions of this optimal investment model, including making per period resources depend on health. Although optimal investments affect the shape of mortality, they do not fundamentally change it–mortality remains highest at young and old ages. The model also implies that lifetime investment and initial stocks are complementary, and there are also strong "dynamic" complementarities between investments at different ages, as in Cunha and Heckman (2007). We discuss implications for compensation of populations that experience negative shocks.

This paper is organized as follows. We first describe the data and the basic observations that motivate our model. We then describe the model and its properties in its simplest form. To assess the fit of the model we estimate its parameters and study their evolution. We also investigate how shocks affect the evolution of health and mortality. Then we investigate implications for optimal investments. We finish by considering some applications of the model, and how it can be used to understand findings in the literature.

1 Basic mortality patterns from the Human Mortality Database

To study mortality over the lifetime we make use of the Human Mortality Database (HMD) which provides mortality rates by age, birth year, country and gender. These are constructed using birth counts based from birth certificates, death counts from death certificates, as well as population estimates

derived from censuses. Age-specific mortality rates become available the year a country enters the database, which occurs when the census and vital registration system is deemed to be almost 100% complete. These data constitute the highest quality and largest data available to study cohort mortality. The most significant limitation of the data is that it does not account for migration patterns, but no data are available to correct mortality rates appropriately.

In this paper we use the life tables of six countries (Belgium, Denmark, Iceland, France, Netherlands, Norway, and Sweden) with complete mortality data for a large number of cohorts. We look at men and women born between 1860 and 1940. For all these cohorts we observe mortality rates from birth to age 90, except for post-1920 cohorts who can only be followed to age 80. This results in 83,934 observations–there are a few missing data points, particularly for war years.[2]

We focus specific on France for convenience. It has the longest time series of cohorts that can be followed from birth to age 90, except for Sweden. The data pertain to very large populations, the 1860 birth cohort had almost 900,000 individuals in France, but less than 200,000 in the other counties. Finally France had 2 very deadly wars and high maternal mortality which we also study.

In France, life expectancy at birth increased from 30 to 45 from 1800 to 1900, and reached 80 in 2004 (Pison, 2005). Figure 1 plots the log of the mortality rate by age for all cohorts of men and women born in France between 1860 and 1940. It shows that (the log of) mortality almost like a tick mark: it starts very high early in life, plummets to very low but variable levels in adolescence and young adulthood, and then rises with age starting in middle age. If we examine the data by decade as in Appendix Figure 16, we see that this pattern is the norm across all cohorts. Appendix Figures 16 show that although there is some variation across countries, the shape of mortality is also very similar across countries for a given cohort.

Several other features of the shape of mortality are noteworthy. First after middle age, log mortality rises almost linearly with age–this regularity was first noted by Gompertz in 1820 and has since led to the "search for a unified law of mortality". Moreover the slope after age 45 appears to have remained relatively unchanged. Mortality decreases across cohorts result in downward shifts of the curves–in old age these declines appear to shift the lines in almost parallel fashion. What is remarkable about this feature is that one might have hypothesized when infant mortality fell (infant mortality in France fell from roughly 17 percent in 1860 to about 8 percent in 1940), it left frailer individuals alive, and could have resulted in higher mortality in older ages. But this is not what we observe, the curves do not cross–the mortality rate in old age is lower for cohorts with lower infant mortality rates, as has been noted by Finch and Crimmins (2004).

Second, the greatest deviations (in logs or proportional terms) from the tick-

[2]There are a maximum of 14,236 observations per country. France, Belgium and Denmark there are a few missing data points. In the case of France many correspond to war years. There is no almost data for 1914, 1939, 1943, 1945, 1946. There area afew other missing points.

mark shape occurs during reproductive ages. For both men and women, there are visible "spikes" corresponding to war years, as can be seen for cohorts born around 1920 who experienced WWII during ages 19-25. In the absence of wars, for example for the 1860 cohorts, there is a visible rise in mortality after age 15, which demographers refer to as a "hump" (Preston et al. 2000). Two causes of death have been documented to account for a large fraction of deaths for ages 15-45 in non-war times: maternal mortality and "external" causes, which include accidents and violent deaths (including suicide and murder).

Lastly for the most recent 1940 cohort, there is almost no humps and there are no visible spikes. This cohort has not experienced a war. In most Western nations there has been substantial declines in violents deaths in the last century (Pinker, 2011), consistent with a decrease in the bump for men. Also maternal mortality disappeared after 1930, also possibly explaining the disappearance of the bump for women for the most recent cohorts. In France maternal mortality between 1850 and 1890 is estimated to be around 5 per 1,000 births (Bardet et al. (1981)) and to have remained at that level until the 1920s, it fell rapidly after the mid 1930s (Loudon 1992). And indeed for women born in 1940 (having children between 1960 and 1980) there is no apparent "excess mortality", consistent with the almost complete elimination of maternal mortality.[3] The tick-shape is most clearly visible for the most recent cohort. This observation motivates our basic model which seeks to describe "natural" mortality in the absence of "external" causes that depend on choices such as whether to have children or events like war.

2 A parsimonious model of health and death

In this section we provide a characterization of health and mortality based on frailty, in the spirit of Vaupel et al. (1979) (and similar to the idea of vitality in recent work by Li and Anderson, 2013). But here the distribution of frailty is dynamic over the lifetime, similar to models of in-utero shocks (Bozzoli et al., 2009 and Bruckner and Catalano, 2007). Health is treated like a stock, affected by investments and subject to depreciation, as in models of human capital (Grossman 1972 and Cunha and Heckman 2007). This basic model can predict the evolution of mortality in the absence of external causes–we examine the role of external causes later.

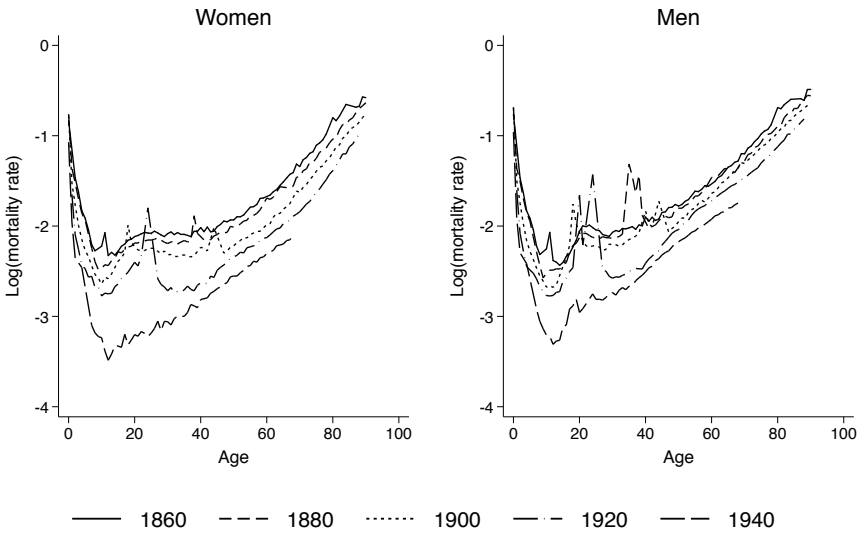
2.1 A model of "natural mortality"

Assume individuals are born with an initial health level H_0. This initial health endowment differs across individuals in the population and follows a normal distribution $N(\mu_H, \sigma_H^2)$.[4] Every period the environment provides average resources denoted by I that can increase or decrease H. Individuals in the environment

[3] in 1970 maternal mortality was around 28 per 100,000 (Bouvier-Colle et al. 2008)

[4] Birth weights and other traits measured at birth follow a normal distribution Wilcox and T RUSSELL (1983).

Figure 1: Mortality rates in France, for cohorts born 1860-1940



Data: Human Mortality Database

are more or less lucky and experience an idiosincratic shock ε_t to their resources. For example in a stationary environment I characterizes the mount of food that a given country produces, but a given person might receive less if for instance rain was unusually low in their location. The variance of ε_t captures how unequal the distribution of resources within the population is. These idiosyncractic shocks are assumed to be i.i.d. and the are drawn from a normal distribution $N(0, \sigma_\varepsilon^2)$ every period. Finally the health stock is subject to depreciation every period–in other words there is a negative shock every period akin to a user cost, reflecting cumulative death cell and organ damage. This depreciation increases more than linearly with age, at rate δt^α. Together these forces determine the evolution of the health stock.

People die when their stock of health first crosses a threshold $\underline{H}$, which is fixed throughout the lifetime and identical for all individuals. Formally let $D_t = \mathbb{I}(H_t \leq \underline{H})$ denote the random variable equal to one if the individual dies in period t, and define the mortality rate at time t as $MR_t = E(D_t) = P(D_t = 1 | D_{t-s} = 0 \, \forall s < t)$. Therefore we have that the population's health and mortality can be characterized by the following dynamic system

$$\begin{aligned} H_0 &\sim N(\mu_0, \sigma_0^2) \\ H_t &= H_{t-1} + I - \delta t^\alpha + \varepsilon_t \\ \varepsilon_t &\sim N(0, \sigma_\varepsilon^2) \\ MR_t &= P(H_t < \underline{H} | H_{t-s} > \underline{H}, \forall s < t-1) \end{aligned} \tag{1}$$

with $\delta \in (0,\infty), \alpha \in (0,\infty), I \in \mathbb{R}$.[5]

In this model health is a latent **unobserved** construct that determines observed mortality. Figure 17 illustrate the dynamic relationship between population health and mortality rates implied by this model for the first two periods (and Appendix X gives the mathematical expression for the corresponding mortality rates). The initial distribution is normal. In the first period it moves right (if$I$ is positive and larger than the aging term) and gets wider (because of ε_t). Then the individuals to the left of the threshold, die (these individuals were either born frial or had large negative shocks). The mortality rate (the fraction of individuals that die in the first period) is given by the area under the curve below the threshold. Individuals with low initial health or large shocks die. In the second period this truncated distribution moves right again. And the population receives a new shock, generating mortality again. Notice that in the absence of a shock there would be no deaths in period 2 – or in any period thereafter until the depreciation term becomes large enough to push the distribution below the threshold. This illustrates that stochastic nature of the process is essential to generate mortality at every age, and it is one key feature that differentiates this model from previous ones like the Grossman (1972) model. An implication is that the distribution of health at any age (and therefore the mortality rate) is a function of the entire history of shocks and investments.

The second key feature of the model is the accelarating aging component, which eventually moves the distribution closer and closer to the threshold, guaranteeing the eventual death of the entire population. The continuous-time analogue of our model would be a brownian motion with a nonlinear drift and kill, where death occurs when the value of H first hits a threshold set at zero. These kinds of models are used to model companies' default probability and to price securities in finance (eg Lando 2004). This literature has established that these models do not admit a closed form solutions for the parameters, except for the particular case of constant, linear, drift. Here the drift is not constant: it increases at an increasing rate with age, similar to what was first proposed by Grossman, and consistent with biological models of senescence (Armitage-Doll, 1954 ot Pompei Wilson, 2002) In addition we are also tracking the distribution of such brownian motions, rather than individual ones.

This model has 7 parameters. But notice that the expression for mortality is just the standard expression for the Probit model, which requires a scale and location normalization: the threshold $\underline{H}$ and the standard deviation of the initial distribution σ_H^2 are not identified: at any age, we can substract $\underline{H}$ and divide by σ_H^2 on both sides of the expression determining the probability of dying and leave the mortality rate unchanged. So without loss of generality we set $\underline{H} = 0$ and $\sigma_H^2 = 1$.[6] Thus in its simplest form, this model characterizes the biological evolution of health and mortality of a cohort by age using 5

[5] We could impose some restrictions on these parameters. For example the share of individuals that survive to reproductive ages is never been observed to be much below fifty percent–this would appear to be a requirement of species that do not disappear.

[6] More precisely we need to normalize 2 out of three parameters. We find it more intuitive to normalize the threshold rather thanthe initial mean, but this choice is arbitrary.

(rescaled) parameters: one for initial conditions (μ_0), two govern the aging process (δ, α), and two characterize the effects of environment, in the form of average investments (I) and the variance of these investments or shocks (σ_ε^2). We then intrepret μ_H as the distance from the threshold of the initial distribution in standard deviations of the initial distribution. All other parameters are also expressed in "standard deviation" units, except for α which is "scale free"– it does not depend on the initial distribution.

Although we do not observe health, we can sometimes observe disease and disability rates which are also functions of health. The model has implications for the age profile of morbidity, where we define morbidity as having a level of health that is above the dying threshold but below some other arbitrary threshold. Conceptually this definition captures both temporary (acute) and permanent (chronic) conditions. The key difference between morbidity and mortality is that mortality is an absobing state whereas morbidity is not. If we denote H_d the morbidity threshold, then the morbidity rate at a give age is the number of individuals that are alive whose health is below that threshold

$$DR_t = Pr\left(0 < H_t \leq Hd\right)$$

2.2 The behavior of health and mortality over the lifetime

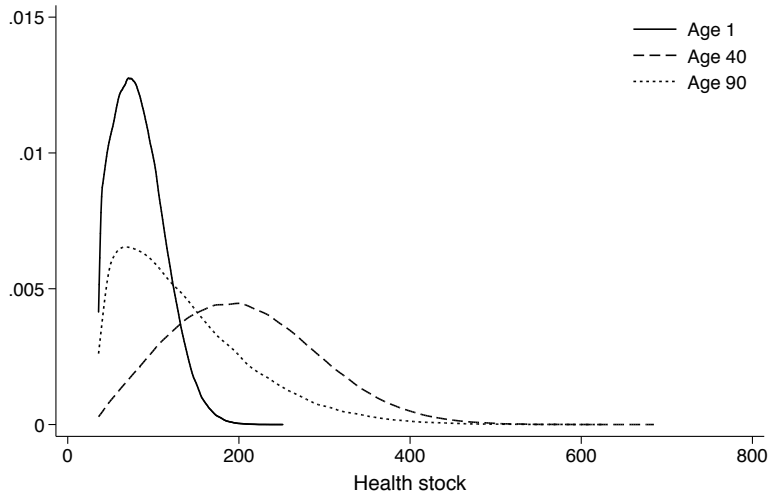
We now describe the behavior of this model and then go on to analyze the effect of changes in each of its underlying parameters. Let $\hat{H}_t \equiv \mathbb{E}\left[H_t \mid H_t > 0\right]$ denote the average health in the living population with age t and $\sigma_{H_t} \equiv Var\left[H_t \mid H_t > 0\right]$ the variance of health among the living.

Proposition 1 basic features of the model: see appendix for proofs

1. $\lim_{t\to\infty} Pr\left(H_t = 0\right) = 1$. Everyone dies with probability 1.
2. $\frac{\partial MR_t}{\partial t} \leq 0$ if $t < t_1$ and $\frac{\partial MR_t}{\partial t} \geq 0$ if $t \geq t_1$. For sufficiciently high I (relative to $\sigma_\varepsilon^2$and σ_H^2) mortality rates declines (up to age t_1) and then increases with age.
3. $\frac{\partial \hat{H}_t}{\partial t} \leq 0$ if $t < t_2$ and $\frac{\partial \hat{H}_t}{\partial t} \geq 0$ if $t \geq t_2$. The average health of the living increases and then decreases with age.
4. $\frac{\partial \sigma_{\hat{H}_t}}{\partial t} \leq 0$ if $t < t_3$ and $\frac{\partial \sigma_{\hat{H}_t}}{\partial t} \geq 0$ if $t \geq t_3$. The variance of health among the living increases and then falls.

Figure 2 illustrates (for a specific set of parameters, roughly matching the 1860 Belgium cohort mortality) the evolution of the distribution of the health stock as cohorts age in this model. This distribution at age 1 is truncated at the threshold, it moves right and broadens until age 40. Then it starts moving left and eventually becomes triangular at the threshold. At any given age after infancy and before old age, the distribution of health is very close to a normal distribution despite truncation, because it is approximately equal to a sum of

Figure 2: The evolution of the health distribution over lifetime



normal distributions. This is consistent with the observation that health related stocks like heights, which grown from birth until maturity are close to normally distributed (Limpert et al 2001).

