Long-Run Economic Perspectives of an Ageing Society

Research Report 34

Editors:

Holger Strulik

Gottfried Wilhelm Leibniz Universität Hannover

Carl-Johan Dalgaard

Department of Economics, University of Copenhagen

Fidel Perez-Sebastian

Department of Economics, University of Alicante

Alexia Prskawetz

Wittgenstein Centre for Demography and Global Human Capital (IIASA, VID/ÖAW, WU), Vienna Institute of Demography/Austrian Academy of Sciences and Institute of Mathematical Methods in Economics, Vienna University of Technology (TU)

March 2012

INSTITUT FÜR DEMOGRAPHIE

ÖSTERREICHISCHE AKADEMIE DER WISSENSCHAFTEN

Forschungsbericht Nr. 34

Forschungsprojekt:
Long - Run Economic Perspectives of an Ageing Society, LEPAS. Grant Agreement: SSH-2007-3.1.01- 217275. Theme: Socio-economic sciences and humanities: The impact of European demographic changes in Europe.

Projektleitung am Institut für Demographie der Österreichischen Akademie der Wissenschaften:
Prof. Dr. Alexia Prskawetz

Forschungsfördernde Institution:
European Commission, Seventh Framework Programme

Die Arbeit unterliegt ausschließlich der Verantwortung des Instituts für Demographie und wurde der phil.-hist. Klasse nicht vorgelegt.

Medieninhaber und Herausgeber:
Institut für Demographie
Vienna Institute of Demography
Österreichische Akademie der Wissenschaften
A-1040 Wien, Wohllebengasse 12-14
Tel.: +43-1/51581/7702
Fax: +43-1/51581/7730
E-mail: vid@oeaw.ac.at
www.oeaw.ac.at/vid

Vertrieb:
Verlag der Österreichischen Akademie der Wissenschaften
A-1011 Wien, Postgasse 7, Postfach 471
Tel.: +43-1/51581/3401-3406
Fax: +43-1/51581/400
e-mail: verlag@oeaw.ac.at
homepage: http://www.verlag.oeaw.ac.at/

ISBN 978-3-7001-7302-1

Long-Run Economic Perspectives of an Ageing Society

Holger Strulik Carl-Johan Dalgaard

Fidel Perez-Sebastian Alexia Prskawetz

Preface

Nearly 25 percent of people in the European Union in 2030 will be above age 65, up from about 17 percent in 2005. Europe's old-age dependency ratio (the number of people age 65 and older compared with the number of working-age people ages 15-64) could more than double by 2050, from one in every four to fewer than one in every two. This is an unprecedented phenomenon with potentially very important implications on society that are far from being well understood.

Researchers with public finance background frequently deliver gloom predictions of the impact of ageing. For example, Feldstein (2006), among others, argues that increasing longevity and diminishing birth rates will slow down economic growth in European economies by decreasing the growth of capital and by weakening the productivity of the labor force. Investigations using the EU Commission's own macroeconomic model QUEST II estimate that ageing causes the EU's average potential income growth rate to drop from 2.1 percent annually to 1.3 percent. The growth rate of income in per capita terms is predicted to decrease by 0.4 percentage points (Pichelmann and Roeger, 2004).

However, as stated in the European Commission's Communication (2006) Europe's demographic future provides not only challenges but also opportunities. In particular scholars of economic growth frequently emphasize the positive effects on long-run economic performance that are provoked by increasing human longevity (e.g. Ehrlich and Lui, 1991, Barro and Sala-i-Martin, 2004). The seeming contradiction is resolved when one recognizes that the public finance view emphasizes the social burden aspect that originates from an increasing dependency ratio whereas growth economists emphasize the positive dynamic feedback effects that individual ageing brings about. For example, they argue that the prospect of increasing longevity (i) increases the incentive to save for old age and thus creates the incentive for higher investment and capital accumulation, which in turn enhances growth (see, for example, Kalemli-Ozcan et al.,

2000, Chakraborty, 2004, Zhang and Zhang, 2005) and (ii) increases the incentive for education because individuals expect a longer working age during which human capital accumulation pays off. (Ehrlich and Lui, 1991, Boucekkine et al., 2002, Zhang et al., 2003, Blackburn and Cipriani, 2002).

A serious problem with the available economic studies on ageing and longevity, however, is that they are built upon an overly simplistic conception of the individual ageing process. Most of the economic literature focusses exclusively on the mortality aspect of ageing which is either modeled as an exogenously given probability to die before the next stage of life is reached or as an exogenously given expected time of death. The life sciences, in contrast, understand the process of ageing very differently as senescence that is as the gradual deterioration of body and mind. The missing notion of ageing as senescence in dynamic macroeconomics constituted the idea and starting point of the LEPAS project.

We believe that in order to fully understand the effect of ageing on the economy it is necessary to develop an economic theory of ageing that takes into account the endogenous evolution of human frailty and disability. Only if we understand both the biological and economic forces behind the changes in the functional status of human beings during their lives will we be able to analyze the economic determinants of successful ageing. Moreover, a solid modeling of ageing is required to study its feedback effects on economic growth, innovation, health expenditure, and retirement.

The LEPAS project thus set out three years ago to survey the natural science literature of ageing with a special focus of what can be learned and adopted for economic modeling. The result of this endeavor is presented in Part I of the present study. The LEPAS partners then started to implement the notion of physiological ageing in economic life-cycle models. This ambitious project was subdivided into work packages with distinct special emphases: health and income (WP 4), ageing and retirement (WP 5), ageing and immigration (WP 6), ageing, health and education (WP 7) ageing, health

and public health care (WP 8). Part II of the present study provides an overview of our results with a special focus on European countries and European policy.

The LEPAS partners organized two international workshops at which they discussed their ideas with distinguished scholars from the field of economics as well as from the natural sciences. In particular our project benefitted from comments and suggestions from James Carey, Arnold Mitnitski, and Leonid Gavrilov and Natalia Gavrilova. Papers and proceedings of these workshops (situated in Vienna, 2010, and in Alicante, 2011) are available at the project website: `http://www.lepas-fp7.de`. The website contains also detailed background papers for the chapters of Part II of this study as well as short summaries (policy briefs) and discussion papers of interesting other – partly still ongoing – research of the LEPAS project that did not find its way in the present book.

We, as the local project leaders, want to thank the members of the locals teams, Katharina Werner, Sebastian Vollmer, and Timo Trimborn from Leibniz University Hannover (LUH), Klaus Prettner, Michael Kuhn, and Stefan Wrzaczek from Vienna Institute of Demography (VID), Pablo Selaya from University of Copenhagen (UCPH) and Maria D. Guilló and Alex Pérez-Laborda from the University of Alicante (UA) for three years of dedicated research assistance. Our work has also benefitted from comments and discussion by David Canning, Karl Dietrich, Gustav Feichtinger, Oded Galor, Michael Grimm, Franz Hof, Chad Jones, Soeren Leth-Petersen, Wolfgang Lutz, Patrick Puhani, and Engelbert Theurl.

This research was funded by the European Commission within the project "Long-Run Economic Perspectives of an ageing Society" (LEPAS) in the Seventh Framework Programme under the Socio-economic Sciences and Humanities theme (Grant Agreement: SSH7-CT-2009-217275)

Hannover, March 2012.

References

Barro, R.J., and Sala-i-Martin, X., 2004, *Economic Growth*, 2nd ed., MIT Press, Cambridge, MA.

Blackburn, K. and Cipriani, G.P., 2002, A model of longevity, fertility, and growth, *Journal of Economic Dynamics and Control* 26, 187-204.

Boucekkine, R., de la Croix, D. and Licandro, O., 2002, Vintage human capital, demographic trends, and endogenous growth, *Journal of Economic Theory* 104, 340-375.

Chakraborty, S., 2004, Endogenous lifetime and economic growth, *Journal of Economic Theory*, 116, 119-137.

Commission of the European Communities, 2006, Commission Communication: The demographic future of Europe - from challenge to opportunity, Brussels.

Ehrlich, I. and Lui, F.T., 1991, Intergenerational trade, longevity, and economic growth, *Journal of Political Economy* 99, 1029-59.

Feldstein, M.S., 2006, The effects of the Ageing European Population on Economic Growth and Budgets: Implications for Immigration and Other Policies," NBER Working Paper 12736.

Kalemli-Ozcan, S., Ryder, H.E., and Weil, D.N., 2000, Mortality decline, human capital investment and economic growth, *Journal of Economic Development* 62, 1-23.

Pichelmann, K. and Roeger, W., 2004, The EU Growth Strategy and the Impact of Ageing, Review of International Economics 12, 213-232.

Zhang, J. and Zhang. J., 2005, The effect of life expectancy on fertility, saving, schooling and economic growth: theory and evidence, *Scandinavian Journal of Economics* 107, 45-66.

Zhang, J., Zhang, J., Lee, R., 2003, Rising longevity, education, savings, and growth, *Journal of Development Economics* 70, 103-117.

Contents

List of Figures

Part I.

The Mechanics of Ageing and Death: A Primer for Economists

1. Measuring Ageing

by Holger Strulik

1.1. The Force of Mortality

Ageing is defined as the intrinsic, cumulative, progressive, and deleterious loss of function that eventually culminates in death (Arking, 2006, Masoro, 2006). Here we focus on ageing of organisms (mostly humans) but everybody who is driving a car will recognize that the definition of ageing applies to non-living matter as well. Actually, living and inanimate systems are ageing in a similar way, a phenomenon to which we return later. Sometimes ageing and senescence are used synonymously (Masoro, 2006), sometimes senescence is reserved for the period of obvious functional decline in the later years of an animal's life-span (Arking, 2006).

Ageing can be measured by directly accounting for the accumulated deficits, a concept to which we return later, or by its consequences for mortality, which we discuss next. Demographers employ several indices of mortality. For example, the probability to survive to age x, the life-expectancy at age x, or the probability of dying in the age interval $x + \Delta x$. Of these, the probability to die is the most suitable measure of ageing because it provides the *age-specific* impact on mortality independently from age-specific events at other age groups. To see this note that, for example, the probability of dying in the age interval (0,1) clearly affects the survival probability to any age $x > 1$ as well as life expectancy at age 0, but not the other way around.

Biologists and gerontologists are also emphasizing that individual ageing must be understood as an event-dependent not as a time-dependent process (Arking, 2006). Hence, according to this view there is no such thing as a "biological clock"; the probability of dying is just a measure of how members of a population (a species, a population of country) age *on average*.

Given a sample of a population and the observation (from a life table) that a number S_x thereof survives to age x and a number $d_x = S_x - S_{x+\Delta x}$ dies between age x and $x + \Delta x$, the probability to die in age interval $x + \Delta$ is given by $q_x = d_x/S_x$. The discrete-time measure has the inconvenient side-effect that the probability to die depends on the length of the age-interval x. In order to get rid of this problem we take the continuous limit

$$\mu(x) = \lim_{\Delta x \to 0} = \frac{S(x) - S(x + \Delta x)}{\Delta x S(x)} = -\frac{\dot{S}(x)}{S(x)}$$

which provides the *force of mortality*: the conditional probability to die at age x given survival up to age x. Note that age x is conceptualized as continuous and that $\mu(x)$ is not bounded from above by one.

From the empirical data the force of mortality can be approximated using, for example, the method of finite differences on $\mu(x) = -\mathrm{d}(\log S(x))/\mathrm{d}x$:

$$\mu(x) = \frac{1}{2\Delta x} \left[\log S(x - \Delta x) - \log S(x + \Delta x)\right].$$

The notion of the force of mortality should be familiar not only to engineers (as the failure rate) but also to economists as the hazard rate; the hazard here is to expire at age x. The "perpetual youth" models built on Blanchard (1985) and Yaari (1968) investigate the special case $\mu(x) = \lambda$ for all x. The assumption that the probability to die at age x is independent from x, constitutes, in fact, the definition of a non-ageing organism in biology (Arking, 2006). Thus, the most popular economic models that are taking the finiteness of human life into account are inherently unsuitable to display and discuss

aspects of human ageing.[1]

1.2. The Gompertz-Makeham Formula

The force of mortality for humans (and other organisms) increases with age in a particular way. This is shown exemplarily for US Americans in Figure 1.1, in which dots represent the actual age-specific force of mortality. Human life can be roughly subdivided into two periods: an initial *phase of development* ranging from birth to puberty during which the probability to die decreases (from A to B in Figure 1.1), immediately followed by the *phase of ageing* during which the probability to die increases. Notice that almost nowhere along the human life cycle is the probability to die constant.[2]

The most striking feature of Figure 1.1 is the long period of life, ranging from about 30 to 90 years of age, for which age and the force of mortality are log-linearly related. This relationship is known as Gompertz law after actuary Benjamin Gompertz (1825) who first observed and stated it formally: $\log \mu(x) = a + \alpha x$, or equivalently,

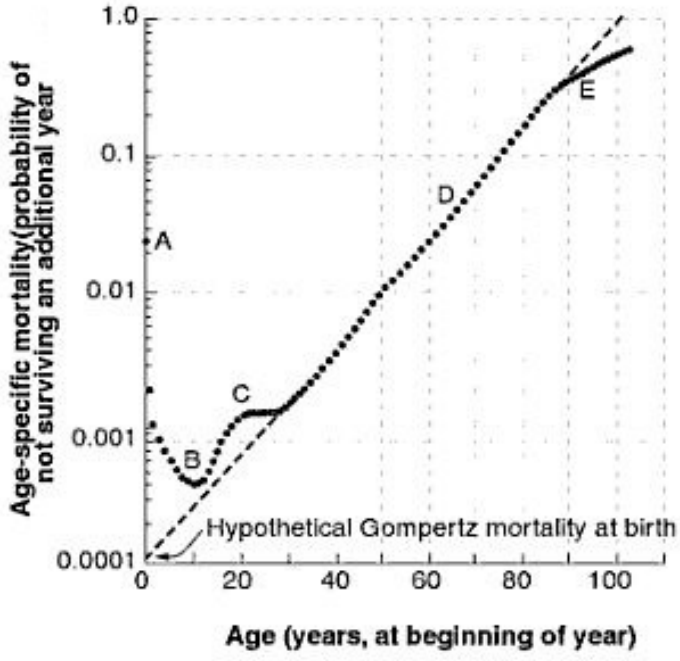
$$\mu(x) = R \exp(\alpha x). \tag{1.1}$$

For humans the estimate of α is around 0.1 implying that the probability to die doubles approximately every 10 years. The Gompertz formula was further improved by William Makeham (1860) who noticed that not all causes of death are age related. Consequently, a constant, reflecting an age-unrelated force of mortality, is added to yield the famous Gompertz-Makeham formula.

[1] This is of course acknowledged by the developers of the model by coining the "perpetual youth" expression.

[2] This rough division of human life is possibly good enough an approximation for economists interested in ageing. Yet at a closer look the picture gets more complicated and puzzling for evolutionary biologists and gerontologists. The trough of the mortality rate is actually not constant. It was falling over the last century and is now reached in fully developed countries around age 10, i.e. several years before sexual maturity (Milne, 2006). Moreover, ageing understood as the increasing loss of bodily function starts at least at birth if not earlier since infants are already subject to somatic mutations and telomere shortening (Frenck et al. 1998). See also Chapter 2.

Figure 1.1.: Age-specific Mortality Rate: United States 1959-1961



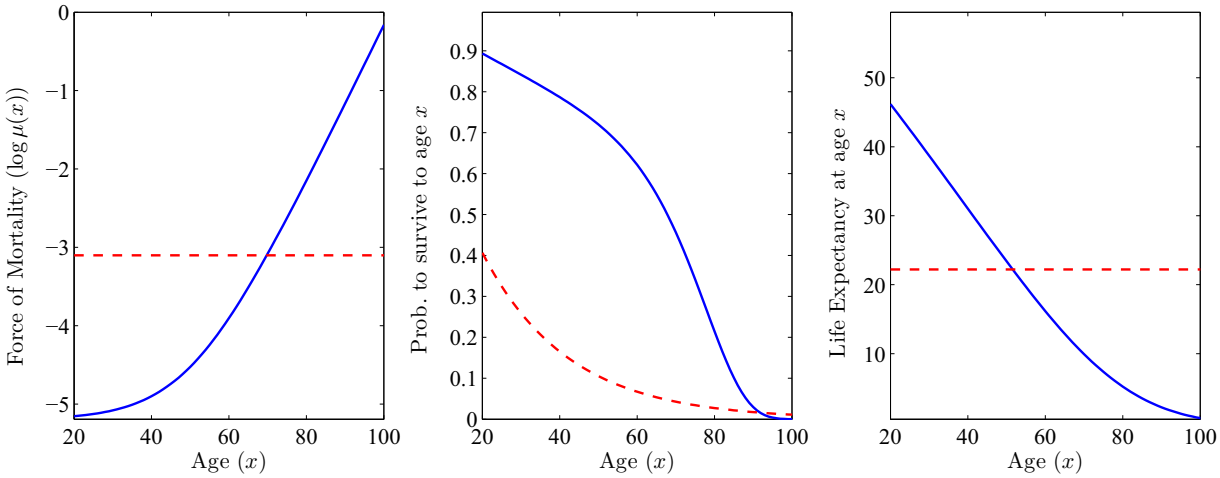
Dots indicate data points, the straight line shows the Gompertz estimate. Figure taken from Arking (2006).

$$\mu(x) = A + R\exp(\alpha x). \tag{1.2}$$

Subsequently, many researchers from various disciplines have tried to further improve the formula to little avail (see Gavrilov and Gavrilova, 1991, Ch. 2 for a detailed discussion). Taking both simplicity and precision into account, the Gompertz-Makeham formula is to the present day the most appropriate, concise, and widely used formal description of ageing (Olshansky and Carnes, 1997). Not only for humans are its parameters estimated with great precision also species as different as yeast, fruitflies, rats, and horses, have been shown to age according to the Gompertz-Makeham formula. The estimated coefficients, of course differ greatly, reflecting the large variation in life-span across species (Arking, 2006, Gavrilov and Gavrilova, 1991).

The fact that the Gompertz-Makeham law so precisely describes ageing from young adulthood onwards makes it a convenient tool to describe mortality of a representative individual in dynamic overlapping generations models, i.e. within a modeling framework which usually neglects decision making during childhood and assumes that a household's choices are made by persons at the onset of young adulthood. For these kinds of analyses the Gompertz-Makeham law (7.1) clearly shows that there exists no upper limit to human life-span. Formally, $\mu(x)$ is finite for finite x; there exists no age x at which $\mu(x)$ converges towards a pole. Or more vividly put "no matter how old one is, the probability to die on the next day is never unity". This statement may sound like platitude but, in fact, until recently many biologists were convinced that there exists a maximum human life-span (e.g. Fries, 1980). With respect to economic overlapping-generations modeling the Gompertz.Makeham law implies that there is no capital T marking "the end of the second period", i..e. the end of life.

Figure 1.2.: Ageing according to Gompertz-Makeham vs. Probabilistic Death



Solid lines: Estimates from Swedish life tables: $A = 0.00552$, $R = 0.000033$, $\alpha = 0.1013$ (Gavrilov and Gavriolova, 1991), own calculations. Dashed lines: hypothetical perpetual youth scenario: $\lambda = 0.025$.

Figure 1.2 compares mortality according to Gompertz-Makeham estimated for Swedish men alive 1901-1910 with mortality of an individual of perpetual youth. The panel on

the left-hand side shows the log of the force of mortality. It can be seen that the Gompertz part becomes dominating at ages around 50, from which on the force of mortality appears to be log-linearly increasing in age. The perpetual youth model clearly overestimates mortality of young adults and underestimates it for the old. Notice, by comparing Figure 1.1 and 1.2, how the Makeham-amendment leads to a much better approximation of μ for young adults. While the Gompertz formula approximates human mortality reasonably good for ages between 40 and 80 years, the Makeham amendment yields a good approximation for ages between 15 and 90 years (Carnes et al., 2006).

The middle panel of Figure 1.2 shows the probability to survive to age x (or the fraction of the original cohort size still alive at age x). It is obtained as the solution of $\dot{S}(x) - \mu(x)S(x) = 0$ given that $S(0) = 1$, implying for Gompertz-Makeham

$$S(x) = \exp\left(-Ax - \frac{R}{\alpha}\left(\exp(\alpha x) - 1\right)\right) \tag{1.3}$$

and, trivially, for the perpetually youth agent $S(x) = \exp(-\lambda x)$. According to Gompertz-Makeham, $S(x)$ is concave: the probability to survive decreases relatively with age at young ages and strongly at old ages. The perpetual youth model predicts just the opposite, i.e. steeply falling survival prospects at young ages.

The panel on the right hand side of Figure 1.2 shows life-expectancy (expected remaining years to live) at age x

$$e(x) = \frac{\int_x^\infty S(a)\mathrm{d}a}{S(a)}.$$

For the perpetual-youth agent life-expectancy is independent of age (and equals $1/\lambda$) while, actually, according to Gompertz-Makeham life-expectancy is almost linearly falling with age in the range of 20 to 70 years, after which it levels off.

At the upper age range, for the *oldest old*, the Gompertz-Makeham formula loses

precision. In contrast to what one might have expected, at ages of above 90 the force of mortality increases *less* than log-linearly with age . Again, this phenomenon is not unique to humans but visible in other species as well. Various amendments to the original formula have been suggested to take the oldest old into account (e.g. Perks, 1932). If this leveling off become so strong that the force of mortality stops to increase, the oldest old become indeed non-ageing. The possibility that – once a certain age has been reached – we turn out to be not exactly immortal but non-ageing has, of course, inspired research and debate in biology, demography, and gerontology (Wachter and Finch, 1997, Oeppen and Vaubel, 2002, Carnes and Olshansky, 2007).

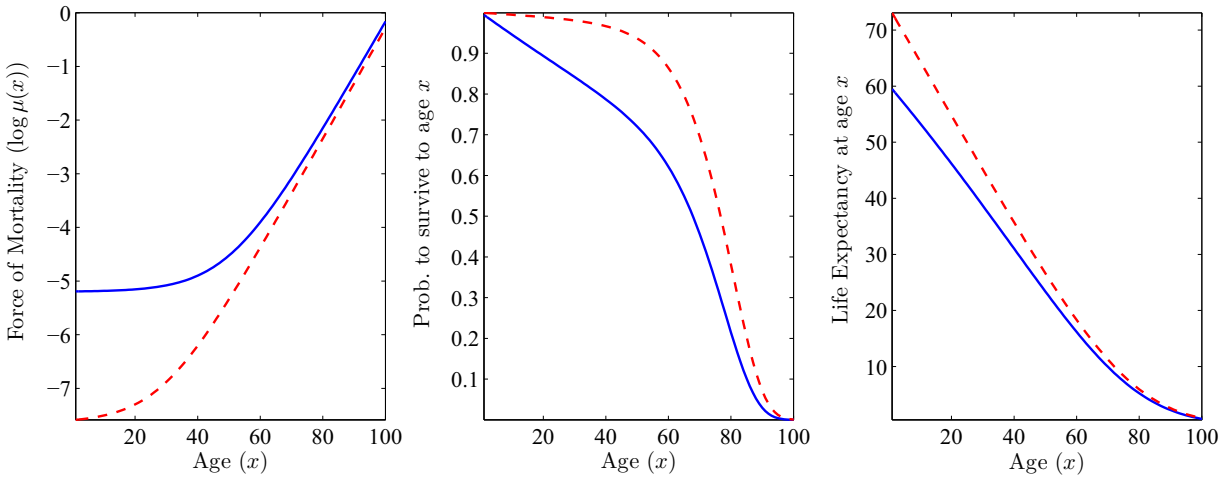
For a correct assessment note, however, that at best the force of mortality reaches a high plateau. It does not relapse to the level of background mortality A. Speaking of *negligible senescence* (Finch and Austad, 2001) with respect to the oldest old is literally correct but may nevertheless be misleading. Nowhere is the probability to die as high as for the oldest old. Intellectually, the biological foundation of a constant mortality rate is relatively easy to grasp (we return to that in section 5). Intellectually more fascinating is the Gompertz-Makeham law on which we focus now.

1.3. Stability of the Gompertz-Makeham Law: History and Science Fiction

Over the last century human life expectancy at birth has increased by more than 20 years in most of the fully developed countries (Riley, 2001). It is interesting to investigate how these huge improvements of human longevity have affected the Gompertz-Makeham law. For this purpose the notion is helpful that the Makeham-parameter A is not an inherently stochastic influence on human survival (as suggested by the perpetual youth model) but rather reflects the age-unrelated forces of mortality, i.e. the background mortality or extrinsic mortality (Carnes and Oshansky, 2007). In particu-

lar we expect prevention, eradication, or cure of age-unrelated diseases to be reflected in A. We expect progress with respect to the ageing process itself to be reflected in a change of Gompertz-parameters R and/or α.

Figure 1.3.: Ageing In Sweden 1901-1910 and 1983



Parameters estimated from Swedish life tables for men in 1901-1910 (solid lines) and in 1983 (dashed lines) Estimates from Gavrilov and Gavriolova (1991) and own calculations.

By estimating the Gompertz-Makeham law using life tables for a sample of countries over a very long horizon, Gavrilov and Gavrilova (1991) were able to show that the age-dependent parameters R and α vary across countries and across sexes but are strikingly stable over time. With contrast, the background risk parameter A has been found to be sex independent and (for most countries) continuously falling over time. For Sweden, for example, the parameter A fell over the last century by about one order of magnitude from 0.55 percent to 0.048 percent. From this Gavrilov and Gavrilova conclude that so far in human history advances in longevity manifested themselves almost exclusively in reduced background risk. Human ingenuity has not yet been able to manipulate much the biological mechanism behind the ageing process, captured by the age related parameters R and α.

Carnes et al., 1996, follow a different approach by trying to partition for a popula-

tion sample extrinsic causes of death (i.e. through exposure to risk) and intrinsic causes and by estimating R and α for the subsample of intrinsic deaths. This way they document a downward trend for R over the last 40 years for 5 fully developed countries. From this they conclude that the "mortality signature" for humans has already been modified. Unfortunately they did not let the data decide whether the extrinsic Makeham parameter is still significantly present so that the two studies are not comparable. Actually, as the authors admit, it is often difficult to partition deaths in extrinsic and intrinsic categories. Getting killed by a road accident, for example, is at least partly intrinsic since, ceteris paribus older victims are more likely to succumb to their injuries. The age distribution of pedestrians killed by road accidents parallels the distribution of deaths from all causes (Comfort, 1979).

Moreover, there appears to be evidence for an impact of nutrition and health in infancy and early childhood on the Gompertz parameter R. Note that an increase in R would be visible diagrammatically by a downshift of the $\log(\mu)$–curve at all ages whereas an increase in A shifts the curve predominantly at young ages such that curves for different A converge at high ages as shown in Figure 1.2. Finch and Crimmins (2004) observe – by eyeballing – such a parallel downshift of the mortality curve for Sweden over the last century and argue in favor of a causal effect of inflammatory exposure and nutrition during early childhood on this shift.

Some of the implications of the stability of the Gompertz-parameters can be discussed with help of Figure 1.3, which shows for Swedish men the estimated Gompertz curve for 1901-1910 (solid lines, replicating Figure 1.2) and for 1987 (dashed lines), together with its implication for the probability to survive to age x and life-expectancy at age x. The Figure impressively shows that so far technological, economic, and cultural progress has predominantly improved the survival probabilities for young people. Life-expectancy at age 80 in 1983 differs little from what it has been at the beginning of the 20th century.

The structural stability of the Gompertz-parameters is also helpful to explain the popular ideas of "rectangularization" and "compression of morbidity" (Fries, 1980). The middle panel of Figure 1.3 conveys the information that compared to the beginning of the century a higher share of Swedish men reaches an old age of, say 70 years, and expires then more quickly during their last years before death. Visually, the survival curve becomes closer to rectangular over time. On average, later born persons spend more of their life in a relatively healthy condition, an observation that has inspired the notion of a "compression of morbidity". Across individuals, however, there is no compression of morbidity. In fact, the panel on the left showing the force of mortality supports the opposite prediction: over time less young and intrinsically healthy people are killed from exogenous forces implying a rising share of the population displaying age-related frailties and chronical illness.

Figure 1.3 is also useful to shed light on the often controversially led discussion about the *human life* span. It has already been shown that strictly speaking such a thing as *maximum* human life-span does not exist. Even under time-invariance of the Gompertz-Makeham parameters, the simple fact that the "sample size" of people who ever lived on earth is continuously rising lets us expect that the maximum ever-observed life-length will rise as time proceeds.

Finch and Pike (1996) suggest to define T as the estimated age at death of the last survivor of a population of given finite size. This definition allows to calculate explicitly how life-span depends on sample size N. Plugging $S(x) = 1/N$, i.e. the probability to be the last man standing out of N, into (1.3) and ignoring background risk ($A = 0$) provides

$$\frac{1}{N} = \exp\left(-\frac{R}{\alpha}\exp(\alpha T) + \frac{R}{\alpha}\right) \quad \Rightarrow \quad T = \frac{1}{\alpha}\log\left(\frac{\alpha}{R}\log N + 1\right).$$

Given their estimates of the Gompertz parameters Finch and Pike obtain life-span as 105 years for $N = 10^3$ and 114 for $N = 10^7$. It turns out, that – without taking an

oldest-old amendment into account – the Gompertz law predicts actually life-span of local human and animal populations quite well. For example, it predicts for the sample size of US females in 1980 $T = 115$ years while it was actually 112 years. Note also the strong inverse relationship between T and α, suggesting large potential improvements of T for small improvements of the pace of mortality α.

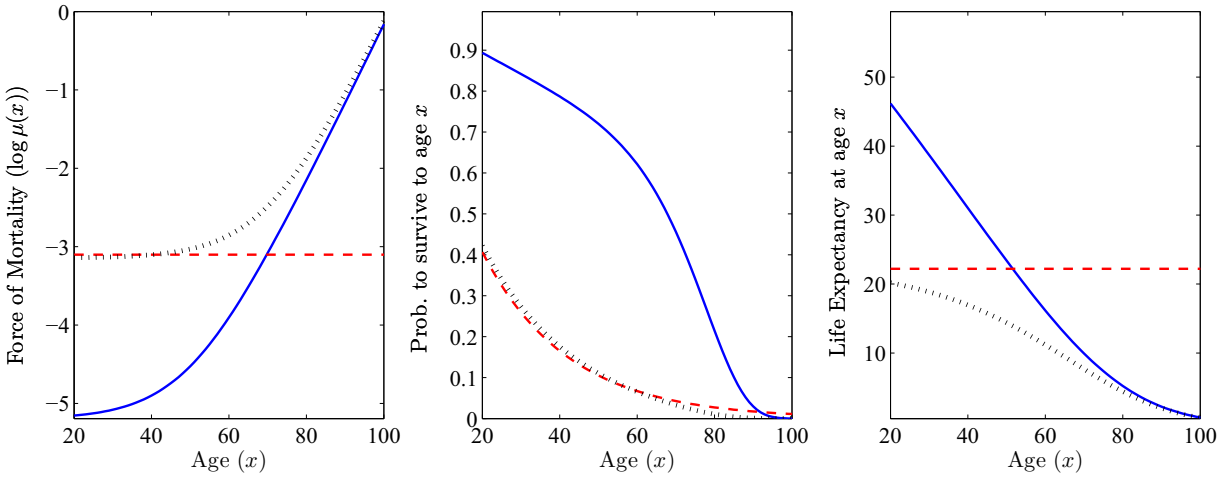
For constant α, however, life-span does not improve much through changes of N or R. Sensitivity of life-span with respect to background risk A can be read off from Figure 1.3. Inspection of the right and middle panel clarifies that the statement that "life-expectancy at birth (or at age 20) went up tremendously during the last century" is supported by the same data that lead to the conclusion that life-expectancy at age 100 or life-span (x for $S(x) = 1/N$) has not improved much.

The observation that increasing life-expectancy was overwhelmingly brought forth by declining background mortality inspired the following experiment. Suppose the Gompertz-Makeham law holds and the Gompertz parameters are indeed stable over time. To fit the curve to life-expectancy at birth in medieval England (1330-1479) which was about 22 years (Clark, 2007). we set $A = 0.043$. The implied predictions of the Gompertz-Makeham model are shown by dotted lines in Figure 1.4, which for better comparison also re-iterates from Figure 1.2 mortality for Sweden in the 20th century (solid lines) and the perpetual youth scenario (dashed lines).

The middle panel of Figure 1.4 clearly shows that survival according to Gompertz-Makeham under the high medieval background mortality is visibly not very different from the prediction of the "perpetual youth" model (dotted lines). The force of mortality, shown in the left panel, follows almost a straight line until about age 50, i.e. until an age to which less than 10 percent of population survive. In other words, although intrinsically present and not different from today, the mechanics of ageing are not visible. The large majority of individuals dies from extrinsic causes. The notion that survival probabilities were about the same in human paleontology as in medieval

England (Clark, 2007, Gurven and Kaplan, 2007) inspires a first attempt towards an *explanation* of human ageing: under high extrinsic mortality there was little evolutionary pressure for non-ageing humans.[3]

Figure 1.4.: Ageing in the 20th-century and in Pre-Modern Times vs. Probabilistic Death



Solid lines: Estimates from Swedish life tables as in Figure 1.2. Dashed lines: hypothetical perpetual youth scenario: $\lambda = 0.025$ as in Figure 1.2. Dotted lines: Gompertz-Makeham with elevated background risk $A = 0.43$.

The time-(in)-variance of the Gompertz-Makeham parameters can also be used to assess predictions about *future* gains in human life-expectancy. If future life-expectancy could be derived from interpolating historic trends, one can indeed arrive at the conclusion that life-expectancy at birth will improve by about 3 month per year of birth (Oeppen and Vaupel, 2002). The Gompertz-estimates, however, lead to a completely different conclusion. Because A is already close to zero for most developed countries, there will be little more future gains in life-expectancy as long as technological progress improves almost exclusively background mortality and future life-expectancy at birth will unlikely exceed 85 years (Fries, 1980, Carnes and Olshansky, 2007).