Figure 3 shows that the model reproduces the age-profile of mortality well: (log) mortality starts high and plummets to very low levels by adolescence. It remains low and variable until around age 40, and then it starts rising linearly with age. The initially high infant mortality rate is mostly a result of many infants born with low health endowment, though there are also unlucky babies with large negative shocks. In childhood, mortality rates depend mostly on the the variance of the shock, and the size of the mean investment level which pulls the distribution away from the threshold. But eventually the depreciation process becomes larger than the investment and an increasing number of individuals fall below the threshold.

The figure also shows the evolution of health and morbidity. Over the lifetime, health and mortality are moving in opposite directions. Average population health increases and reaches a peak late in mid-life. This pattern is consistent with the evolution of self-reported health by age for recent US (Deaton and Paxson, 1998) and UK cohorts (Contoyannis et al. 2004). The variance of health increases and then falls, due to selective mortality. This behavior is similar to the behavior of consumption: cohort consumption inequality increases with age, so long as shocks to consumption are not perfectly correlated across individuals (Deaton and Paxson 1994, Deaton and Paxson, 1997). Finally morbidity is u-shaped, being high among children, reaching low levels from ages 20 to 60, and increasing thereafter. Contemporary data also show that hospitalization days

(a rough proxy for morbidity) are indeed u-shaped.[7]

There have been several attempts in the demographic literature to generate a unified model of mortality. Our model differs in several important dimensions. Many models, like Gompertz', only account for mortality after a certain age (Li and Anderson, 2013). Therefore these models do not lend themselves to a formal exploration of how early conditions affect mortality later in life. A very popular model proposed by Heligman and Pollard (1980)uses 8 parameters to describe the probability of dying at a given age. More recently Sharrow and Anderson (2016) propose a 6 parameter model and Palloni and BeltrÃ¡n-SÃ¡nchez (2016). However these models are difficult to interpret and it is not clear how to use them to investigate the long term cumulative effect of insults.

3 The determinants of mortality

The five parameters in our basic model affect the shape of the mortality curve in different ways. Proposition 2 summarizes the main qualitative insights.

3.1 Basic comparative statics for mortality

Proposition 2: basic comparative statics (see appendix 2 for proofs)

1. $\frac{\partial MR_t}{\partial I} \leq 0$, $\frac{\partial MR_t}{\partial \mu_0} \leq 0$. Increasing the investment Ior the average health at birth μ_0 unambigously decreases mortality at all ages.[8]

2. $\frac{\partial^2 MR_t}{\partial I \partial \mu_0} \leq 0$. Investment and health at birth are complements.

3. $\frac{\partial MR_t}{\partial \delta} \geq 0, \frac{\partial MR_t}{\partial \alpha} \geq 0$. Increasing any of the aging parameters, δ or α, unambigously increases mortality at all ages.

4. Changes in $\sigma_\varepsilon^2$and σ_H^2 can increase or decrease the mortality rate at a given age.

5. $\frac{\partial MR_t}{\partial \sigma_{Ht}^2} \geq 0$ if $\delta t^\alpha \leq I$. On impact, an increase in σ_H^2 before the middle age increases the mortality rate.

6. For some t_σ, $\frac{\partial MR_{t+s}}{\partial \sigma_{Ht}^2} < 0, \forall s > t_\sigma$. Ultimately, an increase in σ_H^2 generates selection and reduces mortality in the very old age. [9]

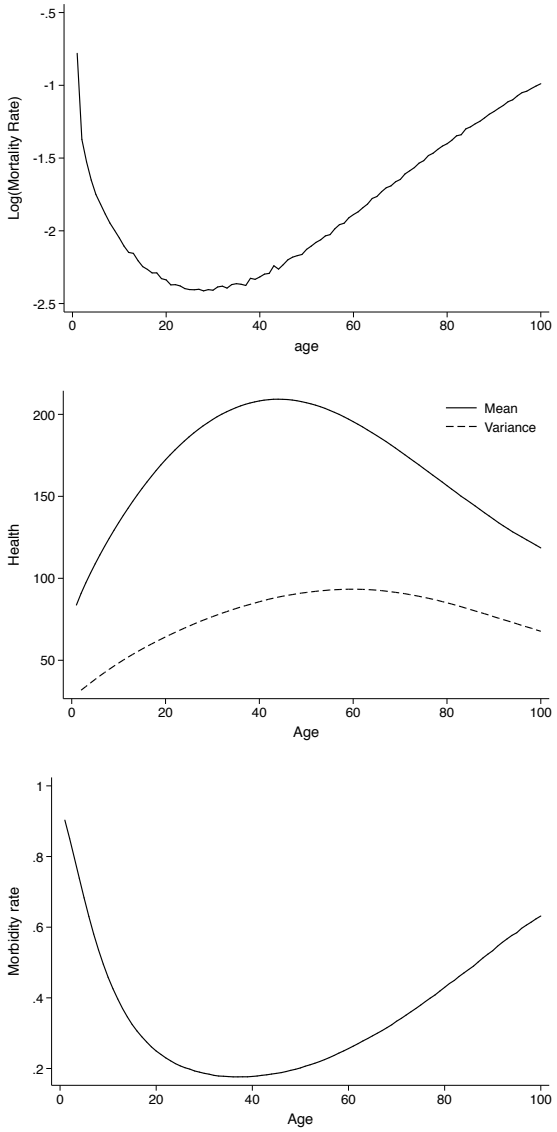
We illustrate these effects graphically for a specific set of parameters, chosen to roughly describe the Belgian 1860 cohort. Figure 19 shows how changing each parameter by 50% changes the age-specific mortality rates. The figure plots the difference in the mortality rate, relative to the baseline model. So each point corresponds to the change in mortality (new mortality rate minus

[7] hospitalization rates by age for the US are available here https://ftp.cdc.gov/pub/Health_Statistics/NCHS/NHIS/SHS/2014_SHS_Table_P-10.pdf

[8] Changing the threshold also affects mortality rates negatively throughout the lifetime.

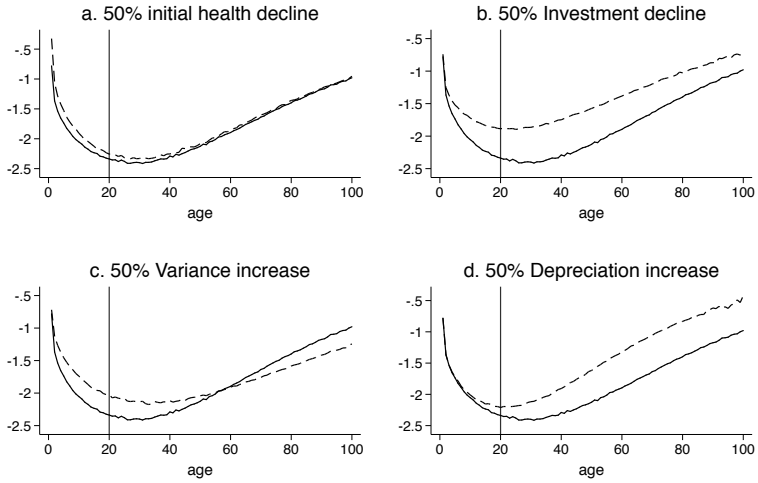
[9] Typically, in our simulation, the crossover in mortality rates occurs around age XX

Figure 3: Age profile of population health and mortality



Note: Simulated data for a population of 500,000 individuals. We assume H_0 is normally distributed with mean 68 and standard deviation of 34. The threshold for death $\underline{H}$ is 36 generating an infant mortality rate of about 17%. Shocks ε_t are drawn every period from a $N(0, 16)$, the rate of depreciation is δ=0.04, and the exponent on age is α =1.3. The level of investment I_t is constant at 4.5.

Figure 4: Comparative statics for log mortality



Note: Simulated data for two population of 500,000 individuals each. The figures show the effect of changes relative to the baseline model, which is simulated using the same parameters we used for Figure 3.

baseline mortality rate). A positive change corresponds to an increase resulting from the simulated change in the parameter. The effects are presented both in levels and in percentage terms, since these are two common ways in which mortality is analyzed.

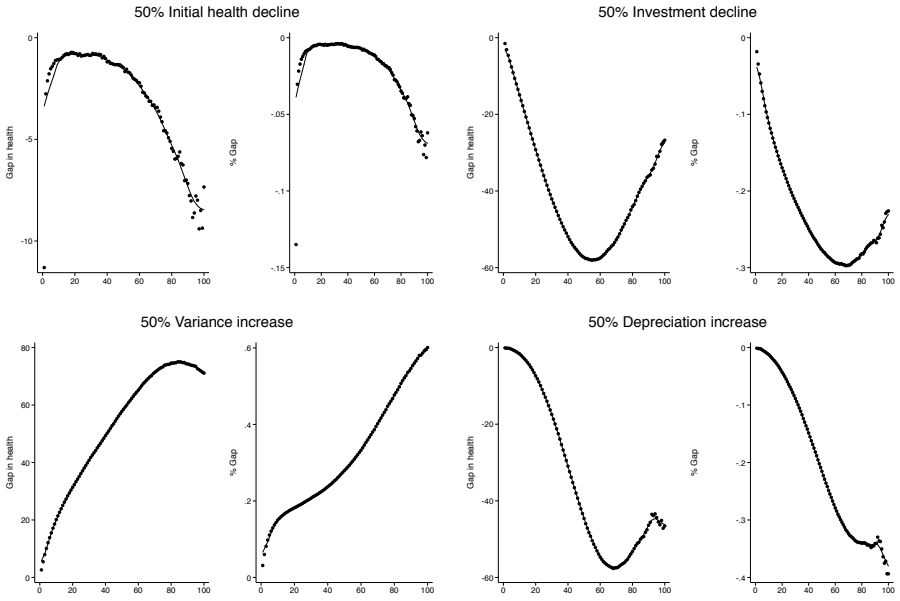
Lowering initial health results in markedly higher infant and adult mortality though the effects decline with age (in both absolute and percentage terms).[10] Lowering the average annual investment results in higher mortality at all ages (panel b). But the effects are not monotonic in age. In levels the effects are u-shaped: they are large at birth, decrease in middle age, and start increasing monotonically after middle age. But in percentage terms the effects are hump-shaped, increasing until middle ages and declining ithereafter.

Increasing the variance of the random shocks has ambiguous effects results in a "cross over". The population with high variance has higher mortality at younger ages but lower mortality at older ages (in both levels and percentage terms). This occurs because when the variance is higher many more die initially. But in the population with greater variance, many individuals are also the lucky recipients of greater positive shocks, and these individuals will live longer as a result. A variance shock can be conceptualized as a "infectious disease shock with immunity" – that is a shock that kills many but makes a few individuals hardier later in life.

Finally increasing the depreciation rate δ results in higher level of mortality

[10] Increasing the threshold throughout the lifetime has a similar effect.

Figure 5: Comparative statics for health



Note: Simulated data for a population of 500,000 individuals.

all throughout life, but the effects are imperceptible for many years, and then rise rapidly with age. In very old age the effects decline in percentage terms.[11]

Figures 5 illustrate the effects of changing the parameters on the average health of the living. A lower initial initial stock (panel a) lowers health at all ages, but the pattern is hump-shapped, with a large initial decline initially, almost no impact for many years thereafter and larger impacts as individuals age (in levels and percentage terms). Lower investments (panel b) also lower the average health at all ages. But the effects is u-shaped: it increases from birth to old age, and then starts declining once mortality starts rising. Increasing the variance of shocks (panel c) increases the health of the surviving population at all ages, and the effects get bigger with age. Lastly increasing the depreciation rate (panel d) lowers health at an increasing rate with age, except for old age.

In utero shocks. The results in the this section can be used to interpret a large literature that has investigated the effect of in utero shocks, such as famines, pandemics, wars, and stress (Almond and Currie, 2011 and Almond et al., 2017). Negative in utero shocks are equivalent to lower initial mean levels of health. The model predicts that these shocks will result in markedly higher infant and adult mortality and lower health throughout the lifetime. This is entirely consistent with the empirical evidence from this literature. What our model implies that has not been noted before is that the age when the effects

[11] Changing α has similar effects but we do not show them here. In our estimations we find that alpha has remained unchanged.

of the shocks are measured will have a significant impact of the size of the estimated effect. For instance effects of the shock on mortality will not be large in old age, but there will be large and growing effects on disability in old age. Because these effects are present throughout the lifetime, cross-sectional estimates will necessarily underestimate the full effect of these shocks, which are best summarized by how they affect (health adjusted) life expectancy.

Socio-economic status and mortality. A large literature documents large and persistent differences in health and mortality by education, income and other permanent markers of socio-economic status such as occupation and race (Cutler et al., 2012). If education and incomes are indicative of higher average resources throughout the lifetime (I), then the model predicts that mortality rates gaps (or "gradients") that are large at birth, fall to zero among young adults and then grow with age. In logs (or percentage terms) the gaps will be positive but fall after middle age. This results in log-mortality curves that are roughly parallel but slowly converge at very old ages after age 45. This is consistent with findings from Chetty et al. (2016) on the relationship between earnings at age 40 (a measure of permanent income) and mortality thereafter. This prediction is also consistent with the literature that has documented that the effects of education on mortality fall with age, in percentage terms (Hummer and Lariscy 2011).

The model also predicts that among the living the health gap between rich and poor will in fact rise with age and then fall, following a hump-shape. These predictions are consistent with evidence in Case et al. (2002) or Currie and Stabile (2003), who show that the gaps in self-reported health status between those born in poor and rich families grow with age, and decline after 65.

3.2 Adding accidents to the model

In our baseline model, mortality is purely driven by health and luck. However, many deaths, like accidents and homocides, are purely accidental and are likely to strike regardless of the health status of an individual. To account for these "extrinsic" causes of death in the simplest possible way, we extend our baseline model with an "accident shock".

Suppose that a random fraction $\kappa \in [0,1]$ of the population is killed by an accident in every period. This accident rate is assumed to be constant over the lifetime and independent of health. Each individual experiences i.i.d. shock ν_t drawn uniformly between 0 and 1 every period. Then model becomes:

$$\begin{gathered} H_0 \sim N(\mu_H, \sigma_H^2) \\ H_t = 1_{\{\nu_t > \kappa\}} \cdot [H_{t-1} + I - \delta t^\alpha + \varepsilon_t] \\ \varepsilon_t \sim N(0, \sigma_\varepsilon^2)\,,\ \nu_t \sim \mathcal{U}[0,1] \\ MR_t = P(H_t < \underline{H} | H_{t-s} > \underline{H}, \forall s < t-1) \end{gathered} \tag{2}$$

Accidents increase mortality rates at all ages, but more so during reproductive ages (Appendix Figure X). Around reproductive ages, when biological causes of death are dampened, accidents become the dominant source of death (in percentage terms). This is consistent with empirical evidence that external

causes of death (unintentional injury, suicide and homicide) account for a larger share of mortality during these ages today.[12] Around birth and in old age, when the average health stock is low, accidents account for a diminishing fraction of deaths, because many individuals die due to natural causes anyway (competing risks). The behavior of the model with the addition of accident-deaths is unaffected. This extended model shares all the features shown in Proposition 1. Intutively this random accident rate places a floor in the level of mortality that is constant by age: if all health related deaths were eliminated, then we would observe this accident rate at every age and its level would uniquely determine the longevity of the population ($1/\kappa$).

4 Model estimation and results

4.1 Simulated methods of moments

To estimate the parameters for a given cohort we use the simulated method of moments. We simulate mortaliy rates for a given set of parameters and compare the simulated rates to the observed rates. We then iterate until we find the parameters that can best predict the data, that is those that result in the smallest prediction error. More precisely we chose the parameters by minimizing the distance between actual survival and predicted survival at each age:

$$\min_{\theta} \|SR_t - SR_t(\theta)\|$$

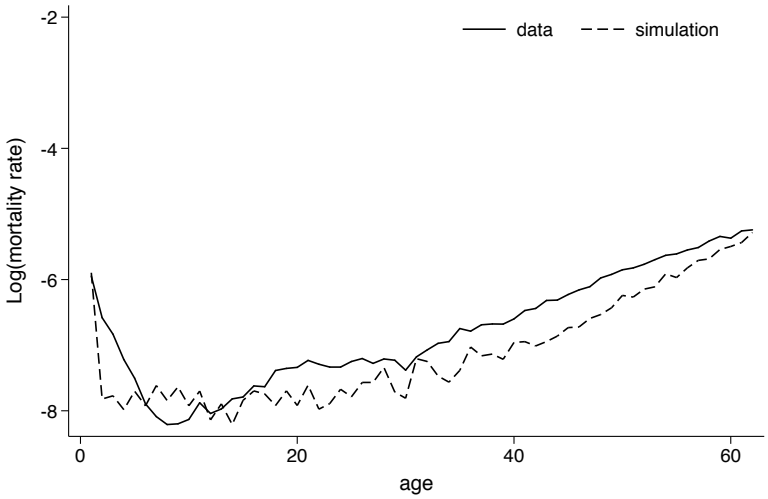
where $\theta = \{\alpha, \delta, I, \mu_0, \sigma_\epsilon, \kappa\}$.