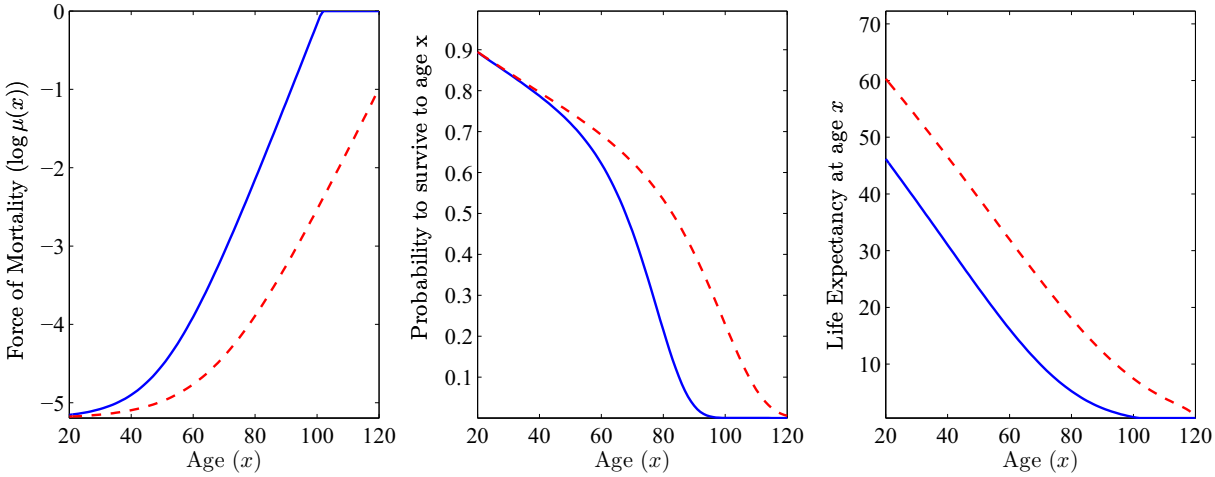
One problem with the stability hypothesis for the Gompertz parameters is that if improvements in mortality were exclusively through declining A and A is by now

[3] Evolutionary theories of ageing are discussed in Chapter 3.

close to zero in fully developed countries, we should observe convergence behavior. Age-specific mortality across fully developed countries should converge over time and the age-distribution of deaths within countries should converge towards a limiting distribution. Yet, Wilmoth (1997) cannot find such convergence behavior in the data.

In any case, future technological progress may make it possible (e.g. through gene therapy) to reduce the Gompertz parameter α and it is illuminating to discuss briefly how this would be reflected by the mortality statistics. Technological progress operating through α would be very different from anything observed so far: it would effectively slow down the ageing process. Note that curtailing or even abolishing major causes of death (like, for example, cancer or diabetes) would *not* be expressed as reduction of α. Without a reduction of α any successful fight against one bodily impairment entails inescapably the occurrence of another impairment in the near future.[4]

Figure 1.5.: Ageing Before and After the Wonka-Vite



Solid lines: historical data as in Figure 1.2. Dotted lines: hypothetical reduction of α from 0.101 to 0.08.

Figure 1.5 shows the consequences of the introduction of a "Wonka-Vite" (Dahl, 1987, Kirkwood, 1999) i.e. a hypothetical pill that reduces α from 0.101 to 0.008. The difference to the observed historical evolution of ageing shown in Figure 1.3 is striking.

[4] A health-economic model in this spirit has been developed by Becker and Phillipson (1998).

The Wonka-Vite effectively increases life-expectancy at *every* age by about 15 years and thus indeed raises human life-span. While it does not produce a rectangularization of the survival curve or a compression of morbidity, it effectively increases the health status at every age, implying the prediction of less age-related deaths at *every* age and a lower share of unhealthy (morbid) individuals in the population.

1.4. Species-Specific Invariants: The Strehler–Mildvan Correlation

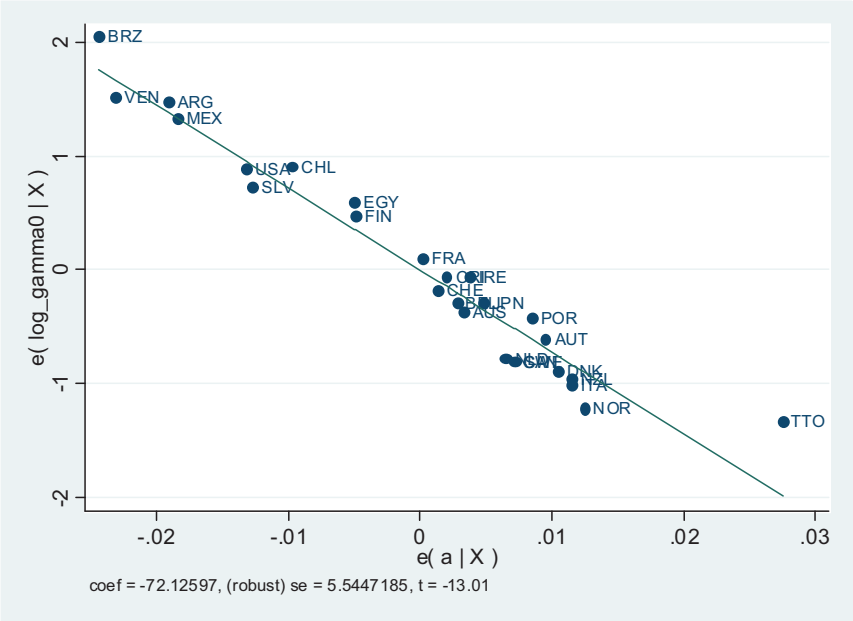
While the Gompertz-parameters α and R seem to be invariant over time they differ actually across sexes and across countries. On average, people age in different ways depending on their sex and provenience. On average, across countries, women face a lower R and a higher α (Garilov and Gavrilova, 1991), indicating that women have an initial advantage of lower ageing, which is with rising age eventually caught up by men. In fact, there is a strong inverse (log-linear) association between R and α known as the Strehler-Mildvan correlation or the compensation effect of mortality (Strehler and Mildvan, 1960, Gavrilov and Gavrilova, 1991). Figure 1.6 shows the association between these parameters for 30 countries.

The presence of the Strehler-Mildvan correlation has inspired the search for a species-invariant component of ageing and a new definition of human life-span (Gavrilov and Gavrilova, 1991). To see this, suppose the estimated Strehler-Mildvan correlation is given by

$$\log R_i = \log M - \alpha_i T \tag{1.4}$$

where for $i = 1, \ldots, n$ the R_i and α_i are the country- and sex-specific realizations of the Gompertz parameters and M and T are the invariant coefficients of the Strehler-Mildvan correlation. Thus $\mu_i = A_i + R_i \exp(\alpha_i x)$ is the country and sex-specific force of

Figure 1.6.: The Strehler-Mildvan Correlation



Data for 31 countries from Strehler and Mildvan (1950), Parameter a (measured along the abscissa coincides with α in the text. Parameter $\gamma 0$ (measured in logs along the ordinate) coincides with R in the text.

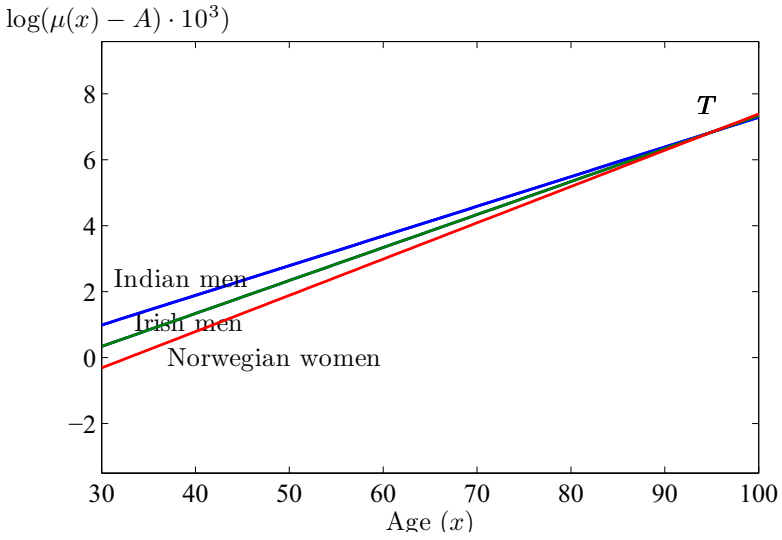
mortality. Inserting into this the Strehler-Mildvan correlation provides:

$$\mu_i - A_i = M \exp\left[\alpha_i (x - T)\right]. \tag{1.5}$$

Observe that the modified Gompertz-Makeham law implies a fixed point. It predicts that – controlling for country-specific background risk – men and women across countries share a common force of mortality M at age T. This implication is visualized for a sample of three life-time tables in Figure 1.7. Since the focal point (T, M) is assumed approximately by all observations it can be interpreted as a *species-specific* constant. For humans the point estimate is $T = 95$ years. Gavrilov and Gavrilova (1991) have shown that such a species-invariant focal point is also observable within animal societies for different populations of drosophila and of labor rats. The time invariance of

the Gompertz law and the Strehler-Mildvan correlation suggests that – controlling for country-specific background risk (degree of economic development) humans share a common mechanism of ageing, a common stochastic process according to which individual bodies lose function over time and bodily failures and impairments are accumulated.

Figure 1.7.: The Strehler-Mildvan Correlation Across Countries



Data from Gavrilov and Gavrilova (1991) and own calculations.

A general pattern emerging from the Strehler-Mildvan correlation is that economically more advanced countries are characterized by a lower R and higher α, i.e. by a lower initial mortality rate and faster speed of ageing. This implies that although lifespan T seems to be unaffected by economic advances, the mortality rate is everywhere, i.e. for all ages below T, lower in the economically advanced country. An important conclusion from this is that ageing quickly (having a high α) is actually not a burden but an indication of superior fitness. This notion will be further substantiated in the next section. Finally, note that if R falls continuously and α rises such that lifespan T remains constant, we would nevertheless expect that life-expectancy is continuously

improving and that the record of the highest so far observed age at death (the conventional measure of lifespan) is rising as well.

1.5. Reliability Theory

The fact that all humans appear to age according to a common formula, which we moreover share with fruitflies, rats, and other animals, motivates the search for a common process that drives ageing. The quest for such a process is challenging because, as emphasized already, there exists no biological clock. Trying to explain ageing following a line of reasoning by stating that humans age because their organs (e.g. the cardiovascular system) age, and that organs age because the tissue they are made of ages etc. will turn out to be utterly futile. At some point a micro-level will be reached that consists of non-ageing entities, for example atoms. Eventually, we want to explain why a system ages that consists of non-ageing elements.[5]

In explaining ageing systems biologists do not have to start from scratch. They can built upon a subdiscipline in engineering, reliability theory, which is concerned just with this particular problem, i.e. the problem how complicated mechanical systems consisting of non-ageing elements (like cars) are increasingly losing function over time so that the failure rate, i.e. the probability of the expiry of the system increases with age (Barlow and Proschan, 1975). The task for biologists is to modify and extend the available theory such that it is capable of motivating the Gompertz-Makeham law (and possibly other important phenomena of human ageing, like the Strehler-Mildvan correlation or the non-ageing of the oldest-old). There are several, sometimes complementing, reliability theories available (Gavrilov and Gavrilova, 1991, Novoltsev, 2006, Finkelstein, 2008). Here we try to explain the idea by describing two particularly straightforward models in detail after which we briefly mention existing extensions.

[5] Of course, the claim that atoms do not age is justified only within a reasonable time-frame. If the time frame gets long enough, even the universe ages.

The presentation follows Gavrilov and Gavrilova (1991) who were the first to integrate reliability research into the realm of biology.

Consider an organism constructed of n non-ageing blocks. Non-ageing means that the failure rate λ is constant over time. Given age x the probability of a block to fail is $1-\exp(-\lambda x)$. Blocks are connected in parallel and the organism lives as long as at least one block is in order. The probability that the organism expires before age x is given by $F(x) = [1 - \exp(-\lambda x)]^n$ and the probability to survive to age x is $S(x) = 1 - F(x)$. The unconditional probability to die at age x is thus given by $\mathrm{d}F/\mathrm{d}x = -\lambda n \exp(-\lambda x)[1 - \exp(-\lambda x]^{n-1}$ and the force of mortality is

$$\mu(x) = \frac{\lambda n \exp(-\lambda x)[1 - \exp(-\lambda x]^{n-1}}{1 - [1 - exp(-\lambda x)]^n}.$$

Taking a Taylor-approximation around zero, implying $1 - \exp(-\lambda x) \approx \lambda x$, the expression simplifies to $\mu(x) \approx n\lambda^n x^{n-1}$ and using L'Hospital's rule we obtain $\lim_{x\to\infty} \mu(x) = \lambda$.

The simple model is thus capable of explaining ageing: the force of mortality $\mu(x)$ is increasing with age x. Ageing is explained as a *loss of redundancy* over time. This notion of ageing as accelerated loss of organ reserve is in line with the mainstream view in the medical science. For example, initially, as a young adult, the functional capacity of human organs is estimated to be tenfold higher than needed for survival. (Fries, 1980).

The model is also useful to explain mortality of the oldest-old. With increasing age the organism loses redundancy until survival depends, in the limit, just on the survival of the last functioning non-ageing block. Thus the organism converges towards the constant mortality rate λ of its non-ageing elements. This is a general result from reliability theory: the rate of ageing, i.e. the age-dependent component of the mortality rate (the failure rate), is increasing in the complexity (redundancy) of the system. We can expect that, for example, light bulbs and bacteria, age at a much lower rate than cars or humans. Seen this way, ageing appears as a positive trait. It occurs because

complex systems start out at a high level of redundancy and are thus less likely to expire at young age.

The problem with the simple model is that the derived $\mu(x)$ at young ages does not follow the Gompertz-Makeham law. It follows – similar to the failure rate of mechanical systems – a Weibull distribution. In order to describe the ageing process of humans, the model has to be made "more human". Next we consider one of several possible extensions.

Suppose an organism consists of m *irreplaceable* blocks, i.e. blocks connected in series such that the organism dies if one block fails. Each block consists of n elements, connected in parallel with age-independent failure rate λ. Following the computations from above we know that the failure rate of a block is $n\lambda^n x^{n-1}$ for small x and approaches the constant λ for large x. Because blocks are connected in series (each of them being essential), the failure rate of the organism equals the sum of failure rates of blocks i.e. $m \cdot n\lambda^n x^{n-1}$ for small x and $m\lambda$ for large x.

Next suppose that many elements are initially defect. The probability of an initially functioning element is given by q. The failure rate of a block with i initially functioning elements is thus $\mu_b(i) = i\lambda^i x^{i-1}$ for small x and $\mu_b(i) = \lambda$ for large x. Blocks, ordered according to their number of initially functioning elements, are binomially distributed. We approximate the Binomial with a Poisson distribution

$$P(i) = c \cdot \exp(-k)\frac{k^i}{i!}$$

where $k \equiv nq$ is the mean number of initially functioning elements and c is a normalizing constant ensuring that the sum of probabilities equals one. The failure rate of the system, computed as the sum of the failure rate of blocks, is then obtained as

$$\mu(x) = \sum_{i=1}^{n} mP(i)\mu_B(i) = mc \cdot \exp(-k)\sum_{i=1}^{n}\frac{k^i}{i!}\mu_B(i).$$

Inserting $\mu_B(i)$ we obtain that for young age (small x) the failure is approximately

$$\mu(x) \approx R \sum_{i=1}^{n} \frac{(k\lambda x)^{i-1} \cdot i}{i!} = R \cdot \left[\sum_{i=1}^{\infty} \frac{(k\lambda x)^{i-1}}{(i-1)!} - \sum_{i=n+1}^{\infty} \frac{(k\lambda x)^{i-1}}{(i-1)!} \right].$$

with $R \equiv mc\lambda k \exp(-k)$. Now consider a complex, redundant organism with a large number of elements. In the limit, for $n \to \infty$, the last term in brackets converges to zero. The first term in brackets simplifies to $\exp(k\lambda)$.

Inserting into $\mu(x)$ the $\mu_B(i)$'s for large x we get the failure at high age approximately as

$$\mu(x) \approx mc \exp(-k) \sum_{i=1} n \frac{ki}{i!} \lambda = mc\lambda \exp(-k) \left[\sum_{i=1}^{\infty} \frac{k^i}{i!} - \sum_{n+1}^{\infty} \frac{k^i}{i!} \right].$$

Again the last term in brackets is approximately zero for large n. The first term equals $\exp(k)$ such that the expression simplifies to $\mu(x) = cm\lambda$. Collecting results we have

$$\mu(x) \approx \begin{cases} R\mathrm{e}^{\alpha \cdot x} & \text{for small } x \\ c \cdot m \cdot \lambda & \text{for large } x. \end{cases} \tag{1.6}$$

with $\alpha = k\lambda$. The organism ages according to Gompertz law. For large ages (the oldest old) the force of mortality converges to a high plateau.

Taking log's of R we get $\log R = \log(cmk\lambda) - k$ and inserting $\alpha = k\lambda$ we arrive at

$$\log R = \log M - \alpha T \tag{1.7}$$

with $M = \alpha mc$ and $T = 1/\lambda$, which is the Strehler-Mildvan correlation. The reliability model is capable to generate the most important regularities of human ageing.

For an interpretation of parameters, first note that T is uniquely pinned down by λ, the age-independent failure rate of an element. If λ is a species-specific constant, then the model predicts a unique focal point (life-span) T for the species. Across species, T depends inversely on the robustness of its non-ageing elements. Let's reasonably

assume that m, the number of irreplaceable blocks, is also a species-dependent constant. Then all variation within a species results from variation of k, the mean number of initially functioning elements.

The parameter k is a compound parameter $k = nq$. Consider first variation in n, the number of elements per block. For humans these differences could exist across countries, because their citizens are on average of different size. This could in principle come through country-specific diet and/or Darwinian selection because optimal body-size depends on ambient temperature and the disease environment. The model then predicts that $\alpha = k\lambda = nq\lambda$ is small in countries were people are on average relatively short (say, India) and R is small in countries were people are on average tall (say, the Netherlands). While the model predicts the same life-span for Dutchmen and Indians, it also predicts that the Dutch start out at lower mortality rates at younger years and are ageing faster. In fact, the Dutch's headstart is caught up by the Indians just at T implying that everywhere below T survival prospects are better and life-expectancy is higher in the Netherlands. Generally, the model predicts in line with the evidence a positive association between body size and life-expectancy. The explanation is redundancy. Larger humans have a larger organ reserve to life off.[6]

Next, hold n constant and consider variation of q, the probability for an element to be initially functioning. Here it is probably more reasonable to consider within country variation over time (say Indians now and 50 years ago, or Dutchmen now and 300 years ago). The idea here is that available nutrition (for mother and child) and disease exposure early in life have shaped q. Taking the historic improvement in nutrition and health into account, the model then predicts that both Indians and Dutchmen are ageing faster now but starting out much better at younger age. Consequently, survival prospects and life-expectancy have improved at any age. Actually, both Indians and Dutchmen are also predicted to be taller because of the generally improved conditions

[6] Chapter 2 of this review discusses the interaction of body size and mortality in detail.

(Fogel, 1994) so that mortality would have improved because of the simultaneous and amplifying effect of q and n on $k = nq$. But even holding body size (n) constant, the model predicts through better initial functionality "fetal origins" (Barker, 1995, Black et al., 2007), i.e. lasting consequences of early events in life on later mortality and morbidity.

While the model does a fairly good job in explaining basic mechanisms of human ageing, a better approximation of the actual ageing process can be achieved by adding more details, i.e. more complexity. Important extensions of the basic model include cascading effects, i.e. the phenomenon that failure of one element of the system entails failure of other elements (Gavrilov and Gavrilova, 1991) and mechanisms of imperfect (cell-) maintenance and repair (Finkelstein, 2008). The most recent model of Milne (2009) includes also a mechanism for redundancy expansion in early life and is complex enough to redraw with great precision actually observed mortality data.

A common characteristic of all reliability-based models is that organisms are conceptualized as complex systems consisting of essential parts (e.g. organs, tissue) connected in series, which are in turn built of smaller entities connected in parallel. Parallel connectivity means that every reliability theory is built upon the idea of redundancy. Another common theme is a stochastic failure rate for the basic entities. The notion of ageing as driven by a "natural" stochastic process helps to explain the "unfair" nature of human fate, i.e. why we actually observe large differences of ageing on the individual level. Reliability theory can explain why individuals raised under equal conditions and/or built from the same genes (monozygotic twins) can age and eventually die in a very different ways. At the same time the models provides a toehold to explain how population- (e.g. country-) specific characteristics and the environment early in life have a bearing on aggregate ageing behavior of entire populations and/or sub-populations. In short, it integrates "good luck" as another important driver of longevity besides "good genes" and "good behavior". By now the new view that age-

ing and death are probably best conceptualized as accidental stochastic shocks on the molecular level is firmly established in the natural sciences and has entered the biology textbooks (Arking, 2006).

1.6. Ageing, Frailty, and Morbidity

Although age is such a powerful predictor of mortality on the aggregate level, it is a relatively poor predictor on the individual level and biologists are tirelessly emphasizing that ageing should not be conceptualized as time-dependent but as an event-dependent process. We thus turn next to the measurement of how ageing is *expressed* in individuals and the question how – on the aggregate level – the expression of ageing is related to age. These questions are frequently addressed under the heading of *biomarkers*. A biomarker in general is an indicator of the biological state of an organism. In bio-gerontology a set of biomarkers is an indicator of the *biololigcal age* of an individual. On the individual level biomarkers allow us to disentangle chronological age and individual age, i.e. to understand why a 40 year old person feels like being 60 and vice versa. On the aggregate level it is interesting to investigate how biomarkers develop with age and how chronological age and biological age are related.

There exists no unified theory how biomarkers for ageing should look like and consequently the literature has followed different approaches (e.g. Comfort, 1979, Borkan and Norris, 1980, Shock, 1981, Weale, 1997, Jackson et al., 2003). A common theme is the development of an index, i.e. a whole set of biomarkers measuring physiological (and sometimes also the cognitive) variables followed by an investigation of how the developed index correlates with age and how well it predicts mortality. Typical age- and mortality-related biomarkers are systolic blood pressure, grip strength, forced expiratory volume, white blood cell count, hearing loss, and reaction time.

Next we survey one indicator, the so called frailty index, in more detail because

for this indicator the literature went farthest to explore its relation to mortality. The frailty index has been established and investigated by Mitniski and Rockwood and several coauthors in series of articles (e.g. Mitniski et al, 2002a, 2000b, 2005, 2010). The following exposition is mainly based on Mitnitski et al., 2002a).

The frailty index counts for a large sample of individuals the bodily impairments which are actually present out of a list of potential impairments, ranging from mild deficits (reduced vision, incontinence) to near lethal ones (stroke, cancer). Mitnitski and Rockwood consider a sample of 15 to 79 years old Canadians and 38 impairments. They then show that the relative number of deficits, i.e. the frailty index D, of an individual correlates with age in the exponential Gompertz-Makeham fashion. Specifically, deficit accumulation evolves with age x according to

$$D_i(x) = B + Q_i \exp(\beta_i x) \tag{1.8}$$

where $i = \{m, w\}$ is an index taking sex-specific differences into account.

Given the large sample size, the result can be interpreted as the probability of having D percent of all possible frailties at the age of x. The parameters of the frailty accumulation law are estimated with great precision. The point estimate and the 95 percent confidence interval for Canadian men are $B = 0.02 \pm 0.001$, $\log Q_m = -5.77 \pm 0.06$, and $\beta_m = 0.043 \pm 0.001$ with an R^2 of 0.97. Women are estimated to face the same initial frailty B and a lower β ($\beta_w = 0.031 \pm 0.001$) and a higher Q ($\log Q_w = -4.62 \pm 0.06$). This means that men are estimated to start out better and then to accumulate frailties at higher speed i.e. to age faster than women.

The fact that men are initially less impaired and accumulate deficits faster implies that the sex-specific frailty curves are meeting each other at some age T_D. Following the Streher-Mildvan approach we get

$$D_i = B + D_T \cdot \exp\left[\beta_i (x - T_D)\right]. \tag{1.9}$$

Implying that men and women share the same frailty $B + D_T$ at age T_D. While the fact that such an intersection exists is not surprising, the striking result here is that the estimate of $T_D = 94 \pm 2$ years coincides with the focal point of mortality (life-span) estimated by Gavrilov and Gavrilova (1991). The result suggests – at least for Canadians – that the mechanisms behind frailty accumulation and mortality are probably very similar. It supports, indirectly the reliability theory of ageing and mortality.

Using $T = T_D$, the strong association between mortality and frailty, both apparently governed by a Gompertz-Makeham law, can be made even more visible by solving (1.9) for T_D and substituting the result for T in (7.1) using the definition $M \equiv R\exp(\alpha T)$. This provides the positive functional association between mortality and frailty as

$$\mu_i(D_i) = A + M \cdot \left(\frac{D_i - B}{D_T}\right)^{v_i}, \tag{1.10}$$

where $v_i \equiv \alpha_i/\beta_i$.

Apparently the association between frailty and mortality is non-linear. Recalling that α is about 0.1 the exponent v_i is larger than unity, suggesting a convex shape. In words, another bodily impairment affects mortality relatively little when frailty is low compared to another impairment when frailty is already high. An intuitive explanation could be that less severe frailties, like impaired vision or backpain, are on average taken up earlier in life compared to (near) lethal ones like cardiac infarction or stroke.

The striking result here is that the exponent is lower for men than for women reflecting the fact that the force of mortality increases faster for women (α is larger for women) than for men while frailties are accumulated faster by men (β is larger for men). It implies that for ages below $T_D = T$, women display more frailties than men but men are more likely to die.

The fact that (Canandian) men have a comparative advantage in frailty accumulation can be straightforwardly explained by reliability theory. Men are on average larger and have thus a larger organ reserve, a larger redundancy n (more bone mass, more muscle

muss) to wear off during ageing. Male bodies are thus more reliable than female and are predicted to fail later in life.

The puzzling effect here is that women are nevertheless predicted to have a comparative advantage in mortality. One possible explanation for this could be found by taken into account the size and singularity of organs. Eyes, for example, usually come in pairs of about the same size for men and women. This opens the door for q, the probability of an initial functioning element, to have an impact. Suppose q is larger for women. Recalling that the mean probability of a block (say an organ) to fail is given by the compound qn, it is imaginable that women have a comparative advantage when size n of the organ does not matter much for function (e.g. the cardiovascular system) whereas men have a comparative advantage when size is important (e.g. muscle strength and bone mass). Consistently with the mortality-frailty data, the model would then predict that men are on average more likely to develop, for example, a blocked neck artery (a stroke) but are less likely to suffer from back pain and osteoporosis.

The discussion leads us, finally, to a distinction between loss of function (measured by the frailty index) and morbidity and their specific interaction with age and mortality. While the frailty index is just a counter of personal health experiences, the term morbidity is commonly used with an emphasis on the *quality* of being unhealthy, in particular with relation to disability. Although the evidence is sometimes conflicting, there exist several studies documenting for the U.S. and other developed countries a decline of disability among the old.[7]

The fact that the old, given age, experience less disability today than past generations entails no contradiction to the adamant loss of bodily function with age predicted by reliability theory. Both phenomena are compatible once technological progress is taken into account. For example, before the invention of eyeglasses or hip replace-

[7] Note that declining disability among the old does not imply a compression of morbidity at the population level. The rising population share of old and potentially disabled persons works in the opposite direction. See Freedman et al. (2002) Crimmins (2004), and Parker and Thorslund (2007) for surveys on trends of disability and morbidity among the old.

ment, short-sightedness or arthritis may have caused disability while today these age-dependent impairments are at best mild nuisances. Seen this way, disability prevalence among the old is not a good indicator of ageing but a good indicator of medical technological progress (and, possibly, of efficiency and equality of health care provision).

Technological progress, however, has also driven a wedge between mortality and frailty by postponing death from age-related causes. Contemplating the fact that Mitnitski and Rockwood's frailty-counter advances by one if for, example, a person has *survived* a heart attack or a stroke, may lead to the conclusion that the frailty index is actually a better measure of ageing than the force of mortality. In any case it highlights the double role of medical technological progress. To summarize, extrapolating past technological trends we expect more people to survive towards a state of less functionality in less redundant bodies. But we also expect less disability, since loss of bodily function will be increasingly repaired by technology. With contrast, a yet to come biomedical progress affecting the accumulation rates α and β (as, hopefully, epi-genetic advancements) would have a very different quality. It promises improvement of *both* measures of ageing, i.e. lower mortality and higher body functionality at any given age.

The Gompertzian accumulation of frailties suggests for (health-) economists a straightforward improvement of modeling of health deterioration with age. Normally, health is introduced as a state variable similar to human capital evolving like $\dot{H} = I - \delta H$, where a dot signifies a derivative with respect to age and I is health investment, see e.g. Grossman (1972) or Ehrlich and Chuma (1990). Holding δ constant the model predicts, counterfactually, that health loss is largest when the state of health is best (δH is largest when H is highest). Of course, health researchers are aware of this contradiction of the facts, and try to repair the health equation by introducing an age-dependent depreciation rate $\delta(x)$. But how exactly should the deprecation function look like so that the model provides an acceptable approximation of real health deterioration? An intuitive

and readily implemented approach is available by Gompertzian frailty accumulation.

Suppose actual health is defined as best attainable health minus accumulated frailties, $H = \bar{H} - D$ where $\bar{H}$ could be the state of health of a healthy 10 year old. This implies $\dot{H} = -\dot{D}$. Taking the derivative in (1.9) with respect to age and ignoring sex-specific indices we get

$$\dot{D} = \beta Q \exp(\beta x) = \beta(D - B). \tag{1.11}$$

Inserting this into $\dot{H}$ and re-substituting D provides

$$\dot{H} = \beta \left(\bar{H} - H - B \right). \tag{1.12}$$

Equation (1.12 constitutes) a simple linear differential equation for health, which – in line with the facts – predicts that health loss is small at a good state of health and increasing losses are predicted when the health stock deteriorates. In fact, the model is, at least for Canadians, readily calibrated with the available data. After all that has been discussed so far it is also clear where successful health investments would enter the equation. After the Wonka-Vite, they might reduce the β. Up to now, health investment reduces at best the B because it has not yet the power to slow down the pace of ageing,

1.7. Conclusion

This chapter has shown a strong log-linear association between age and the force of mortality, known as the Gompertz-Makeham law, and its implications for survival and life-expectancy. It has used the formula to discuss the history and science fiction of human ageing. It has then shown another strong association, the Strehler-Mildvan correlation, and developed its implication for human life-span and the idiosyncracies of the ageing process across countries and sexes.

These observations have set the stage for an explanation of ageing using reliability

theory from engineering, a theory that explains ageing of organism and other systems constructed of non-ageing elements by the intricate interplay of stochastic processes operating in series and in parallel, i.e., in short, by the depletion of initial functional redundancy within the system. The theory was then confronted with an empirical measure of functional loss, the frailty index. Again a strong association between the log-linear accumulation of frailties with age and the Gompertz-Makeham law emerged. The overall conclusion is thus the notion of ageing as the increasing loss of bodily function with age. Although this loss of function can be best conceptualized as being purely stochastic at the cell level it manifests itself as a "law of ageing" at the aggregate level. This has lead to some first suggestions for the modeling of health deterioration and maintenance for the LEPAS project as well as for health economics in general.

Some important aspects ageing have not been addressed in this chapter, for example how evolution has shaped human ageing, how ageing relates to physiological processes, in particular to metabolic activity, and how different elements of bodily function age differently, i.e. how different aspects of human capital (e.g. muscle strength, different functions of the brain) deteriorate with age. These questions are in detail addressed in the following chapters.

1.8. References

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

Barker, DJP, 1996, Fetal origins of coronary heart disease, *British Medical Journal* 311, 171-174.

Barlow, R.E. and Proschan, F., 1975, *Statistical theory of reliability and life testing: Probability models*, Holt, Rinehart and Winston, New York.

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

Borkan, G.A and Norris, A.H, 1980, Assessment of biological age using a profile of physical parameters, *Journal of Gerontology* 35, 177-184.

Carnes, B.A. and Holden, L.R. and Olshansky, S.J. and Witten, M.T. and Siegel, J.S., 2006, Mortality partitions and their relevance to research on senescence, *Biogerontology* 7, 183–198.

Carnes, B.A. and Olshansky, S.J., 2007, A realist view of ageing, mortality, and future longevity, *Population and Development Review* 33, 367- 381.

Comfort, 1979, *The Biology of Senescence*, 3rd edition, Elsevier, New York.

Crimmins, E.M., 2004, Trends in the health of the elderly, *Annual Review of Public Health* 25, 79–98.

Dahl, R, and Foreman, M., 1987, *The Complete Adventures of Charlie and Mr Willy Wonka*, Puffin Books.

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

Finkelstein, M.S., 2008, *Failure Rate Modeling for Risk and Reliability*, Springer, London.

Finch, C.E. and Pike, M.C., 1996, Maximum life span predictions from the Gompertz mortality model, *Journal of Gerontology* 51, 183–194.

Finch, C.E and Austad, S.N, 2001, History and prospects: symposium on organisms with slow ageing, *Experimental Gerontology* 36, 593–597.

Finch, C.E. and Crimmins, E.M, 2004, Inflammatory exposure and historical changes in human life-spans, *Science* 305, 1736-1739,

Fogel, R.W., 1994, Economic growth, population theory, and physiology: The bearing of long-term processes on the making of economic policy, *The American Economic Review* 84, 369–395.

Freedman, V.A. and Martin, L.G. and Schoeni, R.F., 2002, Recent trends in disability and functioning among older adults in the United States: a systematic review, *JAMA* 288, 3137-3146.

Frenck, R.W. and Blackburn, E.H. and Shannon, K.M., The rate of telomere sequence loss in human leukocytes varies with age, *Proceedings of the National Academy of Sciences of the United States of America* 95, 5607-5610.

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

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

Gavrilov, L.A. and Gavrilova, N.S., 2001, The reliability theory of ageing and longevity, *Journal of theoretical Biology* 213, 527–545.

Gompertz, B., 1825, 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* 115, 513–583.

Grossman, M., 1972, On the Concept of Health Capital and the Demand for Health, *Journal of Political Economy* 80, 223-255.

Gurven, M. and Kaplan, H., 2007, Longevity among hunger-gatherers: a cross-cultural examination, *Population and Development Review* 33, 321-365.

Jackson, S.H.D. and Weale, M.R. and Weale, R.A., 2003, Biological age—what is it and can it be measured?, *Archives of Gerontology and Geriatrics* 36, 103–115

Kirkwood, T.B.L., 1990, The Disposable Soma Theory of Ageing, in: *Genetic Effects on Ageing II*, D.E. Harrison, (Ed.), Caldwell, N.J.: Telford Press, 9-19.

Kirkwood, T.B.L.,, 1999, *Time of Our Lives: The Science of Human Ageing*, Oxford University Press, Oxford,

Makeham, W.M., 1860, On the law of mortality and the construction of annuity tables, *Journal of the Institute of Actuaries* 8, 301–310.

Masoro, E.J., 2006, Are age-associated diseases and integral part of ageing?, in: Masoro, E.J. and Austad, S.N., (eds.), *Handbook of the Biology of Ageing*, Academic Press.

Medawar, P.B., 1952, *An unsolved problem of biology*, Lewis, London.

Milne, E.M.G., 2006, When does ageing begin, *Mechanisms of Ageing and Development* 127, 290-297.

Milne, E.M.G., 2008, The natural distribution of survival, *Journal of Theoretical Biology* 255, 223-236.

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

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

Novoseltsev, V.N, 2006, Mathematical modeling in biology: Systems capable to live and die, *Automation and Remote Control* 67, 835–855.

Perks, W., 1932, On some experiments in the graduation of mortality, *Journal of the Institute of Actuaries* 63, 12-57.