Ideally the objective function that we choose to target for estimation would not affect the estimated parameters. This is not the case however. If we minimize the error in the level of mortality, we give equal weight to errors at all ages and penalize large deviations in levels (at very young and very old ages). If we minimize the errors in logs, then we minimize deviations in percentage terms–this effectively gives more weight to errors that are large relative to baseline, so in effect it weights reproductive ages more heavily. Ultimately the choice matters. Not only for estimation (different objective functions weight mortality at different ages differently), but because this choice implicitly reflects some welfare criteria.

We target the survival curve, which is equivalent to matching life expectancy. Life expectancy weights early ages very heavily because early deaths result in large losses in terms of years lived, and conversely it weights mortality in old ages less. This is a commonly used criteria in epidemiology and the results are easily interpretable: we can summarize the fit of the model based on how far the predictions are from observed life expectancy, and compare our fit to alternative models using this metric.

[12] https://www.cdc.gov/injury/images/lc-charts/leading_causes_of_death_age_group_2014_1050w760h.gif

Figure 6: 1945 French women's mortality profile



Note: Estimation using MATLAB's particleswarm routine. Life expectancy until age 62 is 60.2 in the data and 60.1 in our simulation.

Although standard errors can be bootstrapped, we do not report them here. Because each curve is traced out from complete population data the standard errors are effectively negligible, as as been noted elsewhere (HonorÃ© and Lleras-Muney, 2006). Other estimation details are in the Appendix. In this non-linear model mortality rates at all ages are a function of all the parameters, and conversely the parameters are functions of the mortality rates at all ages. But the comparative static results give intuition about which mortality rates identify the parameters of the model.

4.2 Results for a baseline cohort

We first apply our estimation procedure to a single cohort with the smallest excess mortality in reproductive ages: French women born in 1945. The results in Figure 6, show that we can predict the evolution of mortality rates over the lifetime well, particularly for women. The parameters are shown in Appendix table 1. Our estimates result in a life expectancy between 0 and age 62 of of 60.1–the actual number if 60.2. So the model provides a very good fit.

However visual inspection suggests that there is remaining error in estimation. There are at least three sources of error. First we estimate a model that assumes a stationnary environment with constant annual resources. But if health resources are related to GDP, then resources are growing over the lifetime along with GDP. Indeed if we allow I to grow at a fix rate (of 2%) then our fit improves. However since we do not know how to map health resources to GDP (we consider this below) we consider these only suggestive. Second if health

expenditures are "optimally allocated" over the lifetime then our assumption of a constant investment is also incorrect–we also address this later on. Lastly, the log curve reveals that there is still visible excess mortality during reproductive ages. We study this next.

4.3 Estimates for primates.

Our natural baseline model is not unique to humans and so far simply describes biological disease and aging processes. Many papers have suggested that primates and human have similar mortality profiles, particularly in old age. In this section we use data for primate populations to estimate our model. The data is taken from Kohler et al. (2006) who draw on the International Species Information System, a network of North American and European zoos. Kohler et al report period tables for several of phylogenic groups in captivity.[13] The Ape group comprises four species: Gorilla, Orangutan, Siamang, and White-cheeked gibbon. The survival rates for apes are computes for a population of 2,069 animals in total.

While our estimation procedure for humans relied on the cohort mortality rates at every age, Kohler et al report the survival rates only for ages 1, 2, 5, 10, 15, 20, and 25. We use just these 7 moments in our estimation.[14] Figure shows that the fit is remarkably good, though this is not surprising given that we have 7 moments and 5 parameters. But the estimated parameters in Table 1 are of interest. Collectively apes have much lower initial mean health, and substantially larger variance in annual resources. Most interestingly, α, the only scale free parameter, is very similar across humans and apes.

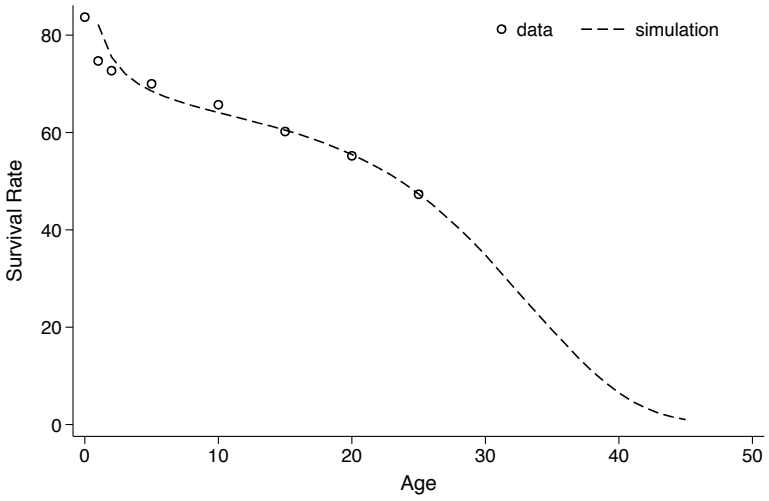
5 Understanding shocks: Accounting for mortality during reproductive ages

Our basic model does not account for the large variation in mortality rates observed during reproductive ages. Most cohorts display what looks like a smooth "hump" in reproductive ages. There are also large "spikes" during war years, which are also most visible during the reproductive years. Also there are differences between men and women. The spikes are larger for men and correspond to war years. And the bumps have different shapes, they are appear to last longer for women, and be more concentrated in the early reproductive years for men. If we interpret these as "exogenous shocks" then the model should be able to rationalize these as changes in the underlying parameters. In this section we investigate this.

[13] Unfortunately these are not cohort tables, but arguably the gap between cohort and period tables is much less problematic than for humans, particularly for animals in captivity. There are some cohort data for apes that have been recently collected but unfortunately these do not include data on infant mortality.

[14] Kohler et al. also report the 1 week survival rate, for which we do not have a direct counterpart in our discrete time annual model.

Figure 7: Survival curve for apes



Source: Kohler et al. for the life table of the Apes group.

5.1 Simulations

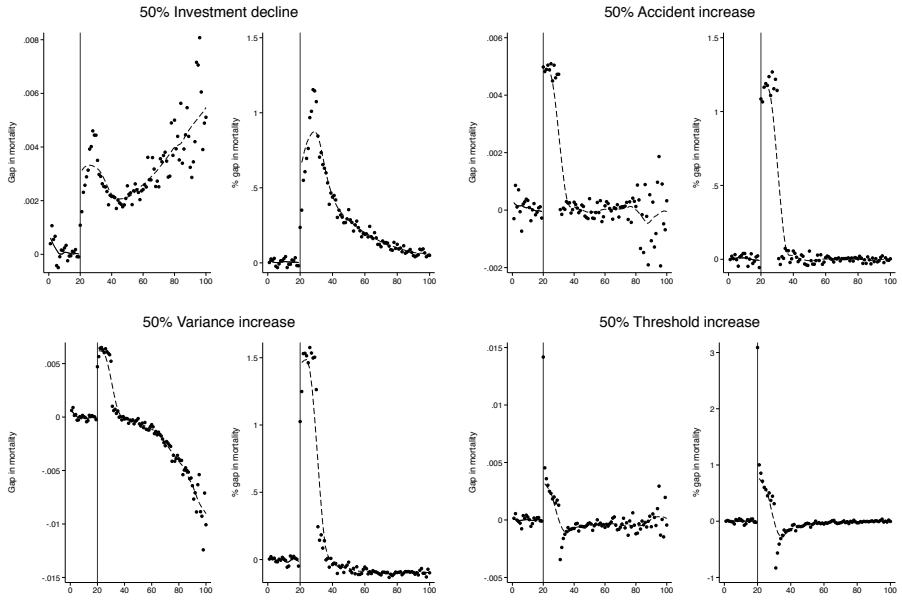
To explore whether changes in parameters can generate these patterns, we simulate 4 types of temporary shocks: an increase in the variance, a decrease in the annual investment level, an increase in the threshold and finally an increase in the accident rate.[15] We start with a baseline model without accidents. Figure 8 illustrates the effects of a shock starting at age 20 and lasting 10 years.

Each type of shock leaves a unique imprint on the mortality profile of the affected cohort. Temporary decreases in investment levels generate spikes in mortality, similar to those observed for wars. When investment falls, mortality rates start to rise, they peak the last year of the shock and they fall back thereafter. But mortality rates remain elevated throughout the lifetime (relative to the counterfactual of no shock) thus generating "scaring". By contrast, temporary increases in accidents temporarily increase mortality but have no permanent effects: mortality goes back to its initial path immediately after the shock ends.

Increases in the variance result in a sustained increase in mortality during the shock period. But after the shock ends, mortality falls below couterfactual levels–this is the result of having individuals with large positive shocks. Finally, an increase in the threshold appears to generate "harvesting"– earlier deaths for the frail. It results in very high mortality in the first year of the shock. But it starts dropping before the shock ends. Once the shock ends, it dips

[15] We do not simulate an increase in the aging parameters because they would not result in a large immediate spike in mortalit–their effects would be most visible at older ages (results available upon request).

Figure 8: Effect of exogenous temporary shocks at age 20



Note: Simulations for two populations of 500,000 individuals each. The shock lasts ten years. The baseline parameters are the same as in Figure 2.

below counterfactual mortality and then rises back up and converges to its counterfactual level. This is because all frail individuals are killed when the threshold first increases. And when the threshold is restored to its original (lower) level mortality falls substantially because there are very few individuals close to the threshold.

5.2 The effects of WWII

To test the predictions of the model we analyze the effects of wars. Wars correspond to the largest temporary change in mortality rates over time across cohorts for individuals aged 20.[16] We focus on WWII because individuals who were young at the onset of WWI also experienced large mortality increases in mortality in their 40s and 50s during WWII. Also WWII is the longest conflict in our sample lasting six years. The main disadvantage of looking at WWII is that there is very little data on mortality by age during this period–in fact we have information for 1940, 41, 42 and 44 but not for 1939, 43, 45 and 46. to estimate the model we use the same procedure we employed to estimate parameters for apes.

WWII is estimated to have killed around 600,000 individuals in France, about 1.4% of the 1939 population.We hypothesize that this war can be best

[16] Appendix Figure X shows the evolution of mortality rates in France for age 0, 20, and 40.

understood as a change in the average level of (health-related) resources or inputs, I. GDP, food supplies and sanitary conditions declined substantially during the war.[17] Infant mortality rates, which are very sensitive to these inputs, rose substantially, as can be seen in Appendix Figure 18.

Appendix Figure 20 shows the mortality profiles of French cohorts born in 1920 who were 19 when the war started. We also show the profiles for Sweden–the only country in the sample that did not participate in the war. Compared to Sweden mortality rates in France are substantially more elevated during the war years, and they remain elevated for many years after the war, similar to the simulated effect of an investment decline. However these comparisons are imperfect because Swedish mortality rates before the war are different than the French's. We assess the effect of the war using our model instead.

We estimate the structural parameters of the 1920 cohort explicitely allowing for a shock lasting six years. We estimate the model four times, corresponding to the four type of shocks simulated above. If the war is like a decrease in the annual investment, then this model should provide the best fit for the data. This is precisely what we find (see Appendix Table 2). In 9 we show the results for the best fitting model along with the "counterfactual" mortality–what we would have observed in the absence of the war. Using the counterfactual mortality curve we can estimate the "true" life expectancy-cost of the war. We find that the war lowered life expectancy by approximately 5.5 years for the 1920 cohort.

5.3 Accounting for maternal mortality and reproductive age mortality in non-war times

In peaceful times, there is still visible excess mortality during reproductive ages that manifests itself in a bump. We now investigate this.

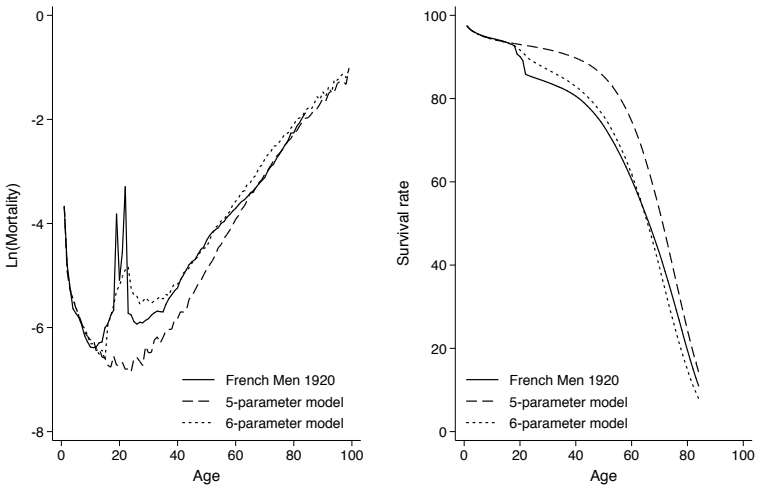
For women, maternal mortality was one of the largest causes of death during reproductive ages in the past, and it was high in France in 1860. Maternal mortality risk is only incurred with childbirth. Appendix Figure 22 shows that in fact the age profile of fertility is strongly correlated with the extent of excess mortality during reproductive ages. We have no data on the age at maturity or for fertility rates by age for the oldest cohorts in the data. So instead we consider a model that changes the parameters arbitrarily at age 13. Although this is couterfactual, it provides a natural starting point.[18]

It is unclear how to best conceptualize maternal mortality risk in terms of the parameters of the model. Loudon (2000) argues that historically poor hygiene and obstetric practices were mostly responsible for infections (spesis) and

[17] GDP declined substantially during the war–this is shown in appendix Figure X. Moreover Occhino et al. (2006)estimate that between 20 to 55% of GDP was appropriated by Germans every year of the occupation. There was also a substantial decrease in the availability of food–food rationing began in 1940. There was also a deterioration in sanitary conditions in France. For example diphteria cases among school-aged children rose per 100,000 increased from 32.3 (in 1940) to 110 in 1943 (Stuart 1945).

[18] We investigate the simplest case here. We can in fact do better by allowing the onset of maturity to follow a normal distribution in the population. But this would add two more parameters to all our estimations.

Figure 9: WW2 as an investment shock. French Men born in 1920.



hemorrage–the main reasons why women died during childbirth. In countries where poor obstetric practices were widespread, maternal mortality was large and it killed both rich/healthy and poor/unhealthy women. Based on this evidence we hypothesize that maternal mortality can be thought of as an increase in the accident rate and should be largest during prime reproductive ages. But again we estimate four models (increase in threshold, increase in accident rate, decrease in annual resources, and increase in variance) to see which one matches the data best.

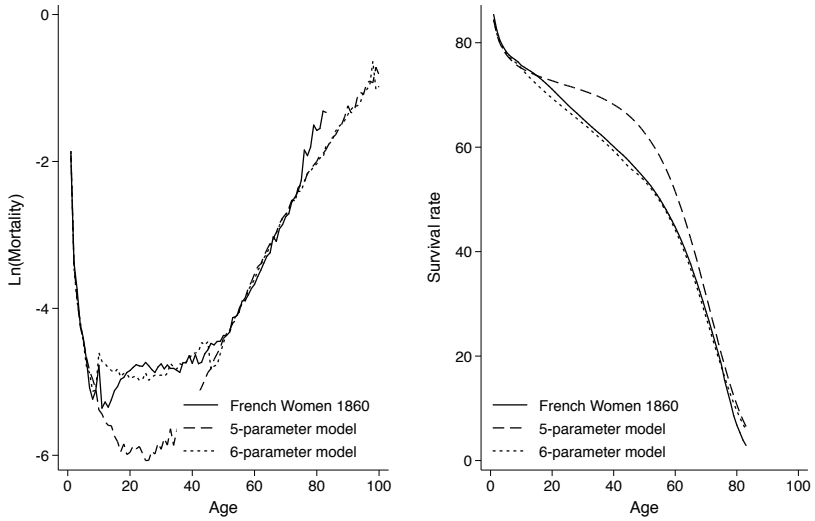
Appendix Table 3 shows the results from this exercise. It shows that although all models improve the fit, accidents improve the fit the most. Figure 10 illustrates the results for an increase in accidents along with counterfactual mortality. Maternal mortality is estimated to have lowered female life expectancy by around 5.7 years for the 1870 cohort. If we allow for the onset of maturity to be normally distributed then out model results in further improvements–this can be seen in Appendix Figure X for the 1870 cohort for whom we estimate the onset of maturity to be 14.5 years of age on average.

6 Evolution of parameters over time

We now estimate the parameters for successive cohorts born from 1860 to 1940, accounting for both wars and reproductive age mortality. Appendix Table 4 reports estimates for selected cohorts, and summary statistics across cohorts. We succesfully replicate life expectancy increases over time, and level differences between men and women.

Figure 7 show the evolution of the parameters across birth cohorts when we

Figure 10: Accounting for maternal mortality and excess mortality among women

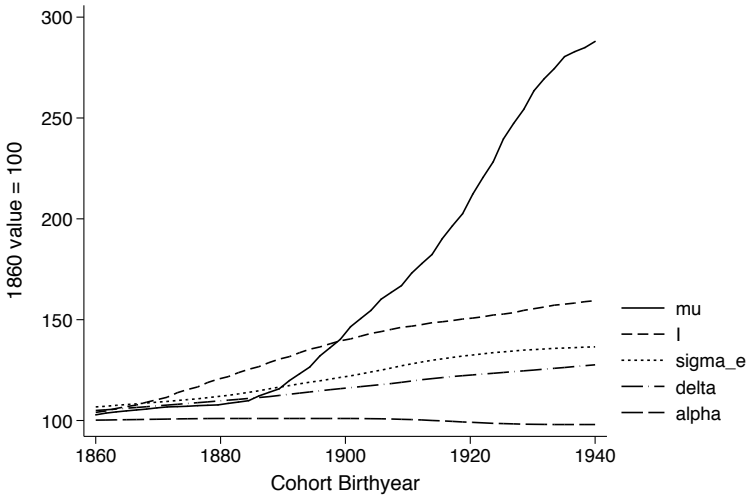


estimate the parameters without accounting for reproductive age mortality. It shows some remarkable patterns. Initial health only starts rising around the turn of the century. This parameter is hard to interpret as purely driven by initial health as it could capture improvements in utero as well as shocks that affect only those under age 1. But the results are entirely consistent with the intuition that changes in initial conditions are mostly pinned down by infant mortality: Appendix Figure 18 in the appendix shows that infant mortality in France started dropping sharply around the turn of the century, and demographic work has attributed these declines to improvements in sanitation and water quality. . The correlation between estimated initial levels and infant mortality rates is high.

The level of annual investment I increases throughout the period showing that lifetime resources were increasing for all cohorts since 1860, at a rather constant rate. This is consistent with increasing levels fo GDP per capita throughout and the estimated correlation between I and GDP is also high. We estimate that the variance of annual shocks has increased but the baseline accident rate has fallen.

We estimate two parameters for aging, one which is a function of the scale of the intial distribution (δ) and one which is not (α). δ increases a little bit following mean investment increases. On the other hand the parameter governing the acceleration of depreciation with age is constant throughout, consistent with it capturing a biological parameter. It is conforting that this parameter does not change despite changes in initial conditions and investments–this pa-

Figure 11: Evolution of estimated parameters, France 1860-1940 cohorts



Note: Local polynomial smoothing.

rameter is "scale free" (unaffected by our scale and location normalization) and thus should not vary with the changes in initial health of annual resources.

Figure (TBD shows the evoluation of parameters if we account for wars and humps in estimation (the parameters are shown in Appendix Table 4). Accounting for the shocks significantly affects the parameter estimates–most significantly we find that the estimated annual investment is much larger once shocks are accounted for. But all parameters are affected. This underscores the importance of fully accounting for the history of shocks. However the main qualitative conclusions about the trends in these parameters are unaffected.

Men and women (TBD).

Secular trends in morbidity and mortality. The historical literature analyzing secular trends in morbidity and mortality has found that morbidity and mortality have fallen in tandem (Costa, 2005). This is consistent with our finding that investments have increased, because greater investments lead to both improved health and lower mortality. This is indeed the interpretation of the historical evidence by Fogel (1994), who documents that the improved availability of food led to increases in heights, and to declines in morbidity and mortality.

7 Implications for optimal investment profiles

For now we have assumed a constant investment profile over the lifetime. But would that be an optimal allocation of resources over the lifetime? In this section we show that a social planner concerned with maximizing the life-expectancy

of a population would choose an investment profile that generates patterns of mortality with striking similar shapes of the ones studied in the previous sections. We then go on to re-estimate our parameters alowing for investment to be chosen optimally to assess how much our basic parameters are affected, and evaluate the life expectancy gains resulting from optimization.

First we develop notation to describe the problem that would face benevolent social planner. The survival function tracks the probability of surviving over time. It is naturally expressed as a function of the cdf of health in the population. The probability of surviving until the end of period t is $S_t = 1 - F_t(0)$. Life expectancy at birth is conveniently related to the survival function

$$LE = \sum_{t=1}^{\infty} S_t$$

Several observations are in order. First, in practice, this is a finite sum.Second, contrast this concept with the "period" life expectancy usually computed. If the distribution where stationary over time, then the two concepts would coincide. But as the data shows and our estimates corroborate, the mortality rates are not stationary.

Now suppose that instead of keeping I constant that the social planner can choose an investment path $\mathcal{I} = \{I_t\}$ that is age-dependent. Also assume that the budget (B) over the lifetime is fixed but that the planner can move resources over time periods costlessly, as if a perfect annuity were available.[19] The optimization problem takes the form

$$\begin{aligned} \max_{\mathcal{I}} LE(\mathcal{I}) &= \max_{\{I_t\}} \sum_{t=1}^{\infty} S_t(\mathcal{I}) \\ s.t. & \quad \sum_{t=1}^{\infty} I_t \cdot S_t(\mathcal{I}) \leq B \end{aligned}$$

The first order conditions for this maximization are given by:

$$\sum_{s=t}^{\infty} \frac{\partial S_s(\mathcal{I})}{\partial I_t} - \lambda \left[S_t(\mathcal{I}) + I_t \frac{\partial S_t(\mathcal{I})}{\partial I_t} + \sum_{s>t}^{\infty} I_s \frac{\partial S_s(\mathcal{I})}{\partial I_t} \right] = 0$$

where is the Lagrange multiplier and it represents the shadow cost for the social planner of an additional year of life expectancy.

The FOC imply that on the optimal investment path, the marginal effect of increasing investment at a given age must be equalized across all ages. Increases in life expectancy (the first term on the left-hand), must be balanced by the losses incurred by having to tighten the budget at subsequent period to keep the budget balanced (the term in brackets).

All the terms in the bracket are positive.

[19] This is a standard assumption in this type of models, for example see Murphy and Topel (2006).

7.1 Timing of optimal investments, polynoms

A full nonparametric approach for the optimal investment profile over the lifetime would require optimizing over a hundred or so parameters (one for each age) for each cohort. In the absence of a closed-form solution, this is impractical. It is also not feasible since we have 90 data points: if we allow for a unique investment level at every age we are under-identified (we would have 90 data points and 95 parameters to estiamte). Instead, we follow a lower-dimensional sieves estimation method.

We start with by approximating the investment profile over age with a second order polynomial. We impose the constraint that the total spending per cohort is the same as in the constant investment case. Given a budget B we run a grid search to find the quadratic investment profile that maximizes the life expectancy of the cohort.

The results of this exercise are in Figure 12. We find that a U-shape investment profile is optimal to maximize the average life-expectancy in the population (panel a). Notice that although our original model sets I to be constant in levels, in percentage terms, relative to the baseline level of health at a given age, I was already U-shaped. What we find is that the optimal investment is even more U-shaped – that is, it transfers additional resources to the young and the old, away from the middle-aged individuals.

Panel b shows the mortality curve before and after optimization–it has the same basic shape we have observe. But there are gains from optimization: these are shown in panels c and d (levels and percent changes). When we move from a fixed to an optimal profile then we decrease child and elderly mortality. Interestingly health care expenditures by age in most countries actually follow this age-profile.

These results show that optimal health investments are largest when health is at its lowest, that is at very young and very old ages. This is consistent with empirical findings Wagstaff (1986).

7.2 Other Properties of investments

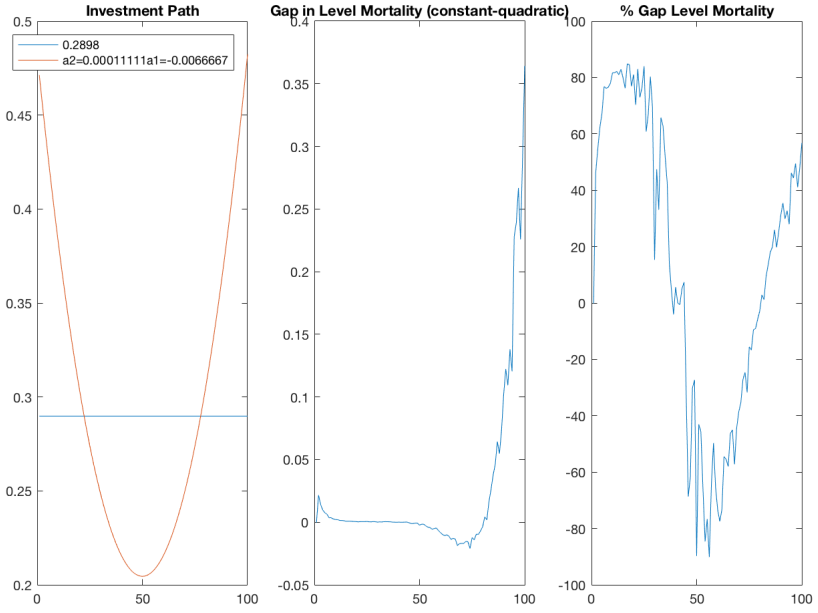
7.2.1 Dynamic complementarities

Each investment profile, $\mathcal{I} = \{I_t\}$, generates a sequence of distributions of health, $F_{H_t}(\mathcal{I})$, and its associated mortality rates, $MR_t(\mathcal{I})$. How are investment decisions at different ages related? We show in the following proposition that investment in health are dynamic complements.

Proposition 4 Optimal investments in health are dynamic complements throughout lifetime. $\forall t_1, t_2\, t_1 < t_2$, $\frac{\partial^2 MR_{t_2}(\mathcal{I})}{\partial I_{t_1} \partial I_{t_2}} \leq 0$

The complementarity arise through the health accumulation process. A higher investment at time t pushes the distribution of health up moving more individuals away from the threshold. Additional investments make it more unlikely that negative shocks will push individuals below the threshold.

Figure 12: Optimal Investment Levels by Age



Data: France, Women 1870

7.2.2 Compensation

We can now study optimal compensation: if a cohort suffers from an unlucky shock and the planner wishes to compensate them so that survivors can enjoy the same mortality rates they would have in the absence of the shock, how can the planner achieve this, and how does this vary with the level and timing of the shock? Our results imply that negative shocks need to be more than compensated for, because of complementarity. In other words a decrease in investments needs to be followed by an increase investment level that is greater than than the loss to give the survivors the mortality profile they would have experienced in the absence of a shock.

- Currie versus Chetty.

7.2.3 Extension 1: optimization when budgets depend on health.

We have solved the optimization problem under the (strong) assumption that resources are not a function of population health. But if food and other resources are produced rather than taken from the environment, health is likely to impact resources by affecting the work capacity of the population. Indeed nutrition levels and disease rates have been shown to affect productivity and wages (Thomas et al., 2004). They also affect inputs into wages such as cognition and education (Field et al., 2009). Many empirical studies report a strong correlation between income and health (Cutler et al., 2012). While our baseline model embeds the effect of resources of health, a causal link going in the other direction is also at play: people who get sick suffer a drop in income (Smith, 1999).

The simplest way to incorporate this channel in our setting is to assume that people whose health lower than some disability or disease level, h_D , are unable to participate in production and thus generate zero income while people whose health is high enough generate income w .

$$\begin{aligned} \max_{\mathcal{I}} LE\left(\mathcal{I}\right) \quad &= \quad \max_{\{I_t\}} \sum_{t=1}^{\infty} S_t\left(\mathcal{I}\right) \\ s.t. \qquad & \sum_{t=1}^{\infty} I_t \cdot S_t\left(\mathcal{I}\right) \leq B = \sum_{t=1}^{\infty} w\left[1 - F_{H_t}(h_d;\mathcal{I})\right] \cdot S_t\left(\mathcal{I}\right) \end{aligned}$$

Again here we assume either an annuity market, or some pooling across cohorts in a stationary environment such that it is the cohort budget that matters, not the the within-cohort, per-period one. The first order condition is:

$$\sum_{s=t}^{\infty} s \cdot \frac{\partial S_s\left(\mathcal{I}\right)}{\partial I_t} + \lambda \left[\sum_{s=t}^{\infty} \left\{ w\left[1 - F_{H_t}(h_d;\mathcal{I})\right] - I_t \right\} . \frac{\partial S_s\left(\mathcal{I}\right)}{\partial I_t} - \sum_{s=t}^{\infty} \left\{ w \frac{\partial F_{H_s}(h_d;\mathcal{I})}{\partial I_t} + 1 \right\} \cdot S_s\left(\mathcal{I}\right) \right] = 0, \, \forall t > 1$$

or

$$\sum_{s=t}^{\infty} s \cdot \frac{\partial S_s(\mathcal{I})}{\partial I_t} = \lambda \left[\sum_{s=t}^{\infty} \left\{ w \frac{\partial F_{H_s}(h_d;\mathcal{I})}{\partial I_t} + 1 \right\} \cdot S_s(\mathcal{I}) + \sum_{s=t}^{\infty} \{ I_t - w [1 - F_{H_t}(h_d;\mathcal{I})] \} \cdot \frac{\partial S_s(\mathcal{I})}{\partial I_t} - \right] \forall t > 1$$

First notice that $\frac{\partial F_{H_s}(h_d;\mathcal{I})}{\partial I_t} < 0$ for all t. (cf Proposition 1) and $\frac{\partial S_s(\mathcal{I})}{\partial I_t} > 0$.

The resulting optimal investments by age, and associated mortality profiles are shown in Figure X. TBD.

7.2.4 Extension 2: optimizations when inputs into health are required.

A second important simplification in our analysis so far is that we have assumed that it is possible to directly manipulate or choose I. But I is measured in this model in health units. Individuals cannot in fact directly chose levels of I. Instead, like in the Grossman model, they chose inputs into health: they decide how much food exercise medicine alcohol to consume, based on how much utility they derive from these items directly and on how these affect their health.