Philipson, T.J. and Becker, G.S., 1998, Old-age longevity and mortality-contingent claims, *Journal of Political Economy*106, 551–573.

Oeppen, J. and Vaupel, J.W., 2002, Broken limits to life expectancy, *Science* 296, 1029–1031.

Olshansky, S.J. and Carnes, B.A., 1997, Ever since Gompertz, *Demography* 34, 1-15.

Parker, M.G. and Thorslund, M., 2007, Health trends in the elderly population: getting better and getting worse, *Gerontologist* 47, 150–158.

Riley, J.C., 2001, *Rising life expectancy: a global history*, Cambridge University Press, Cambridge.

Rockwood, K. and Song, X. and MacKnight, C. and Bergman, H. and Hogan, D.B. and McDowell, I. and Mitnitski, A., 2005, A global clinical measure of fitness and frailty in elderly people, *Canadian Medical Association Journal* 173, 489-495.

Rockwood, K. and Rockwood, M.R.H. and Mitnitski, A., 2010, Physiological redundancy in older adults in relation to the change with age in the slope of a frailty Index,, *Journal of the American Geriatrics Society* 58, 318–323.

Shock, N.W., 1981, Indices of functional age, in: *Ageing: A Challenge to Science and Society*, D. Danon et al. (eds), Oxford University Press, Oxford.

Strehler, B.L. and Mildvan, A.S., 1960, General theory of mortality and ageing, *Science* 132, 14-21.

Wachter, KW and Finch, CE., 1997, *Between Zeus and the Salmon: The Biodemography of Ageing*, Washington, DC: National Academy of Sciences.

Weale, R.A., 1997, Human biological decline and mortality rates, *Mechanisms of Ageing and Development* 97, 55-72.

Williams, G.C., 1957, Pleiotropy, Natural Selection, and the Evolution of Senescence, *Evolution* 11, 398–411.

Wilmoth, J.R., 1997, In search of limits, in: Wachter, KW and Finch, CE. (eds) *Between Zeus and the Salmon: The Biodemography of Ageing*, Washington, DC: National Academy of Sciences, 38-63.

Yaari, M.E., 1965, Uncertain lifetime, life insurance and the theory of the consumer, *Review of Economic Studies* 32, 137-1350.

2. Metabolism and Longevity

by Carl-Johan Dalgaard

In this chapter we discuss theories that relate metabolism (thus, indirectly, stature) to longevity. In essence, therefore, these are theories that highlight the importance of the environmental factors during life in explaining ultimate longevity. A sensible initial question to ask, however, is whether this avenue of influence is likely to be important. Indeed, one may imagine that genes, or influences while *in utero*, could be (much) more important determinants of life span.

Hence, we begin by briefly touching upon work aimed at examining the importance of genes, or environmental influence while being in the mother's womb, for life expectancy. That is, we look at the importance of "pre-birth" determinants of longevity. Then we turn to a discussion of "post-birth" determinants of life expectancy; the object of main interest in the present context.

2.1. Pre-birth Determinants of Longevity

There are two separate issues worth considering before we turn to the possible link between metabolism during life and longevity. The first issues is whether longevity is determined genetically. If so then an evolutionary approach is called for, rather than an approach which emphasizes the impact from the environment. The second issue

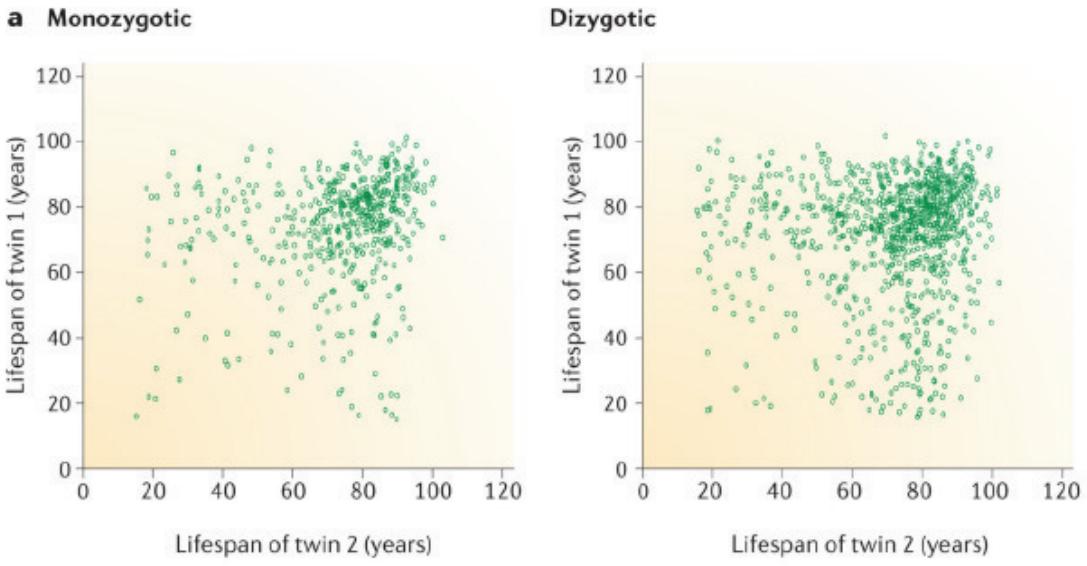
is whether the important environmental line of influence is that while *in utero*. If it is, then this too would suggest a change in focus away from post-birth determinants of longevity.

2.1.1. Genes

During the last roughly 150 years human life expectancy has risen tremendously. If measured in terms of the evolution of the country-year pair of highest life expectancy world wide, human life expectancy has risen by about 1/4 of a year, per year, on average (Oeppen and Vaupel, 2002). This comes to an increase of more than 30 years over the period in question. 150 years is, in evolutionary terms, all but an eye blink. Hence, the "growth record" in life expectancy might in itself suggest that other forces, beyond evolution, hold considerable explanatory power as drivers of human life expectancy.

Naturally, genetic traits may nevertheless be strong determinants of variation in life expectancy *within* populations and between countries.

Figure 2.1.: The Role of Genes in Determining Longevity



Source: Christensen et al., 2006.

One way to examine the first issue is to examine the longevity of "identical twins" (or, "Monozygotic twins"). If life span is genetically determined one would expect to see a positive correlation between age at death within twin pairs. Figure 2.1 shows the correlation between between identical twins, and twins originating from two separate eggs ("Dizygotic twins"). There is a visually obvious positive correlation. At the same time there is a lot of variation, or noise, to be seen in the figure suggesting that life span is also affected by non-genetic factors, such as life style and luck. Indeed, Christensen et al. (2006) argue that genes account for around 1/4 of the variation in within country populations. Hence, the genetic make-up of individuals does seem to matter, and by extension, that there is ample room for genes to influence cross-country differences in life expectancy (see also Galor and Moav, 2007).

2.1.2. Environmental Influences *in utero*

Another possibility is that life expectancy is highly influenced by external factors while *in utero*. This idea, sometimes referred to as "the fetal origins hypothesis", has had some influence on research economics. For instance, Van den Berg et al. (2006) document that children born during business cycles booms, as compared to the immediate adjacent bust, live longer. In a similar vein, Doblhammer and Vaupel (2001) document that the birth month seems to matter to longevity; children born (in the Northern Hemisphere) during October-December live longer than children born between April and June. Intriguingly, the results are shifted about 6 month when countries in the Southern Hemisphere are considered, suggesting an impact from the changing seasons. The authors document that these seasonal effects tend to disappear over time when looking at recent cohorts.

In economics the typical bio-marker for early-life influence is the birth weight of children; several studies have used this variable to examine adult outcomes while re-

ferring to the fetal origins hypothesis as theoretical foundation.[1] The basic idea which is exploited in economics is that a low birth weight signals initial health, as determined by factors that influenced the foetus.

But a strong causal link between body size at birth and the potential for successful ageing can be questioned. In an interesting paper Christensen et al. (1995) examine the life span consequences of being born as a twin rather than being "single born". If birth weight matters to longevity *per se* twins should feature higher mortality than single born offspring (later in life), as they on average are considerably smaller at birth; on average twins are about 900 g lighter at birth. Figure 2.2 shows the mortality patterns for twins and the general population in Denmark. Aside from mortality early in life, which is rather noisy due to relatively few deaths, it seems clear that there is not any particular difference in death rates to be seen. Hence, one may wonder whether birth weight really matters as such. Perhaps a low birth weight is correlated with other factors that influence longevity in a causal manner?[2]

These concerns notwithstanding, most medical researchers seem to accept the notion that early life influences does have an effect on mortality patterns later in life. The exact magnitude may be hard to assess precisely, but in their survey Vaupel et al. (1998) suggest that about 1/4 of the variation in longevity (within populations) may be accounted for by early life influences.

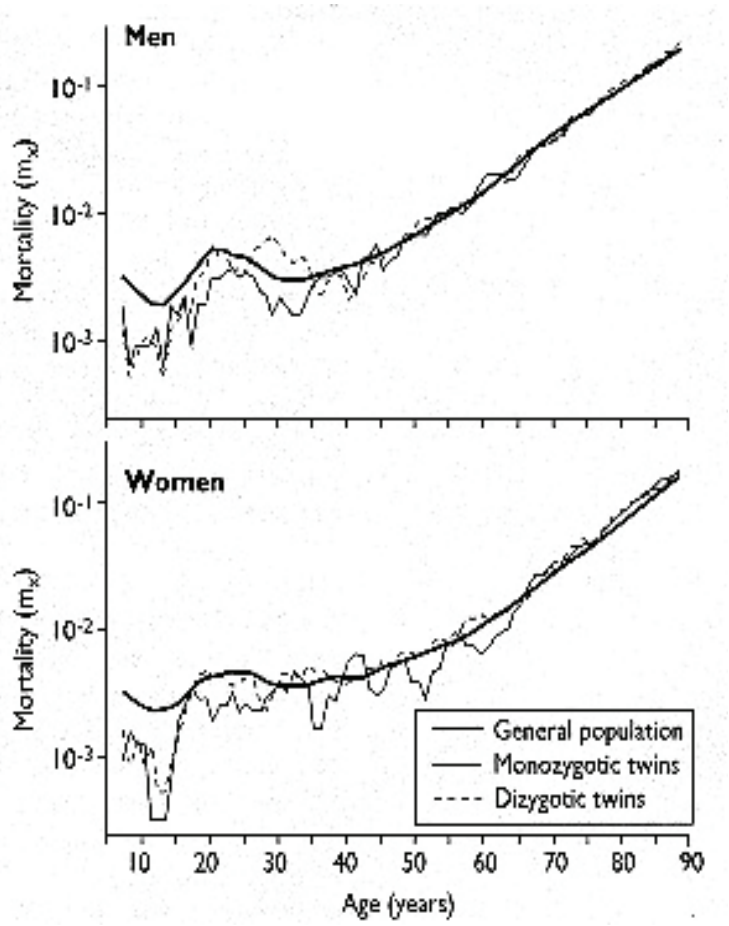
2.1.3. The Quantitative Significance of Pre-birth Factors

In summary of the discussion above, the best available evidence suggests that about 1/2 of the variation in longevity within populations is due to genetic factors, and early life influences *in utero*; roughly divided equally between the two. This leaves another

[1] See Black et al. (2007) and the references cited therein.

[2] Black et al. (2007) exploit variation in birth weight *between* twins in their study. Nevertheless, given the underlying theory for correlation between birth weight and outcomes later in life, it remains difficult to square a causal interpretation of birth weight with the lack of a difference in longevity between single born children and twins, as documented by Christensen et al. (1995).

Figure 2.2.: Mortality Schedules for Twins and Background Population



Source: Christensen et al. (1995).

50% of the variation which then must be due to environmental influences during life.

2.2. Post-birth Determinants of Longevity

2.2.1. The Rate of Living Theory

The rate of living theory can be traced back to the writings of Aristole, who observed:[3]

"*A lesser flame is consumed by a greater one, for the nutriment, to wit the smoke, which the*

[3] As quoted in Speakman (2005).

former takes a long period to expend is used up by the big flame quickly." Hence, the basic notion is that there is a certain "amount of living" available, which can be used up quickly or slowly. This biological idea has influenced economic theory. In particular, the basic theory of optimal capital utilization asserts that a more intensive use of machinery leads to higher user cost of capital through accelerated capital depreciation and thus lowered capital longevity. This is a clear application of the rate of living theory, albeit to a non-biological setting.[4]

There is biological evidence to suggest a - roughly - "fixed amount of living". We begin with the observation (which stems from the literature on allometric scaling)[5] that maximum life span scales approximately with body mass, m, in accordance with

$$T^{\max} = T_0 m^{1/4},\ T_0 > 0. \tag{2.1}$$

That is, comparing different species life span rises with body mass in accordance with the above log-linear specification. Figure 2.3 illustrates this regularity, which holds for mammals as well as birds.[6] The basic message, that larger individuals tend to live longer, have long since found its way to economics research (see e.g., Fogel, 1994).

Next, we need to invoke another scaling relationship. Namely that between *the heart rate*, h, of a mammal (at rest) and its body mass. Empirically, the heart rate declines with the size of the animal in question: $h = h_0 m^{-1/4}$ (e.g., West and Brown, 2005). As a result, *total* heart beats per life time is

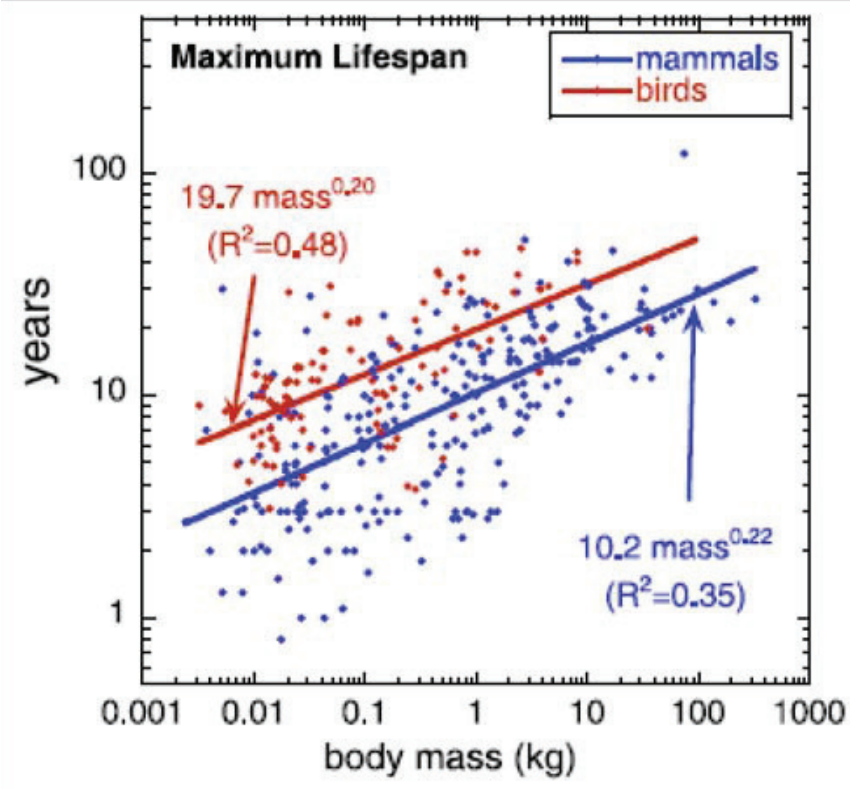
$$T^{\max} \cdot h \approx T_0 m^{1/4} h_0 m^{-1/4} = T_0 h_0,$$

[4] The classical reference is Taubman and Wilkinson (1970). The basic machinery has since then found use in business cycle research (e.g., Greenwood, Hercowitz and Huffman, 1988; Burnside and Eichenbaum, 1996) as well as in growth theory (e.g., Dalgaard, 2003; Dalgaard and Hansen, 2005; Chatterjee, 2005).

[5] "Allometric scaling" is a technique used in biology to study how selected biological variables of an organism correlate with the *size* of the organism. The size of the organism is usually summarized by its body mass. See Brown et al. (2004) for a survey.

[6] "Maximum life span" is naturally measured with some error. Usually proxies are found in recordings of life spans of animals in captivity.

Figure 2.3.: Life Span and Body Size



Source: Hulbert et al. (2007).

and thus size independent. A shrew, a human, and an elephant has about the same number of heart beats per life time. Taken literately, about 955.787.040. Superficially, this is consistent with the notion of a "given amount of living", which can be expended fast or slowly, consistent with the rate of living theory.

Extreme care should be taken in not over-interpreting this finding. For one thing the relation is approximate. Second, "maximum life span" is a difficult concept, and one may worry that especially humans may be able to manipulate it technologically. Hence, the above result should not be taken to suggest that life is exactly like a coupon ticket. But the fact that heart beats per life time is *roughly* mass independent (at least among mammals) may indicate the existence of some type of bio-mechanical constraint that characterizes most animals on the planet, and serves to limit life spans in a unify-

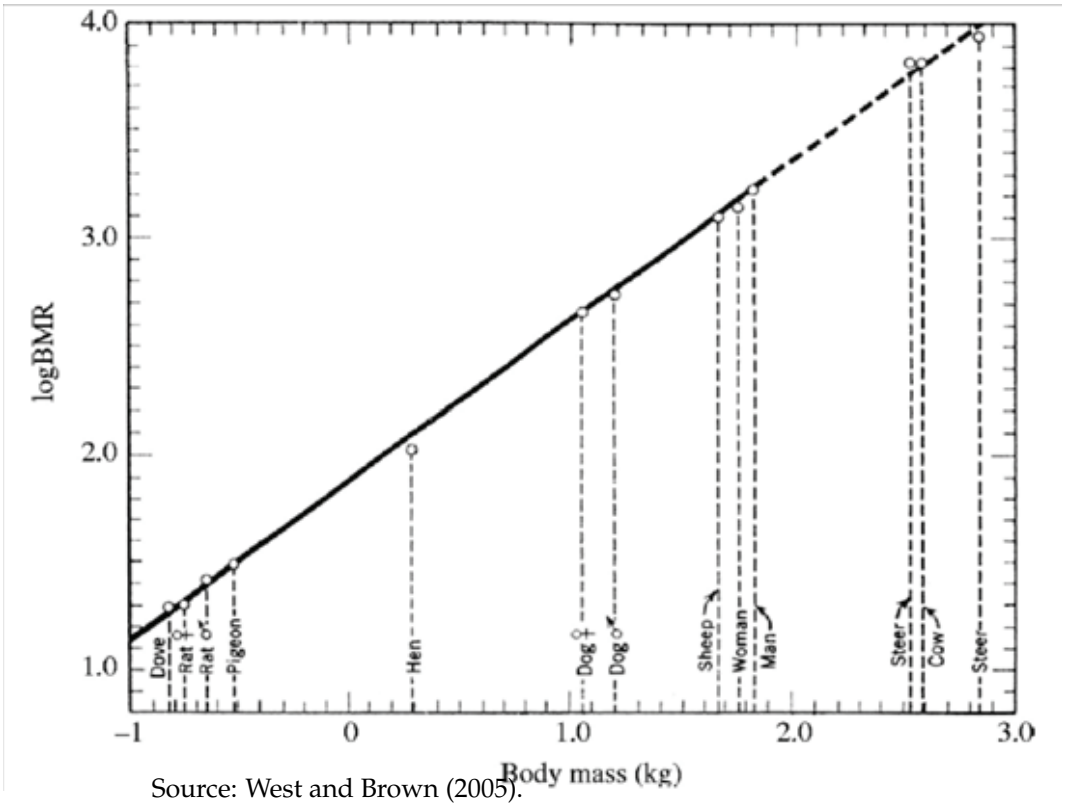
ing way across species.

The quest for a candidate bio-mechanical constraint takes us to metabolism. To understand why, one needs to introduce a third regularity, known as Kleiber's (1932) law:

$$B = B_0 m^b, \; B_0 > 0, b = 3/4, \tag{2.2}$$

where B is the basal metabolic rate (BMR) and B_0 is a species-dependent constant.[7] Thus, drawn on log-log paper the energy-body mass relationship is linear with slope of 3/4, see Figure 2.4.

Figure 2.4.: Kleiber's Law



Source: West and Brown (2005).

A slope of 3/4 has been verified by Brody (1945) for almost all terrestrial animals

[7] "Metabolism" refers to the biochemical processes by which nutrients are transformed into energy, which allows the organs of the body (i.e. ultimately the cells of the body) to function. The "basal metabolic rate" is defined as the amount of energy expended while at rest

yielding the famous "mouse-to-elephant curve". Recently biologists and physicists have managed to provide a theoretical explanation for Kleiber's law. A living organism needs to feed its cells. For that purpose energy and material is transported through hierarchically branching networks like blood vessels in mammals. The network in use, however, is not of arbitrary structure. Given that organisms have evolved through natural selection, it must be one that minimizes energy used for transport, *i.e.* one that minimizes hydrodynamic resistance. West et al. (1997) have shown that organisms, viewed as energy transporting networks that minimize energy dissipation, fulfill Kleiber's law. The interesting observation, from the point of view of the above discussion, is that the network model provides a solid foundation for a central implication of Kleiber's law: *Energy needs per unit of body mass is declining in body size* ($B/m = B_0 m^{-1/4}$). Another way to put this fact is that each cell apparently *needs to work less hard* in larger animals, as larger networks are more efficient. Perhaps this could have bearing on the fact that a longer life span is possible in larger animals, as greater body size might then translate into a slower rate of *cell* ageing? The case in favor of this assertion would seem to gain strength if the process of metabolism itself has certain life shortening implications. "The free radical theory" provides just such a causal link.

2.2.2. Free Radical Theory

The free radical theory of ageing is usually attributed to Gerschman et al (1954) and Harman (1956). The basic idea is that ageing is an a consequence of toxic *by-products* of metabolism; the generation of so-called radical oxygen species (ROS) which causes cellular damage. In reduced form, the prediction is that a higher rate of metabolism (per cell) leads to a faster accumulation of damage and thus ageing.[8]

There is evidence that the main thrust of the theory might be important in the hu-

[8] See Speakman (2005) for a detailed description of the biochemical processes at work.

man species. For one thing, it is by now fairly well established that restricting energy consumption (without causing malnutrition) slows ageing (e.g., Ramsey et al., 2000). Moreover, the free radical theory may also explain the declining death rates among the oldest old. In a recent study Ruggerio et. al (2008) examines a sample of Canadians, who were followed over a 40 year period. The study reveals two important facts: (i) BMR declines with ages, and at an accelerated rate at older ages, (ii) BMR is a predictor of time of death; higher BMR leads to an earlier time of death. Hence, the study corroborates the link between the rate of metabolism and ageing, and, moreover, provides some evidence that the theory may be able to account for the (off hand puzzling) declining death rates among the oldest old.

In order to see how the free radical theory might produce a link between life span and body size, with metabolism as the mediator, consider the following *simplified* version of "damage equation", discussed in Chapter 1.

$$\dot{D}(t) = \mu D(t), \; D(0) \text{ given.} \tag{2.3}$$

As before, death is assumed to occur when the damage index reaches a critical level: $\bar{D}$. The rate at which damage is accumulated is μ. Now, suppose this rate is *determined* by the rate at which the average cell is damaged, which, according to the free radical theory depends on metabolism per cell.[9] Then we have

$$\mu = \mu_0 \frac{B}{N},$$

[9] Going one step back to the underlying premises: (i) We are assuming that damage to each cell grows exponentially over time at the rate μ, (ii) that total damage to the organism as a whole is damage to each cell, multiplied by total cells in the body. As long as body mass is constant, and thus the organism has a constant amount of cells, this leads to equation (2.3). Hence, the current considerations are likely more relevant for an intermediate age range, thus excluding children (growing body size), and the very old (declining body size) from consideration. Moreover, if cellular damage is the cause of the empirically observed growing death probability during life, then μ can also be thought of as the growth rate in the death rate in a Gompertz equation. Note that the Gompertz equation is a fair description of the data precisely in the "intermediate" age bracket.

where μ_0 is an appropriately chosen (mass independent) constant, B is basal metabolism of the entire organism, and N the total number of cells in the body. Since the total mass of an organism is the weight of each cell multiplied by the number of cells, we may write

$$\mu_0 \frac{B}{N} \propto \mu_0 \frac{B}{m} \propto m^{-1/4},$$

where the last step employs Kleiber's law (equation (2.2)). Finally, maximum life span, $T^{\max}$, can be seen as the time it takes to reach $\bar{D}$, starting at $D(0)$.[10] Using equation (2.3) and the equation above:

$$T^{\max} = \log\left[\frac{\bar{D}}{D(0)}\right]\mu^{-1} \propto T_0 m^{1/4}. \tag{2.4}$$

Hence, by combining the basic free radical theory, the notion of increasing biological damage (frailty) and Kleiber's law, one is able to reproduce the basic biological correlation between body mass and (maximum) life span, see equation (2.1).

In order to turn the above into a dynamic model of longevity, it is necessary to consider the dynamics of body growth. A general model of body growth is developed by West et al. (2001); this theory has recently been introduced into the economics literature by Dalgaard and Strulik (2010). The point of departure for the West et al. model, is the following energy conservation equation:

$$E(t) = \kappa N + \psi \dot{N},$$

where $E(t)$ is energy intake during a period, κN are energy needs for maintenance of cells (N) and ψ captures the energy costs of cell creation. Since each cell has a certain

[10] We are necessarily dealing with maximum life span as we are recording the time it takes until the organism *must* collapse due to accumulated damage. In practise, of course, any given mammal may die before then, due to disease, predation or bad luck.

weight, ω say, we can convert the right hand side into body mass:

$$E(t) = \frac{\kappa}{\omega} m(t) + \frac{\psi}{\omega} \dot{m}(t) \Leftrightarrow \dot{m} = \frac{\omega}{\psi} E(t) - \frac{\kappa}{\psi} m(t).$$

In West et al. energy intake is assumed to equal energy requirements as given by Kleiber's law; Dalgaard and Strulik determine it through optimization. In the former case, however, the final dynamic equation governing body size becomes (using equation (2.2) in the equation above):

$$\dot{m} = a m^{3/4} - b m(t), \tag{2.5}$$

with a and b being appropriately defined parameters. Biologically, a and b are species specific, and thus map into species specific asymptotic body sizes. Moreover, by equation (2.4), they will map into species specific longevity. In human populations, it may be reasonable to assume that asymptotic body size is somewhat amendable to an optimally chosen trajectory for energy intake, and thus subject to optimization. In addition, however, it will be meaningful to consider scenarios where $T^{\max}$ is affected by technological forces (broadly defined) in addition to biological forces.

2.2.3. Teleomere Theory

The idea of Harman (1956), discussed above, that metabolism produces oxygen free radicals which generate cell damage, suggests that something like a biological clock may exist: the process of converting nutrient intake into muscular work effort inevitably leads to ageing. The later work of Heyflick (1965) can also be viewed as pointing to the existence of a biological clock, but via a seemingly different mechanism; Heyflick documented that human cells ceases to replicate after a certain amount of cell doubling. In other words: there seems to be an upper limit to the reproductive life span of cells: the so-called "Heyflick limit".

A likely proximate determinant of the Heyflick limit is the gradual reduction in "teleomere length" . Telomeres are protein caps that are located at the end of chromosomes. The role of these end-caps is to protect the intregrity of the chromosomes during cell division. However, as cells divide the length of the telomeres decline. Eventually their length reach a critical size, after which cell division ceases; the Heyflick limit is reached (see Olovnikov, 1996).

A layman's way to think of the role of telomeres is the following analogy. Think about a DNA sequence as a written text on a piece of paper; DNA replication is then conceived as a process whereby the text is photocopied, and then the copy is copied and so forth. The telomeres are then analogous to the white spaces in between individual letters and words. Now, every time the text is copied the letters of the text will tend to become increasingly blurred; the distance between words and letters will seem shorter. Eventually, the text becomes unreadable; further copying is pointless. Much like the white spaces in between letters and words is crucial for the readability of a text, teleomeres are critical for a meaningful process of cell division.

Some attempts have been made to model this process mathematically. For instance, Aviv et al. (2003) postulate an equation governing the evolution of telomere length, which (in continous time) can be written

$$\dot{L}(t) = -\phi_1 \cdot \left[\phi_2 \frac{\dot{m}(t)}{m(t)} + \phi_3 m(t) \right],$$

where $\dot{L}$ is the change in telomere length from one instant to the next, m is body mass and $\phi_1 - \phi_3$ are positive parameters. Hence, this equation captures that periods of body growth, which are associated with rapid cell division, work to run down telomere length. Combing this equation with equation (2.5) provides a simple model of the evolution of telomere length, for initial body size and telomere size given. Finally, Aviv et al. (2003) postulate a link between telomere length, and the probability of remaining disease free; the probability declines as L shrinks. The authors argue that the model is

able to shed light on various phenomena, like the life shortening effect of obesity.

While body growth, and thus cell division, is one way in which telomere length is reduced, it is increasingly understood that there is another force that also leads to telomere shortening: oxidative stress (von Zglinicki et al., 2000). Hence, the same mechanism that leads to cell damage, as discussed above, also has the effect of speeding up cell senescence. From a formal perspective; with this notion on telomere erosion in mind, the theoretical approach from the last section may be a reasonable reduced form representation here as well.

There is indeed evidence to suggest that teleomere shortening is empirically relevant in the context of human longevity. For one thing, accelerated telomere shortening has been shown to be the result of smoking and obesity (Valdes et al., 2005), known sources of oxidative stress and accelerated ageing. Epel et al. (2004) find evidence that *psychological* stress is significantly associated with heightened oxidative stress and shorter telomere length, and Kimura et al. (2008) show that telomere length indeed holds predictive power vis-a-vis mortality, in a sample of elderly twins.

2.3. References

Aviv, A. and Levy, D. and Mangel, M., 2003, Growth, telomere dynamics and successful and unsuccessful human ageing, *Mechanisms of ageing and development* 124, 829–837.

Black, S.E. , Devereux, P.J. and Salvanes, K.G., 2007, From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes, *Quarterly Journal of Economics* 122, 409–439.

Burnside C., M. Eichenbaum and Rebelo, S., 1996, Sectoral Solow Residuals. *European Economic Review* 40, 861-69.

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

Brown, J., Gillooly, J., Allen, A., Savage, V.M. and West, G., 2004, Towards a metabolic theory of ecology, *Ecology* 85, 1771-89.

Chatterjee, S., 2005, Capital utilization, economic growth and convergence, *Journal of Economic Dynamics and Control* 29, 2093-2124.

Christensen, K. and Vaupel, J.W. and Holm, N.V. and Yashin, A.I., 1995, Mortality among twins after age 6: fetal origins hypothesis versus twin method, *British Medical Journal* 310, 432-36.

Christensen K, Johnson T.E., and Vaupel J.W., 2006, The quest for genetic determinants of human longevity: challenges and insights, *Nature Reviews Genetics* 7, 436-448.

Dalgaard, C-J., 2003, Idle Capital and long-run productivity, *The B.E. Journal of Macroeconomics*: Vol. 3 : Iss. 1 (*Contributions*) article 6.

Dalgaard, C-J. and Hansen, J.W., 2005, Capital utilization and the foundations of club convergence, *Economics Letters* 87, 145-152.

Dalgaard, C-J., and Strulik, H., 2010, The physiological foundation of the wealth of nations. Working Paper (University of Copenhagen).

Doblhammer, G. and Vaupel, J.W., 2001, Lifespan depends on month of birth, *Proceedings of the National Academy of Sciences of the United States of America* 98, 2934-39.

Epel, E.S., Blackburn, E.H., Lin, J., Dhabhar, F.S., Adler, N.E., Morrow, J.D., and Cawthon, R.M., 2004, Accelerated telomere shortening in response to life stress, *Proceedings of the National Academy of Sciences* 101, 17312-15.

Fogel, R.W., 1994, Economic growth, population theory, and physiology: The bearing of long-term processes on the making of economic policy, *American Economic Review* 84, 369-395.

Galor O. and Moav, O., 2007, The Neolithic Revolution and contemporary variations in life expectancy. Working Papers 2007-14, Brown University.

Greenwood J., Hercowitz Z., and Huffman, G.W., 1988, Investment, capacity utilization and the real business cycle, *American Economic Review* 78, 402-17.

Gerschman, R., Gilbert, D.L., Nye, S.W., Dwyer, P., and Fenn, W.O., 1954, Oxygen poisoning and x-irradiation: a mechanism in common, *Science* 119, 623-626.

Harman, D., 1956, Aging: a theory based on free radical and radiation biology, J*ournal of Gerontology* 11, 298-300.

Hayflick, L., 1965, The limited in vitro lifetime of human diploid cell strains, *Experimental Cell Research* 37, 614-636.

Hulbert A.J., Pamplona, R., Buffenstein, R., and Buttemer, W.A., 2007, Life and death: Metabolic Rate, membrane composition,and life span of animals, *Physiological Reviews* 87, 1175–1213.

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

Kimura, M., Hjelmborg, J.B., Gardner, J.P., Bathum, L., Brimacombe, M., Lu, X., Christiansen, L., Vaupel, J.W., Aviv, A., and Christensen, K., 2008, Telomere length and mortality: a study of leukocytes in elderly Danish twins, *American Journal of Epidemiology* 167, 799-806.

Oeppen, J. and Vaupel, J.W., 2002. Broken limits to life expectancy. *Science* 296, 1029–1031.

Olovnikov, A.M., 1996, Telomeres, telomerase, and ageing: origin of the theory, *Experimental gerontology* 31, 443–448.

Ramsey, J., Harper, M-E., and Weindruch, R., 2000, Restriction of energy intake, energy expenditure and ageing, *Free Radical Biology & Medicine* 29, 946-968.

Ruggerio, C., Metter, E.J., Melenovsky, V. , Cherubini, A., Najjar, S.S., Ble, A., Senin, U., Longon, D. L., and Ferrucci, L., 2008, High basal metabolic rate is a risk factor for mortality: The Baltimore longitudinal study of aging, *Journal of Gernotology: Medical Science* 63A, 698-706.

Speakman, J.R., 2005, Body Size, energy metabolism and life span. *Journal of Experimental Biology* 208, 1717-1730.

Taubman P. and M. Wilkinson, 1970, User cost, capital utilization and investment theory, *International Economic Review* 11, 209-15.

Valdes, A.M., Andrew, T., Gardner, J. P., Kimura, M., Oelsner, E., Cherkas, L.V., Aviv, A., and Spector, T.D., 2005, Obesity, cigarette smoking, and telomere length in women, *The Lancet* 366, 662-664.

Vaupel, J.W., Carey, J.R., Christensen, K., Johnson, T.E., Yashin, A.I., Holm, N.V., Iachine, I.A., Kannisto, V., Khazaeli, A.A., Liedo, P., Longo, V.D., Zeng, Y., Manton, K.G., Curtsinger, J.W., 1998., Biodemographic trajectories of longevity, *Science* 280, 536.