Suppose that spending $C(h)$ unit of money delivers h units of health. Then the planner's problem can be refomulated as follows:

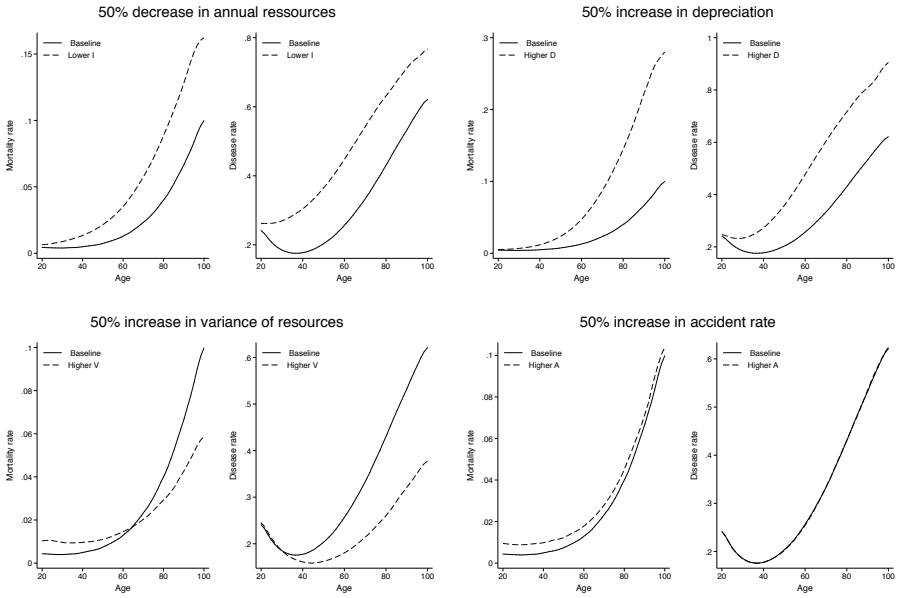
$$\begin{aligned} \max_{\mathcal{I}} LE(\mathcal{I}) \quad &= \quad \max_{I_t} \sum_{t=1}^{\infty} S_t(\mathcal{I}) \\ s.t. \quad & \sum_{t=1}^{\infty} \int_0^{\infty} C(I_t(h)) \, dF_{H_t}(h;\mathcal{I}) \leq B \end{aligned}$$

Here we must assume that the production function is concace, otherwise individuals could buy their immortality. The resulting investment profile is not much affected so long as this concavity is the same for every age. However this model is still too simple: we are only allwoing one input into health. But since we have no data to estimate this model for the cohorts we consider we leave this to future research.

8 Applications

Recent increases in white mortality in the US. Case and Deaton (2015) documented that in the US mortality rates among whites ages 40 and above have started to rise, particularly among those with low education. Case and Deaton (2017) further show that the age profile of mortality and morbidity have become steeper with each successive birth cohort born after 1950. They investigate possible reasons and conclude that "The data are consistent with long-run processes influencing outcomes, rather than contemporaneous shocks affecting health." They propose this health decline is caused by one (or two) latent negative factor that affect all cohort member when they enter the labor market and stays with them for the rest of their lives.

Figure 13: Permanent changes at age 20



To investigate whether our model can replicate these findings we simulate the effects of permenent changes in the parameters that occur at age 20. Figure 13 show the results for mortality and morbidity. Lower annual resources or larger depreciation fit the facts: they result in steeper age profiles. Higher accident rates, higher variance and higher threshold (not shown) do not. Case and Deaton argue that "The data are consistent with long-run processes influencing outcomes, rather than contemporaneous shocks affecting health". Our model is entirely consistent with their conclusions, as there is no temporary change, whether at age 20 or 40, that can produce what Case and Deaton observe. An just like them we find that there are two unobserved underlying factors that can possibly explain the patterns: lower annual resources and higher depreciation. The model is explicit about why lower resources lead to poor health and higher mortality. What might cause higher deterioration rates? Several hypothesis are available in the literature, the two most prominent ones concern the effect of low rank and its effects on the immune system (Sapolsky, 2004) and another is the weathering hypothesis that suggests that continuous exposire to hazardous environments results in accelerated aging (Geronimus et al., 2006).

Income and GDP puzzle. Preston first showed that life expectancy rises with GDP, albeit at a declining rate among the richest countries. Yet many studies have demonstrated that changes in GDP are not strongly predictive of changes in life expectancy, and moreover in the short run, recessions are linked

to lower (not higher) mortality (Ruhm 2000).[20] Our model can be used to think about why short and long term effects of income differ. We provide two possible resolutions of this puzzle.

The first explanation is suggested by the simulations above. Increases in the variance of resources increase mortality in the short term but decreases it in the long term. If recessions lower both the level of resources (I) and the variance of its distribution (σ_e) then the model can generate the pattern we observe. If the decline in income is not too large relative to the decline in the variance, then in the short term mortality will fall. In the long term mortality will rise however. This is due to 3 effects. First there is scaring caused by short term decline in resources (as with the war case). The second is that variance changes generate selection effects: fewer frailer individuals die in the short run. Lastly lower variance also generates fewer lucky recipients of large positive shocks which also causes mortality to rise in the long term.

A second possible resolution of the puzzle comes from thinking more carefully about the production of health. We treat I as a constant–it represents annual health investments. But as noted in Section X, individuals cannot directly buy I, it needs to be produced with inputs like food and exercise. Cutler et al. 2016 argue that recessions increase both good and bad inputs into health. In bad economic times food consumption falls but alcohol consumption falls too, and so does pollution. The net short term impact of these changes in inputs on health is ambiguous. And their long term impacts on I is also ambiguous and can differ in sign–this depends a great deal on whether the changes in inputs are temporary or permanent. Cutler et al. 2016 use the model provided here to provide mathematical derivation of these results in their appendix.

Period age and cohort effects. Our model places parametric restrictions on the evolution of health and mortality by age. A parametric model of health and mortality by age has many advantages. A long literature in demography and economics has struggled to separately identify age, period and cohort effects, which are not non-parametrically identified. Because age effects are parametric here, cohort and period effects can be separately and non-parametrically identified. This is in fact illustrated here: we provide separate estimates by cohort, and we also estimate period effects like wars, separately for each cohort, and at different ages.

The persistance of health and other implications for individual panel data analysis. This model has some attractive features relative to the standard and widely used Grossman model. We do not assume a fixed horizon–the age at death is naturally determined by health and health investments.

It also has implications for analysis using individual panel (longitudinal) data which track individual income, health and mortality over time. Our model predicts that changes in resources and changes in health will be related but only weakly, aprticularly when health is measured in discrete units. But measures of permanent resources will be strongly predictive of mortality. (TBD).

[20] For a longer discussion of this evidence see Cutler et al. 2006

9 Conclusion

This paper proposes a simple model of the evolution of health and mortality over the life course. This model is inspired by the basic observation that the age-profile of mortality is remarkably constant over time and cohorts. If health leads to mortality then we can learn about the underlying evolution of health by observing mortality rates. The basic model has five parameters and can be easily simulated. It can approximate quite well the mortality profile of cohorts born 1860 to 1940.

We use the model to study the short and long term effect of shocks occuring at different points in the lifetime. And to study the cumulative effect of permanent circumstances over the life course. This model can be used to investigate many interesting questions that we have not considered. For example if can be used to study correlations in health across generations in an over-lapping generations setting. It can also be used to think about the age profile of mortality among the oldest old. Its implications for wages, consumption and health care expenditures can be improved and taken to data for more recent cohorts.

This paper has some important limitations. First and foremost health is treated as an unobserved latent variable–we only demonstrate that our model of health delivers a mortality age-profile that is consistent with observed cohort mortality. Secondly there is a scale and location normalization that cannot be avoided–there are 2 parameters of the model that cannot be identified. This makes it is difficult to interpreting parameters across cohorts and countries: increases in initial health really could come from lower threshold or changes in standard deviation. Lastly our model does not provide closed form solutions and must be estimated using numerical methods. These methods are very sensitive to initial conditions and methodology.

References

Almond, Douglas and Janet Currie, "Killing me softly: The fetal origins hypothesis," *The Journal of Economic Perspectives*, 2011, *25* (3), 153–172.

_ , _ , and Valentina Duque, "Childhood Circumstances and Adult Outcomes: Act II," Technical Report, National Bureau of Economic Research 2017.

Bardet, Jean-Pierre, K.-A. Lynch, G.-P. Mineau, Michael Hainsworth, and Matthew Skolnick, "La mortalitÃ© maternelle autrefois: une Ã©tude comparÃ©e (de la France de l'ouest Ã l'Utah)," in "Annales de dÃ©mographie historique," Vol. 1981 PersÃ©e-Portail des revues scientifiques en SHS 1981, pp. 31–48.

Bouvier-Colle, M.-H., F. Pequignot, and E. Jougla, "Mise au point sur la mortalitÃ© maternelle en France: frÃ©quence, tendances et causes," 2008.

Bozzoli, Carlos, Angus Deaton, and Climent Quintana-Domeque, "Adult height and childhood disease," *Demography*, 2009, *46* (4), 647–669.

Bronikowski, Anne M., Jeanne Altmann, Diane K. Brockman, Marina Cords, Linda M. Fedigan, Anne Pusey, Tara Stoinski, William F. Morris, Karen B. Strier, and Susan C. Alberts, "Aging in the natural world: comparative data reveal similar mortality patterns across primates," *Science*, 2011, *331* (6022), 1325–1328.

Bruckner, Tim and Ralph Catalano, "The sex ratio and age-specific male mortality: evidence for culling in utero," *American Journal of Human Biology*, 2007, *19* (6), 763–773.

Carnes, Bruce A., S. Jay Olshansky, and Douglas Grahn, "Continuing the search for a law of mortality," *Population and Development review*, 1996, pp. 231–264.

Carriere, JACQUES F., "Parametric models for life tables," *Transactions of the Society of Actuaries*, 1992, *44*, 77–99.

Case, Anne and Angus Deaton, "Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century," *Proceedings of the National Academy of Sciences*, 2015, *112* (49), 15078–15083.

_ , Darren Lubotsky, and Christina Paxson, "Economic status and health in childhood: The origins of the gradient," *The American Economic Review*, 2002, *92* (5), 1308–1334.

Chetty, Raj, Michael Stepner, Sarah Abraham, Shelby Lin, Benjamin Scuderi, Nicholas Turner, Augustin Bergeron, and David Cutler, "The association between income and life expectancy in the United States, 2001-2014," *Jama*, 2016, *315* (16), 1750–1766.

Contoyannis, Paul, Andrew M. Jones, and Nigel Rice, "The dynamics of health in the British Household Panel Survey," *Journal of Applied Econometrics*, 2004, *19* (4), 473–503.

Costa, Dora L., "Causes of improving health and longevity at older ages: a review of the explanations," *Genus*, 2005, pp. 21–38.

Cunha, Flavio and James Heckman, "The Technology of Skill Formation," *The American Economic Review*, 2007, *97* (2), 31–47.

Currie, Janet and Mark Stabile, "Socioeconomic Status and Child Health: Why Is the Relationship Stronger for Older Children?," *The American Economic Review*, 2003, *93* (5), 1813–1823.

Cutler, David, Angus Deaton, and Adriana Lleras-Muney, "The determinants of mortality," *The Journal of Economic Perspectives*, 2006, *20* (3), 97–120.

Cutler, David M., Adriana Lleras-Muney, Tom Vogl, S. Glied, and P. C. Smith, "The Oxford Handbook of Health Economics," 2012.

_ , Wei Huang, and Adriana Lleras-Muney, "Economic Conditions and Mortality: Evidence from 200 Years of Data," Technical Report, National Bureau of Economic Research 2016.

Deaton, Angus S. and Christina H. Paxson, "The effects of economic and population growth on national saving and inequality," *Demography*, 1997, *34* (1), 97–114.

_ and _ , "Aging and inequality in income and health," *The American Economic Review*, 1998, *88* (2), 248–253.

_ and Christina Paxson, "Saving, growth, and aging in Taiwan," in "Studies in the Economics of Aging," University of Chicago Press, 1994, pp. 331–362.

Field, Erica, Omar Robles, and Maximo Torero, "Iodine deficiency and schooling attainment in Tanzania," *American Economic Journal: Applied Economics*, 2009, *1* (4), 140–169.

Finch, Caleb E. and Eileen M. Crimmins, "Inflammatory exposure and historical changes in human life-spans," *Science*, 2004, *305* (5691), 1736–1739.

Fogel, Robert W., "Economic Growth, Population Theory, and Physiology: The Bearing of Long-Term Processes on the Making of Economic Policy," *The American Economic Review*, 1994, *84* (3), 369–395.

Geronimus, Arline T., Margaret Hicken, Danya Keene, and John Bound, "âWeatheringâ and age patterns of allostatic load scores among blacks and whites in the United States," *American journal of public health*, 2006, *96* (5), 826–833.

Gompertz, Benjamin, "A Sketch of an Analysis and Notation Applicable to the Estimation of the Value of Life Contigencies," *Philosophical Transactions of the Royal Society of London*, 1820, *110*, 214–332.

_ , "On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies," *Philosophical transactions of the Royal Society of London*, 1825, *115*, 513–583.

_ , "A Supplement to Two Papers Published in the Transactions of the Royal Society," On the Science Connected with Human Mortality;" The One Published in 1820, and the Other in 1825," *Philosophical Transactions of the Royal Society of London*, 1862, *152*, 511–559.

_ , "On one uniform law of mortality from birth to extreme old age, and on the law of sickness," *Journal of the Institute of Actuaries and Assurance Magazine*, 1871, *16* (5), 329–344.

Grossman, Michael, "On the concept of health capital and the demand for health," *Journal of Political economy*, 1972, *80* (2), 223–255.

Heligman, Larry and John H. Pollard, "The age pattern of mortality," *Journal of the Institute of Actuaries*, 1980, *107* (01), 49–80.

HonorÃ©, Bo E. and Adriana Lleras-Muney, "Bounds in competing risks models and the war on cancer," *Econometrica*, 2006, *74* (6), 1675–1698.

Hummer, Robert A. and Joseph T. Lariscy, "Educational attainment and adult mortality," in "International handbook of adult mortality," Springer, 2011, pp. 241–261.

Kohler, Iliana V., Samuel Preston, and Laurie Bingaman Lackey, "Comparative mortality levels among selected species of captive animals," *Demographic Research*, 2006, *15*, 413–434.

Li, Ting and James Anderson, "Shaping human mortality patterns through intrinsic and extrinsic vitality processes," *Demographic research*, 2013, *28*, 341–372.

Loudon, Irvine, "The transformation of maternal mortality.," *BMJ: British Medical Journal*, 1992, *305* (6868), 1557.

_ , "Maternal mortality in the past and its relevance to developing countries today," *The American journal of clinical nutrition*, 2000, *72* (1), 241s–246s.

Murphy, Kevin M. and Robert H. Topel, "The value of health and longevity," *Journal of political Economy*, 2006, *114* (5), 871–904.

Occhino, Filippo, Kim Oosterlinck, and Eugene N. White, "How Occupied France Financed Its Own Exploitation in World War II," Technical Report, National Bureau of Economic Research 2006.

Palloni, Alberto and Hiram BeltrÃ¡n-SÃ¡nchez, "Demographic Consequences of Barker Frailty," in "Dynamic Demographic Analysis," Springer, 2016, pp. 147–176.

Pinker, Steven, *The better angels of our nature: Why violence has declined*, Vol. 75, Viking New York, 2011.

Pison, Gilles, "France 2004: Life expectancy tops 80 years," *Population & Societies*, 2005, (410), 1.

Preston, Samuel, Patrick Heuveline, and Michel Guillot, "Demography: measuring and modeling population processes," 2000.

Ruhm, Christopher J., "Are recessions good for your health?," *The Quarterly journal of economics*, 2000, *115* (2), 617–650.

Sapolsky, Robert M., *Why zebras don't get ulcers: The acclaimed guide to stress, stress-related diseases, and coping-now revised and updated*, Macmillan, 2004.

Sharrow, David J. and James J. Anderson, "Quantifying Intrinsic and Extrinsic Contributions to Human Longevity: Application of a Two-Process Vitality Model to the Human Mortality Database," *Demography*, 2016, *53* (6), 2105–2119.

Smith, James P., "Healthy bodies and thick wallets: the dual relation between health and economic status," *The journal of economic perspectives: a journal of the American Economic Association*, 1999, *13* (2), 144.

Thomas, Duncan, Elizabeth Frankenberg, Jed Friedman, Jean-Pierre Habicht, Nathan Jones, Christopher McKelvey, Gretel Pelto, Bondan Sikoki, James P. Smith, Cecep Sumantri, and others, "Causal effect of health on labor market outcomes: Evidence from a random assignment iron supplementation intervention," *California Center for Population Research*, 2004.

Vaupel, James W., Kenneth G. Manton, and Eric Stallard, "The impact of heterogeneity in individual frailty on the dynamics of mortality," *Demography*, 1979, *16* (3), 439–454.

Wagstaff, Adam, "The demand for health: some new empirical evidence," *Journal of Health economics*, 1986, *5* (3), 195–233.

Wilcox, Allen J. and IAN T RUSSELL, "Birthweight and perinatal mortality: I. On the frequency distribution of birthweight," *International Journal of Epidemiology*, 1983, *12* (3), 314–318.