Van den Berg, G.J., Lindeboom, M., Portrait, F., 2006, Economic conditions early in life and individual mortality, *American Economic Review* 96, 290–302.

von Zglinicki, T., Pilger, R., and Sitte, N., 2000, Accumulation of single-strand breaks is the major cause of telomere shortening in human fibroblasts, *Free Radical Biology & Medicine* 28, 64–74.

West, G. and Brown, J.H., 2005, Review: The origin of allometric scaling laws in biology from genomes to ecosystems: towards a quantitative unifying theory of biological structure and organization, *Journal of Experimental Biology* 208, 1575-1592.

West G.B., Brown, J. H., Enquist, B.J., 1997, A general model of the origin of allometric scaling laws in biology, *Science* 276, 122-126.

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

3. Evolutionary Theories of Ageing

by Michael Kuhn and Alexia Prskawetz

This chapter is devoted to evolutionary theories of ageing that aim for an explanation of the "genetic architecture" of the life history with a view to understanding why ageing and senescence occur. Life histories are defined as cycles of maturation, fertility and mortality. A key question is to understand why ageing, i.e. decreasing fertility and increasing mortality with age, has evolved, although it is clearly bad at the individual level. We begin by reviewing the 'classical' evolutionary theories of ageing: mutation accumulation, antagonistic pleiotropy and disposable soma theories. To some extent, these theories fail to explain, however, decreasing mortality at young ages and/or post reproductive survival, phenomena that occur in many species, in particular in birds and mammals, including humans. Newer theories seek to address these issues by incorporating transfer-giving to children and more detailed modeling of the physiological mechanisms underlying the process of ageing. We consider first an extension to mutation-selection theory including inter-generational transfers and then turn to discuss life history models that explicitly include optimization of a fitness criterion as an evolutionary objective function. We discuss how these models have been used to derive realistic mortality patterns and other features of ageing.

3.1. Introduction

This chapter explores approaches from evolutionary biology to explain the phenomenon of ageing. Formalizations of the evolutionary advantage of ageing date back to Weismann (1891) who argued that ageing might help to keep population size in bounds and avoid overcrowding as well as to allow for generational replacement and hence adaptation to changing environments. However his theory was dismissed rather soon. The theory by Weismann only offers a purpose of ageing but no insight into the mechanisms of ageing.

More recently, Partridge and Barton (1993, p. 305) argue that ageing[1] may be seen as an evolutionary paradox: "If organisms can function well in youth, why can they not continue to do so in old age?" Since the genetic contribution to future generations is reduced by ageing, natural selection should therefore work against the force of ageing. Similarly Kirkwood (2002, p.737) state that "evolution theory argues against programmed ageing, suggesting instead that organisms are programmed for survival, not death."

While mechanistic explanations relate the process of ageing to damages of cells, tissues, organs, etc., evolutionary theories of ageing aim to explain the variation across species of avoiding or reducing the damages, i.e. why ageing may (or may not) occur.

Evolutionary theories help to explain the "genetic architecture" of the life history, i.e. when maturation, fertility and death occur. Against the backdrop of increasing age at childbirth and continuing improvements in old age survival among humans in industrialized countries, evolutionary theories of ageing may help to explain the underlying mechanisms of these changes. Moreover, by generating insights into these mechanisms, evolutionary theories of ageing may help to understand the constraints

[1] While Partridge and Barton use the term ageing and senescence interchangeably, Baudisch 2008, 2009 (based on Medawar 1952) distinguishes between the two concepts. While ageing can represent maintenance, deterioration or improvement with age, senescence denotes deterioration or decay, i.e. increasing mortality and/or decreasing fertility with age.

and opportunities of ageing in the future.

Following Partridge and Barton (1993) two explanations of the evolution of ageing can be distinguished. Both explanations are built on the assumption that the force of selection declines with age. The **mutation accumulation** theory of ageing (Medawar 1952, Hamilton 1966) argues that detrimental mutations that show up only late in life (i.e. after the age of reproduction) accumulate and are less likely selected out by nature. By contrast, deleterious mutations early in life (before the age of maturity) will be selected. For organisms that die young, the force of selection to oppose mutations within the genome that lead to deleterious effects in old age is therefore low. The second evolutionary explanation of ageing is based on an **optimality criterion**. In these models, forces of evolution are assumed to yield optimal life-histories in terms of an optimal trade-off of fertility and mortality, within the constraints of specific intrinsic physiological and extrinsic environmental resources.

Optimality theories of ageing include the **antagonistic pleiotropy model** by Williams (1957) and the **disposable soma theory** by Kirkwood (1977).

Antagonistic pleiotropy is similar to the mutation accumulation theory of Medawar but introduces life-history trade-offs. It assumes that some genes exhibit effects on fertility and mortality that turn from positive to negative (or vice versa) with progressing age. Such genetic trade-offs are seen as an important cause of ageing. For instance, if the evolutionary aim is to have a high number of offspring, large positive effects at younger ages might have a more important effect as compared to detrimental effects at later ages where survival and fertility is lower. Hence, natural selection declines with age.

The disposable soma theory is built on the assumption that longevity requires investment in somatic maintenance and repair and these investments compete with investments in growth and reproduction. Hence, genes that affect the maintenance and durability of the soma will also influence longevity. Kirkwood (1977) explicitly distin-

guishes between non reproducible (somatic) and reproductive tissues. While the former only serves a single generation it is the latter that determines the cell lineage (germ line) accounting for variability across generations. The soma therefore only helps to transport the genetic codes across generations. Since the costs of repair of the soma are too high, evolution trades off senescence of the soma with persistence of the germ line.

In life-history optimization models evolutionary success is typically measured by fitness or reproductive success of a genotype, i.e. the intrinsic rate of population increase r as implicitly defined by the Lotka equation

$$1 = \int_0^\infty e^{-ra} l(a)\, m(a)\, da, \tag{3.1}$$

where $l(a)$ and $m(a)$ denote the survival function and fertility rate, respectively, that apply to age a and the term e^{-ra} discounts later born offspring by the population growth rate. The unique r that solves this equation is termed Lotka's r or alternatively the intrinsic population growth rate r (i.e. intrinsic to the given survival and fertility schedule). In stationary populations where the optimal growth rate is zero, optimal trajectories of fertility and mortality can be derived by maximization of the net reproduction rate (the expected number of offspring per individual) instead as given by:

$$R = \int_0^\infty l(a)\, m(a)\, da. \tag{3.2}$$

Partridge and Barton (1996) suggest to apply the concept of reproductive value (Fisher 1930) to measure senescence. Closely linked to fitness, the reproductive value at age a measures the remaining reproductive contribution of an individual at age a and is given by

$$v(a) = \frac{e^{ra}}{l(a)} \int_a^\infty e^{-rx} l(x)\, m(x)\, dx, \tag{3.3}$$

where $l(x)$ and $m(x)$ denote the survival function and fertility rate, respectively, that apply to age x. Future reproduction is discounted by the intrinsic population growth rate r through the term e^{-rx}. The constant term $\frac{e^{ra}}{l(a)}$ indicates the survival to age a.[2] As shown in Baudisch (2008, p. 6 ff), senescence occurs when the reproductive value declines with age. Since senescence involves the age specific dynamics of both mortality and fertility, it is the relation between those two processes that will determine whether senescence occurs. More specifically, if the relative change in mortality is greater than the relative change in fertility at a specific age, senescence will happen (Baudisch 2008, p. 8).[3]

Empirical tests of the evolutionary theories of ageing are numerous (see Partridge and Barton 1993 for a review) and commonly employ model organisms such as fruitflies. Organisms that are faced with low extrinsic risks of mortality and/or show up increasing fertility with age, are those that exhibit high maximum lifespan. These characteristics obviously reflect selection for high levels of maintenance and hence for long life spans. However, distinguishing between the specific theories is more difficult.

After this brief review of the 'classical theories' of evolutionary explanations of ageing, in the following subsections we introduce more recent contributions that aim to explain decreasing mortality at young ages and/or post reproductive survival. These phenomena occur in many species, in particular in birds and mammals, including humans, and cannot be explained by the classical theories introduced so far. Newer theories seek to address these issues by incorporating transfer-giving to children and more detailed modeling of the physiological mechanisms underlying the process of ageing. We consider first an extension to mutation-selection theory including intergenerational transfers and then turn to discuss life history models that explicitly include optimization of the reproductive value as an evolutionary objective function. We

[2] Note that for $r = 0$ the reproductive value at age $a = 0$ equals the net reproduction rate R.

[3] Taylor et al. (1974) have shown that maximization of the net reproduction rate is mathematically equivalent to maximimizing the ultimate rate of increase r, i.e. fitness.

discuss how these models have been used to derive the scope for negative senescence and realistic U-shaped mortality patterns.

3.2. The Role of Intergenerational Transfers in Mutation-Selection Equilibrium

Lee (2003) seeks to provide an explanation for both the U-shaped mortality schedules and post-reproductive survival that are found in many species, including humans.[4] In so doing he explicitly accounts for the relevance of inter-generational transfers, in particular of transfers from older individuals (parents but possibly also grand-parents) to the young. Such transfers, to be understood in a wide sense as the provision of care and food (potentially even within utero), improve the quality of the offspring in terms of their capacity to survive, reproduce and produce resources (e.g. by foraging).

In the presence of transfer-giving a quantity-quality trade-off arises with respect to children not dissimilar to the one studied by Becker and Lewis (1973): A larger number of children through higher fertility and/or lower child mortality implies that each surviving child receives a lower volume of transfers. But then fertility and mortality during childhood take on an ambiguous role for the maximization of fitness as in (3.1). Consider the death of a child during its early development. Here, the loss of the investments accumulated in that child is small and likely to be over-compensated by the resources freed up for improving the quality of the remaining offspring. Thus, *early child mortality* may well be a beneficial genotype that is *selected in*. In contrast, the death of an adolescent at the point of turning into an adult would constitute a considerable loss of reproductive and productive capacity. Furthermore, as such an individual is

[4] The life-history models by Kaplan and Robson (2002, 2009) and Robson and Kaplan (2003, 2007) as well as those by Chu and Lee (2006) and Chu et al. (2008) also include intergenerational transfers. However, as these models follow the optimal life history approach we defer their discussion to subsection 3.3.

no longer prone to receive transfers, its death would not free up any additional resources. Consequently, *adolescent mortality* should be a trait that is *selected out*. Thus, the quality-quantity trade-off can, indeed, explain decreasing mortality over the period of childhood up to the point of adolescence. Likewise, by implying a quality disadvantage, too high a level of fertility may be a trait that is selected out within species for whom transfers towards their children are important.

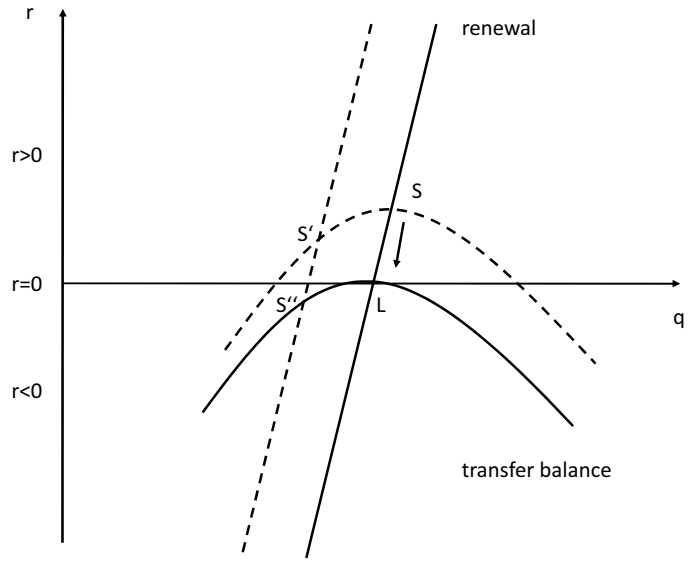
At the other end of the life-line, the provision of transfers would imply a contribution to fitness even by post-reproductive individuals. Thus, in contrast to the earlier theories that cannot explain extended periods of post-reproductive survival, this phenomenon is consistent with the role of transfer giving. Indeed, *mortality* should be *selected out* for those older individuals that are still able to contribute towards transfers.

Building on this intuition, Lee (2003) studies a population equilibrium in terms of net population growth, r, and the quality of individuals, q, when quality can be enhanced by intergenerational transfers (see Figure 3.1).[5] Equilibrium points in (q, r)-space (e.g. S, S′, S″, L) are defined by the intersection of a renewal schedule and the locus at which transfers balance out across generations at each point in time. The positive slope of the renewal schedule follows as a higher average quality in the population allows higher survival $l(a)$ and/or fertility $m(a)$ in the fitness equation (3.1), thus allowing for a higher intrinsic rate of population growth, r. The quality-quantity trade-off implies a concave shape of the transfer balance. A higher rate of growth, r, brings with it a larger share of young individuals in the population. As young individuals receive rather than contribute transfers, fewer resources are available per capita and quality is lower. Hence for high levels of quality the transfer balance is downward sloped. Starting from low levels, however, improvements in q allow greater productivity and thus the maintenance of a balanced budget even at an increased rate of population

[5] Lee (2003) himself refers to the 'level of consumption' (γ) rather than 'quality'.

growth.

Figure 3.1.: Tradeoff Between Population Growth and Quality of Individuals



Source: adapted from Lee (2003).

In the short-run, population equilibrium may imply positive, negative or zero population growth as long as the transfers are balanced (i.e. any of the equilibria S, S', S'' and L are feasible). Here, a population carrying a certain mutation would gradually replace the original population if and only if it exhibits stronger growth. In the long-run, only zero growth, $r = 0,$ is feasible, however. Any shrinking population ($r < 0$) would eventually vanish, whereas sustained growth at $r > 0$ is unfeasible as it would at some point hit the carrying capacity of the ecosystem. At that point the transfer balance shifts downwards, where for any level of quality the reduction in available resources stifles growth (e.g. the shift from S to L). Lee (2003) examines the effects on short-run equilibrium growth and quality of a variety of mutations in fertility and mortality, which in some cases run counter to the 'classic' selection theory. For instance, starting from S and for a given transfer balance, a mutation leading to greater fertility would

in the new equilibrium S' result in both lower quality and growth.[6] More prominently, however, Lee (2003) finds that in a long-run equilibrium with $r = 0$, selection can only work through the transfer balance and not through the 'classical' channels of renewal. Selection should ensure that the long-run equilibrium is characterized by an optimal transfer-structure.[7] This implies that any long-run equilibrium point must lie at the peak of the balanced budget line (as e.g. L). But then any further mutation is inconsequential unless it affects the budget balance.[8] This leads Lee (2003) to the conclusion that transfers, not births, shape the process of ageing in social, i.e. transfer-giving, species.

Lee (2008) simulates the evolutionary process for a single-sex hunter-gatherer society in the presence of transfers. The underlying model is somewhat more general than the one in Lee (2003) and allows for non-stable age-structures and, importantly, for different structures of kinship. As it turns out, the results of Lee (2003) are largely confirmed, namely that the presence of transfers explains declining child-mortality and post-reproductive survival. Kinship, however, plays a crucial role. More specifically, the impact of transfer giving on the mortality pattern is the more pronounced the closer is kinship as measured by the degree to which resources are shared within the kin group only. If resources are shared at population level, then the 'classical' result reemerges: mortality is flat up to the age of sexual maturity and there is no post-reproductive survival.

Despite their prominent role in qualifying the classical evolutionary theory of ageing, Lee's (2003, 2008) contributions have a number of drawbacks arising from the fact

[6] In fact, an increase in fertility would also shift downwards the transfer budget, implying that q and r will lie at even lower levels, as e.g. in S".

[7] This reasoning begs an interesting analogy: Optimality theories of ageing are equivalent to social planner solutions in economics; whereas theories of mutation-selection balance are more about the adaptive processes that may (or may not) lead to an optimal long-run equilibrium (as e.g. in Lee 2003). Thus, to a degree they resemble the 'decentralized' solution in economics.

[8] Consider, e.g. an upward shift of the renewal line (e.g. due to an increase in fertility) starting from L. The new equilibrium compatible with the transfer balance is S", where $r < 0$. The same obviously holds for a downward shift in renewal.

that by taking a macro-perspective they make a number of ad-hoc assumptions about the physiological underpinnings (e.g. the relationships between mortality, fertility and production and between all of these and 'quality'). Thus, the trade-offs involved with the resource allocation within the individual remain obscure and the optimal allocation of resources is not characterized. A precise statement of the age-schedules of fertility, mortality and transfers that would maximize fitness in (3.1) is also ruled out.

A more detailed characterization of the resource allocation and its age-structure can be gained by the use of optimal life-history models. These models focus at the optimal long-run equilibrium (as e.g. point L in Figure 3.1) straight away without asking, however, as to whether or not such a state can be attained. The optimal age-schedules of fertility, mortality and, where relevant, transfer-giving are then derived as the solution to a problem where fitness in (3.1) is maximized subject to physiological and transfer constraints.

3.3. Optimal Life-history Theories of Ageing

Optimal life-history theories are based on the trade-off between reproduction and survival. From an evolutionary point of view, the strategy that yields the highest fitness (in terms of maximizing lifetime reproductive success) will be selected. The specific assumption on the trade-off will determine the age pattern of mortality and fertility that is chosen by evolution. An excellent review on how different assumptions on the trade-off between survival and reproduction shape the patterns of ageing (allowing for senescence, sustenance, or inverse senescence) is given in Baudisch (2009). "Crucial points in the model assumptions are linearity vs. non-linearity in the trade-offs, inclusion or exclusion of mediating variables that determine either or both mortality and fertility, endpoint conditions of the problem's time horizon, future returns to current investment reflected in the potential for indeterminate growth, and constraints on

the qualitative shape of mortality and fertility patterns." (Baudisch 2009, p. 9).

3.3.1. The Case for Negative Senescence

In the following we review a model by Vaupel et al. (2004) that introduces an intermediate variable (the physiological state size) that determines fertility and mortality. Size itself evolves dynamically over the life-history and is determined by optimal investment into the maintenance of size and reproduction. Depending on the specific dependence of fertility and mortality on size, various patterns of ageing are possible.

Based on optimization models of life-history strategies Vaupel et al. (2004) present an evolutionary explanation of age specific mortality and fertility. In particular, they question Hamilton's theory (Hamilton 1966) that senescence (defined as age-related deterioration of the organism) cannot be avoided. Observations from human, animal and plant populations provide evidence against Hamilton's theory since mortality declines from age of conception to age of sexual maturity and may thereafter exhibit increasing, decreasing or stable age trajectories (see Figure 1 in Vaupel et al. 2004). Hence, mortality may decline and fertility may increase after the age of reproductive maturity. Such a pattern would constitute negative senescence.

To conciliate negative senescence with an evolutionary optimal strategy the authors refer to a model of size-dependent mortality. Under conditions of increasing size with age and negative dependence of mortality on size, negative senescence (i.e. decreasing mortality with age) could be the result. The evolutionary model chosen involves maximization of the net reproduction rate at age zero, R, as given by equation (3.2).

The authors assume that an individual's size (encompassing its vitality and strength) determines its resources that are split between growth and maintenance on the one hand and reproduction on the other hand. Starting from the *maintenance mode*, i.e. a state where size and hence mortality and fertility are constant over age, they derive conditions under which it would be optimal to temporarily invest more in growth

and sacrifice reproduction for the sake of larger size and hence higher reproduction and lower mortality later on. The temporary phase of higher investment in growth coincides with the period of negative senescence. Formally, the condition for negative senescence is an increase in the reproductive value of the latter strategy (investing temporarily in growth) as opposed to the maintenance strategy.

The authors next introduce a more general framework assuming that size is a function of age and its change over time is determined by size specific deterioration (higher size implies more complexity and consequently higher depreciation) together with age specific investments (as measured by a specific fraction $0 \leq \pi(a) \leq 1$ of the age specific size). Hence, damage and repair determine the stage of the size variable. The remaining investments $1 - \pi(a)$ determine reproduction. Mortality is assumed to be inversely related to the size of the organism. Unless the authors assume a non-linear fertility function (i.e. the control $\pi(a)$ enters the fertility function in a non-linear way), negative senescence does not occur. The only possible outcome is a strategy of first putting all investment into growth and then switching to the maintenance strategy at the onset of reproduction. Though negative senescence is not the standard outcome it is definitely not positive senescence that can be observed in the theoretical model. This is because the maintenance mode implies constant size and hence unchanging mortality. To obtain positive senescence the authors introduce a model with determinate growth as opposed to the models of indeterminate growth considered so far. For this purpose they introduce a second state variable in addition to size, describing the functionality of the body. The product of the size and functionality variable denotes the vitality of the individual.

In summary, the theoretical models outlined in Vaupel et al. (2004) imply that (i) senescence is likely for species that obtain their maximum size around the age of maturity and for which fertility declines with age. (ii) If fertility does not decline with age and if maximum size is attained near but not at the age of maturity, senescence is

not likely, and finally (iii) negative senescence may be characteristic for species where fertility increases with age and maximum size of the species is obtained far above the age of maturity.

In terms of the trade-off between reproduction and maintenance the findings are that inverse senescence is related to concave trade-offs, convex trade-offs lead to senescence and linear trade-offs imply sustenance as characterized by constant mortality and fertility (see Baudisch 2009b, page 9).

3.3.2. Explaining U-shaped Mortality and Post-reproductive Survival: Models Including Intergenerational Transfers

In the following, we focus on a number of life-history models drawing on intergenerational transfers as a key factor in explaining U-shaped mortality profiles and post-reproductive survival. As it turns out, these models yield important lessons for ageing.[9]

Kaplan and Robson (2002) and Robson and Kaplan (2003) seek to explain the co-evolution of brain size and longevity as it has occurred in primates, including humans. The brain is interpreted as a capital stock that requires a costly investment in life's early stages and yields returns later on, flowing e.g. from an improved capability in foraging. Mortality renders uncertain the returns to investments in brain size, similar to investments in financial or human capital within economic life-cycle models with uncertain survival (e.g. Ehrlich 2000, Becker 2007). In analogy to these models, the desire

[9] U-shaped mortality is also explained by uncertainty with respect to mortality (Sozou 1998, Sozou and Seymour 2003) and size effects (Tuljapurkar and Boe 1993, Tuljapurkar 1997). In the former case, the gradual resolution of initial uncertainty with respect to (e.g. extrinsic) mortality leads to increasing investments into survival and thus declining mortality during childhood. When size enhances fertility or mortality, then the fitness returns from an individual increase in its size, implying again declining mortality during the growth phase. Tuljapurkar and Boe (1993) and Tuljapurkar (1997) also show that stochastic fertility may lead to multiple equilibria in the sense of several different phenotypes generating the same (maximal) level of fitness. In this case, mortality may exhibit a relatively flat profile late in life.

to self-insure against the mortality risk then generates an incentive to invest in survival, where improved survival boosts, in turn, investments into 'brain' capital. Brain size and survival turn out to be complements, again very much in analogy to the complementarities between health and education (Becker 2007). Due to this complementarity one would, indeed, expect brain size and longevity to coevolve. In particular, as the productivity of the brain increases with the complexity of the environment, it follows that species living in more complex environments develop both larger brain sizes and lower mortality. Indeed, these patterns are confirmed by empirical evidence.[10]

Kaplan and Robson (2002) and Robson and Kaplan (2003) start out from a typical life-history model with a fitness objective as in (3.1) but then go on to show that the problem is equivalent to an economic problem in which an individual's life-time energy surplus is maximized subject to a per period production and survival function. The period production of energy increases in brain size and is used to 'finance' investments in survival. The remaining (net) energy can be transferred over time and is converted into fertility. The brain is treated as a capital asset, serving to increase per period productivity. In Robson and Kaplan (2003) investments into the brain are subject to increasing costs and are, therefore, stretched out over several periods (of childhood). The model is solved for the optimal steady-state level of capital, implying an investment path during childhood, and for the optimal path of investment into survival. They show that the solution implies, indeed, a U-shaped mortality schedule. The mortality decline during childhood is driven by an increase in the value of life (or survival) over the period in which there are ongoing investments in brain size. Here, the value of life is to be understood as the discounted expected net energy surplus over the remaining life course, which increases in the capital stock. Mortality reductions carry on for a certain period after maturity of the brain as long as learning by-doing, say, leads to

[10] Kaplan an Robson (2002) give an illustrating example: The human brain, for instance, exceeds the brain size of our closest relatives, chimpanzees and gorillas by a factor of four, while our lifespan is longer by a factor of about two.

further increases in productivity. When productivity declines from some age onwards, this implies a decrease in the value of life and, thus, declining investments into survival.[11] Thus, while not inevitable, the onset of ageing is determined by reductions in productivity, as triggered by reductions in the quality of brain-capital.

The model helps to explain the rectangularization of mortality in humans relative to primates, where for a given level of mortality the human mortality profile exhibits a steeper decline (increase) during youth (old age). The reason lies in the value of life, the change of which over time is driven by the current net-surplus lost with the progress of time. As for humans both the costs and returns from investments into brain size are larger, the net surplus during youth is prone to be smaller (or more negative) than that for chimpanzees, say, implying a lower depreciation of the value of life. In contrast, for older individuals the loss of net-surplus per life-year (in terms of productive capacity) is positive and larger for humans than for chimpanzees. This implies an overall steeper mortality profile for humans (albeit at a lower level for all ages). More generally, this insight has the interesting implication that the speed of ageing, as measured by increasing mortality, is importantly governed by the functionality of the body.

However, this leaves unanswered the question as to what precisely is governing the functionality of the body. The assumption by Kaplan and Robson (2002) and Robson and Kaplan (2003) that the productivity of the human brain declines from some age onwards is intuitive and evidently confirmed. However, as the authors themselves point out , this is not sufficient from the perspective of life-history modeling. A priori it is not clear why evolution should arrange the body in a way that functionality (of the brain or other parts) declines with age. Indeed, Vaupel et al. (2004) show that it may as well be optimal to maintain functionality at a steady level for all times and, thereby, fore-

[11] The notion of value of life has been used extensively in the modelling of mortality risks over the life-cycle (see e.g. Shepard and Zeckhauser, 1984; Murphy and Topel, 2006). In these models, the economic value of life corresponds to the willingness to pay for a decline in mortality. Notably, it is subject to a similar pattern as the evolutionary value of life in Kaplan and Robson (2003): Increases (declines) in earnings capacity during the early (late) career boost (lower) the value of life.

stall senescence. Thus, declining functionality would somehow have to be explained as an optimal strategy. This problem is addressed in Robson and Kaplan (2007) and Kaplan and Robson (2009), where a life-history model similar to Robson and Kaplan (2003) is amended to include not only the size of the body (or brain) but also its quality. Drawing on the disposable-soma theory (e.g. Kirkwood 1977, 1997), they assume that 'somatic' (i.e. body) cells are subject to damage, reducing their functionality/quality, but may be repaired (or even improved) at a cost (per cell). Thus, while the quality of the body may be maintained at full functionality, the cost of doing so increases with body size. This triggers a quantity-quality trade-off, where large bodies are more productive in generating energy but are also more costly in repair. In contrast, 'germ' cells, serving as the vehicle to pass on the genetic code to offspring, can be maintained at full functionality at negligible cost. The optimal evolutionary investment strategy is then to build up a certain body size and keep it constant for the remaining life. During the initial phase of growth, the marginal benefit but also the marginal cost from quality improvements increase jointly, triggering an ambiguous path of quality investments. At higher ages where body size is fixed, rising mortality depresses the marginal benefit from quality investments (related to the value of life). For a given marginal cost of quality this leads to a decline in maintenance effort and, thus, to a decline in cell quality. This, in turn, depresses the productivity of the somatic capital, triggering a reduction in investments into survival and, thus, an increase in mortality.

Senescence is therefore induced for two reasons: In the presence of a germ line, the quality of which can be maintained at virtually no cost, the only reason to maintain the quality of somatic capital lies in its role for providing energy. However, as it is efficient to grow to a (relatively) large body size, the associated maintenance cost renders it efficient to run down the quality of somatic capital. While being plausible these results hinge on the assumptions about the depreciation of quality and the cost of maintenance. Indeed, for certain specifications of the cost function, it may well be the case

that permanent maintenance of quality is efficient. Thus, a deeper understanding yet of the physiological properties is called for.

As Kaplan and Robson (2009) point out, two assumptions are to some extent critical. First, the irreversibility of growth implies that somatic capital cannot be shrunk in order to lower the cost of maintenance. Second, inter-generational transfers towards offspring are crucial in an environment in which resources cannot be stored and in which initial productivity is low. Similar to Lee (2003) these transfers provide the only plausible explanation for post-reproductive survival.[12]

Nevertheless, the analysis of the flow of transfers and their role within evolution remains somewhat implicit in Kaplan and Robson (2002) and Robson and Kaplan (2003, 2007). A more detailed take on the role of intergenerational transfers is provided in the models of Chu and Lee (2006) and Chu et al. (2008). Chu and Lee (2006) focus on the coevolution of intergenerational transfers and longevity. They consider a discrete-time life history model where a fitness term similar to (3.1) is maximized by choice of an age-dependent strategy of fertility, maintenance and growth subject to a linear (energy) budget constraint. Total available energy depends on body-size. From the solution to the dynamic programming problem they conclude that transfer-giving is the more likely to evolve the greater survival and the greater the increase with age in the efficiency of energy production (e.g. by way of forageing). Conversely, if transfer-giving evolves this enhances improvements in survival up to the age at which transfers are made. Chu and Lee (2006) do not endogenize the age-structure of transfer giving, they rather focus on the preconditions for a transfer to occur at some given age and on its effects on mortality. Thus, to some extent the model remains open ended.

[12] For humans, who can at least to some extent make conscious decision about reproduction and the allocation of resources, the question emerges as to how evolution 'implements' an optimal strategy. Here the literature on the evolutionary foundation of preferences provides some answers (Robson and Samuelson, 2011). In particular, it is asked which age profile for the rate of time preference is consistent with the evolutionary optimum (Hansson and Stuart, 1990; Rogers, 1994; Sozou, 1998, 2009; Sozou and Seymour, 2003; Robson and Samuelson 2007, 2009; Robson and Szentes 2008; and Kageyama, 2011).

Chu et al. (2008) consider a continuous-time version of the model in Chu and Lee (2006) with a more general (possibly non-linear) energy budget constraint. They also endogenize intergenerational transfers in the sense that they can be chosen freely as long as they balance out across generations at each point in time. They formally identify the reason for declining juvenile mortality that was more intuitively proposed in Lee (2003): namely as the early death of a juvenile sets free a larger set of resources to be used for surviving offspring than a late death, this lowers the value of 'early' relative to 'late' juvenile survival.[13] Chu et al. (2008) identify a further effect driving towards declining mortality during young ages: to the extent that young individuals are already self-supporting (at least partially so) it pays to sacrifice early-life survival for early-life growth, as the latter serves to relax the energy budget constraint.

In contrast to Robson and Kaplan (2003), the models by Chu et al. (2008) and Kaplan and Robson (2009) explicitly optimize over fertility, which allows them to provide a more precise account of the determinants of adult mortality. Again, the benefit from survival can be summarized in the value of life. As it turns out, given an optimal fertility schedule, the value of adult life now falls into two distinct age-specific components: the remaining reproductive value and the value (in fitness terms) of transfers still to be made to juveniles.[14] The former corresponds to Hamilton's (1966) value of survival for the purpose of reproduction, whereas the latter provides an explicit measure of the value of post-reproductive survival.

We can summarize the most important lessons for an understanding of ageing as follows. (i) The speed of ageing, as measured by the rate of mortality (and its change),

[13] To some extent this is the flipside of the effect identified in Robson and Kaplan (2003), where an increased stock of capital - as accumulated by way of transfers - raises the value of (juvenile) life.

[14] Thus, the value of life falls into a reproductive part and a productive (economic) part. Again, there is an interesting analogy to economic life-cycle models. There, the value of life has conventionally been determined under the assumption of egoistic preferences, excluding any value assigned to descendants. This value of 'current' life corresponds to the 'economic' part in the evolutionary value of life. However, once dynastic preferences a la Becker and Barro (1988) are considered, the value of life is amended by a reproductive part, measuring the individual's value of progeny (Birchenall and Soares, 2009; Kuhn et al., 2010).

depends on three driving factors: First, the remaining reproductive value. Second, the degree to which the adult individual still contributes towards investments into descendants. Third, the productivity, in terms of generating energy surplus, of the somatic capital (body, brain size and quality) and the pace at which it depreciates. Bodies with greater productivity tend to age later but at a greater pace. (ii) From a life-history perspective, the process of ageing cannot be understood fully without an equal understanding of the process of fertility and investments during childhood. This is obvious for the first two determinants. But even the third determinant cannot be understood without an early life perspective, as both the size and the quality of the somatic capital is determined during childhood. Thus, an analysis of the late stages of life must remain incomplete unless it draws at least to some extent on the processes during early life.

3.4. References

Baudisch, A., 2008, Inevitable ageing? Contributions to Evolutionary-Demographic Theory. Demographic Research Monographs. Springer, Berlin, Heidelberg.

Baudisch, A., 2009, How ageing is shaped by trade-offs. MPIDR Working Paper WP 2009-043. Rostock, Germany.

Becker, G.S., 2007, Health as human capital: synthesis and extensions, *Oxford Economic Papers* 59, 370-410.

Becker, G.S. and Barro, R.J., 1988, A Reformulation of the Economic Theory of Fertility, *Quarterly Journal of Economics* 103, 1-25.

Becker, G.S. and Lewis, H.G., 1973, On the Interaction between the Quantity and Quality of Children, *Journal of Political Economy* 81, S279-S288.

Birchenall, J.A. and Soares, R.R., 2009, Altruism, fertility, and the value of children: Health policy evaluation and intergenerational welfare, *Journal of Public Economics* 93, 280–295.

Chu, C.Y.C and Lee, R.D., 2006, The co-evolution of intergenerational transfers and longevity: An optimal life history approach, *Theoretical Population Biology* 69, 193-201.

Chu, C.Y.C, Chien H.-K. and Lee, R.D., 2008, Explaining the optimality of U-shaped age-specific mortality, *Theoretical Population Biology* 73, 171-180.

Ehrlich, I., 2000, Uncertain lifetime, life protection, and the value of life saving, *Journal of Health Economics* 19, 341-367.

Fisher, R.A., 1930, The fundamental theorem of natural selection. In: The Genetical Theory of Natural Selection, Clarendon Press, Oxford, 25-30.

Hamilton, W.D., 1966, The Molding of Senescence by Natural Selection, *Journal of Theoretical Biology* 12, 12-45.

Hansson, I. and Stuart, C., 1990, Malthusian Selection of Preferences, *American Economic Review* 80, 529-544.

Kageyama, J., 2011, On the intertemporal allocation of consumption, time preference, and life-history strategies, *Journal of Bioeconomics* 13, 79-95.

Kaplan, H.S. and Robson, A.J., 2002, The emergence of humans: The coevolution of intelligence and longevity with intergenerational transfers, *PNAS* 99, 10221-10226.