Appendix A1: Figures

Appendix A2: Proofs omitted in the text

Mortality rates in periods 1 and 2

In the first period the (infant) mortality rate MR_1 is given by

$$MR_1 = P(H_1 \leq \underline{H}) = P\left(H_0 + I(Y_1, B_1) - \delta + \varepsilon_1 \leq \underline{H}\right)$$

$$= P(\varepsilon_1 \leq \varphi_1) = F(\varphi_1)$$

where $\varphi_1 = \underline{H} - I(Y_1, B_1) + 1 - H_0$ captures the threshold for dying in period 1 in terms of the random shock. Investments lower this threshold (lower mortality) and depreciation increases it (increases mortality).

Consider now the probability of dying at age $t = 2$. This is given by the probability that the stock falls below $\underline{H}$ at age 2, conditional on having survived to age 2, which can be expressed as:

$$\begin{aligned} MR_2 &= E(D_2 = 1 | D_1 = 0) = P(H_2 < \underline{H} | H_1 > \underline{H}) \\ &= \frac{P(H_2 < \underline{H}, H_1 > \underline{H})}{P(H_1 > \underline{H} | g_1, g_2)} = \frac{P(\varepsilon_2 < \varphi_2 - \varepsilon_1, \varepsilon_1 > \varphi_1)}{1 - F(\varphi_1)} \\ &= \frac{K(\varphi_2, \varphi_1)}{1 - F(\varphi_1)} \end{aligned} \tag{3}$$

where $\varphi_2 = \underline{H} - I(Y_1, B_1) - I(Y_2, B_2) + 1 + 2^{\alpha} - H_0$ captures the threshold for dying in period 2, and $K(\varphi_2, \varphi_1) = \int_{\varepsilon_1 = \varphi_1}^{\infty} \int_{\varepsilon_2 = -\infty}^{\varphi_2 - \varepsilon_1} f(\varepsilon_1) f(\varepsilon_2) d\varepsilon_1 d\varepsilon_2$. $K(\varphi_2, \varphi_1)$ is the density right above the old threshold and below the new threshold, that is the fraction of survivors who dies as a result of a new small shock. It is increasing in the current threshold $\dfrac{\partial K}{\partial \varphi_2} > 0$ and decreasing in the past threshold $\dfrac{\partial K}{\partial \varphi_1} < 0$. The denominator is the fraction of survivors, and it is a negative function of the previous thresholds (because $F'(\varphi_1) > 0$).

Proof for Proposition 1 basic features of the model

1) Everyone dies

Consider the process $\{H_t^*\}_{t=1}^{\infty}$, defined by $H_0^* = H_0 \sim \mathcal{N}(\mu_0, \sigma_0)$ and the recurrence relation:

$$H_t^* = H_{t-1}^* + I - \delta \cdot t^{\alpha} + \epsilon_t \ , \ \epsilon_t \sim \mathcal{N}(0, \sigma_{\epsilon})$$

Notice that for $t > 1$, we can reformulate the mortality process in our basic model as:

$$H_t = \begin{cases} H_t^* & \text{if } H_t^* > 0 \text{ and } H_{t-1} > 0 \\ 0 & \text{otherwise} \end{cases}$$

Figure 14: Age profile of mortality of women born in France between 1860 and 1940, by decade

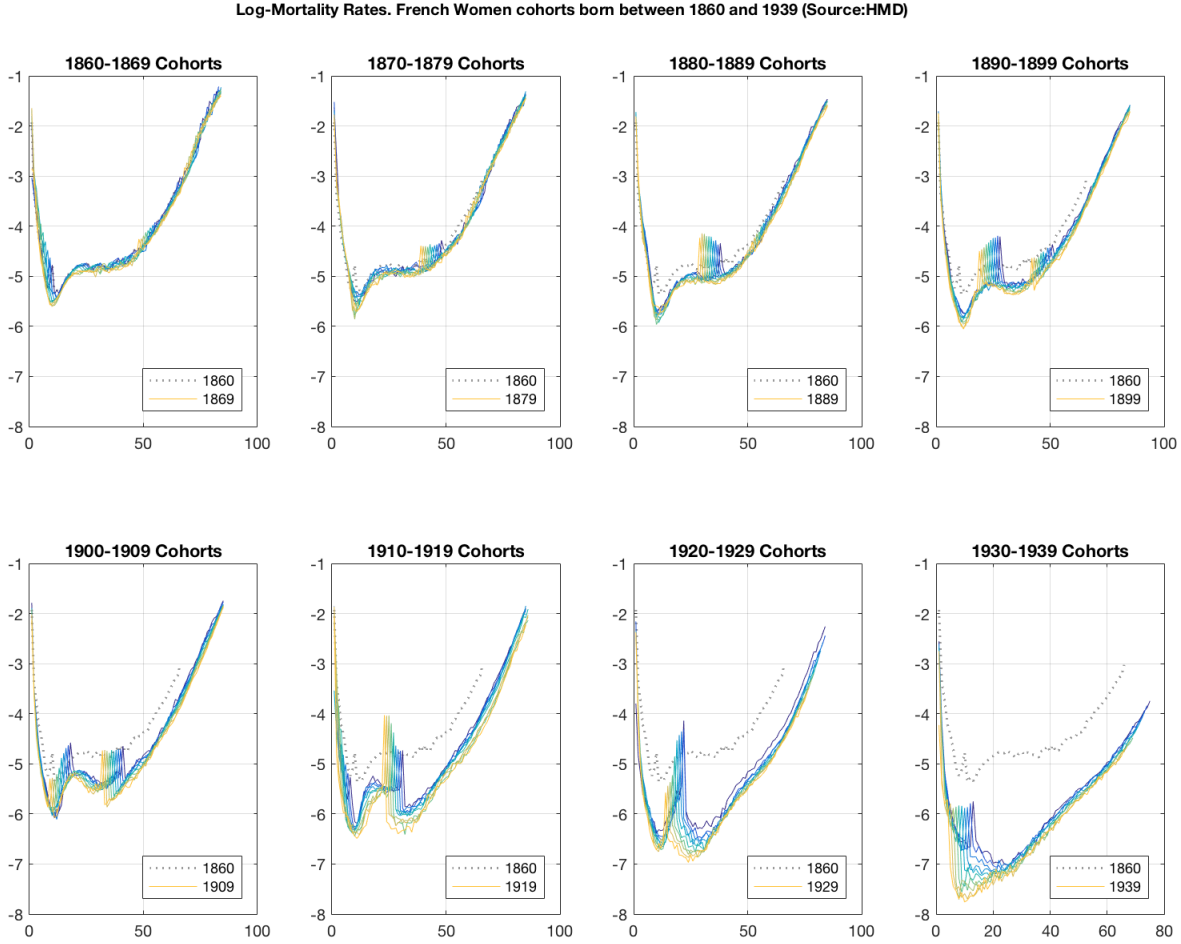


Figure 15: Age profile of mortality of men born in France between 1860 and 1940, by decade

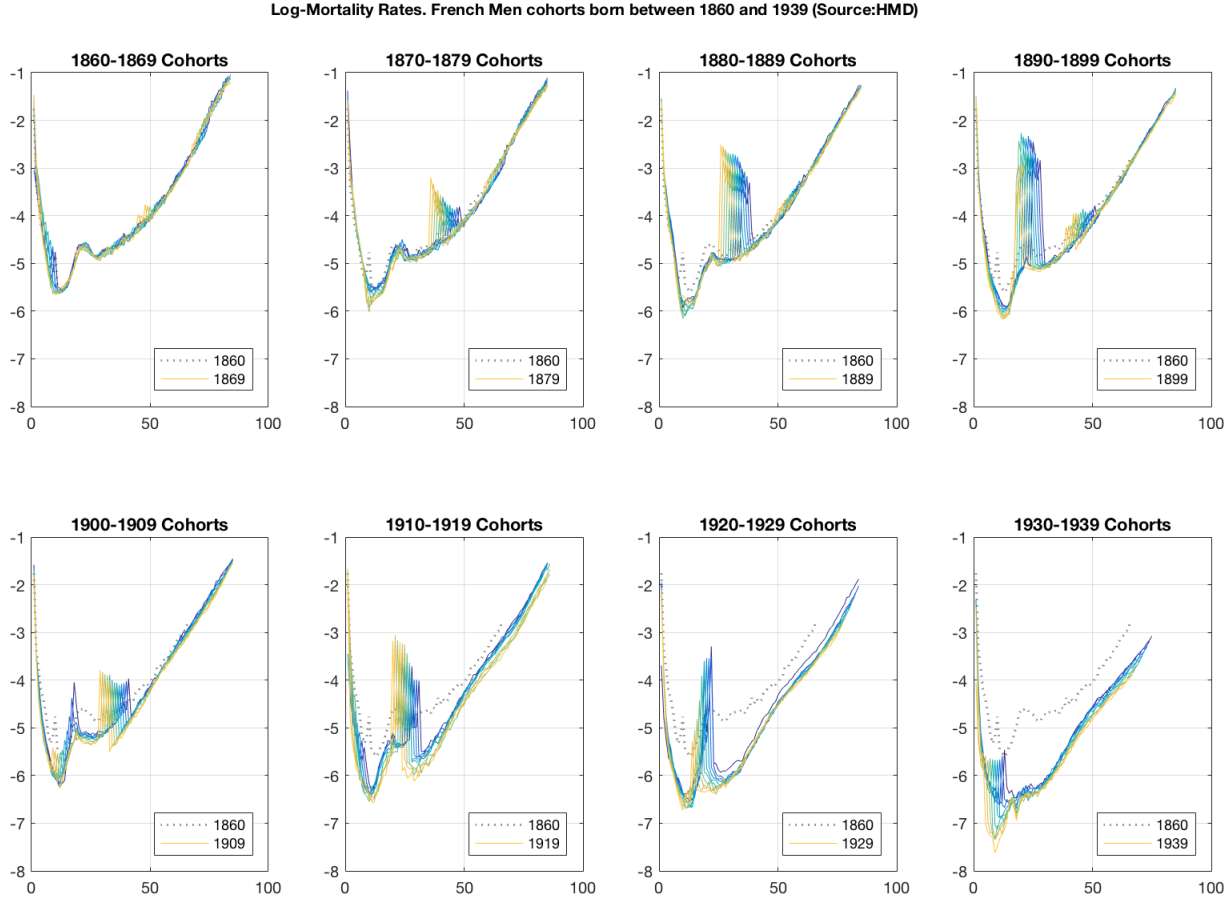
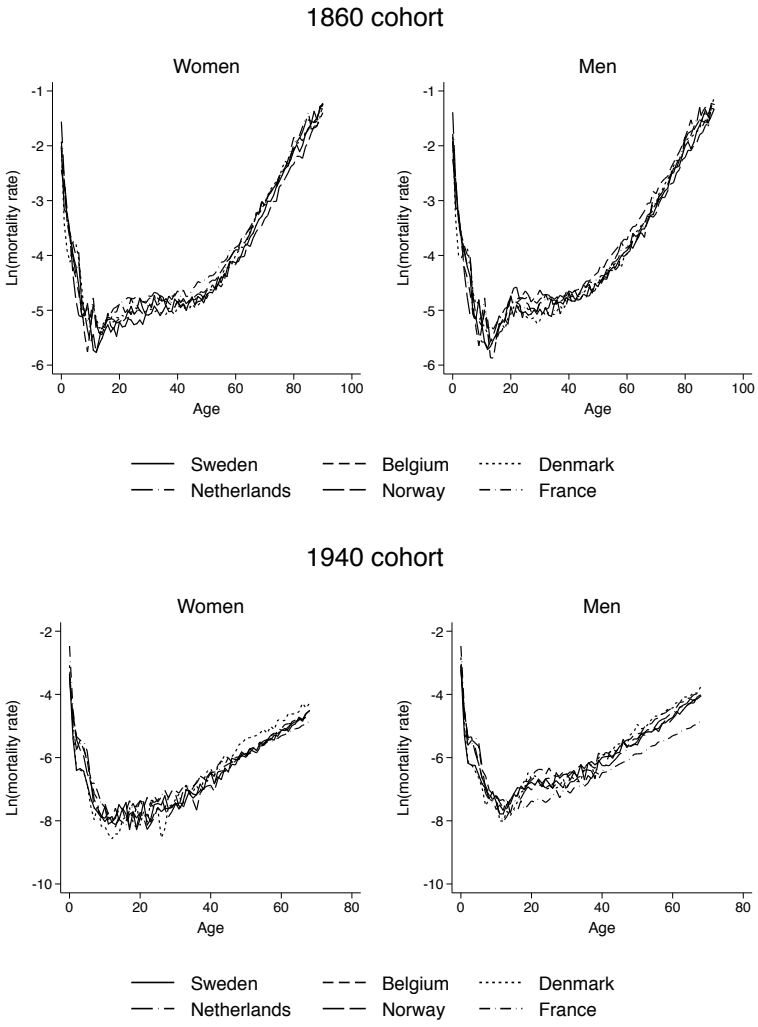
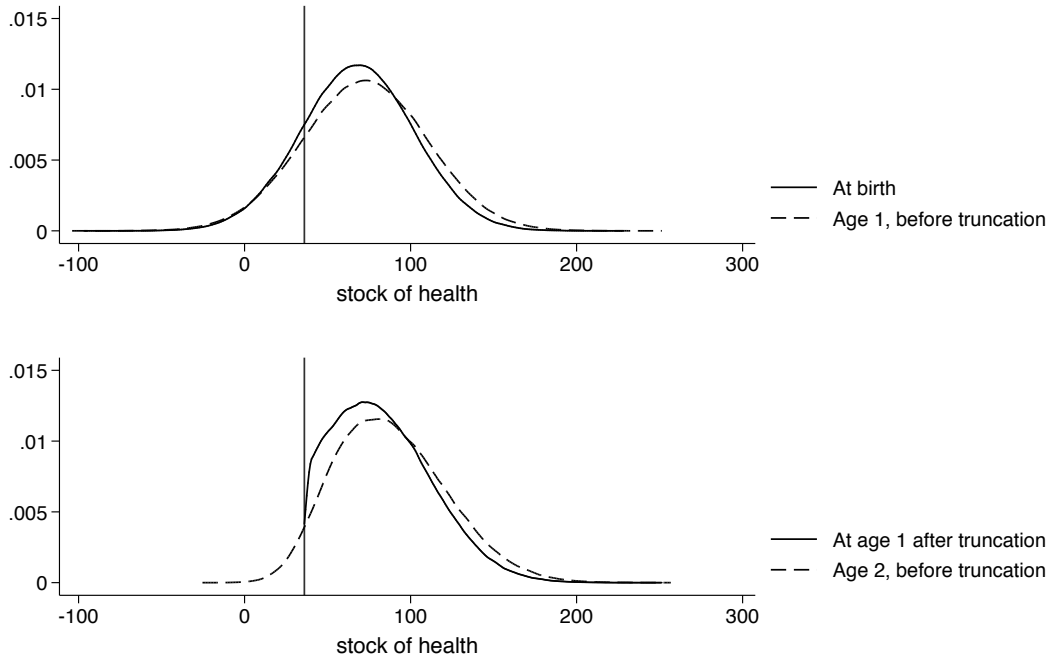