Kaplan, H.S. and Robson, A.J., 2009, We age because we grow, *Proc. R. Soc. Lond. B* 276, 1837-1844.

Kirkwood, T.B.L., 1977, Evolution of ageing, *Nature* 270, 301-304.

Kirkwood, T.B.L., 1997, The origins of human ageing, *Phil. Trans. R. Soc. Lond. B* 205, 531-546.

Kirkwood, T.B.L., 2002, Evolution of ageing, *Mechanisms of Ageing and Development* 123, 737-745.

Kirkwood, T.B.L. and Austad, S.N., 2000, Why do we age?, *Nature* 408, 233-238.

Kuhn, M., Wrzaczek, S. and Oeppen, J., 2010, Recognizing progeny in the value of life, *Economics Letters* 107, 17-21.

Lee, R.D., 2003, Rethinking the evolutionary theory of ageing: Transfers, not births, shape senescence in social species, *PNAS* 100, 9637-9642.

Lee, R.D., 2008, Sociality, selection, and survival: Simulated evolution of mortality with intergenerational transfers and food sharing, *PNAS* 105, 7124-7128.

Medawar, P.B, 1952, Uniqueness of the individual, in: *An Unsolved Problem of Biology*, H.K. Lewis, London, 44-70.

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

Partridge, L. and Barton, N.H., 1993, Optimality, mutation and the evolution of ageing, *Nature* 362, 305-311.

Partridge, L. and Barton, N.H., 1996, On measuring the rate of ageing, *Proc. R. Soc. Lond. B* 263, 1365-1371.

Robson, A.J. and Kaplan, H.S., 2003, The evolution of human life expectancy and intelligence in hunter-gatherer economies, *American Economic Review* 93, 150-169.

Robson, A.J. and Kaplan, H.S., 2007, Why do we die? Economics, biology, and ageing, *American Economic Review* 97, Papers & Proceedings, 492-495.

Robson, A.J. and Samuelson, L., 2007, The evolution of intertemporal preferences, *American Economic Review* 97, Papers & Proceedings, 496-500.

Robson, A.J. and Samuelson, L., 2009, The evolution of time preference with aggregate uncertainty, *American Economic Review* 99, 1925-1953.

Robson, A.J. and Samuelson, L., 2011, The evolutionary foundation of preferences, in Benhabib, J., Jackson, M.O., and Bisin, A. (eds.), *Handbook of Social Economics*, New York: Elsevier Press.

Robson, A.J. and Szentes, B., 2008, Evolution of time preference by natural selection: Comment, *American Economic Review* 98, 1178-1188.

Rogers, A.R., 1994, Evolution of time preference by natural selection, *American Economic Review* 84, 460-481.

Shepard, D.S. and Zeckhauser, R.J., 1984, Survival versus consumption, *Management Science* 30, 423-439.

Sozou, P.D., 1998, On hyperbolic discounting and uncertain hazard rates, *Proc. R. Soc. Lond. B* 265, 2015-2020.

Sozou, P.D., 2009, Individual and social discounting in a viscous population, *Proc. R. Soc. Lond. B*, online publication.

Sozou, P.D. and Seymour, 2003, Augmented discounting: interaction between ageing and time-preference behaviour, *Proc. R. Soc. Lond. B* 270, 1047-1053.

Taylor, H.M., Gourley, R.S., Lawrence, C.E. and Kaplan, R.S., 1974, Natural selection of life history attributes: An analytical approach, *Theoretical Population Biology* 5, 104-122.

Tuljapurkar, S., 1997, The evolution of senescence, in: Wachter, K.W. and Finch, C.E. (eds.), *Between Zeus and the Salmon: the Biodemography of Longevity*, Washington DC: National Academy Press.

Tuljapurkar, S. and Boe, C., 1993, No Senescence Before Its Time: Life History and Mortality, Paper presented at the annual meeting of the Population Association of America.

Vaupel, J.W., Baudisch, A., Dölling, M., Roach, D.A. and Gampe, J., 2004, The case for negative senescence, *Theoretical Population Biology* 65, 339-351.

Weismann, A., 1891, *Essays Upon Heredity and Kindred Biological Problems*, Vol.1, second ed., Clarendon Press, Oxford.

Williams, G.C., 1957, Pleiotropy, natural selection, and the evolution of senescence, *Evolution* 11, 398-411.

4. Ageing and Human Capital Formation

by Alejandro Perez-Laborda and Fidel Perez-Sebastian

4.1. Introduction

Life expectancy is one of the best indicators of human welfare. Indeed, it is a main component of the Human development index (e.g., see Anand and Sen 2000). Life expectancy is thought as well to be a main determinant of an individual's investment decisions. It is then important to understand how longevity is determined. Key questions that we would like to answer are: Why do we age and ultimately die? Why has longevity increased so much over time? Is this process going to stop and, if so, which age is the ceiling? Why does longevity vary so much across countries? What is the impact of longevity changes on economic outcomes?

To address the last of these questions, the tools offered by economics can be sufficient. However, understanding the feedback from the economic environment into life expectancy might require knowledge from other disciplines. At his respect, biology along with other fields such as psychology and neuroscience represent natural allies for economics. This survey tries to serve as a compendium of knowledge mainly from biology useful for economists to understand the process and consequences of ageing.

The idea that economics can learn from biology is an old one. Alfred Marshall, one of the most influential economists, in his 1890 book *Principles of Economics* (Marshall 1961) thought that Biology and not Physics should be the "Mecca of the economist". He went further, and declared that Economics "is a branch of Biology broadly interpreted" because economies are constantly changing.[1] However, as Mokyr (2006) defends, with the exception of the application of the evolutionary perspective to a variety of economic problems, not much has been done regarding the interaction of biologists and economists.

One area in which Biology can definitely offer useful knowledge is on the nature of human ageing. The concept of ageing in economic theory reduces to having a finite life span, defined as the number of years during which an individual lives. A reason for having such a simple approach to ageing is that life-expectancy is constructed as the average life span in a given population group. Ageing in life sciences is, however, defined as the intrinsic, cumulative, progressive, and deleterious loss of function that eventually culminates in death (Arking 2006, Masoro 2006). Economic approaches then only focus on the mortality aspect of ageing, and forget about senescence – the gradual deterioration of body and mind. This missing notion of ageing as senescence is a severe shortcoming to having a truly complete economic theory of human-capital evolution along the life cycle.

Authors such as McFadden (2008) emphasize that economists have not given enough attention to the process of human capital depreciation and maintenance, that is, to the process of senescence in the words of life sciences. A natural question to ask is how this depreciation occurs, whether it is just an exogenous consequence of biological forces or, on the contrary, can be controlled. This missing piece is becoming increasingly disturbing given the ageing process of society, especially in the developed world. Advancing in the understanding of the impact of the ageing phenomenon on the economy requires

[1] See Hodgson (1993) for a more detailed discussion on Marshall's ideas about the relationship between these two sciences.

the comprehension of the different dimensions of senescence. Only then, a successful study of its feedback effects on economic growth, innovation, health expenditure and supply, retirement decisions and intergenerational solidarity will be possible.

This chapter puts the problem into perspective, treating ageing as the decay of human capital. It briefly reviews different approaches to human capital accumulation that make evident the lack of attention that economists have given to the determinants of its depreciation. The chapter also surveys evidence that clearly points out that the decay of human skills and health is influenced by the environment, usage, and maintenance practices. The conclusion is that the evolution of human capital during the life cycle is badly described by existent theories that treat depreciation or, for our purposes, ageing as purely exogenous.

4.2. Senescence, a Missing Piece in Human Capital Theory

The traditional economic approach to capital accumulation is the same for all forms of capital. Capital accumulates through investment i, and suffers an exogenous rate of depreciation δ. Obviously, the specific interpretation given to each variable in the equation depends on the type of capital considered.

In the case of human capital accumulation, this traditional approach establishes that

$$\dot{h} = i - \delta h; \tag{4.1}$$

where

$$h = g(h^h, h^s). \tag{4.2}$$

The variable h is defined as a person's human capital stock, and represents a composite measure of factors that affect labor input efficiency, being the main ones physical health

(h^h) and skills or abilities (h^s). Its accumulation, as Schultz (1961) among many others emphasize, depends on investments related to nutrition, years of schooling, and self-instruction. More importantly, depreciation is thought to be a consequence of skill obsolescence and the deterioration of both health and abilities, being this deterioration a biological process over which people have little control (Case and Deaton 2005). That is, the process of senescence is a fundamental part of human capital depreciation that is generally assumed to be exogenous in economic models.

A more specific example of the use of equation (4.1) and some of its caveats is offered by Grossman (1972). The Grossman model is still a cornerstone on the study of human-decisions of investment in health.[2] His view is equivalent to the traditional approach to capital accumulation with investment being composed of medicines. As a consequence, perfect repair of the biological effects of ageing, that is, eternal life is possible through investment. To get people death in this model, you need to increase exogenously the depreciation rate with age.

More generally, the literature has long recognized that human capital is a more multidimensional entity than physical capital. For this, even though expression (4.1) might be relevant for the accumulation of tools, it can represent a misleading approximation for capital embodied in humans. Main refinements incorporate insights from Phycology and Neuroscience into the investment side. Some of these refinements have to do with the existence of different types of abilities.

From the viewpoint of neuroscience, h^s is a vector of cognitive (IQ levels) and non-cognitive abilities (self-regulation, motivation, time preference, forward looking, etc...).[3] One of the earliest versions of a human-capital specification was influenced by the work Jensen (1969). Authors like of Jensen (1969) and Hernstein and Murray (1994)) argue that cognitive skills are the most important component of human capital because

[2] See, for example, Case and Deaton (2005) and Sanso and Aísa (2006) for recent contributions that build on Grossman (1972).

[3] See, for example, Heckman *et al.* (2007).

the IQ explains well both labor and socioeconomic outcomes.[4] The introduction of this concept into Economics is attributed to the signaling literature. Papers in this vein such as Arrow (1973), Spence (1973) and Stiglitz (1975) forget about health and assume that schooling, a direct consequence of the inherited IQ level, only serves as a filter to signal more able individuals. In this way, an individual's human capital can be written as a time function of her constant cognitive ability θ, which transforms function (4.2) in

$$h = g(h^s(\theta)) = g(\theta) \tag{4.3}$$

For most part of the economic literature, however, human capital has a clear cumulative nature, in which innate endowments only represent inputs of production. Belief in this concept is based on its capacity to explain many features of the earning distributions that a static conception of human ability can not. Work by Becker (1964), Ben Porath (1967), Mincer (1974), Griliches (1977), Becker and Tomes (1979, 1986), Aiyagari *et al.* (2002), Caucut and Kumar (2003), and Seshandri and Yuki (2004) emphasize the role of investment and family background in determining schooling and income levels. The accumulation of abilities for this approach could be described as

$$\dot{h} = f(i, h, \theta) - \delta h; \tag{4.4}$$

where h can also be influenced by parental characteristics.

Recent developments in neuroscience have challenged the concept of time invariant genetically acquired abilities, pushing the literature towards a more general concept of life-cycle skill formation. Cunha *et al.* (2005), Cunha and Heckman (2007), and Heckman (2007) summarize these findings and propose a model that can account for them. The basic premise is that human capital is multidimensional, composed of health and

[4] Empirical studies that establish that cognitive abilities are strong predictors of schooling attainment, wages, and social behaviors like smoking and drug use include Cawley, Heckman and Vytlacil (2001) and Hernstein and Murray (1994).

cognitive and non-cognitive skills, but that not all ability dimensions are equally malleable at each moment in time. IQ levels stabilize around 9 years of age, whereas non-cognitive abilities have been shown to be influenceable till at least age 20. In addition, early stages are more productive in producing skills because there is a skill multiplier effect of early investment.

In Heckman's (2007) formalization of these ideas, human capital h is a function of health and mental capacity θ. In particular, expression (4.2) is reformulated as $h = g(h^h, h^s(\theta)) = g(h^h, \theta)$; where θ is now a vector with different components of cognitive and non-cognitive abilities. These stocks accumulate following

$$\dot{\theta} = f(i, h^h, \theta) - \delta\theta, \tag{4.5}$$

$$\dot{h}^h = z(i, h^h, \theta) - \delta h^h. \tag{4.6}$$

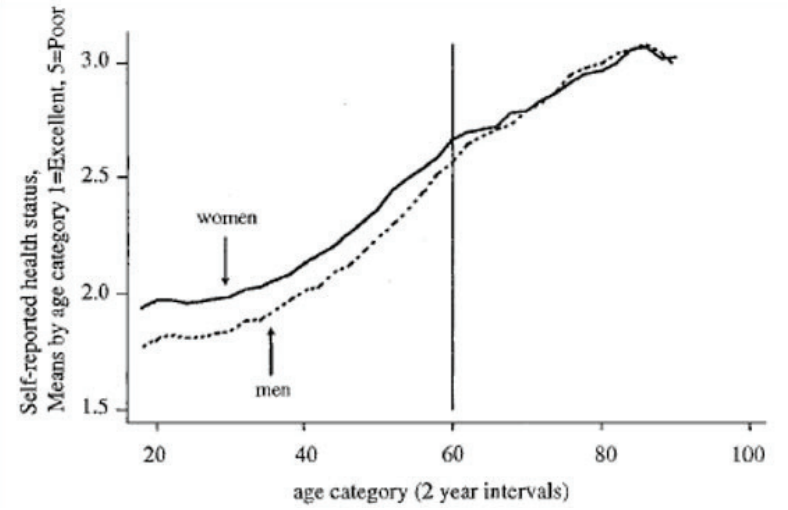
These authors consider that there are dynamic complementarities if skills produced at one stage raise the productivity of investment at subsequent stages, $\partial^2 f(i, h^h, \theta)/\partial i \partial \theta > 0$. In addition, the concept of self-productivity is also introduced when $\partial f(i, h, \theta)/\partial \theta > 0$: skills produced at one stage augment the skills at later stages.

We can deduce from the above discussion that the human capital literature has devoted already some effort to importing insights from other disciplines to have a more complete view of skill formation. Economists have given, however, less attention to the process of human capital depreciation and maintenance. Important questions to ask are whether the depreciation rate of the different dimensions is an exogenous consequence of ageing, as economic theory generally assumes, or can be controlled through work, study, and behavioral choices, and whether it is predictable or random.

McFadden (2008), among others, provides numerous examples that illustrate why a better understanding of human capital depreciation is important. A crucial one for the study of the impact of ageing is the following. Consider accumulation equation (4.1)

in which an individual's investment can consist of schooling when young, and experience and on-the-job training when adult. The constant rate of depreciation implies that human capital over the life cycle achieves its maximum at a fairly young age. This is inconsistent with wage profiles that do not peak until later in life. A possible explanation for this behavior lies on incentives that surround labor contract negotiations. Another one, in which we are more interested, is the misspecification of the evolution of senescence.

Figure 4.1.: Self-reported Health Status by Age and Sex



Source: National Health Interview Survey 1986-2001.

Let us think of human capital as in expression (4.2), that is, divided in physical health and skills, with the latter composed of cognitive and non-cognitive abilities. There is some consensus that indeed physical health declines with age after becoming around 20, for the simple reason that all organs are subject to biological worsening. For instance, there is a decrease in reserve capacity of the lungs, kidneys and heart, also muscle and bone mass decline, reflexes reduce, cartilage erodes and fluid thickens.[5] Figure 4.1, extracted from Case and Deaton (2005), illustrates this point. The Figure

[5] See, for example, Agnusdei *et al.* (1998), Porter *et al.* (1995), and Young (1997).

shows self-reported health status against age starting at 18 years. We see a clear negative trend in perceived health status in men and women that becomes steeper after age 30.

Time trends of physical decay are, however, subject to heterogeneity and, more importantly, are affected by environmental issues and human decisions (Harper and Marcus 2006). Sensory impairments can be often easily compensated or accommodated. Physical strength and functional capacity can be improved with exercise and a well-balanced diet. They also depend on factors that are not directly related to age, like education and other socioeconomic aspects that determine an individual's lifestyle.

Focusing now on the brain, which is responsible of our abilities, the evolution along the life cycle of its capacity is subject to a lot of plasticity and individual heterogeneity. Even though it is generally accepted that cognitive abilities change with age, Schaie has shown in a number of contributions that intellectual decay is not universal, old age is not necessarily a period of cognitive or intellectual decline.[6]

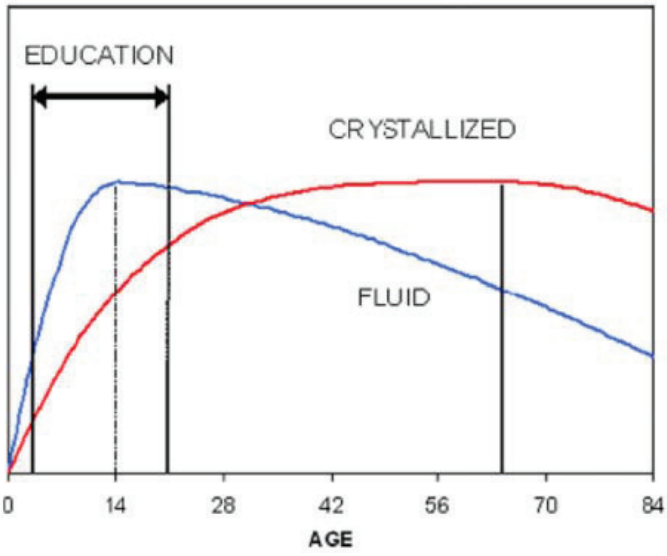
Generally defined, cognitive ability is the capacity to perform higher mental processes of reasoning, remembering, understanding, and problem solving (Bernstein *et al.* 2006). Psychologists, like Horn and Cattell (1966), Horn (1970), and Horn and Donaldson (1976), propose a bi-dimensional approach that split cognitive skills into "crystallized" and "fluid" abilities. Crystallized abilities consist of mental skills acquired by the individual through learning and experience, and are thought to improve with age. Fluid abilities are, on the contrary, more related to genes, and involve the capacity of information processing, attention, and memory. These last abilities, connected with some particular areas of the brain, are subject to biological worsening.[7]

Figure 4.2 shows the typical evolution of fluid and crystallized abilities along the life cycle. Fluid skills start deteriorating at around age 14 with the psychological ageing of the brain. Crystallized skills, on the other hand, continue growing well beyond the end

[6] Schaie (1974, 1989), Baltes and Schaie (1976), and more recently Schaie (2005, 2008).
[7] This is in line with neurological studies about brain shrinking. See for example Raz (2004).

of formal education until about 65 years of age; the subsequent depreciation is a consequence of the appearance of illnesses such as cerebrovascular and brain degenerative diseases.

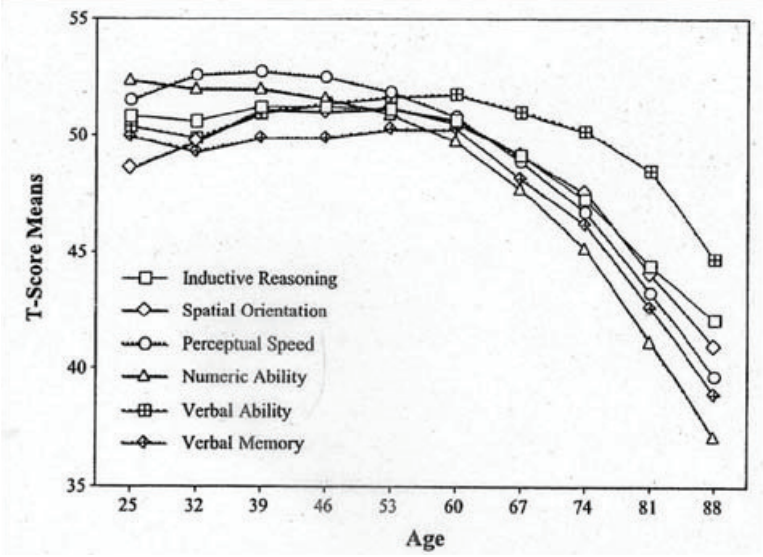
Figure 4.2.: Fluid and Crystallized Cognitive Abilities



The different dimensions of cognitive abilities generally involve both crystallized and fluid components. This has led the literature to adopt a more multidimensional approach. Schaie (1974), for example, divide cognitive abilities in six primary categories: inductive reasoning, spatial orientation, numeric ability, verbal ability, verbal memory, and perceptual speed. The evolution of these categories are plotted in Figure 4.3 that uses data from the Seattle longitudinal study. It reveals the same message as Figure *4.2*. Categories that require mainly fluid skills, like numeric ability, go down constantly during adulthood; whereas the ones in which crystallized skills possess a higher weight, like verbal ability, peak much later in life.

The depreciation rate of brain abilities can be affected as well. Cognitive skill levels depend on genetics, but as Shaie and Willis (1986a,b,c) document, their decline is often also a matter of disuse – "use-it or loose-it". Moreover, LeCarret *et al.* (1996), Adam *et*

Figure 4.3.: T-scores for Different Cognitive Ability Categories and Age Groups



al. (2006), Scarmeas and Stern (2003), and Mazzonna and Peracchi (2009) provide evidence that cognitive ability decay can be delayed through higher levels of education in young age, and occupational and leisure activities in old age. Other socioeconomic variables also show power to delay brain decline, like living in an stimulating environment (Gribbin *et al.* 1980) and home environment and parental influence during childhood (Knudsen et al. 2006, Cunha and Heckman 2007, and Shephard 1997a). All this has a close connection to the concept of "cognitive reserve" (Stern 2002), which establishes that higher levels of cognitive skills prevent neuropathological damage through different cognitive strategies of brain network recruitment and brain resource optimization.

4.3. Conclusion

The above discussion leads us to a simple conclusion that has a powerful implication. Human capital is a complex product that changes with age, augments through invest-

ment in nutrition, education and training, and endogenously decays influenced by the environment, usage, and maintenance. It is then clear that equation (4.1) with the traditional assumption of an exogenous constant rate of decay misses fundamental features that shape the evolution of human capital, and that become more and more important as people age.

Insights from biology, if successfully imported into economics, can help to fill the gap and complete our theories of human capital evolution along the life cycle.

4.4. References

Adam, S., Bay, C., Bonsang, E., Germain S., and Perelman, S., 2006, Occupational activities and cognitive reserve: A frontier approach applied to the Survey of Health, Ageing and Retirement in Europe (SHARE), CREPP WP 2006/05, University of Liege.

Agnusdei, D., Civitelli, R., Camporeale, A., Parisi, G., Gennari, L., Nardi P., and Gennari, C., 1998, Age-related decline of bone mass and intestinal calcium absorption in normal males, *Calcified Tissue International* 63, 197-201.

Aiyagari, S.R., Greenwood J., and Seshadri, A., 2002, Efficient investment in children, *Journal of Economic Theory* 102, 290-321.

Anand, S., and Sen, A., 2000, The income component of the Human Development Index, *Journal of Human Development* 1, 83-106.

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

Arrow, K.J., 1973, Higher education as a filter, *Journal of Public Economics* 2, 193-216.

Baltes, P. B., and Schaie, K.W., 1976, On the plasticity of intelligence in adulthood and old age: Where Horn and Donaldson fail, *American Psychologist* 31, 720-725.

Becker, G.S., 1964, *Human Capital*, Columbia University Press (New York).

Becker, G.S., and Tomes, N., 1979, An equilibrium theory of the distribution of income and intergenerational mobility, *Journal of Political Economy* 87, 1153-1189.

Becker, G.S., and Tomes, N., 1986, Human capital and the rise and fall of families, *Journal of Labor Economics* 4, S1-S39.

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

Bernstein A., Penner, L.A., Clarke-Stewart, A., and Roy, E.J., 2006, *Psychology* (6th ed.), Houghton Mifflin, Boston.

Case, A., and Deaton, A., 2005, Broken down by work and sex: How our health declines, chapter 6 in D.A. Wise (Ed.), *Analyses in the Economics of Ageing*, Chicago University Press for NBER, v. 10.

Caucutt, E., and Kumar, K., 2003, Higher education subsidies and heterogeneity: A dynamic analysis, *Journal of Economic Dynamics and Control* 27, 1459-1502.

Cawley J., Heckman J.J., and Vytlacil, E.J., 2001, Three observations on wages and measured cognitive ability, *Labour Economics* 8, 419-442.

Cervilla, J.A., Prince, S., Joels, S., Lovestone, S., and Mann, A., 2000, Long-term predictors of cognitive outcome in a cohort of older people with hypertension, *British Journal of Psychiatry* 177, 66-71.

Cunha, F., and Heckman, J.J., 2007, The technology of skill formation, *American Economic Review* P&P, 31-47.

Cunha, F., Heckman, J.J., Lochner, L.J., and Masterov, D.V., 2005, Interpreting the evidence on life cycle skill formation, *Handbook of Economics of Education*, Elsevier (Amsterdam, Holland).

Gribbin, K., Schaie, K.W., and Parham, I.A., 1980, Complexity of life style and maintenance of intellectual abilities, *Journal of Social Issues* 36, 47–61.

Griliches, Z., 1977, Estimating the returns to schooling: Some econometric problems, *Econometrica* 45, 1-22.

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

Harper, S., and Marcus, S., 2006, Age-related capacity decline: A review of some workplace implications, *Ageing Horizons* 5, 20-30.

Heckman, J.J., 2007, The technology and neuroscience of capacity formation, *Proceedings to the National Academy of Sciences* 104, 13250-13255.

Hernstein, R.J., and Murray, C., 1994, *The Bell Curve: Intelligence and Class Structure in American Life* Free Press (New York, USA).

Hodgson, G.M. (1993), The Mecca of Alfred Marshall, *Economic Journal* 103, 406-415.

Horn, J.L., 1970, Organization of data on life-span development of human abilities, in: L.R Goulet and P.B. Baltes (Eds.), *Life-span Development Psychology*, Academic Press (New York, USA).

Horn, J.L. and Cattell, R.B., 1966, Refinement and test of the theory of fluid and crystalized intelligence, *Journal of Educational Psychology* 57, 253-270.

Horn, J.L., and Donaldson, G., 1976, On the myth of intellectual decline in adulthood, *American Psychologist* 3, 701-719.

Jensen, A.R., 1969, How much can we boost IQ and scholastic achievement?, *Harvard Educational Review*.

Knudsen, E.I., 2004, Sensitive periods in the development of the brain and behavior, *Journal of Cognitive Neuroscience* 16, 1412-1425.

Le Carret, N., Lafont, S., Letenneur, L., and Dartigues, F., 2003, The effect of education on cognitive performances and its implication for the constitution of the cognitive reserve, *Developmental Neuropsychology* 23, 317-337.

Marshall, A., 1961, *The Principles of Economics*, 9th ed., with annotations by C. W. Guillebaud, Macmillan (London, UK).

Masoro, E.J., 2006, Are age-associated diseases and integral part of ageing?, in Masoro, E.J., and S.N. Austad, (eds.), *Handbook of the Biology of Ageing*, Academic Press.

Mazzonna, F., and Peracchi, F., 2009. "Ageing, cognitive abilities and retirement in Europe," CEIS Research Paper 152, Tor Vergata University, CEIS, revised 04 Dec 2009.

McFadden, D., 2008, Human capital accumulation and depreciation, *Review of Agricultural Economics* 30, 379-385.

Mincer, J., 1974, *Schooling, Experience and Earnings*, National Bureau of Economic Research (New York).

Mokyr, J., 2006, Economics and the biologists: A review of Geerat J. Vermeij's Nature: An Economic History, *Journal of Economic Literature* 44, 1005-1013.

Porter, M.M., Vandervoort, A.A., and Lexell, J., 1995, Ageing of human muscle: structure, function and adaptability. *Scandinavian Journal of Medicine and Science in Sports* 5, 129-142.

Sanso, M., and Aísa, R.M., 2006, Endogenous longevity, biological deterioration and economic growth, *Journal of Health Economics* 25, 555-578.

Scarmeas, N., and Stern, Y., 2003, Cognitive reserve and lifestyle, *Journal of Clinical and Experimental Neuropsychology* 25, 625-63.

Seshadri, A., and Yuki, K., 2004, Equity and efficiency effects of redistributive policies, *Journal of Monetary Economics* 51, 1415-1447.

Stiglitz, J.E., 1975, The Theory of screening, education and the distribution of income, *American Economic Review* 65, 283-300.

Schaie, K. W., 1974, Translations in gerontology – From lab to life: Intellectual functioning, *American Psychologist* 29, 802-807.

Schaie, K.W., 1989, Individual differences in rate of cognitive change in adulthood, In V.L. Bengston and K.W. Schaie (Eds.), *The Course of Later Life: Research and Reflections*, Springer (New York), 65-85.

Schaie, K.W., 2005, *Developmental Influences on Adult Intellectual Development: The Seattle Longitudinal Study*, Oxford University Press (New York, USA).

Schaie, K.W., 2008, Historical processes and patterns of cognitive ageing, In S.M. Hofer and D.F. Alwin (Eds.), *Handbook on Cognitive Ageing: Interdisciplinary Perspective*, 368–383.

Schaie, K. W., and Willis, S.L., 1986a, Can decline in adult intellectual functioning be reversed?, *Developmental Psychology* 22, 223-232.

Schaie, K. W., and Willis, S.L., 1986b, Practical intelligence in later adulthood. In R. J. Sternberg & R. K. Wagner (Eds.), *Practical Intelligence: Origins of Competence in the Everyday World*, 236-268, Cambridge University Press (New York).

Schaie, K. W., and Willis, S.L., 1986c, Training the elderly on the ability factors of spatial orientation and inductive reasoning, *Psychology and Ageing* 1, 239-247.

Schultz, T.W., 1961, Investment in human capital, *American Economic Review* 51,1-17.

Shephard, R.J., 1997, Curricular physical activity and academic performance, *Pediatric Exercise Science* 9, 113-126.

Spence, A.M., 1973, Job market signaling, *Quarterly Journal of Economics* 87, 355-374.

Stern, Y., 2002, What is cognitive reserve? Theory and research application of the reserve concept, *Journal of the International Neuropsychological Society* 8, 448–460.

Young, A., 1997, Ageing and physiological functions, *Philosophical Transactions: Biological Sciences* 352, 1837-1843.

Part II.

Integrating Ageing into the Economic Life Cycle Model

5. Understanding Differences in Longevity across the EU Member States

by Carl-Johan Dalgaard and Holger Strulik

Executive Summary. Longevity varies considerably across the post-enlargement EU member states. Evaluated around the year 2000 life expectancy at age 20 varied by roughly a decade while income per worker varied by about a factor of eight. With respect to income, convergence within the members of the European Union is to be expected in the years to come. Naturally, the question arises whether and how much the convergence of income will help to close the gap in longevity. Moreover, beyond income, what are the most effective ways at raising longevity within the EU? In order to address these questions the Lepas Project has developed a physiologically-founded economic model of health demand over the life-cycle, drawing on recent advances in the modeling of ageing and longevity from the natural sciences. The main conclusions summarized in this chapter are:

- Convergence of *income* per capita (labor productivity) across the EU member states can be expected to close a considerable fraction the life expectancy gap. But productivity advances are not the most powerful determinant of longevity.

- The most powerful determinant of increases in longevity is improvement in health *efficiency* (medical technological progress). Health efficiency should be the main policy target.

- Targeting *prices* (subsidies on wages or prices in the health sector) is a relatively ineffective way of increasing longevity.

A detailed background paper to this chapter is available as Dalgaard and Strulik (2011).

5.1. The Facts

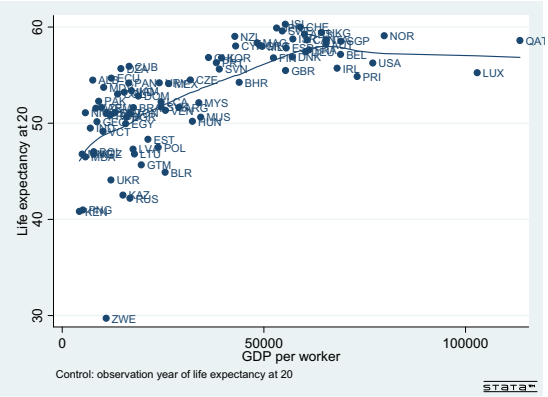
In a seminal contribution Samuel Preston (1975) documented a striking empirical fact: In a cross-section of countries higher levels of income per capita are associated with greater life expectancy; the curve that fits the data best – dubbed "the Preston curve" by later research – is concave. This discovery served to promote the idea that prosperity leads to greater longevity, and that income inequality works to lower average longevity; the latter being a consequence of the observed concavity of the mapping between income and life expectancy. A second observation made in the paper was that the curve shifts upwards over time, implying greater longevity at all levels of income per capita. Preston hypothesized that these shifts represented improvements in health technology (broadly defined), and noted that the shifts accounted for the lion's share of global improvements in longevity over time. These are powerful ideas, which continue to be influential. As Bloom and Canning (2007, p. 498) observe: "Samuel H. Preston's classic paper, 'The Changing Relation between Mortality and Level of Economic Development', published in 1975, remains a cornerstone of both global public health policy and academic discussion of public health."

Figure 5.1 A depicts a recent – modified – version of the Preston curve for a cross-section of countries, focusing on labor productivity (GDP per worker in PPP$) rather

than GDP per capita, and life expectancy at age 20 rather than at birth. Despite these slight differences in the variables involved, the regularity is essentially that recorded by Preston: citizens of countries that are wealthier tend to be healthier. Moreover, the association between productivity and adult life expectancy is positive and non-linear, featuring a "flattening" at the top.

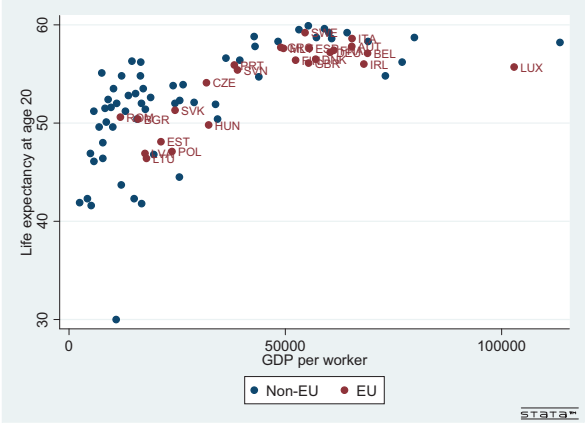
Figure 5.1.: Life expectancy and Labor Productivity

A: The modified Preston curve



The figure shows the cross-country link between life expectancy at age 20 and GDP per worker cirka the year 2000. *Notes:* The line is estimated semi-parametrically, with year of data collection for life expectancy being the linear control. Labor productivity is significant (p-value of 0.000). See Dalgaard and Strulik (2011) for details on the estimation algorithm. See Appendix for data sources.

B: Highlighting EU



The figure shows male life expectancy at age 20 vs. GDP per worker in PPP$, inside and outside EU, circa the year 2000: 85 countries. *Notes*: Life expectancy at age 20 is not recorded for all countries in the world, for which reason the sample only includes 25 out of the 27 current EU member States: Cypress and Netherlands are missing.

As can be discerned from Figure 5.1 B, there is considerable variation in the data within the European Union (EU); both in terms of labor productivity levels and life expectancy. In 2000 GDP per worker varied by a factor of eight (from about 12,000 PPP$ in Romania to roughly 100,000 in Luxembourg), and life expectancy (recorded around the same time) varied by a remarkable 13 years (from ca. 46 years in Lithuania to a little in excess of 59 in Sweden).

The strong income gradient, visually obvious from Figure 5.1, suggests that higher

income enables more direct investments in health, which slows down the ageing process and prolongs life. It is thus tempting to infer that convergence in income levels across the EU member states would entail convergence in longevity. However, reflecting on the identity of the countries at the extremes of the two distributions leads one to the conclusion that there may be more to it than that: longevity is neither the lowest in Romania nor the greatest in Luxembourg. Hence, one might suspect that other factors could be responsible for the apparent link between longevity and income: Health technology or human capital for instance. To make matters worse, it seems impossible to rule out *a priori* that the positive correlation could be attributed (fully or partially) to an impact from longevity on income, i.e., the reverse line of causality.

Accordingly, in order to make progress it is necessary to try to unravel the forces that have shaped the Preston curve. For that purpose we suggest a theory-driven approach; by modeling the ageing process, and the decision to invest in life prolonging initiatives, it is possible to examine the extent to which the income-to-health channel can account for the Preston curve by way of model simulation. Subsequently we then attempt to gauge the likely impact from income convergence on the distribution of health within the EU.

5.2. An Economic Life Cycle Model with Physiological Ageing

The analysis conducted in Dalgaard and Strulik (2011) aims to distil the logic of the income-to-longevity channel. The theory builds on two elements. First, a sound modeling of the ageing process for a representative member of society, drawn from the natural science literature. The second element is the standard economics modeling of intertemporal choice of consumption for the representative consumer, yet extended by the additional option to invest in health, which serves to slow down ageing. We

describe these two elements in turn.

5.2.1. Modeling Ageing

Following the underlying reasoning of reliability theory we think of ageing as being characterized by increasing *frailty* (Gavrilov and Gavrilova, 1991; Chapter 1 in this book). That is, as the redundancy of the human organism shrinks we become more fragile. An empirical measure of human frailty has been developed by Mitnitski and Rockwood and various coauthors in a series of articles (e.g., Mitnitski et al, 2002a,b; 2005). As humans age they develop an increasing number of disorders, which Mitnitski et al. (2002a) refer to as "deficits". Some of these deficits may be viewed as rather mild nuisances (e.g., reduced vision) while others are more serious in nature (e.g., strokes). Nevertheless, the notion is that when the number of deficits rises the body becomes more frail. A frailty index can then be estimated as the proportion of the total potential deficits that an individual has, at a given age.

Mitnitski et al. (2002a) show that the following equation fits data on the proportion of deficits, $D(t)$, of the representative individual at age t very well:

$$D(t) = E + Be^{\mu t}.$$

This "law of increasing frailty" explains around 95% of the variation in the data, and its parameters are estimated with great precision. The parameter E turns out to be common for men and women; using a data set encompassing 66,589 Canadians, aged 15 to 79, Mitnitski et al. (2002a) estimate E to 0.02, with a standard error of 0.001. The parameters B and μ are, however, gender specific. For Canadian men (women) $\log(B)$ is -5.77 ± 0.06 (-4.63 ± 0.06), while μ is 0.043 ± 0.001 (0.031 ± 0.001). Interestingly, very similar estimates for B and μ are obtained on data for Australia, USA and Sweden (Rockwood and Mitnitski, 2007). Hence, in these four developed countries (in spite

of differences in samples, the precise contents of the frailty index etc.) the average individual accumulates 3-4% more deficits from one birthday to the next.

We can restate the law of increasing frailty in flow form by differentiating with respect to age:

$$\dot{D}(t) = \mu (D(t) - E), \tag{5.1}$$

where E works to slow down the speed of deficit accumulation. In order to see that the influence of E in (5.1) is consistent with Mitinitski and Rockwood's equation for the *level* of deficits, integrate (5.1) and insert the initial condition $D(0) = D_0$ to get the solution $D(t) = (D_0 - E)\,\mathrm{e}^{\mu t} + E = D_0 \mathrm{e}^{\mu t} - E(\mathrm{e}^{\mu t} - 1)$. Since $\mathrm{e}^{\mu t} > 1$ for all $t > 0$, a larger autonomous component E implies less deficits at any given age t. Note also that the compound parameter $(D_0 - E)$ corresponds to Mitnitski et al.'s estimate of B. In the natural science literature the parameter E is interpreted as capturing the impact of non-biological factors on deficit accumulation (Mitnitski et al., 2002a). Accordingly, we will assume that E is amendable to change by way of deliberate investment.

Specifically, the following parsimonious refinement of the process of deficit accumulation is employed in Dalgaard and Strulik (2011):

$$\dot{D}(t) = \mu (D(t) - a - Ah(t)^{\gamma}), \tag{5.2}$$

where $D(0)$ is given. The parameter a captures environmental influence on ageing beyond the control of the individual (less pollution, say, implying a higher value for a), the parameters $A > 0$ and $0 < \gamma < 1$ reflect the efficiency by which investments in health initiatives are converted into tangible results vis-a-vis reductions in speed of ageing. Accordingly, A may naturally reflect the state of health technology, but also health institutions which influences how much of the health spending that is channeled into health initiatives rather than, say, administration and the like. While A refers to the general power of health expenditure in maintenance and repair of the human body,

the parameter γ specifies the degree of decreasing returns of health expenditure.

By way of contrast to E, the parameter μ – impressed by its empirical constancy across developed countries – is considered to be a physiological parameter. In the remaining we will refer to this physiological parameter as *the force of ageing*, as it drives the inherent and inevitable process of human ageing. While the force of ageing is exogenous, it is unlikely to be universally constant.

In order to capture death, we need to invoke an upper boundary to deficit accumulation, $\bar{D}$. In the analysis below the representative individual remains alive as long as $D(t) < \bar{D}$. Direct evidence on the existence of an upper boundary for D is found in Rockwood and Mitnitski (2006). Observe that equation (7.1), along with the restriction that $D(t) < \bar{D}$, provides a complete description of ageing until death. In this process, chronological age does *not* play a role in itself. While the model developed below concerns optimal ageing and death of a representative agent of a cohort, it is nevertheless worth noting that this formulation is in concordance with a central point made by biologists and gerontologists: individual ageing is *not* time-dependent. This follows since $\dot{D(t)}$, by (7.1), is only influenced by current investments and accumulated deficits; chronological age t plays no independent role.

5.2.2. Intertemporal Choice

The modeling of the representative individual's economic choices follows the standard approach in dynamic economics. In each period of life the individual receives a wage income and, if wealth is positive, an interest income. This income can either be spent on non-health consumption, it can be saved for future purposes, or it can be invested in health, which, as explained above, serves to reduce the speed of ageing and thus delays the time of death.

The representative individual derives utility from taking initiatives that lower the speed of ageing because it delays the date of expiry, which is ultimately determined

by investments in h. At the same time individuals enjoy other forms of consumption which are not life prolonging. As a result, the individual faces a trade-off between consuming today or making health investments, which allow for greater consumption in the future. Hence, the solution to the optimization problem provides a program for optimal consumption over the life cycle, along with optimal health investments.

The economic model of consumption demand captures two fundamental characteristics of human behavior, *consumption smoothing* and *time discounting*. Consumption smoothing reflects the law of diminishing marginal utility. Consuming a further unit now (e.g. another car) provides more utility but less additional utility than the last consumption unit. Facing such a concave mapping from consumption into utility at every instant in time, consumers essentially prefer a smooth path of consumption expenditure over time. Consumption smoothing provides an incentive for health expenditure. Instead of using additional income to consume more now (and running into diminishing marginal utility) some income can be invested into health in order to prolong life and thus adding more periods over which consumption expenditure can be stretched (e.g. another year of driving pleasure). Time preference, on the other hand, captures human impatience. Individuals prefer to consume a given unit of goods now more than later on. For obvious reasons we expect the rate of time preference to be close to the interest rate, i.e. the price of present consumption in terms of future consumption.

Given these fundamental trade-offs in human behavior it is a priori unclear how individuals would optimally like to allocate their life-time income on consumption, savings, and health investments now and in the future and how long they thus want to live given their life-time budget constraint and facing the force of ageing. In order to answer this question we calibrated the model of consumer behavior to US data. That is, the parameters of the model are chosen such that the model matches behavior of the average US male citizen exactly in several dimensions such as in terms of labor income, life expectancy, evolution of frailty over life and more. With the calibrated

model in hand it is possible to conduct experiments, whereby parameters are changed after which the consequences for ageing and longevity can be assessed.

5.3. Policy Experiments

5.3.1. Policy Experiment 1: Income

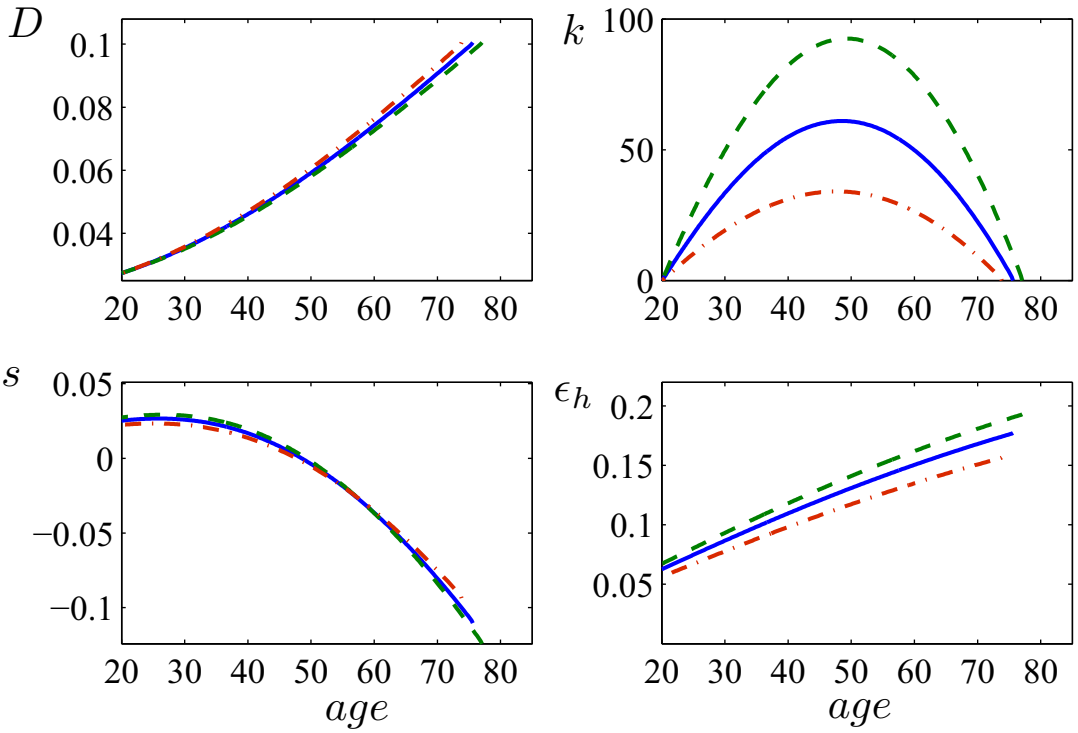
Figure 5.2 shows how the representative individual reacts if his income is perturbed. The blue line shows life-trajectories for the representative US individual for frailty (deficit accumulation) D (top left panel), capital accumulation k (top right panel), the savings rate s (bottom left), and the health expenditure share ϵ_h (bottom right) from the age of 20 to death. The green (dashed) line is associated with an increase of labor income by $1/3$, the red (dashed-dotted) line depicts the reaction to a reduction of w by $1/3$ (in all the experiments below, "green" is associated with increases, and "red" with reductions in the parameter of interest). As can be seen from the figure, the consequence of higher income is an increase in longevity, peak wealth, and the share of health spending.

The intuition for why higher income leads to longer life is found in the desire to smooth consumption. With increasing income, health spending rises more than regular consumption. This occurs as the incentive to smooth the latter is relatively stronger due to diminishing per period marginal utility. Higher income therefore leads to a larger adjustment in the level of health spending compared to non-health consumption. The end result is a slower speed of ageing, and a longer life.

The issue of main interest, however, is the quantitative impact on longevity. As seen from the top left hand side corner of the figure, the impact is modest though not inconsequential. An increase of income of $1/3$ (achievable in a generation with an income growth rate of about 1% per year) translates into an increase in longevity of 1.7 years; the reduction involves a fall in longevity of 1.6 years. If we convert the

Figure 5.2.: Health and Wealth over the Life Cycle

Variation of Labor Income



Green (dashed): wage income increases by 1/3. $\Delta T = 1.47$, implied elasticity 0.07). Red (dotted): wage income decreases by 1/3 ($\Delta T = -1.78$, implied elasticity -0.09).

impact into an elasticity - the elasticity of longevity with respect to income - we find it to be around 0.08, implying that a 100 percent increase of income would result in an 8 percent increase of life-expectancy. This effect is close to econometric estimates (e.g., Pritchett and Summers, 1996). Below we return to the issue of whether the magnitude of this effect is sizable enough to account for the Preston curve.

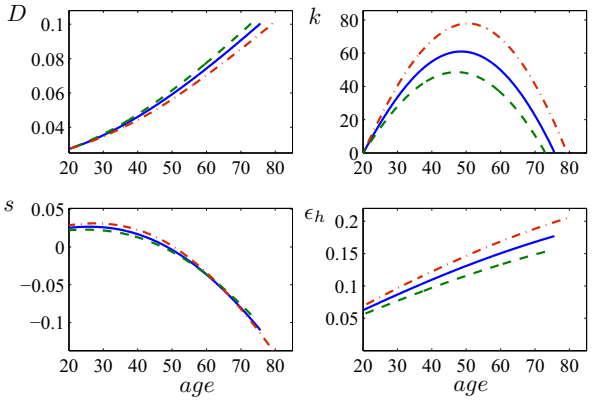
5.3.2. Policy Experiment 2: Health Costs

The next experiment concerns health costs; the relative price of h. Here we consider a doubling of the relative price of health goods. Rising relative health prices is a realistic scenario; the price index of medical care has risen faster than the price index of GDP (Cutler et al., 1998).

As is clear from Panel A of Figure 5.3, when the relative price of health increases, individuals substitute towards regular consumption. As a result, the health share declines. With less health investments, savings (s) decline as well. The end result of a doubling of the relative health price for longevity is a reduction by 2.5 years. This amounts to a longevity-price elasticity of 0.09. While this elasticity is of roughly the same numerical size as the income elasticity it is well to bear in mind that the potential for variation is far smaller; income levels vary across the world easily by a factor of 30 or more, far towering any cross-country variation in price levels.

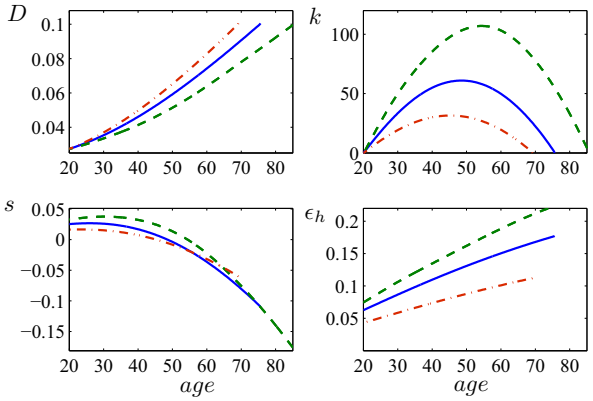
Figure 5.3.: Health costs and Medical effectiveness

A: Variation of Health Costs



Green (dashed):doubling p, $\Delta T = -2.7$, implied elasticity -0.09). Red (dotted): halving p ($\Delta T = 3.6$, implied elasticity 0.13).

B: Variation of General Medical Effectiveness (A)



Green (dashed): 33 percent increase of A ($\Delta T = 9.8$, implied elasticity 0.53). Red (dotted): 33 percent decrease of A ($\Delta T = -6.4$, implied elasticity -0.35).

5.3.3. Policy Experiment 3: Health Technology

In Figure 5.3, panel B, we depict the impact of health productivity, i.e. general medical effectiveness (parameter A in equation 2) on longevity. In the experiments above, the impact from the parameter of interest were indirect. For instance, an increase in income translates into both higher health spending and higher consumption. Medical technologies (or the productivity of health investments more broadly), however, have

a direct impact on the evolution of deficits and therefore on longevity. A larger impact is therefore to be expected.

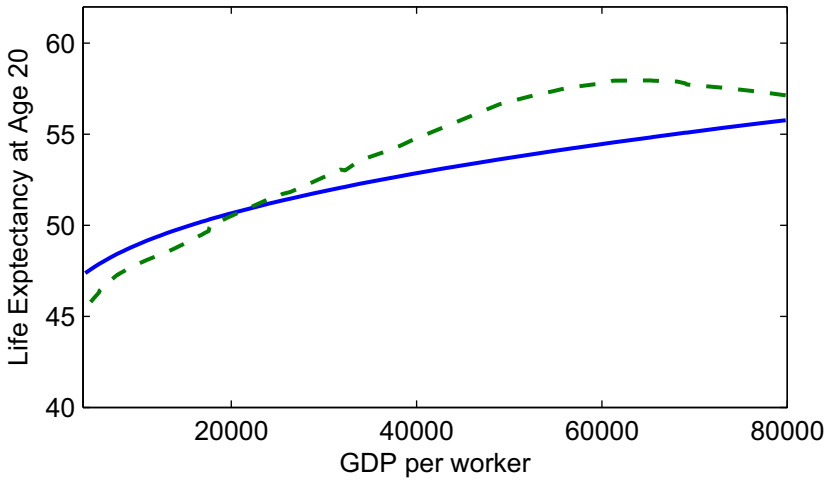
In Panel B we examine the impact from increasing A by 1/3; an increase by the same factor as that which we analyzed in terms of income. If health productivity rises by 33% the consequence is an increase in longevity of nearly a decade. The implied elasticity is about 1/2. This is a very large effect, suggesting that the impact from improvements in health productivity easily may have towered that of rising income. It also suggests that improvements in health efficiency have a large potential to improve life expectancy.

5.4. The Preston Curve: Is Income All That?

As discussed above, a key issue is whether the income gradient, captured by the so-called Preston curve, reflects a causal impact of income on longevity mediated by health investments. In order to examine the importance of the pure income channel, in accounting for the Preston curve, we feed the income levels (PPP GDP per worker in 2000) for individual countries through the model, keeping the parameters we calibrated to US data fixed. Figure 5.4 allows for a visual comparison between the model's predictions regarding life expectancy in 2000, and the empirically estimated income gradient from the cross section in 2000, whereas Table 5.1 provides some summary statistics. Accordingly, the solid line represents the best fit of the income gradient to the data displayed in Figure 5.1; it captures fitted values from estimating the model $l = f(y) + \epsilon$, where l is longevity, y is income and the function $f()$ is determined by the data. See Dalgaard and Strulik (2011) for details.

As can be seen from the figure, the model (solid line) does a pretty good job at matching the Preston curve (dotted curve). As seen from the table, the model does underestimate the range of life expectancies spanned by the Preston curve somewhat, but at average or median income the actual income gradient and the one predicted

Figure 5.4.: The Model vs. the Preston Curve



The figure compares the empirically estimated Preston curve (dotted) to the Preston curve predicted by the model (solid). Source: Dalgaard and Strulik (2011).

Table 5.1.: Summary statistics, Model vs. Data

	Data	Model
LE(min)	45.4	47.4
LE(max)	58.0	55.8
LE(mean)	52.4	51.6
LE(median)	51.5	51.3
R^2		0.78

The R^2 is the coefficient of determination: the estimated income gradient constitute data, whereas the model's prediction constitutes "fitted values". The reported number thus reflects the fraction of the variation in life expectancy along the income gradient that the model can account for.

by the model essentially coincide. The table also provides a summary measure of the goodness-of-fit of the calibrated model in the form of the coefficient of determination: the R^2 of the income-to-longevity link comes to 0.78, which means that 78% of the variation in life expectancy along the Preston curve is accounted for by our model.

While the bulk of the variation in the observed link between life expectancy and income is thus accounted for, there are systematic deviations between the Preston curve and predicted life expectancy, which account for the remaining 22% of the variation.

There are two complementary explanations for the systematic nature of the deviation; the first is based on the assumption that "omitted variable bias" is influencing the empirically estimated income gradient, whereas the second would pertain to the case where reverse causality is thought to be important.

Consider the latter case first. Reverse causality arises if changes in life expectancy instigates changes in income. For instance, suppose that people who live longer also generally are healthier at any point in time during their lives, and therefore are capable of exerting more labor market effort generating higher labor productivity. If so, then the data depicted in figure 1 may not reflect an impact of income on longevity at all, but rather, the influence of longevity on labor productivity. Now, if reverse causality is an issue, one may think of the estimated income gradient as the outcome from the interaction of two separate underlying schedules: an income-to-longevity schedule and a longevity-to-income schedule. The former is in theory captured by our model, which by construction does not admit the reverse line of causality. Now, if the longevity-to-income schedule has a *steeper* slope than the income-to-longevity schedule in the income–life-expectancy space, then the estimated Preston curve will feature a slope that is strictly larger than the income-to-longevity schedule. As a result, our model (capturing only the income-to-longevity mechanism) should overestimate life expectancy at the bottom of the income distribution and underestimate it at the top.

Alternatively, suppose reverse causality is not an issue. If so, then we would interpret deviations between the Preston curve and the model's prediction as the result of omitted variables; factors that are correlated with both life expectancy and income. Theoretically, such factors could map into A (efficiency of health investments), p (relative price of health goods), or both. Our results then suggest that the price level of health in efficiency units (p/A), is higher in most of the poorest countries relative to US. But by the same token p/A must then be larger in the US compared to many of the richest countries, suggesting that the US health care system (at least in 2000) was

less efficient than that of many other rich nations. Whether this is true or not is hard to say. But it remains an observable fact that the US constitutes an "outlier" in health expenditures, but not in terms of life expectancy. In any case, *if* omitted variables is the only channel affecting the Preston curve beyond the income–to-longevity channel, the observed difference between the Preston curve and the model's prediction would have to mean that the price of health in efficiency units is higher in many of the poorest places, yet lower in the richest places, relative to its level in the US.

In practise, of course, we have no way of knowing which of the two explanations is more important in accounting for the left-over residual. What we do know is that they both may influence the Preston curve, and that they together account for some 20% of the variation along the income gradient.

In sum, the analysis suggests that the Preston curve largely, but not exclusively, is due to the causal influence form income on longevity; 78% of the variation in life expectancy along the income gradient is accounted for by our model. This insight leads to an important conclusion. People in the poorest countries are dying earlier than citizen's in rich nations. Our analysis suggests that, to a first approximation, the sad reality seems to be that poor people spend less on health because they are poor and live shorter lives because of it. These conclusions mimic Preston's (1975) own conjecture regarding the underlying forces that generate his curve rather well. That is, the nonlinear link between income and life expectancy is to a large extent caused by lower health investments in several dimensions. Changes in relative prices and health technologies do matter. But in order to understand the income gradient, they do not seem to be the main culprits.

What does this imply for the prospects of convergence in life expectancy within the EU? It is a well known feature of the growth process that relative income gaps tend to narrow over time, between countries sharing structural characteristics like investment rates, population growth, institutions etc. This has led to convergence within Europe in

the past (e.g., Barro and Sala-i-Martin, 1991), and this process will presumably continue in the future. So, suppose the poorest EU countries gradually manage to catch-up with the richest EU countries: how would this affect the distribution of life expectancy?

The analysis above allows us to answer this question employing the calibrated "income elasticity" (see above). If income convergence is taking place it implies that income dispersion across EU drops. But how big of a reduction in inequality in life expectancy would this instigate? Our analysis would suggest that the answer is: Not a lot. More specifically, for each percentage point reduction in the standard deviation in (log) income per worker we would expect to see a reduction of 0.01 percent in the standard deviation of (log) life expectancy.[1] Or, to put it differently: if inequality in income is cut in half, inequality of life expectancy is reduced by 0.6%.

The calculation illustrates that convergence in income will enable convergence in life expectancy in the years to come within the EU; but not in a big way. It is also worth bearing in mind, that the convergence process is believed to be relatively slow. A typical estimate for the speed of income convergence across regions would be that roughly 2 % of the income gap is eliminated per year, implying that it will take about 35 years to cut income inequality in half within the EU, and thus attain the above mentioned reduction in inequality in longevity. This serves to underscore the insight from the policy experiments that that health efficiency should be a major policy theme within the EU.

5.5. Concluding Remarks

The present report has summarized progress in the modeling of the economics of ageing. In particular, the theory allows us to examine the impact from various policy ini-

[1] The calculation works as follows. Assume that life expectancy is related to income in the following fashion: $\log(l) = \lambda \cdot \log(y)$, which implies that $std(\log(l)) = \lambda \cdot std(\log(y)$, where $std()$ denotes the standard deviation across EU regions. The elasticity of inequality in longevity with respect to inequality in income is then $\lambda \cdot std(\log(l)) \cdot std(\log(y))^{-1}$. Around 2000 we have $std(\log(l)) = 0.08, std(\log(y)) = 0.6$; inserting $\lambda = -0.09$ into the formula leads to the stated result.

tiative (such as initiatives leading to a reduction of the health inefficiencies, or health prices) on longevity. It also allows us to gauge the impact of income convergence on inequality in longevity at age 20. Our analysis reveals that income convergence will lead to reductions in "health inequality" across EU; but not by much. The key focus area for policy should be health efficiency.

The present research program has been further developed in various ways. First, the analysis shows the need to understand certain physiological characteristics of the European populations better. Specifically, there is a need to understand the speed of ageing as well as socio-economic differences in health deficit accumulation (across Europe), which is an issue that we have addressed empirically. Second, the the theoretical framework presented above has been extended to a scenario where the representative consumer also chooses education and retirement.

5.6. References

Barro, R. and Sala-i-Martin, X., 1991, Convergence across states and regions, *Brookings Papers on Economic Activity* 107-82.

Bloom, D. and Canning, D., 2007, Commentary: The Preston Curve 30 years on: still sparking fires, *International Journal of Epidemiology* 36, 498-499.

Cutler, D. McClellan, M.B., Newhouse, J.P., and Remler, D., 1998, Are medical prices declining? Evidence from heart attack treatments, *Quarterly Journal of Economics* 113, 991-1024.

Dalgaard, C-J and Strulik, H., 2011, Optimal ageing and death: Understanding the Preston curve. Discussion Paper 11-09 (University of Copenhagen).

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

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

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

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

Preston, S.H., 1975, The changing relation between mortality and level of economic development, *Population Studies* 29, 231-248.

Pritchett, L. and Summers, L.H., 1996, Wealthier is healthier, *Journal of Human Resources* 31, 841-868.

Appendix: Data sources

Data on life expectancy at age 20

We use a samples of data on life expectancy at the age of 20 for males in "2000". Specifically, the sample involves observations covering the period 1997–2006. The data is available in the Demographic yearbook for 2006 (2000 sample). Alternatively, the data can be obtained online at the web address:

http://unstats.un.org/unsd/demographic/sconcerns/mortality/mort2.htm

Data on Labor productivity

In the regression we employ GDP per worker (RGDPW) for 2000, from Penn World Tables, Mark 6.3.

6. Understanding Education-Driven Health Inequality across the EU Member States

by Holger Strulik

Executive Summary. Throughout the EU, as well as in other developed countries, there exists a steep socio-economic gradient in health status and life expectancy. In particular, better educated persons live longer. Understanding the gradient and reducing health inequality is regarded as a top priority policy goal by the European Commission.

This chapter summarizes novel insights of the LEPAS project on the education health nexus. We integrate physiological ageing from gerontology as well as educational choice from labor economics into a life cycle model with endogenous health expenditure and spending on unhealthy goods and identify the return to schooling as a main driver of health behavior. Persons who educate more have more precious human capital to protect and thus decide to indulge less in unhealthy behavior and spend more on health. In a calibrated model we explain among other interesting things:

- An increase of the return to schooling that triggers one more year of education motivates health behavior that prolongs life by about half a year.

- The gradient is mildly non-linear; eight years more of education motivate a gain of eight more years of life.
- Medical technological progress explains why the education gradient widens over time. Better educated persons are motivated to utilize new health technologies more heavily. They thus over-proportionally benefit from medical innovations.

The return to schooling is not genetically given but malleable, in particular in early childhood. Improving equality with respect to child care and education opportunities at young ages is thus identified as the most promising policy gateway to promote health equality among adults at all ages within the EU.

A background paper to this chapter provides the details about the physiological-economic modeling of the education gradient. (Strulik, 2011).

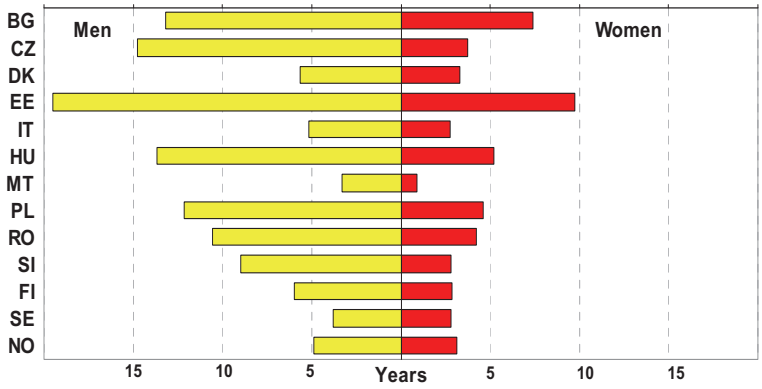
6.1. Stylized Facts and Core Idea

Life expectancy at all ages differs markedly between educational groups in all EU member countries. Figure 6.1, taken from Eurostat (2008) displays this fact for a selection of countries and life expectancy at age 30. Low attainment means lower secondary or less education, high attainment means tertiary education. A similar picture is obtained for the United States where men with any college education could expect to die 7 years later compared to those with only high school or less (Meara et al., 2008). This gap in life-expectancy, also known as the education gradient or just "the gradient", seems to get wider over time.[1]

Life-expectancy is a crude summary statistic of health status. A more detailed picture can be obtained by employing the frailty index. The frailty index counts the proportion

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

Figure 6.1.: Life Expectancy Gaps between High and Low Educational Attainment



Women and Men at Age 30; 2007. Source: eurostat (2008).

of the total potential deficits that an individual has, at a given age. The list of potential deficits ranges from mild ones like impaired vision to severe ones like dementia. The frailty index has been developed by Mitnitski and Rockwood and several coauthors in a series of articles (2002a, 2002b, 2005, 2006, 2007). Mitnitski and Rockwood estimate with an R^2 around 95 percent the rate μ at which health deficits are accumulated. On average, adults in developed countries accumulate 3-4% more deficits from one birthday to the next (see Chapter 1). At the individual level, however, the frailty index reveals marked differences in health status, in particular with respect to education.

A recent study of the LEPAS project, Chatterji et al. (2011), has used the SHARE (2011) data base to compute the frailty index of men and women aged 50+ for a number of European countries. Table 1 shows the result for societies stratified by education. In all 15 European countries for which data was available there exists a strong negative association between education and health deficits. On average, those with tertiary education have accumulated about 50 percent less health deficits than those with only primary education. The Table displays also the well-known fact (see e.g. Chapter 5) that citizens of economically more developed countries are on average healthier. There exists thus a double socio-economic gradient across countries and within countries.

Table 1: The Frailty Index by Education across Europe

2007		Austria	Germany	Sweden	Nether-lands	Spain	Italy	France	Denmark	Greece	Switzer-land	Belgium	Czechia	Poland
By educational level														
Primary	Mean	0.25	0.25	0.2	0.19	0.25	0.23	0.22	0.22	0.19	0.18	0.24	0.22	0.31
	SD	*0.13*	*0.14*	*0.12*	*0.12*	*0.14*	*0.13*	*0.13*	*0.12*	*0.12*	*0.1*	*0.14*	*0.12*	*0.15*
Secondary	Mean	0.2	0.18	0.17	0.15	0.16	0.16	0.19	0.18	0.15	0.14	0.19	0.19	0.24
	SD	*0.12*	*0.12*	*0.1*	*0.09*	*0.09*	*0.1*	*0.1*	*0.11*	*0.1*	*0.08*	*0.11*	*0.11*	*0.13*
Higher	Mean	0.14	0.16	0.14	0.17	0.17	0.15	0.17	0.15	0.12	0.14	0.17	0.17	0.22
	SD	*0.08*	*0.11*	*0.07*	*0.1*	*0.1*	*0.1*	*0.09*	*0.09*	*0.07*	*0.08*	*0.11*	*0.1*	*0.12*

Source: Chatterji et al. (2011).

Tackling health inequality is high on the agenda of the European Commission. This objective has been emphasized in several Communications of the Commission; for example COM (2009) states: "The Commission regards the extent of the health inequalities between people living in different parts of the EU and between socially advantaged and disadvantaged EU citizens as a challenge of the EU's commitment to solidarity, social and economic cohesion, human rights and equality of opportunity."

In order to efficiently address health inequalities it is helpful to fully understand their origins. The Lepas project has therefore developed a theory of ageing that combines research from the medical and economic sciences. It renders a novel view on the education gradient. The core idea is that it can be fully rational for less educated persons to indulge more in unhealthy consumption and to spend less on health because they have less precious human capital to protect.

So far the literature has suggested three different kinds of deeper explanations for the education gradient: common third factors, productive efficiency and allocative efficiency. The "third factor" argument is based on the impact of general attitudes on behavior and becomes particularly intuitive if one thinks of time preference. More patient persons are presumably more willing to delay entry into the workforce for education as well as they are more willing to sacrifice pleasure from unhealthy consumption in exchange for a longer a life. The problem is that, empirically, general attitudes seem to play only a minor role for educational differences in health behavior. Cutler and

Lleras-Muney (2010) estimate that attitudes like time preference account for about 10 percent of health behavior, similar to the contribution of health knowledge, whereas income (access to resources) and cognitive ability account for the greatest shares, each for about 30 percent.

The idea of productive efficiency postulates that less educated individuals "produce" less health out of any given inputs of time and medical care (see e.g. Grossman, 1972, 2000). Allocative efficiency, with contrast, puts the emphasis on the inputs and suggests that less educated individuals use different inputs, presumably because they are less well informed about their "health technology" (see e.g. Kenkel, 1991). The common theme of both ideas is that less educated people behave less efficiently. If they had only access to the health technology and the knowledge of the better educated, they would care more about their health and live longer.

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

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

interpretation as cognitive ability is tempting in light of the above mentioned empirical evidence.