Figure 16: Mortality rates across Europe for 1860 and 1940 cohorts



Data: Human Mortality Database

Figure 17: Health and mortality in the first two years of life



Data from simulations

Figure 18: Trends in Mortality rates by age for France

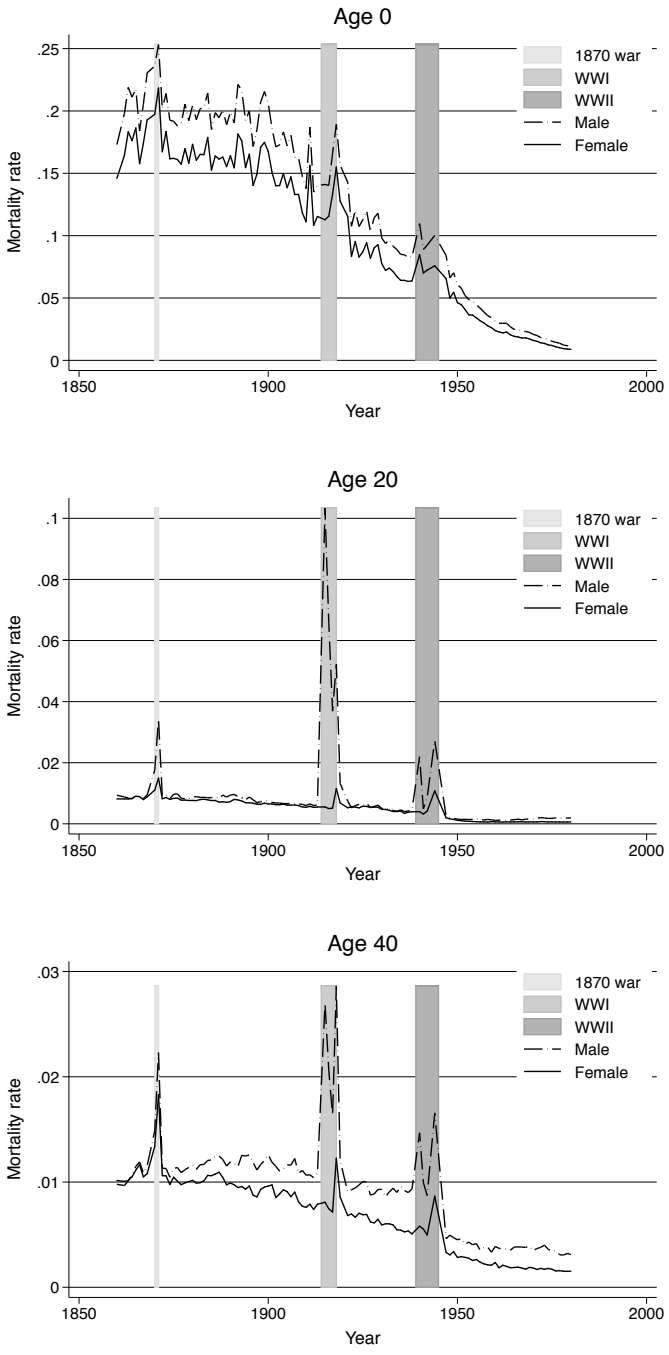
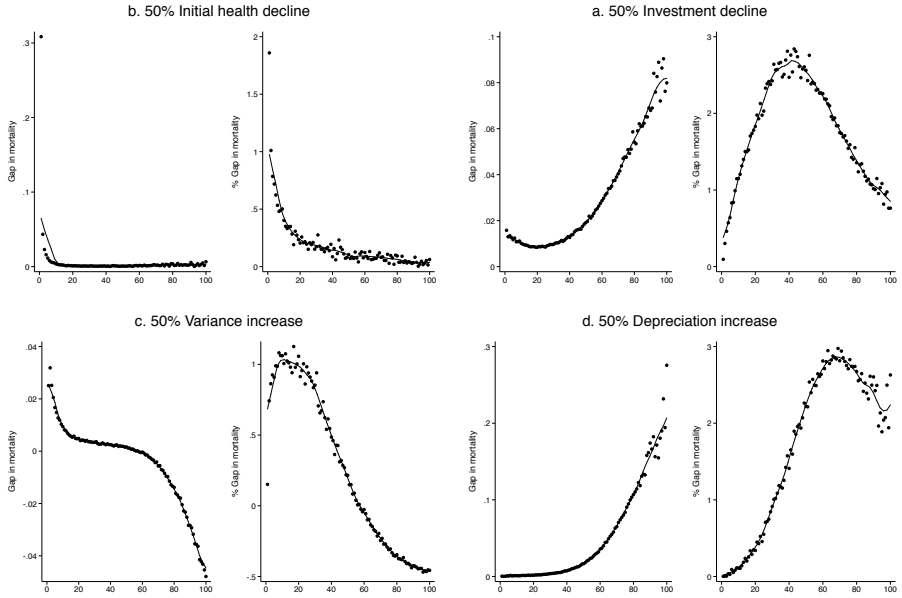


Figure 19: Comparative statics for mortality: level and percentage gaps



Note: Simulated data for two population of 500,000 individuals each. The figures show the effect of changes relative to the baseline model, which is simulated using the same parameters we used for Figure 3.

Figure 20: Evolution of GDP

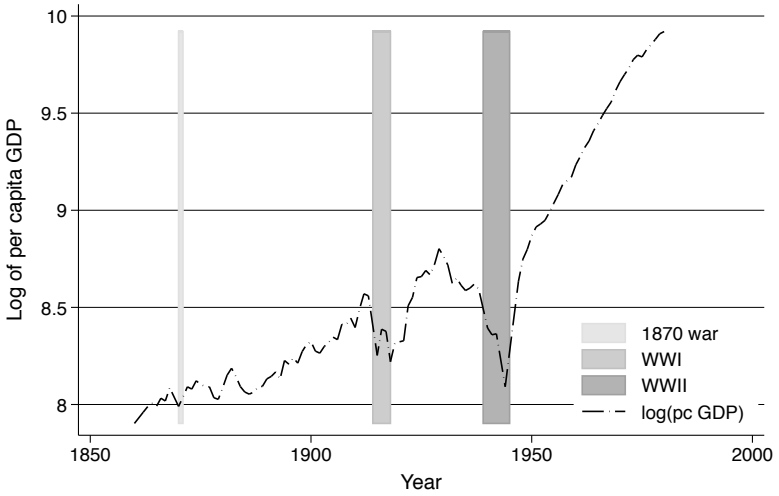
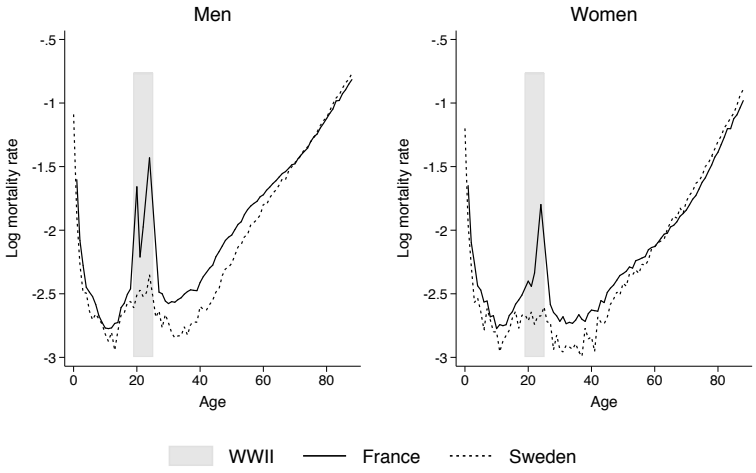


Figure 21: Effect WWII on mortality for 1920 cohort



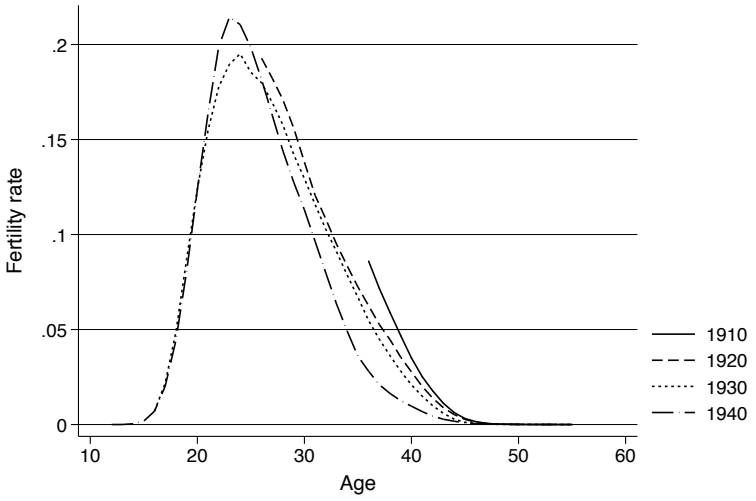
Data: Human Mortality Database

Figure 22: Effect of WWI and WWII on 1895 cohort



Data: Human Mortality Database

Figure 23: Age-Specific Fertility Rates, France 1910-1940 cohorts



Therefore $0 \leq P(H_t > z) \leq P(H_t^* > z)$ for any $z \in \mathbb{R}$.

Now for any $t \geq 2$ H_t^* is normally distributed with mean

$$\mu_{H_t^*} = I \cdot t - \delta \sum_{k=1}^{t} k^\alpha$$

and standard deviation

$$\sigma_{H_t^*} = \sqrt{\sigma_0^2 + t \cdot \sigma_\epsilon^2}$$

As $t \to \infty$, we have $\mu_{H_t^*} \sim I \cdot t - \delta \cdot \frac{t^{\alpha+1}}{\alpha+1}$ and $\sigma_{H_t^*} \sim \sqrt{t} \cdot \sigma_\epsilon$.

Therefore if $\alpha > 0$, $\frac{\mu_{H_t^*}}{\sigma_{H_t^*}} \to -\infty$ as $t \to \infty$.

Finally, $P(H_t^* > z) = 1 - \Phi(\frac{-\mu_{H_t^*}}{\sigma_{H_t^*}}) \to 0$, where Φ is the cdf of the standard normal distribution.

2) U shaped mortality rates

$MR_t = \frac{F_{H_t}(0) - F_{H_{t-1}}(0)}{1 - F_{H_{t-1}}(0)}$

first, as long as σ_ϵ small enough, then $MR_1 > MR_2$

then again we need δ, α to be small enough so that the aging is not too strong, otherwise MR increases at age 2

3) Average health has inverted Ushape

The average health of a cohort at age t is given by

$$\begin{aligned} \mathbb{E}[H_t | H_t > 0] &= \mathbb{E}[H_{t-1} + I - \delta t^\alpha + \varepsilon_t | H_t > 0] \\ &= I - \delta t^\alpha + \mathbb{E}[H_{t-1} | H_t > 0] \\ &= I - \delta t^\alpha + \mathbb{E}[\mathbb{E}[H_{t-1} | H_{t-1} > 0] | H_t > 0] \end{aligned}$$

by the law of iterated expectations.
Hence

$$\begin{array}{rcc} \mathbb{E}\left[H_t \left| H_t > 0\right.\right] - \mathbb{E}\left[H_{t-1} \left| H_{t-1} > 0\right.\right] & = & I - \delta t^{\alpha} \\ & + & \mathbb{E}\left[\mathbb{E}\left[H_{t-1} \left| H_{t-1} > 0\right.\right] \left| H_t > 0\right.\right] \\ & - & \mathbb{E}\left[H_{t-1} \left| H_{t-1} > 0\right.\right] \end{array}$$

The second term is always positive because if we know that the individual will survive the next period as well, because the shock are iid, we can infer that her health today is above average. Suppose we can bound

$$0 \leq \mathbb{E}\left[\mathbb{E}\left[H_{t-1} \left| H_{t-1} > 0\right.\right] \left| H_t > 0\right.\right] - \mathbb{E}\left[H_{t-1} \left| H_{t-1} > 0\right.\right] \leq c + o(t^{\alpha})$$

then the result will follow (except in the middle region for where $I - \delta t^{\alpha}$ is close to c . By definition,

$$\mathbb{E}\left[H_{t-1} \left| H_{t-1} > 0\right.\right] = \int_0^{\infty} x f_{H_{t-1}}(x)\, dx$$

Now,

$$\begin{array}{rcc} P\left(H_{t-1} \leq x \mid H_t > 0\right) & = & P\left(H_{t-1} \leq x \mid H_t > 0, H_{t-1} > 0\right) \\ & = & \frac{P(H_{t-1} \leq x, H_t > 0 | H_{t-1} > 0)}{P(H_t > 0 | H_{t-1} > 0)} \\ & = & \frac{1 - F_{H_{t-1}}(0)}{1 - F_{H_t}(0)} \int_0^x f_{H_{t-1}}(u) \left[1 - \Phi\left(\delta t^{\alpha} - I - u\right)\right] du \end{array}$$

hence

$$\mathbb{E}\left[H_{t-1} \left| H_t > 0\right.\right] = \int_0^{\infty} x f_{H_{t-1}}(x) \left\{ \frac{1 - F_{H_{t-1}}(0)}{1 - F_{H_t}(0)} \left[1 - \Phi\left(\delta t^{\alpha} - I - x\right)\right] \right\} dx$$

and finally

$$\begin{array}{c} \mathbb{E}\left[\mathbb{E}\left[H_{t-1} \left| H_{t-1} > 0\right.\right] \left| H_t > 0\right.\right] - \mathbb{E}\left[H_{t-1} \left| H_{t-1} > 0\right.\right] \\ = \int_0^{\infty} x f_{H_{t-1}}(x) \left\{ \frac{1 - F_{H_{t-1}}(0)}{1 - F_{H_t}(0)} \left[1 - \Phi\left(\delta t^{\alpha} - I - x\right)\right] - 1 \right\} dx \end{array}$$

Fix $\epsilon > 0$, for any t there is an $x(t)$ such that $x > x(t) \Rightarrow \Phi\left(\delta t^{\alpha} - I - x\right) < \epsilon$
for large values of x the term in square bracket goes to 0 very fast, so focus on small values of x.
4) Variance
probably more tricky, but in the same vein
$Var\left[H_t \left| H_t > 0\right.\right]$
5) close to normal?
compare Ht and Ht*, "extra" terms, those who crossed threshold once, twice

6) Log linearity

To prove that the mortality rate is log linear is equivalent to prove that

$\ln MR_t - \ln MR_{t-1} = constant$

Now since $MR_t = \frac{F_{H_t}(0) - F_{H_{t-1}}(0)}{1 - F_{H_{t-1}}(0)}$ this is equivalent to showing that for $t \in [40, 80]$

$$\frac{F_{H_{t+1}}(0) - F_{H_t}(0)}{F_{H_t}(0) - F_{H_{t-1}}(0)} = \frac{1 - F_{H_{t-1}}(0)}{1 - F_{H_t}(0)} + o(t)$$

in continuous time

Proof for Proposition 2 comparative statics Let $a_t = I - \delta t^{\alpha}$. The random variable H_t has a mass point at $z = 0$ but is continuous on $(0, +\infty)$. $F_{H_t}(0)$ is the probability of not surviving until age t while for any $z > 0$, the cdf can be expressed

$$F_{H_t}(z) = \int_{x=0}^{\infty} \Phi\left(\frac{z - x - a_t}{\sigma_\epsilon}\right) f_{H_{t-1}}(x)\, dx + F_{H_{t-1}}(0)$$

Equivalently, after integration by parts, one obtains:

$$F_{H_t}(z) = -\frac{1}{\sigma_\epsilon} \int_{x=0}^{\infty} \phi\left(\frac{z - x - a_t}{\sigma_\epsilon}\right) F_{H_{t-1}}(x)\, dx + F_{H_{t-1}}(0)$$

Hence the mortality rate at age t, which is the probability of dying at age t conditional on surviving until age t, can be written:

$$MR_t = \frac{F_{H_t}(0) - F_{H_{t-1}}(0)}{1 - F_{H_{t-1}}(0)}$$

Suppose that for every t we increase the constant investment level I to some level $I' > I$. Following the expression above, the impact can be be decomposed in two: first, a direct effect on the probability of dying at age t (the numerator) and, second, a compounded effect carried through the distribution of health for those attaining age t.We show that, for any t, both effects go in the same direction: an increase in I simultaneously increases the probability of surviving untill age t (hence increases the denominator) and reduces the probability of dying at age t (the numerator goes down). We prove the following lemma.

Lemma 1 For all t, we have i) $\forall z > 0$, $\frac{\partial F_{H_t}(z; I)}{\partial I} \leq 0$ and ii) $\frac{\partial MR_t}{\partial I} \leq 0$

We prove these inequalities jointly and by induction.

Notice that $\frac{\partial F_{H_t}(\cdot; I)}{\partial I} \leq 0$ signifies that the cdf's are ranked by first order stochastic dominance. The higher the I, the further the distribution is pushed to the right, which decreases the value of the cdf at any point x as I increases.

At $t = 0$: $F_{H_1}(z; I) = \Phi\left(\frac{z - \mu_0}{\sigma_0}\right)$ hence $\frac{\partial F_{H_1}(z; I)}{\partial I} = 0$. $MR_t = F_{H_t}(0) = \Phi\left(\frac{z - \mu_0}{\sigma_0}\right)$ which, again, is non-increasing with I.