6.2. Modeling Unhealthy Behavior, Health Expenditure, and Education over the Life Cycle

6.2.1. Health Deficits Accumulation

Inspired by recent research in gerontology we model a physiologically founded ageing process, according to which ageing is understood as increasing loss of redundancy in the human body. The basics of health deficit accumulation have been set up in Chapter 5 (and in more detail in Dalgaard and Strulik, 2010). Here we maintain from the previous chapter that health deficits accumulation can be delayed through health expenditure and augment this by the assumption that deficit accumulation can be accelerated by unhealthy behavior. Specifically, we propose the following parsimonious refinement of the process of deficit accumulation:

$$\dot{D}(t) = \mu\left[\ D(t) - a - Ah(t)^{\gamma} + Bu(t)^{\omega}\ \right], \qquad 0 < \gamma < 1, \qquad \omega > 1. \tag{6.1}$$

As before, the parameter a captures environmental influence on ageing beyond the control of the individual, the parameters $A > 0$ and $0 < \gamma < 1$ reflect the state of the health technology, and h is health investment. While A refers to the general power of health expenditure in maintenance and repair of the human body, the parameter γ specifies the degree of decreasing returns of health expenditure. The larger γ the larger the relative productivity of cost-intensive high-technology medicine in maintaining and repairing highly deteriorated human bodies. Likewise, the parameter B measures the general unhealthiness of the unhealthy good and the parameter ω measures the degree of "increasing returns" in terms of health damage from excessive consumption.

Increasing returns means that, for example, another whisky is less unhealthy for the occasional connoisseur than for the binge drinker.

We consider the life cycle choices of young adults for which initial frailty (at age 16) is given, $D(0) = D_0$. Furthermore, following, Rockwood and Mitnitski (2006), we assume that a terminal frailty exists at which the individual expires, $D(T) = \bar{D}$.

6.2.2. Education

The modeling of education combines the conventional approach in labor economics (Mincer, 1974) with recent insights on the impact of ageing on productivity. Specifically, we assume that human capital of an individual of age t with s years of schooling is given by (6.2).

$$H(s,t) = \bar{H}\mathrm{e}^{\theta s+\eta(t-s)-\alpha\mu t} - \delta D(t), \tag{6.2}$$

for $t > s$ and $H(t,s) = \bar{H} - \delta D(t)$ otherwise. The length of the schooling period s is a choice variable for individuals. The initial endowment $\bar{H}$ captures predetermined skills and skills acquired through the compulsory schooling period. The parameter θ is the return to education and η is the return to experience (learning on the job). For $\alpha = \delta = 0$ the schooling function boils down to the one used in conventional economic models.

The parameter α controls for the impact of the force of ageing μ on human capital. Most cognitive abilities start to decline between age 20 and 30 (Skirbekk, 2004). The modeling in (6.2) allows to discuss the impact of ageing on cognitive skills and on human capital separately. It captures the fact that the rate of ageing μ and the associated rate of cognitive skill loss appears to all individuals of a population alike and, in particular, independent from occupation whereas the *impact* of the skill loss on wages, measured by α, is occupation specific (Skirbekk, 2004). For example, while economics professors probably experience the same loss of cognitive skills through ageing as race car drivers (and perhaps even a comparable decline of productivity, see Oster and

Hamermesh, 1998), the skill loss has much less severe consequences on their salary, that is, on the value of their human capital. Since α is potentially job-specific it provides an explanation (beyond education) for why holders of certain occupations are healthier (Case and Deaton, 2005). Varying α constitutes an interesting experiment for the calibrated model.

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

6.2.3. Life Cycle Choices

The optimal behavior of individuals over the life-cycle is characterized by several trade-offs. To begin with we assume as for the benchmark model (Chapter 5) that income can be spent on consumption, health expenditure, and saving. In addition we consider now that consumption is divided up into health neutral consumption and explicitly unhealthy consumption like e.g. smoking, alcohol consumption, or the intake of other drugs. While, as before, health expenditure reduces the speed of ageing and postpones the time of death, unhealthy consumption, speeds up ageing and prepones death. Furthermore we assume an exogenous age of retirement and – in line with conventional labor economics – a wage per unit of human capital which is increasing at the rate of aggregate productivity growth.

The resulting optimal allocation of health expenditure over time has been explained

in Chapter 5. Given the model's extension, individuals face now additionally an allocation decision for unhealthy consumption. Quite intuitively, the life cycle optimization produces behavior according to which expenditure on health and on unhealthy goods are negatively correlated, over time as well as across individuals. This means that persons who spend more on health are predicted to indulge less in health-damaging consumption. Demand of unhealthy goods is lower at higher prices and there exists a preemptive price which deters unhealthy consumption. This means that for unhealthy consumption to occur at all, the unhealthy good has to be sufficiently appreciated by the consumer and its price has to be sufficiently low. Optimal behavior furthermore prescribes that expenditure for unhealthy consumption should decrease with age. Intuitively, the damage done by, for example, binge-drinking is relatively harmless at young age when there is still a lot of redundancy in the body. At an advanced age, binge-drinking could be lethal and the model recommends to reduce drinking to an occasional glass of wine.

Optimal schooling behavior requires that the marginal loss from postponing entry in the labor market equals the marginal gain from extending education. For the special case in which health does not matter for human capital, that is for $\alpha = \delta = 0$ it can be shown that optimal education is independent from longevity and coincides with the solution predicted by the conventional model in labor economics. When ageing matters for human capital, however, education and longevity are interdependent through the rate of health deficit accumulation.

6.2.4. Calibration

Given these fundamental trade-offs in human behavior it is a priori unclear how individuals would optimally like to allocate their life-time on schooling and working and their life-time income on unhealthy and health-neutral consumption, savings, and health investments now and in the future. In other words, we want to know what ex-

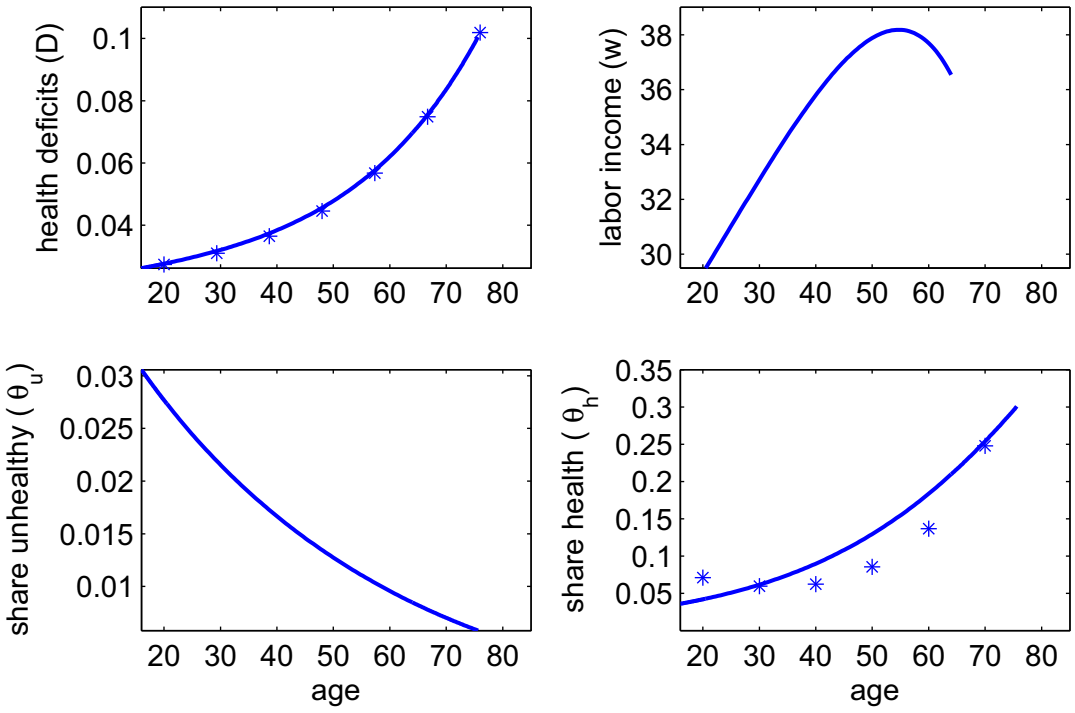
actly determines how long a persons wants to live, facing the force of ageing and the constraints on education and the household budget. In order to answer this question we calibrated the model to US data. That is, the parameters of the model are chosen such that the model matches behavior of an average 16 year old male US American in the year 2000. The initial age is set to 16 years, corresponding to model-age zero, because individuals below roughly the age of 16 are not subject to increasing morbidity (Arking, 2006) and are presumably not well described by the law of increasing frailty. Furthermore, in many states of the US, as well as in many countries around the world, schooling is compulsory up to an age of about 16. This means that there is not really an *individual decision* about education below this age. An implication is that the individual of model age zero has spent already 9 years on compulsory education. The effect of compulsory education are captured by the initial endowment $\bar{H}$.

Details of the calibration can be found in the background paper. With respect to the benchmark model we also had to calibrate unhealthy consumption and the schooling function (6.2). With respect to the latter we took the "consensus value" for the return on schooling of 0.1 (Psacharopoulos and Patrinos, 2004). We set the growth rate of the wage per unit of human capital to an annual rate of one percent, based on US TFP growth in 1995-2000 (Jorgenson et al., 2008). Finally we adjust α and δ such that (i) labor income peaks at age 55 (as observed by French, 2005) and (ii) the representative individual optimally chooses $s = 4.5$ years of extra education, implying in total 13.5 years of education, which corresponds with the US average in the year 2000 (Turner et al., 2007). We adjust the initial unit wage $\bar{w}$ such that total labor income across all working ages equals \$35,320, that is the average annual pay for workers in the year 2000 (BLS, 2011).

Most of the available empirical literature on the consumption of unhealthy goods is about cigarettes and tobacco. We thus took cigarettes as a benchmark case and then proceeded with sensitivity analysis. Preston et al. (2010) estimate that smoking takes

away 2.5 years of life-expectancy of 50 year old US males. We thus set the model parameters such that an individual without health technology dies at age 62 (life expectancy of a 20 year old male in 1900), that actually the calibrated reference American dies at 75.6 (life expectancy of a 20 year old male in 2000) and that he could have lived 2.5 years years longer without unhealthy consumption. The benchmark calibration predicts a value of life of $ 5.7 million for a 20 year old, a value that corresponds well with Murphy and Topel's (2006) estimate of, $ 6.3 million for the value of a statistical life.

Figure 6.2.: Optimal Schooling and Ageing: Basic Run



Solid lines: basic run. Wages in thousands, wealth (capital) in hundred thousands. Stars: data.

Figure 6.2 shows the implied trajectories over the life cycle of the Reference American. Stars in the health deficit panel indicate the actual estimates of Mitnitski and Rockwood (2002). Stars in the lower right panel show the actual age-specific share of health expenditure inferred from Mazzocco and Szemely (2010). The model matches

age-specific health expenditure well until retirement and underestimates it slightly afterwards. According to the data the expenditure share of health is 0.28 for the 60 to 70 year old. The model predicts this expenditure share to be reached at age 73. A different way of describing the same fact is to say that old US Americans are spending somewhat too much on health compared with what the optimal solution suggest.

The lower left panel of Figure 6.2 shows the expenditure share of unhealthy consumption θ_u for the Reference American. When young he is predicted to spend about 3 percent on unhealthy goods. The expenditure share is declining until death. On average the reference American spends $ 403 per year on unhealthy consumption, a figure that squares reasonably well with the $ 319 that Americans spent on average for cigarettes in the year 2000 (BLS, 2002). The upper right panel displays the calibrated invertedly-u-shaped trajectory of labor income across ages.

6.3. Explaining the Education Gradient

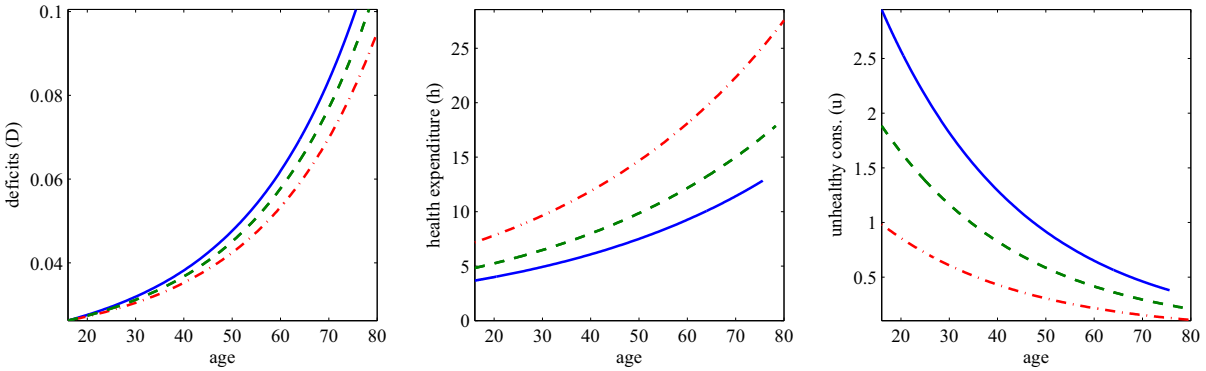
Our benchmark experiment is to increase the return on education such that the reference American is motivated to do one more year of education. Ceteris paribus, he educates a year longer when θ rises from 0.100 to 0.104. This motivates him to reduce unhealthy consumption by 8.8 percent and increase health expenditure by 4.4 percent. As a consequence of the behavioral changes, the better educated person lives half a year longer. The result accords well with the empirical observation that education is positively associated with health through behavior as well as through income and that both channels are about equally important (Cutler and Lleras-Muney, 2010).

The model helps to identify causality, a problem, which has tormented the related empirical literature. The mechanism goes as follows. Higher cognitive skills make education more worthwhile. Better educated persons are endowed with more "precious" human capital, which they care more to protect by indulging less in unhealthy

consumption and by spending more on health. Consequently they live longer.

The effect of education on life-length is non-linear. The more people educate the larger is the return in terms of longevity. Eight years of education on top of the 13.5 from the basic run, that is basically a PhD degree, is predicted to result in a about eight more years of life. This result fits nicely with Meara et al.'s (2008) observation that in 1990 a college graduate could expect to live 8 years longer than a high school dropout of the same age.

Figure 6.3.: Education, Health, and Health Behavior



Health expenditure and unhealthy expenditure in thousands. Blue (solid) lines: basic run (Figure 6.2). Green (dashed) lines: $\theta = 0.12$ (four more years education). Red (dashed-dotted) lines: $\theta = 0.14$ (seven years more education).

The associated evolution of health and health behavior over the life cycle is shown in Figure 6.3. It displays optimal age-trajectories for the reference American (blue lines), a person endowed with $\theta = 0.12$, which takes up four more years of education (green lines), and a person endowed with $\theta = 0.14$ and seven years more education (red lines). The better educated person displays at any given age a better health status (less deficits). The health differences are explained by health behavior. The better educated person spends more on health and less on unhealthy consumption at any given age.

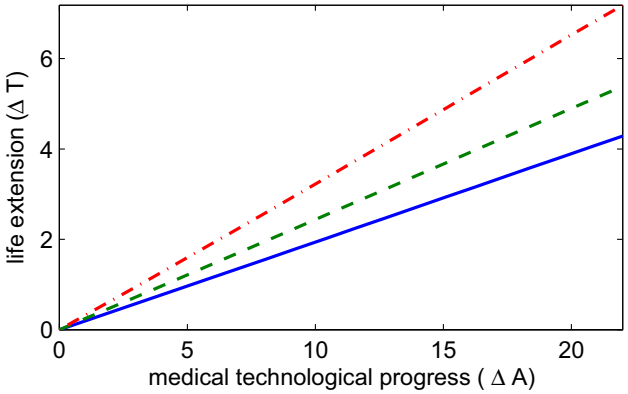
The background paper documents our careful search for other potential drivers of the education gradient. Interestingly, it turned out that two seemingly obvious candidates, namely personal differences in income or time preference cannot explain the ed-

ucation gradient. The only at least partly acceptable alternative mechanism is through occupation specific deterioration of human capital, measured by α. A reduction of α that motivates one more year of education would imply two additional years of life. This means, diagrammatically, that there is an education gradient but its predicted slope is too steep, which means that a lot of the variation in education is left unexplained.

The size of the education gradient predicted by personal differences in the return to schooling is surprisingly robust against alternative specifications of the model. In the background paper we corroborate this claim using alternative setups with respect to wealth endowments (inheritances and bequests), income, the price elasticity and the unhealthiness of the unhealthy good, the weight of health for individual utility, and the introduction of a (rudimentary) public health system. In all cases the education gradient was obtained in the range of 0.3 to 0.65 years with a mean around 0.5 years. Of course, this does not mean that the model's core variables respond insensitively to parametric changes of the environment. It means that the health gradient remained stable *although* the model's predictions for health and life expectancy were quite dramatically modified under the alternative assumptions. For example, in a society in which the reference individual is 50 percent poorer than the reference American, life expectancy is predicted to be seven years shorter (in line with the results from the basic model, see Chapter 5). But the gradient, that is the extra years of life associated with one year more of schooling, is predicted to be 0.51, close to the original 0.5 obtained for the Reference American.

The education gradient, however, responses relatively strongly to medical technological progress. Technological progress is predicted to lead to a steeper education gradient, that is, ceteris paribus, to a more unequal society with respect to health status. The reason is that better educated persons demand more health services in order to protect or repair their human capital.

Figure 6.4.: Medical Technological Progress and the Health Gradient



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

Figure 6.4 shows the longevity gains resulting from alternative increases of A, that is the power of medical technology to reduce health deficits; ΔA is measured in percent of the benchmark run. If medical technology advances at an annual rate of 1 percent (3 percent) the level of A is 20 percent higher after about 18 years (6 years). The solid line shows the predicted longevity for the Reference American (endowed with a return to education of $\theta = 0.1$). The dashed line shows the prediction for a person with $\theta = 0.12$, which educates for 4 years longer and the dash-dotted line reflects longevity of a person with $\theta = 0.14$ and 7.2 more years of education. Although everybody experiences an increase in longevity, the predicted gain of the better educated persons is higher. When A advances by 20 percent the longevity gap between a high school graduate (solid) and a college graduate (dashed) has widened by about 2 years.

6.4. Concluding Remarks

This chapter has summarized progress in the understanding of health inequality driven by educational inequality. The LEPAS project proposes a new view on the education

gradient. Suppressing any explanation based on attitudes, non-cognitive skills, and allocative or productive inefficiency of the uneducated, a large part of the observable education gradient can indeed be motivated by optimal decisions on education and health behavior. The theory has been firmly built on insights from modern gerontology which allowed a robust calibration of a "Reference American". It predicts that a person whose return to education (cognitive skills) motivate one year more of education, spends more on health and less on unhealthy behavior such that he lives about half a year longer. Well educated persons demand relatively more health services in order to protect their precious human capital and thus benefit to a larger degree from health innovations.

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

6.5. References

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

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

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

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

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

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

COM, 2009, Communication from the Commission to the European parliament, the Council, the European economic and social committee and the committee of the regions solidarity in health: Reducing health inequalities in the EU, No. 567, Brussels.

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

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

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

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

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

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

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

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

Eurostat, 2010, Highly educated men and women likely to live longer, Statistics in Focus, 24/2010.

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

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

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

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

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

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

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

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

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

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

Mazzocco, M. and Személy, B., 2010, The decline of the US saving rate and the rise of health expenditure, Working Paper, California Center for Population Research.

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

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

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

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

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

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

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

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

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

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

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

SHARE, 2011, Survey of Health, Ageing and Retirement in Europe, `http://www.share-project`

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

Strulik, H., 2011, Health and Education: Understanding the Gradient, LEPAS Discussion Paper.

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

7. Understanding and Projecting Years of Retirement across EU Member States

by Carl-Johan Dalgaard and Holger Strulik

7.1. Introduction

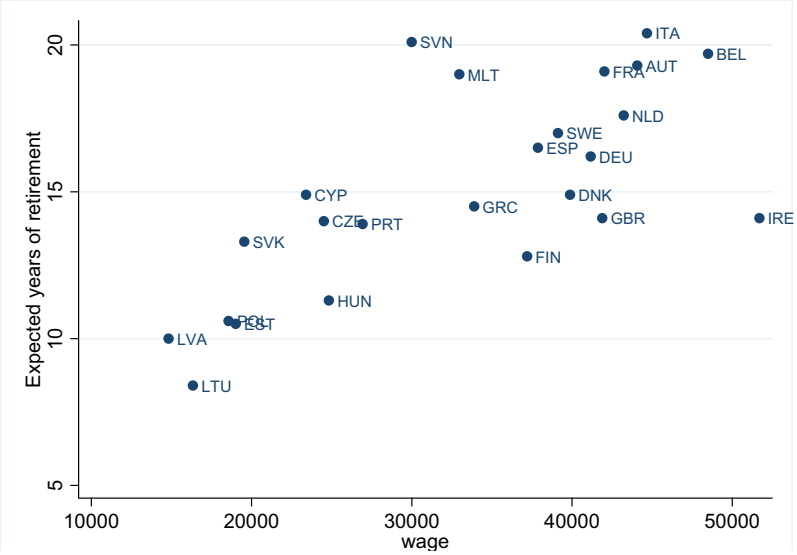
A dominant demographic trend for the 21st century is "ageing". Best estimates suggest that within a few years there will be more people on the planet above the age of 65 than below 5; probably for the first time in human history. Accordingly "ageing populations" are a global phenomenon.

These patterns pose important challenges, in particular for fiscal sustainability. Indeed, a focal point in trying to gauge the impact on government finances from an ageing society involves sorting out how increases in longevity will be distributed between additional years in the labor market and in retirement. A key statistic of interest is (expected) years of retirement; the difference between life expectancy and age of retirement. The present chapter summarizes insights from a research project which can shed light on precisely this issue (Dalgaard and Strulik, 2012).

As a prelude, consider Figure 7.1 below, which shows the simple correlation between

wage income and expected years of retirement (as of age 50) for a selection of EU member states.

Figure 7.1.: Life Time Leisure and Income



The figure shows the difference between (male) life expectancy and expected age of retirement, versus the wage rate. Sources: Age of retirement in 2003: Carone (2005); life expectancy: United Nations (2006, Table 22); wage rate: Penn World Tables 7.0 (Heston et al. (2011) and own calculations; 24 EU countries in 2003. Notes: Luxembourg is an outlier in terms of income and has therefore been omitted from this illustration.

Two regularities immediately stand out from the figure. First, there is considerable variation in the length of retirement across the EU; In Latvia the representative individual (having reached 50 years of age) could expect about 10 years in retirement in 2003, whereas the comparable Italian would be able to look forward to nearly twice that. Second, there is a clear positive correlation between earnings and retirement. Taken at face value, this correlation suggests that economic growth in the years to come will work to prolong the retirement period in the EU, absent reforms.

Yet caution is warranted. One problem with jumping to a quick conclusion on the basis of the above data is that optimal retirement likely is influenced by other forces beyond income. An obvious candidate would be efficiency of the health care system,

which likely is correlated with income across countries and influences optimal retirement choice in its own right. If so, the above correlation would not reflect the impact from wages on retirement. A second problem is that expected years of retirement consists of two underlying terms: life expectancy, and the age of retirement. Both may be influenced by wages, to varying extend. A proper analysis of the retirement burden thus requires a careful analysis of the influence of growth on both longevity and the timing of retirement, together, in order to fully assess the sensitivity of retirement length to prosperity.

In an effort to make progress in these respects we propose a theory-driven approach. Specifically, we augment the baseline model discussed in Chapter 5 by an optimal retirement decision. We then proceed to calibrate the model to a benchmark country for which all relevant parameter values are available. We then subsequently use the model to elicit information about the likely impact of wages on years of retirement, whereby both the age of retirement and death are endogenously determined.[1] In the end this will enable us to gauge the likely evolution of the retirement burden in the future, given reasonable assumption above future wage growth.

7.2. Modeling Ageing and Life-time Labor Supply

The analysis conducted in Dalgaard and Strulik (2012) extends the framework developed in Dalgaard and Strulik (2010) by allowing the age of retirement to be optimally determined alongside life time consumption and longevity. Accordingly, the modeling of the ageing process is as described in Chapter 5. In the next section particular focus is devoted to the novel elements introduced into the analysis and its implications for the "retirement burden".

[1] The simultaneous determination of the age of retirement and longevity is a significant advancement over the existing literature, where life expectancy is treated as a source of *exogenous* influence on optimal retirement. See e.g. Bloom et al. (2003); Heijda and Romp (2009).

7.2.1. Income and Optimal Choice During Life

The modeling of the representative individual's economic life-cycle choices follows the existing literature, except that individuals are subject to physiological ageing. Accordingly, in each period of life the individual receives a wage income and, if wealth is positive, an interest income. In contrast to Dalgaard and Strulik (2010) the wage rate is allowed to change across the life cycle:

$$w(t) = w_0 \left(\mathrm{e}^{\alpha t} - \kappa \mathrm{e}^{\mu t} \right). \tag{7.1}$$

where t is age, α reflects the overall tendency for wages to rise during life as a consequence of learning-on-the-job, whereas μ reflects the deleterious impact from ageing on work effort and thus earnings. As explained in Chapter 5, the parameter μ reflects the relentless process of ageing – the accumulation of health deficits – which leads to reduced physiological capabilities and eventually death. Finally, the parameter κ ensures that the overall path of wages, across the cycle, is hump shaped in keeping with what is known about life time wage income. This empirically meaningful description of wages is one reason why individuals prefer to start life working, and retire late in life; foregone earnings (from not supplying labor) will be greater early in life compared to late in life.

The individuals' income can either be spent on non-health consumption, it can be saved for future purposes, or it can be invested in health, which serves to reduce the speed of ageing and thus delays the time of death. In addition, at each point in time the individual has one unit of time available, which she decides on whether to supply to the labor market or not. Early in life the benefits from supplying labor will exceed its utility costs, as explained below. However, eventually the individual will find it optimal to stop supplying labor in the market; this is the point of retirement.

As in Dalgaard and Strulik (2010) the representative individual derives utility from

consumption per period in life, and from a longer life. In addition, however, Dalgaard and Strulik (2012) introduces disutility from supplying labor, and that the level of disutility from labor supply is increasing in the accumulated health deficits. This assumption captures the idea that labor supply is more strenuous, and thus welfare reducing for people in poor health. It provides another reason why people supply labor early in life, only to retire later in life: early in life the utility costs from working are considered to be smaller than late in life.

Individuals face a trade-off between consuming more today or making health investments, which allow for greater consumption in the future. In addition, however, we now put individuals in front of another trade-off; namely that between working or retirement (leisure). Hence, the solution to the optimization problem provides a program for optimal consumption and health investments over the life cycle, along an optimal timing of retirement. Given optimal health investments the path of bodily deficits over the life-cycle and thus, eventually, the time of death, are determined. As a result, the model delivers optimal years spent in retirement as well.

7.2.2. Optimal Ageing, Retirement and Death

The critical issue under examination is how individuals, according to the theory, respond if the level of wages is increased. That is, what is the influence from growth in income, from one cohort to the next, on optimal retirement *and* (via optimal health spending) optimal expiry. As noted in the Introduction, it is a unique feature of the present framework that the age of retirement as well as longevity are jointly determined within the model.

The key trade-off involved, with regards to optimal retirement, is easily ascertained

from the following first order condition for the underlying optimization problem:[2]

$$\beta_a D(R)^{\nu} = w(R)c(R)^{-\sigma}. \tag{7.2}$$

The condition says that at the age of retirement, R, the disutility from supplying labor $\beta_a D(R)^{\nu}$ (where $D(R)$ represents health deficits at R) must equal the marginal utility gain from labor supply, captured by the marginal utility from consuming ($c(R)^{-\sigma}$) the additional wage income ($w(R)$).

Early in life (i.e., at age $t < R$) the above equation in not fulfilled with equality; rather, the right hand side exceeds the left hand side. As noted above, this is essentially due to the fact that health deficits D are small early in life, and because wages are relatively high. During life, however, individuals suffer physiological decay, which works to increase the costs of labor supply, via rising health, and lowers the gains via (eventually) lower wages. At time R the individual is exactly indifferent between work or leisure, at which point she retires.

Now, consider the consequences of a higher wage level; this corresponds to an experiment in which the wage *profile* of the representative individual is shifted up, implying greater life-time wages.[3] There are three individual channels that influences the individual retirement choice: a health channel; a wage channel and a channel which operates through the level of consumption.

The first channel captures that as wages increase individuals will respond by investing more in health, which lowers health deficits at any given age and thus serves to delay expiry. This "health effect" will work to delay retirement, by lowering disutility from work. The basic intuition behind the intensification of health investments, in the face of wage increases, is covered in Chapter 5 and represents the foundation for the so-called Preston Curve, which depicts a positive (yet concave) association between

[2] See Dalgaard and Strulik (2012) for details on derivations.
[3] Formally, w_0 in equation (1) is shifted up: wages are thus higher at all points in time during life.

income and life expectancy.

The second channel captures a pure income effect: as the wage profile shifts up the alternative costs of leisure rises, enticing the representative individual to stay longer in the labor market. Hence, the income channel will also work to delay retirement.

The third channel, however, renders the overall link between income and retirement ambiguous. Higher wages work to increase the level of consumption, which implies lower marginal utility from additional work effort and concordant wage income. Consequently, the "enjoyment" of the incremental wage addition from staying in the labor market longer, is reduced. This channel then provide individuals with an incentive to retire earlier, thereby increasing their overall utility by increasing leisure rather than consumption.

Given these fundamental trade-offs in human behavior it is a priori unclear how the age of retirement (as well as longevity and thus years of retirement) will change, if income rises. Accordingly, in order to answer this question we calibrated the model o US data. That is, the parameters of the model are chosen such that the model matches behavior of the average US male citizen exactly in several dimensions such as in terms of evolution of frailty, health investments over life and more. With the calibrated model in hand it is possible to conduct experiments, whereby parameters are changed after which the consequences for ageing and longevity can be assessed. In particular, it allows us to gauge the likely impact from wage growth on "the retirement burden" within the EU area.

7.3. Implications for the EU

7.3.1. Income Gradients: The Model and the Data

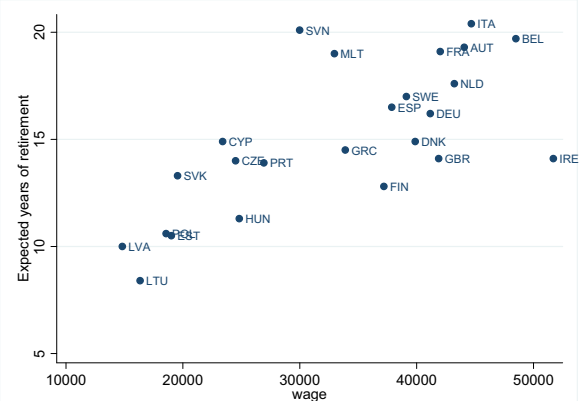
Figure 7.2 depicts the empirically observed link between wage income and years of retirement (Panel A), as well as the model's implications (Panel B). Recall that the model

is not calibrated to match this data, for which reason it constitutes a meaningful "out-of-sample" check of its usefulness.

From visual inspection it seems clear that the model matches the data to a reasonable extent, in that it captures broadly the "position" of the apparent years-of-retirement/income gradient. It does slightly underestimate years of retirement at the high income portion of the sample. But since retirement is unquestionably also influenced by institutional features that the model abstracts from, some deviations from actual experiences would be expected. At the same time, since the model does a rather decent job at mimicking the data, this exercise leaves a clue that observed retirement patters probably are not grossly suboptimal.

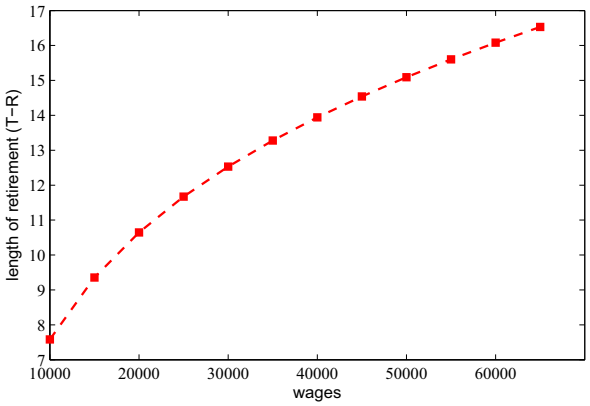
Figure 7.2.: Retirement Years in the EU and in the Model

A: Years of Retirement vs Wages: EU



The figure shows the difference between (male) life expectancy and expected age of retirement, versus the wage rate. Sources: see Figure 7.1

B: Years of Retirement vs Wages: Model



The figure shows link between (male) years of retirement and wages, as implied by the calibrated version of the model developed in Dalgaard and Strulik (2012).

Having observed that the calibrated model matches the "position" of the income gradient reasonably, one may proceed to examine whether it fits the *slope*. From an empirical standpoint the income gradient can be examined by way of regression anal-

ysis. Regressing our calibrated (log) wage level on (log) retirement years we obtain[4]

$$\log(T - R) = -2.58 + .51 \cdot \log(w), \qquad R^2 = 0.6, N = 24, \tag{7.3}$$

which suggests that an increase in wages by one percent translates into an increase in retirement years (length of life, T, minus age of retirement R) by about 0.5 percent (S.E. 0.09). This is not too far from what is predicted by the model, given observed income levels across the EU: here we find an elasticity of about 0.4. In fact, we cannot reject an elasticity of 0.4 statistically. This is interesting as it indicates that the observed correlation between years of retirement, and income, largely reflects the *impact* of income on desired old-age optimal leisure.

Another interesting insight from these results is that the age of retirement must be relatively income *insensitive*. As seen in Chapter 5, higher wages translate into greater health investments and thus longevity. Consequently, since years of retirement increases in the simulation, it must be because age of retirement does not increase to the same extent as longevity. The theoretical interpretation is that an increase in wage has a roughly off-setting effect on the three channels discussed above, given the calibration and level of incomes observed in the EU.