For any $t \geq 1$, suppose that $\frac{\partial F_{H_{t-1}}}{\partial I} \leq 0$ and $\frac{\partial MR_{t-1}}{\partial I} \leq 0$.

Let's first focus on the first claim:

$$\frac{\partial F_{H_t}(z;\, I)}{\partial I} = \frac{\partial}{\partial I}\left[-\frac{1}{\sigma_\epsilon}\int_{x=0}^{\infty}\phi\left(\frac{z-x-a_t}{\sigma_\epsilon}\right)F_{H_{t-1}}(x)\,dx\right] + \frac{\partial F_{H_{t-1}}(0;\, I)}{\partial I}$$

The second term is negative, by assumption, while the first term is equal to

$$\frac{1}{\sigma_\epsilon^2}\int_{x=0}^{\infty}\phi'\left(\frac{z-x-a_t}{\sigma_\epsilon}\right)F_{H_{t-1}}(x)\,dx - \frac{1}{\sigma_\epsilon}\int_{x=0}^{\infty}\phi\left(\frac{z-x-a_t}{\sigma_\epsilon}\right)\frac{\partial F_{H_{t-1}}(z;\, I)}{\partial I}dx$$

Again, by assumption, $\frac{\partial F_{H_{t-1}}(z;\, I)}{\partial I} \leq 0$, which takes care of the rightmost term.

Now, consider the change of variable $u = \frac{z-x-a_t}{\sigma_\epsilon}$. We can rewrite the leftmost term:

$$-\frac{1}{\sigma_\epsilon^3}\int_{u=-\infty}^{\frac{z-a_t}{\sigma_\epsilon}}\phi'(u)\,F_{H_{t-1}}(z-a_t-\sigma_\epsilon u)\,du$$

There are two cases. If $\frac{z-a_t}{\sigma_\epsilon} \leq 0$ then the integrand is always positive as $\phi' > 0$ for negative real numbers, and we conclude that $\frac{\partial F_{H_t}(z;\, I)}{\partial I} \leq 0$. If $\frac{z-a_t}{\sigma_\epsilon} > 0$ then we can split the integral in three terms:

$$\begin{aligned}
& -\tfrac{1}{\sigma_\epsilon^3}\textstyle\int_{u=-\infty}^{-\frac{z-a_t}{\sigma_\epsilon}}\phi'(u)\,F_{H_{t-1}}(z-a_t-\sigma_\epsilon u)\,du \\
& -\tfrac{1}{\sigma_\epsilon^3}\textstyle\int_{u=-\frac{z-a_t}{\sigma_\epsilon}}^{0}\phi'(u)\,F_{H_{t-1}}(z-a_t-\sigma_\epsilon u)\,du \\
& -\tfrac{1}{\sigma_\epsilon^3}\textstyle\int_{u=0}^{\frac{z-a_t}{\sigma_\epsilon}}\phi'(u)\,F_{H_{t-1}}(z-a_t-\sigma_\epsilon u)\,du \\
= \quad & -\tfrac{1}{\sigma_\epsilon^3}\int_{u=-\infty}^{-\frac{z-a_t}{\sigma_\epsilon}}\phi'(u)\,F_{H_{t-1}}(z-a_t-\sigma_\epsilon u)\,du \\
- \quad & \tfrac{1}{\sigma_\epsilon^3}\int_{u=-\frac{z-a_t}{\sigma_\epsilon}}^{0}\phi'(u)\left[F_{H_{t-1}}(z-a_t-\sigma_\epsilon u) - F_{H_{t-1}}(z-a_t+\sigma_\epsilon u)\right]du
\end{aligned}$$

(as $\phi'(-u) = -\phi'(u)$ and the cdf $F_{H_{t-1}}$ is non-decreasing). This proves that $\frac{\partial F_{H_t}(z;\, I)}{\partial I} \leq 0$ for any $z \in \mathbb{R}$.

With that result in hand, it is easy to prove that $\frac{\partial MR_t}{\partial I} \leq 0$. Setting $z = 0$, it follows directly that the denominator decreases with I. Regarding the numerator we have

$$\begin{aligned}
\tfrac{\partial}{\partial I}\left[F_{H_t}(0;\, I) - F_{H_{t-1}}(0;\, I)\right] &= \tfrac{\partial F_{H_t}(0;\, I)}{\partial I} - \tfrac{\partial F_{H_{t-1}}(0;\, I)}{\partial I} \\
&= \tfrac{\partial}{\partial I}\left[-\tfrac{1}{\sigma_\epsilon}\int_{x=0}^{\infty}\phi\left(\tfrac{x-a_t}{\sigma_\epsilon}\right)F_{H_{t-1}}(x)\,dx\right] \\
&\leq 0
\end{aligned}$$

since this is the same integral analyzed at the previous step, with $z = 0$.

By induction, i) and ii) hold for any $t \geq 1$.

1-3 The exact same proof applies for δ and αas their impact on F_{H_t} through the aging function a_t is similar to the effect of I.

1-2 for μ_H same proof, except effect on first period distribution. The successive cdf inherit the first order stochastic dominance property

4 It can be seen right away that this proof will not work for σ_ϵ nor σ_0. Increasing any of these variances, - a mean-preserving spread - will not give rise to the first order stochastic ranking of the cdf's that we have used

5. Increasing MR on impact. The numerator of the MR_t is given by

$$F_{H_t}(0;\, I) - F_{H_{t-1}}(0;\, I) = \int_{x=0}^{\infty} \Phi\left(\frac{\delta t^\alpha - I - x}{\sigma_\epsilon}\right) f_{H_{t-1}}(x)\, dx$$

Since Φ is nondecreasing, if one decreases σ_ϵ at time t, and at this period only, then this expression is necessarily decreasing in σ_ϵ if $\delta t^\alpha \leq I$.

6. This follows from the fact that a higher $\sigma_{\epsilon,t}$ will generate a fatter right-hand tail. For instance, $\lim_{x\to+\infty} \frac{f_{H_t}(x;\sigma_\epsilon)}{f_{H_t}(x;\sigma'_\epsilon)} = 0$. Now if $\sigma_{\epsilon,t}$ is changed only at period . From then on, the distributions are modified through a similar process. It can be shown that the fatter right-hand tail property will be preserved. In the very old age, only the popupulation in the right-tail have survived, hence the result.

to prove that the fatter right-hand tail property is preserved. by inference.

Remark 1 Lemma 1 is actually a subcase of the following result, which is slightly stronger:

Suppose that the level of investment is allowed to change at every period, and denote $\mathcal{I} = \{I_1, I_2, ..., \}$ and $\mathcal{I}' = \{I'_1, I'_2, ..., \}$two investment sequences. The following holds:

$$\forall s \geq 1,\, I'_s \geq I_s \Longrightarrow \forall t \geq 1, \forall z > 0, F_{H_t}(z;\mathcal{I}) \leq F_{H_t}(z;\mathcal{I}') \text{ and } MR_t(\mathcal{I}) \leq MR_t(\mathcal{I}')$$

The mechanics of the proof is almost exactly similar. Increasing investment at any period generates a persistent relation of first-order stochastically dominance in the cdf of health.

Lemma 2 Suppose $\sigma'_\epsilon > \sigma_\epsilon$ then $F_{H_t}(.;\, \sigma_\epsilon)$ (strictly) second-order stochastically dominates $F_{H_t}(.;\, \sigma'_\epsilon)$ for every t.

Proof, by induction.

For $t = 1$, $F_{H_t}(.;\, \sigma_\epsilon) = F_{H_t}(.;\, \sigma'_\epsilon)$, hence $\int_{s=-\infty}^{z} [F_{H_t}(s;\, \sigma'_\epsilon) - F_{H_t}(s;\, \sigma_\epsilon)]\, ds \geq 0$. (not strict at this first period but strict starting from the next period)

For $t = 2$,

$$\begin{array}{rc} & \int_{s=-\infty}^{z} \left[F_{H_2}\left(s;\,\sigma_\epsilon'\right) - F_{H_2}\left(s;\,\sigma_\epsilon\right)\right] ds \\ = & \int_{s=-\infty}^{z}\int_{x=0}^{\infty} \left[\frac{1}{\sigma_\epsilon'}\phi\left(\frac{s-x-a_t}{\sigma_\epsilon'}\right) F_{H_1}\left(x;\sigma_\epsilon'\right) - \frac{1}{\sigma_\epsilon}\phi\left(\frac{s-x-a_t}{\sigma_\epsilon}\right) F_{H_1}\left(x;\sigma_\epsilon\right)\right] dxds \\ = & \int_{s=-\infty}^{z}\int_{x=0}^{\infty} \left[\frac{1}{\sigma_\epsilon'}\phi\left(\frac{s-x-a_t}{\sigma_\epsilon'}\right) \Phi\left(\frac{x-\mu_H}{\sigma_H}\right) - \frac{1}{\sigma_\epsilon}\phi\left(\frac{s-x-a_t}{\sigma_\epsilon}\right) \Phi\left(\frac{x-\mu_H}{\sigma_H}\right)\right] dxds \\ & \int_{x=0}^{\infty} \Phi\left(\frac{x-\mu_H}{\sigma_H}\right) \left[\int_{s=-\infty}^{z} \frac{1}{\sigma_\epsilon'}\phi\left(\frac{s-x-a_t}{\sigma_\epsilon'}\right) ds - \int_{s=-\infty}^{z} \frac{1}{\sigma_\epsilon}\phi\left(\frac{s-x-a_t}{\sigma_\epsilon}\right) ds\right] dx \\ & \int_{x=0}^{\infty} \Phi\left(\frac{x-\mu_H}{\sigma_H}\right) \left[\Phi\left(\frac{z-x-a_t}{\sigma_\epsilon'}\right) - \Phi\left(\frac{z-x-a_t}{\sigma_\epsilon}\right)\right] dx \end{array}$$

.

For $t \geq 2$, since $F_{H_t}\left(x, \sigma_\epsilon\right) = 0$ for all $x < 0$, second order dominance holds trivially for any $z < 0$. So let $z \geq 0$ and suppose that $F_{H_{t-1}}\left(\cdot\,;\,\sigma_\epsilon\right) \succ^2 F_{H_{t-1}}\left(\cdot;\,\sigma_\epsilon'\right)$. We have

$$\begin{array}{rc} & \int_{s=-\infty}^{z} \left[F_{H_t}\left(s;\,\sigma_\epsilon'\right) - F_{H_t}\left(s;\,\sigma_\epsilon\right)\right] ds \\ = & \int_{s=-\infty}^{z}\int_{x=0}^{\infty} \left[\frac{1}{\sigma_\epsilon'}\phi\left(\frac{s-x-a_t}{\sigma_\epsilon'}\right) F_{H_{t-1}}\left(x;\sigma_\epsilon'\right) - \frac{1}{\sigma_\epsilon}\phi\left(\frac{s-x-a_t}{\sigma_\epsilon}\right) F_{H_{t-1}}\left(x;\sigma_\epsilon\right)\right] dxds \\ = & \int_{x=0}^{\infty}\int_{s=-\infty}^{z} \left[\frac{1}{\sigma_\epsilon'}\phi\left(\frac{s-x-a_t}{\sigma_\epsilon'}\right) F_{H_{t-1}}\left(x;\sigma_\epsilon'\right) - \frac{1}{\sigma_\epsilon}\phi\left(\frac{s-x-a_t}{\sigma_\epsilon}\right) F_{H_{t-1}}\left(x;\sigma_\epsilon\right)\right] dsdx \\ = & \int_{x=0}^{\infty} \left[F_{H_{t-1}}\left(x;\sigma_\epsilon'\right) \int_{s=-\infty}^{z} \frac{1}{\sigma_\epsilon'}\phi\left(\frac{s-x-a_t}{\sigma_\epsilon'}\right) ds - F_{H_{t-1}}\left(x;\sigma_\epsilon\right) \int_{s=-\infty}^{z} \frac{1}{\sigma_\epsilon}\phi\left(\frac{s-x-a_t}{\sigma_\epsilon}\right) ds\right] dx \\ = & \int_{x=0}^{\infty} \left[F_{H_{t-1}}\left(x;\sigma_\epsilon'\right)\Phi\left(\frac{z-x-a_t}{\sigma_\epsilon'}\right) - F_{H_{t-1}}\left(x;\sigma_\epsilon\right) \Phi\left(\frac{z-x-a_t}{\sigma_\epsilon}\right)\right] dx \\ \geq & \int_{x=0}^{\infty} \left[F_{H_{t-1}}\left(x;\sigma_\epsilon'\right) - F_{H_{t-1}}\left(x;\sigma_\epsilon\right)\right] \Phi\left(\frac{z-x-a_t}{\sigma_\epsilon'}\right) dx \end{array}$$

(justifications: 2, expression of F_{H_t} used previously, 3: Fubini theorem holds as integrand is absolutely integrable (because of ϕ), 4, F does not depend on s , 5 see below). Since $\sigma_\epsilon' > \sigma_\epsilon$

$$\Phi\left(\frac{z-x-a_t}{\sigma_\epsilon'}\right) - \Phi\left(\frac{z-x-a_t}{\sigma_\epsilon}\right) \begin{cases} \leq 0 & \text{if } z - a_t \geq x \\ > 0 & \text{if } z - a_t < x \end{cases}$$

Now, for any $\eta > 0$, there exist $x\left(\eta\right)$ such that $\Phi\left(\frac{z-x-a_t}{\sigma_\epsilon'}\right)$

Proposition 3 Morbidity The morbidity rates can be expressed in terms of the cdf:

$$\begin{array}{rcl} MbR_t & = & \frac{F_{H_t}(Hmb) - F_{H_t}(0)}{1 - F_{H_t}(0)} \\ & = & 1 - \frac{1 - F_{H_t}(Hmb)}{1 - F_{H_t}(0)} \\ & = & 1 + \frac{F_{H_t}(Hmb) - 1}{1 - F_{H_t}(0)} \end{array}$$

We know from Proposition 2 that increasing Iwill lead to an decrease in $F_{H_t}\left(0\right)$ resulting in an increase in denominator. Using FOSD again $F_{H_t}\left(Hmb\right)$ decreases with I .

Table 1:

Parameter		French Women 1945	French Women 1920	Apes
Investment	I	0.0190		0.7671
Depreciation	δ	0.00139		0.0697
Shock	σ_e	0.0662		0.9572
Aging	α	0.9294		0.8936
Initial condition	μ_0	2.8		0.7818
Accident Shock	κ	3.3 10^-4		

Proposition 4 Dynamic complementarity going back to the proof of Proposition 1

$$\begin{aligned}
\frac{\partial^2 F_{H_{t_2}}(z;\mathcal{I})}{\partial I_{t_1}\partial I_{t_2}} &= \frac{\partial}{\partial I_{t_1}}\frac{\partial}{\partial I_{t_2}}\left[-\frac{1}{\sigma_\epsilon}\int_{x=0}^{\infty}\phi\left(\frac{z-x-I_{t_2}+\delta(t_2)^\alpha}{\sigma_\epsilon}\right)F_{H_{t_2-1}}(x,\mathcal{I})\,dx\right] \\
&+ \qquad \frac{\partial}{\partial I_{t_1}}\frac{\partial F_{H_{t_2-1}}(0;\mathcal{I})}{\partial I_2} \\
&= \frac{\partial}{\partial I_{t_1}}\left[\frac{1}{\sigma_\epsilon^2}\int_{x=0}^{\infty}\phi'\left(\frac{z-x-a_t}{\sigma_\epsilon}\right)F_{H_{t_2-1}}(x,\mathcal{I})\,dx\right]+0 \\
&= \frac{1}{\sigma_\epsilon^2}\int_{x=0}^{\infty}\phi'\left(\frac{z-x-a_t}{\sigma_\epsilon}\right)\frac{\partial}{\partial I_{t_1}}F_{H_{t_2-1}}(x,\mathcal{I})\,dx \\
&\leq 0
\end{aligned}$$

because $\frac{\partial}{\partial I_{t_1}}F_{H_{t_2-1}}(x,\mathcal{I}) \leq 0$ (increasing investment at time 1 creates a FOSD distribution)

And as a consequence the denominator $1-F_{H_{t_2-1}}(0)$ goes up as well.

Proposition 4 mu_0 and I are complement Exactly the same as in Proposition 2 as a change in μ_0 is observationally equivalent to a change in I_1.

Appendix A3: Tables