From a quantitative perspective we may evaluate the elasticity of expected age of retirement with respect to wages by regressing (log) wage level on (log) expected age of retirement. For our EU sample the result is

$$\log(R) = 4.14 + .002 \cdot \log(w), \qquad R^2 = 0.0003, N = 24. \tag{7.4}$$

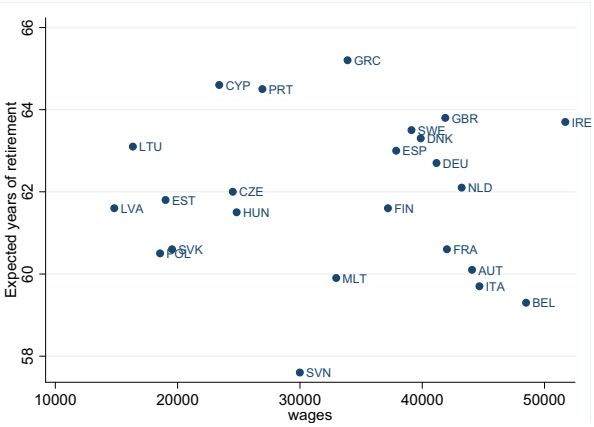
Hence, our point estimate suggest that an increase in wages of about 1 percent increases the age of retirement in the EU (on average) by roughly 0.002 percent; the effect is so small that it is statistically insignificant from zero. Interestingly, the cali-

[4] Luxembourg is omitted in these exercises as this country appears to be an outlier

brated model yields a result which is pretty consistent with this: 0.003. Accordingly, the model roughly matches the data in terms of the slope estimates with respect to both age of retirement, and the difference between longevity and age of retirement. The main mechanism behind the increase in years of retirement, then, is the mechanism discussed in Chapter 5: greater wages increases health investments and prolong life. Figure 7.3 depicts the correlation in the data.[5]

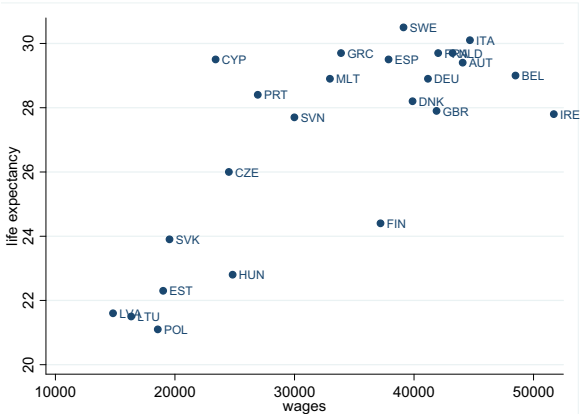
Figure 7.3.: Health and Wealth over the Life Cycle

A: Age of Retirement vs Wages: EU



The figure shows expected age of retirement versus the wage rate: 24 EU countries in 2003. Sources: Age of retirement: Carone (2005). Life expectancy: United Nations (2006). The wage rate: Penn World Tables 7.0 and own calculations.

B: Life expectancy vs Wages



The figure shows life expectancy at age 50 vs. wages. Sources: Life expectancy: United Nations (2006). The wage rate: Penn World Tables 7.0 and own calculations.

7.3.2. Forward Predictions of the Retirement Burden

In a recent report from the EU commission Carone (2005) provides the projected age of retirement and longevity (at retirement) for the EU member countries toward 2050. The projections are conducted by using data on historical exit rates from the labor force, and age-specific mortality rates. Hence, it is of interest to see how different our

[5] It is interesting to note that even life expectancy at age 50 exhibits a concave correlation with wages, reminiscent of the pattern found for age 20 in Chapter 5.

model's predictions would be compared to the (well established but) somewhat more mechanical approach adopted in Carone (2005).

A unique feature of our framework is that both longevity and age of retirement are endogenous to wages. Accordingly, provided one can come up with reasonable projections for income growth, we can predict years of retirement. Expected wage growth until 2050 is obviously also an uncertain proposition. Still, using past experiences a plausible projection would be two percent per annum, plus/minus a percentage point. Given these assumptions about future growth, we can use the model's predicted elasticities for age of retirement and years of retirement, respectively. Table 7.1 reports the results, and those of Carone (2005); they all refer to averages for the EU 25.

Table 7.1.: Projecting Age of Retirement and Years of Retirement

Projections by:	Carone (2005)	Model		
		g=0.01	g=0.02	g=0.03
Mean age of exit, 2050	62,9	61,9	62,1	62,4
Mean years of retirement , 2050	22,1	19	24,8	34,2

Notes: (i) g refers to the assumed annual growth rate in wages in the context of the model. The calculations in column 2-4 are made based on the calibrated models' predictions regarding relevant elasticities; see text for details. (ii) The projections for column 1 are from Carone (2005), Table 4, p. 31, and refer to EU average.

The predictions by our model and those of Carone are rather comparable when it comes to age of retirement, though Carone's projections suggest a slightly higher age of retirement compared to our model's predictions. To match Carone's projection with our model would require an assumed growth rate of wages of roughly four percent per year. Turning to years of retirement, however, we generally find our model to deliver more gloomy predictions (from the point of view of fiscal sustainability). With a growth rate of two percent per year our model would suggest nearly three additional years of retirement, compared to Carone's calculations. If we use a growth rate of three percent, as seen from column 4, we project a significantly longer retirement than

Carone. It is furthermore worth noting that these calculations do not take into account that health efficiency (be that in terms of medical technology or institutions) may improve in the years to come, which may well work to increase years of retirement via its positive impact on longevity.

7.4. Concluding Remarks

The present chapter has summarized progress in the modeling of the economics of ageing. In particular, the theory allows us to examine the influence from wage growth on years of retirement, as both longevity and retirement age are endogenously determined. When applied to the EU member states we find that reasonable growth in wages, in the absence of reforms, can be expected to increase years of retirement considerably in the years to come. Accordingly, there is little to suggest that the weight of the "retirement burden" has been underestimated in past projections. On the contrary, by underestimating the impact of wage growth on longevity, current projections may well underestimate years of retirement and thereby the strain on EU member states government finances in the future.

7.5. References

Bloom, D.E., Canning, D., and Graham, B., 2003, Longevity and Life-Cycle Savings, *Scandinavian Journal of Economics* 105, 319-338.

Carone, G, 2005, *Long-term labour force projections for the 25 EU Member States: A set of data for assessing the economic impact of ageing*, European Commission,

Dalgaard, C.J. and Strulik, H., 2010, Optimal ageing and death: Understanding the Preston curve, Lepas Working Paper 2010-05.

Dalgaard, C.J. and Strulik, H., 2012, A Century of Change: Leisure, Health and Wealth, mimeo (University of Hannover).

Heijdra B. and Romp, W.E.. 2009, Retirement, pensions, and ageing, *Journal of Public Economics* 93 , 586-604.

Heston, A., Summers, R., Aten, B., 2011, Penn World Table Version 7.0, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, May 2011.

United Nations, 2006. Demographic Yearbook. Available at `http://unstats.un.org/unsd/demographic/products/dyb/dyb2006.htm`.

8. Externalities in a Life-Cycle Model with Endogenous Survival

by Michael Kuhn, Stefan Wrzaczek, Alexia Prskawetz, and Gustav Feichtinger

8.1. Introduction

Starting from the seminal work by Grossman (1972) economists have applied the life cycle model to examine how individuals allocate health care and consumption over their life course and what this implies for their health status, mortality and ultimately for longevity (e.g. Ehrlich and Chuma, 1990; Ehrlich, 2000). A related line of literature employs life cycle models to assess an individual's willingness to pay for survival - the value of life - and how it evolves over the life course (e.g. Shepard and Zeckhauser, 1984; Rosen, 1988; Johansson, 2002; Murphy and Topel, 2006). Both strands of literature typically take a positive approach, i.e. they examine the determinants of individual health care choices and valuations but do not question their efficiency. In most of the models efficient life cycle choices are guaranteed anyway, as the individual (i) faces perfect markets and (ii) acts as an isolated decision-maker who is not linked to other individuals (contemporary or future). In the real world neither of the (implicit) assumptions in (i) and (ii) is likely to hold. Markets are typically imperfect or

even missing so that externalities may arise. Likewise individuals are linked to others through altruistic ties but also through externalities. Thus, efficiency is by no means guaranteed.

The research in Kuhn et al. (2011) that is summarized in this chapter seeks to make two contributions to the literature on life cycle models. First, a model is provided in which the individual can by consuming health care reduce mortality and, thereby, shape the schedule of its own (expected) survival. Here, we assume in contrast to Grossman (1972) and Ehrlich and Chuma (1990) that survival is stochastic, an assumption that is also made by Ehrlich (2000) and Hall and Jones (2007). Apart from the different focus of our research question (see below), we depart from the latter two papers by assuming a proportional hazard type mortality schedule (Kalbfleisch and Prentice, 1980) as is underlying an important class of models with realistic demography. As it turns out such a mortality schedule can be nicely modeled in a variety of ways for the purpose of numerical exercises. Second, Kuhn et al. (2011) addresses the extent to which individual health care choices are inefficient in the presence of intra- and intergenerational externalities regarding the provision of health care. Specifically, we focus on the effects of spill-overs related to medical spending.

8.1.1. Externalities in the Provision of Health Care

Typically, individual mortality not only depends on the individual's own consumption of health care but also on the level of aggregate 'activity' within a health care system. One could think of a number of ways in which aggregate activity may either enhance or compromise individual efforts to reduce mortality. First, medical research has identified a positive relationship between volume of (surgical) activity and the quality of care, frequently measured by (lower) mortality (for an overview see Phillips and Luft, 1997). Thus, the effectiveness of individual health care increases in aggregate activity. Conversely, by contributing towards aggregate activity an individual also enhances

the effectiveness of health care targeted at others. As long as these benefits are not internalized in the price of care - and there is no reason to believe they are, whatever the pricing arrangements - such spillovers constitute a positive externality.

Second, in as far as the provision of health care contains public good aspects, such as the provision of medical facilities, hospital bed capacity or emergency services, higher levels of aggregate health care spending may translate into a lower mortality risk at individual level. Similarly, higher aggregate spending levels may lead to greater scope for medical R&D or other quality enhancing activities that would not be lucrative in 'low spending' health care systems. Murphy and Topel (2007), for instance, model an R&D race for a pharmaceutical innovation and show that the overall probability of innovation increases in the share of the social value that the winning firm is able to capture. In our model the prize for innovation would correspond to the winner's share of aggregate health expenditure, thus establishing a link between aggregate health expenditure and individual mortality.

Third, positive spillovers typically arise in the context of preventive activities. The most obvious example relates to vaccination, where individual mortality decreases in the degree to which the population is vaccinated against an infectious disease (for an overview see Philipson, 2000). The same applies to antimicrobial treatment of infectious disease, which is curative from the individual's perspective but also prevents further infections. Other preventive activities that lower both own and other people's mortality include the installation of safety devices such as Anti Lock Breaking in automobiles or fire detectors in tenement flats.

Finally, we may think of measures related to public health such as the cleaning of sewerage, proper disposal of household waste or the reduction of air pollution. Cutler and Miller (2005) show that in the early 20th century nearly half of the total mortality reductions in major US cities can be attributed to the introduction of clean-water technologies, i.e. the filtration and chlorination of water supplies. 'Pure' public health

measures constitute a polar case, where mortality reductions are exclusively due to cumulative expenditure. But even in less extreme cases, the problem of private underprovision arises as long as a part of private health expenditure flows towards a public good (i.e. communal reductions in mortality).

All of the above examples relate to positive spillovers, where higher activity translates into lower mortality. However, in a number of circumstances the converse may be true: aggregate activity may increase individual mortality. Negative spillovers could arise from congestion effects or from microbial resistance against antibiotics. Excessive demand for health care may lead to congestion in the presence of capacity constraints. Hospital crowding, for instance, is likely to hike up mortality due to increased infection risks or due to over-stretched clinical staff lowering the attention afforded to the care of each individual patient and being more prone to committing medical errors. Clement et al. (2008) use DEA techniques to identify congestion, i.e. the production of undesirable outputs (higher risk adjusted mortality for five conditions) together with desired outputs (treatments). They find for year 2000 data that 67 per cent of US hospitals were experiencing some level of congestion amounting to an average efficiency loss of 13 per cent.

Finally, it is well known that microbes tend to develop resistance against antimicrobial treatments. The probability that a resistant microbial strain develops increases in the level of exposure. Thus, individual use of antibiotics tends to curb individual mortality but may, in aggregate, lead to an increased mortality risk due to microbial resistance.[1] In the case of negative spillovers, there is a tendency towards an excessive consumption of care.

The empirical relevance of all of the above mechanisms has been well documented (in the literature referenced). While we are unaware of empirical evidence as to the distortionary effects of these spillovers on the level and pattern of individual health

[1] The positive correlation between intensity of antibiotic use and microbial resistance is well documented empirically (e.g. Cohen and Tartasky 1997; Easterlin 1999).

expenditure, in the light of their prominence, we would expect these effects to be of a non-trivial magnitude.[2]

8.1.2. Life-cycle Implications of Health-related Externalities

The general implications of health-related externalities are straightforward: under-(over-)consumption of health care in the case of positive (negative) externalities. The life cycle aspects of the problem are far more intricate. First, a distinct life cycle pattern of health care spending translates into a distinct pattern at which externalities are generated. Second, through its influence on mortality and ultimately on life expectancy the externality generates an important feedback. This is because changes in life expectancy have a bearing on the individual's aggregate wealth and on the need to spread this wealth over a life span of changing length. Third, the extent / value of the externality is endogenous as it depends on the size and age-structure of the population, the latter being determined by age-specific mortality. Finally, a transfer policy aimed at internalizing the spill-overs needs to reflect the above properties and, therefore, gives rise to a particular age-schedule of the transfer.

We analyze these issues by combining two variants of a life cycle model with endogenous mortality, depending both on individual health expenditure and on a measure of aggregate health expenditure:

1. a 'conventional' life cycle model, where an individual maximizes life time utility without altruistic concerns. This model determines the individual pattern of consumption and health expenditure.

2. an age structured optimal control model, where a social planner maximizes welfare at population level (i.e. individual utilities aggregated over time and age

[2] A number of the externalities discussed (positive and negative) could be avoided through appropriate supply side policies (e.g. hospital congestion). For the purpose of the present research we take their absence as given and ask how the demand for health care could be adjusted to optimal levels.

groups). This model determines the socially optimal pattern of consumption and health expenditure.

Solving and simulating models (1) and (2) and comparing the respective patterns of consumption and health care we can deduce conclusions about the inefficiencies in individual behavior and for which age-groups they arise due to the externality. We derive value of life expressions for the two models, a comparison of which allows to pinpoint the age-specific degree of inefficiency. Furthermore, we combine the two models in order to derive a transfer scheme that, if targeted at the individual, restores the first-best allocation.

To our knowledge, we are the first to integrate externalities into a life cycle model with endogenous survival. This is no trivial undertaking if the spillovers extend *across* different age-groups (or, indeed, cohorts) at any given point in time. By their very nature, the analysis and evaluation of such effects requires a model of the full population such as model (2), and therefore stretches beyond what could be achieved within an individual life cycle model alone. The latter allows to analyze the behavior of a given cohort *along* the time path. However, by construction individual life cycle models are not amenable to an analysis of cross-cohort effects. Our approach provides a consistent and tractable way of analyzing such effects. Technically, the age-structured control model (2) differs from the life cycle model (1) in that the control variables - consumption and health care - and state variables - assets and population size / survival probability - vary in two dimensions - age and time - instead of a single dimension age = time.

The remainder of this paper is structured as follows. The next section provides an overview of the model and our key results. The subsequent and final section discusses some implications for the provision of health care within the ageing societies of Europe.

8.2. Model and Key Results

Consider, first, the individual life cycle model with endogenous mortality. Here, an individual chooses over its life course a stream of age-specific consumption $c(a)$ and health care $h(a)$.

The probability of surviving to age a equals

$$M(a) := \exp\Big(- \int_0^a \mu(s, h(s), \bar{H}(t_0 + s))\, ds\Big).$$

where the mortality rate $\mu(a, h(a), \bar{H}(t_0 + a))$ depends on the individual's age and time of birth t_0, the level of individual health expenditure $h(a)$, and on a measure of health care utilization (or expenditure) at population level $\bar{H}(t_0 + a)$. While individual health care always serves to lower mortality, the impact of aggregate utilization on mortality could be both negative or positive. If mortality falls (increases) with aggregate utilization then we speak of positive (negative) externalities.

Using $u(c)$ to denote period utility from consumption we can then express the life cycle problem for an individual born at t_0 as follows

$$\begin{aligned} \max_{c,h} \quad & \int_0^\omega e^{-\rho a} u(c(a)) M(a)\ da \\ \text{s.t.} \quad & \dot{M}(a) = -\mu(a, h(a), \bar{H}(t_0 + a)) M(a) \\ & \dot{A}(a) = (r + \mu(a, h(a), \bar{H}(t_0 + a)))A(a) + y(a) - c(a) - h(a) \\ & M(0) = 1, A(0) = 0, A(\omega) = 0 \end{aligned}$$

Here, the objective corresponds to the discounted flow of expected life-cycle utility with ρ denoting the rate of time preference. The first equation of motion $\dot{M}(a)$ describes the law of mortality, while the second equation of motion $\dot{A}(a)$ describes the change in the individual's assets A. We follow Yaari (1965) in considering a set-up in which individuals can fully annuitize their wealth by trading actuarial notes at some

gross interest $r + \mu(a, h(a), \bar{H}(t_0 + a))$. This mechanism provides complete insurance against premature and/or late death. We assume that individuals do not leave bequests, implying that $A(\omega) = 0$.

Individuals do not expect to affect (in a measurable way) the level of aggregate health care utilization and therefore take $\bar{H}(t_0 + a)$ as given at each point in time. We can then show that the individual chooses (i) consumption expenditure according to the Euler equation (equating the marginal rate of intertemporal substitution with the compound interest across all periods) and (ii) health expenditure such that

$$\frac{-1}{\mu_h(a, h, \bar{H})} = \psi^P(a). \tag{8.1}$$

The individual equalizes effective marginal expenditure with the private value of life (PVOL) $\psi^P(a)$. The latter, as derived e.g. by Shepard and Zeckhauser (1984), Rosen (1988), Johansson (2002) and Murphy and Topel (2006), captures the willingness to pay for survival through life year a. It amounts to the monetarized expected utility stream over the remaining life course, taking into account changes in the remaining life-time wealth.

We contrast this outcome to the allocation a social planner would choose when maximising the welfare of a whole population. The latter evolves according to the McKendrick equation (see Keyfitz and Keyfitz (1997)),

$$N_a + N_t = -\mu(a, h(a,t), \bar{H}(t))N(a,t) \qquad N(0,t) = B(t), N(a,0) = N_0(a). \tag{8.2}$$

Here, the number of a-year old individuals at time t decreases by the number of deaths $-\mu(a, h(a,t), \bar{H}(t))N(a,t)$ at t, where $\mu(\cdot)$ denotes the mortality rate depending on age a, as described previously. Furthermore, $N_0(a)$ denotes the initial population distribution and $B(t)$ the number of newborns at each instant of time.

The planner chooses age-time profiles of consumption $c(a,t)$ and health care $h(a,t)$ to maximize the discounted sum of instantaneous utility, where the sum is taken across all individuals living within the planning horizon (total utilitarianism), i.e.

$$\max_{c(\cdot),h(\cdot)} \int_0^T \int_0^\omega e^{-\rho t} u(c(a,t)) N(a,t) \; da \; dt \tag{8.3}$$

Here, we note that as the planer considers individuals from different age groups at each point in time we now need to index variables by both a and t. This is in contrast to the individual choice problem, where $a = t$. Solving the planner's problem subject to a budget constraint at cohort level we obtain again the Euler equation for the optimal choice of the consumption profile and the condition

$$\frac{-1}{\mu_h(a,h,\bar{H})} = \psi^S(a,t) + \Theta(a,t) \tag{8.4}$$

with ψ^S and $\Theta(a,t)$ denoting the social value of life (SVOL) and the current value of the externality. While the former corresponds to the value of saving the individual's life, the latter provides the monetary value of the externality bestowed on contemporary individuals by affecting their mortality. Generally, $\Theta(a,t)$ is positive (negative) in the case of positive (negative) externalities, where a decrease (increase) in the mortality levels of others raises (lowers) their expected life-cycle utility. The impact measured by $\Theta(a,t)$ is the stronger, the larger the impact of aggregate health care utilization on mortality, the larger the cohorts that are affected, and the larger the SVOL relating to the members of affected cohorts. To illustrate this, consider for instance the introduction of a new method of heart surgery, e.g. minimally invasive or "off pump", in the 1990s or so: Physicians utilizing such a method would have been subject to strong learning-by-doing effects in the beginning, thereby generating positive externalities. Given that at that time the large baby-boom cohorts were just reaching an age range with increased levels of heart disease, the externality would have affected a large co-

hort. Furthermore, while the value of life is typically decreasing with age due to the reduction in remaining life expectancy, it is usually still of substantive value within the age classes 40-60 (e.g. Shepard and Zeckhauser, 1984; Murphy and Topel, 2006; Hall and Jones, 2007). Hence, one would expect for this example a high positive value of the externality.

The SVOL is defined as

$$\psi^S(a,t) = \psi^P(a,t) + \Omega(a,t) \tag{8.5}$$

and thus includes the PVOL and the net value $\Omega(a,t)$ of externalities the individual is expected to generate over its remaining life course. Hence, SVOL is greater (smaller) than PVOL if the individual's future health behaviour is expected to generate strong positive (negative) externalities.

Comparing the planner's choice of health care with the individual's choice we note two inefficiencies. First, corresponding to the term $\Theta(a,t)$, the individual does not take into account the benefit (harm) it bestows on others through the choice of its current health expenditure. Second, when valuing its own survival, the individual does not take into account the pattern and value of externalities, $\Omega(a,t)$, it will generate over its remaining life course. Typically, individuals will under-spend (over-spend) relative to the social optimum in the case of positive (negative) externalities.

We also derive a tax-subsidy-scheme aimed at optimally internalizing the life cycle externalities. Let $\tau(a, t_0 + a)$ denote a (net) subsidy on each unit of private health care spending or, equivalently, on each unit of private health care consumed. Hence, for each unit of care the individual only spends an amount of $1 - \tau(a, t_0 + a)$. In order to balance the budget in expected terms the government levies a (net) lump-sum tax equal to the amount $\tau(a, t_0+a)h^*(a, t_0+a)$, where $h^*(a, t_0+a)$ corresponds to the socially optimal level of health expenditure for an individual aged a at time t. Note that the lump-sum transfer is entirely exogenous to individual decision making.

In Kuhn et al. (2011) we show that the following transfer corrects the individuals' incentives and generates the efficient allocation of both health care and consumption:

$$\tau^*(a,t) = 1 - \frac{\psi^P(a,t)}{\psi^P(a,t) + \Omega(a,t) + \Theta(a,t)} \tag{8.6}$$

Let us now focus on the intuitive case, where for positive externalities, we have $\tau^*(a,t) \in [0,1]$. Here, the transfer rate represents the share of the full social value of health care that is not internalized by the individual. The optimal transfer is then an increasing function of the current value of the externality, $\Theta(a,t)$, and the value of future contributions towards the externality, $\Omega(a,t)$. If externalities do not play a role, we have $\mu_{\overline{H}}(a,t) \to 0$ and, $\Theta(a,t) + \Omega(a,t) \to 0$. Obviously, private choice then approaches the social optimum and $\tau^*(a,t) \to 0$. In the other polar case mortality can only be reduced through collective expenditure $\bar{H}(t)$, i.e. we have $\mu_h(a,t) \to 0$. In this case, $\Theta(a,t) \to \infty$ and thus $\tau^*(a,t) \to 1$, implying that individuals receive health care free of charge at the point of use but have to pay a lump-sum tax $\tau^*(a,t)h^*(a,t) \to h^*(a,t)$. Indeed, this is precisely the solution we would expect under circumstances where survival is a 'pure' public good. The intermediate cases follow immediately.

If $\Theta(a,t) + \Omega(a,t) < 0$ in the case of negative externalities, the transfer constitutes a tax on health expenditure: $\tau^*(a,t) < 0$. Notably, the tax is no mirror image of the subsidy paid in the presence of positive externalities. In particular, we note that $\tau^*(a,t) > -1$ cannot be ruled out if the negative externality becomes strong enough. In the limiting case, where $\psi^P(a,t) + \Omega(a,t) + \Theta(a,t) \to 0$ taxation at (close to) infinite rates effectively excludes the individual from the consumption of health care.

In a number of European countries high private health expenditure can be deducted from income tax, implying an implicit subsidy. Alternatively, a tax (or subsidy) can be implemented by raising or lowering the VAT on health care services and goods. In most European countries, however, the bulk of health expenditure is either levied by taxes (UK, Scandinavian and Southern European countries) or through statutory

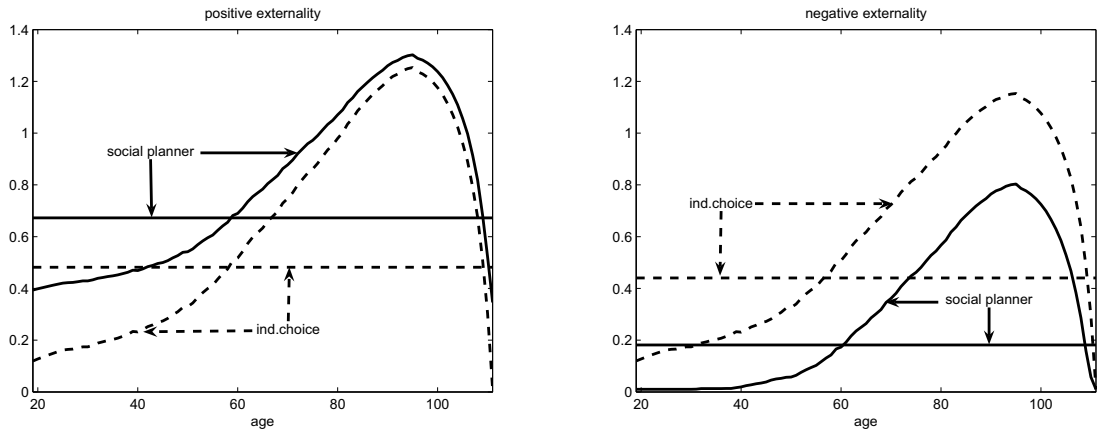
health insurance (Germany, Austria, France). In regard to the latter set of countries we note that we can interpret the transfer scheme in the context of health insurance: In the presence of positive externalities individuals pay an annual premium amounting to $\tau^*(a,t)h^*(a,t)$ and care is then provided subject to a co-payment equal to $1-\tau^*(a,t)$. In reality, of course the design of health insurance contracts follows the trade-off between equalizing consumption across different health states and containing ex-post moral hazard, i.e. the excess consumption of health care free of charge (see e.g. Zweifel et al., 2009). It can be understood intuitively how health insurance should be designed in the joint presence of uncertainty and externalities. With positive externalities the policy-maker seeks to stimulate demand for health care even under certainty and, therefore, generates some 'base-line moral hazard' by setting $\tau^*(a,t) > 0$. In the presence of uncertainty additional moral hazard is then traded-off against insurance, implying that the transfer be set in excess of $\tau^*(a,t)$. From the perspective of externalities, this implies an over-internalization. In contrast, in the presence of negative externalities the policy-maker seeks to discourage the consumption of health care even beyond the level that would be optimal without insurance, implying $\tau^*(a,t) < 0$ and, thus, a co-payment $1-\tau^*(a,t) > 1$. In the presence of uncertainty, the insurance motive would require the co-payment to be reduced, thus implying under-internalization. Note, however, that it is not clear a priori whether insurance with a co-payment below 1 should be provided at all. Indeed, negative externalities militate against the introduction of health insurance.

8.2.1. Numerical Results

To illustrate the results of our model we calculate a numerical example. The influence of individual health expenditures and health care on mortality is modeled according to the proportional hazard level (see Kalbfleisch and Prentice, 1980). We combine this with a constant elasticity of utility function. We use US income and mortality data and set the parameters in a way that the resulting levels of health care and consumption

as well as the value of life corresponds broadly to the scenario considered in Hall and Jones (2007). For a detailed list of parameters and data used we refer to Kuhn et al. (2011). The simulations have been carried out for a steady state (note that the theoretical results are valid also for a transitional path).

Figure 8.1.: Age-specific and Average (Flat Line) Expenditures in the Social Planner and the Individual Choice Model (in Ten Thousands of Dollars)



In the following we present some results for positive (left panel) and negative externalities (right panel). Unsurprisingly, too little (too much) is spent on health care in the case of positive (negative) spillovers, as is graphed in Figure 8.1. More generally, and for any of the cases, the age-profile of health expenditure is hump-shaped. Due to very low base mortality, for the individual there is little to be gained from health care at the youngest ages. While with advancing age the increase in base mortality renders the purchase of health care more and more effective, this is eventually offset for the highest ages, where a falling PVOL and age-related declines in effectiveness lead to a drop in expenditure.[3] This notwithstanding, from a social point of view it is optimal in

[3] A hump-shaped profile of health expenditure stands in contrast to the observation that in most countries (including the US) health expenditure strictly increases with age. The difference arises as our expenditure patterns follow the statistical or *ex-ante* VOL, which typically decreases from some age onwards (see e.g. the numerical exercises in Shepard and Zeckhauser, 1984; Murphy and Topel, 2006). As Philipson et al (2010) argue, however, real health expenditure is driven by the *ex-post* VOL once a life threatening condition has materialized. At this point individuals are typically willing to spend a manifold of the ex-ante VOL. The bunching of life threatening situations at high ages then

the presence of positive spillovers to maintain significant spending levels even for the youngest ages in order to ensure sufficient contributions towards aggregate expenditure.

Figure 8.2.: Difference in Life Expectancy

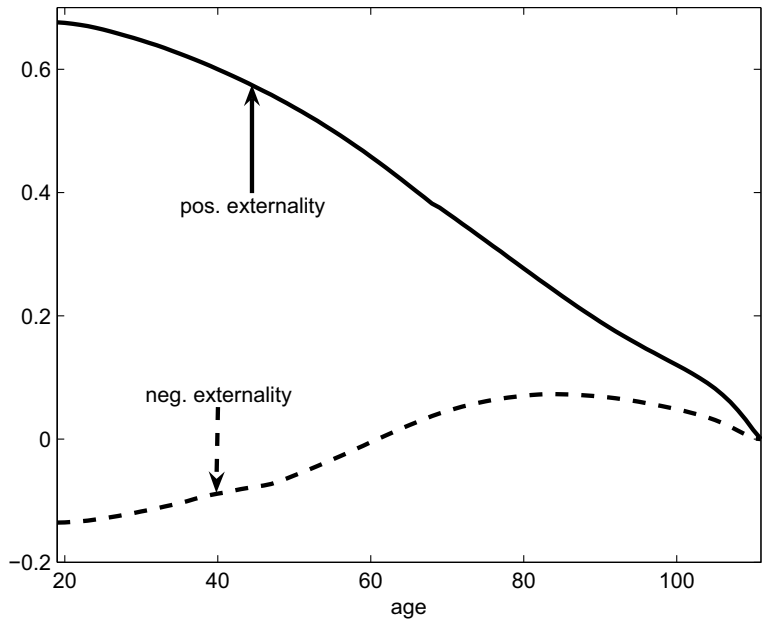
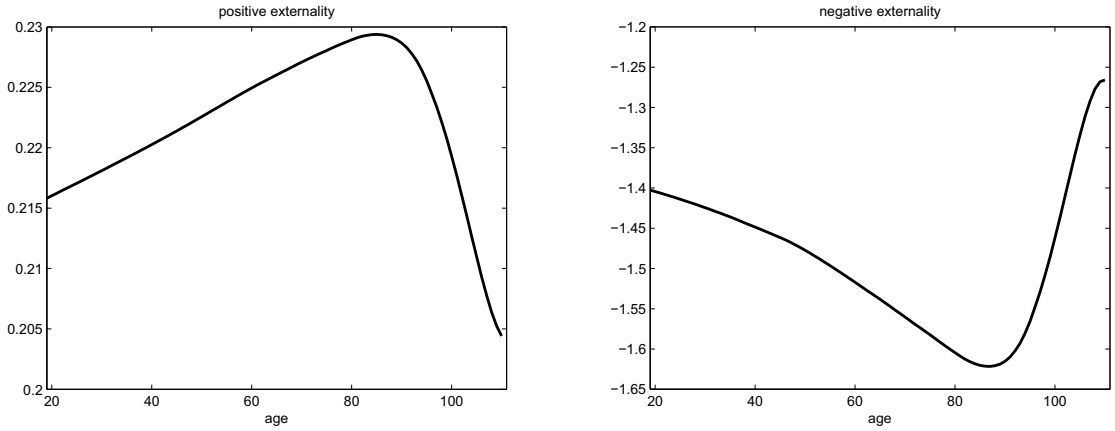


Figure 8.2 plots the net increase of life expectancy at age a that is attainable if the optimal pattern of health expenditure is induced. In the presence of positive externalities (the upper graph), for instance, the life expectancy at age 20 would increase by about 8 months if individuals could be induced to spend optimally. In contrast, the impact on life expectancy of an internalization of negative externalities is small. The reason is that individual efforts to lower mortality and the externality are to a large extent neutralizing each other. individual over-spending in the presence of negative externalities is ineffective to a large extent and only leads to a reduction of consumption: a tread-mill effect. Thus, while the subsidizing of health care in the presence of positive

implies the increasing spending pattern. As our analysis is predominantly of a normative nature and not targeted at a realistic VOL per se but rather how it should be amended in the presence of externalities, the discrepancy between our simulated expenditure and real-world expenditure is of minor consequence.

spillovers would lead to a sizeable increase in life expectancy; the taxation of care in the presence of negative spillovers will predominantly induce a shift from (excessive) health expenditure to consumption without affecting mortality.

Figure 8.3.: Transfer



Finally, consider the age-profile of the optimal transfer, as in Figure 8.3. A first inspection shows that, unsurprisingly, the transfer constitutes a subsidy (tax) in the presence of positive (negative) spillovers. The age-profile of the transfer is driven by changes of the PVOL and the discounted value of future contributions towards the externality, Ω. In the case of positive externalities, the decline in the PVOL suggests a tendency towards an increase in the transfer. Indeed, up to around age 60 this trend is complemented by the increase in the value of future contributions, both widening the wedge between private and social incentives. For higher ages, a sharp decline in the value of future contributions becomes dominant, thus leading to a closing gap in incentives and, thus, to a reduction in the subsidy. For the case of negative externalities the SVOL declines at a higher rate than the PVOL, implying an increasing tendency towards excess spending. The tax should therefore be increased with age, an effect that is reversed only for ages above the mid 80s, where a strong decline in the (absolute) value of the future contributions towards the externality now generates a tendency towards lower taxation.

A further comparison reveals that for negative as opposed to positive externalities the absolute value of the transfer tends to be greater by a little less than an order of magnitude. This reflects two distinguishing features of the two types of externalities. First, while private and social incentives are complementary to each other in the case of positive externalities, they are substitutes in the case of negative spillovers. This implies for the former case that even moderate subsidies are sufficient to complement the private incentive, while for the latter case the policy-maker needs strong taxes to reverse private incentives. Second, it can be shown that the current value of the externality (in absolute terms) is significantly higher in the presence of negative externalities. This reflects the higher levels of mortality, translating into a larger gap between private and social incentives and, therefore, calling for a stronger policy.

8.3. Implications for Ageing Societies

Before discussing some of the policy implications of our analysis we should highlight a number of important caveats. Fist, our analysis is predominantly of a theoretical nature and, therefore, highly stylized on a number of important dimensions. By assuming that the spillovers flow through current aggregate health care expenditure we presume (i) that all age-groups contribute in a symmetric way and, more importantly, (ii) that there are no cumulative effects of spending, as would be present in the context of stock externalities. Furthermore, we abstract from all institutional detail that is prone to shape the incentives. It is easy to see that the model is therefore an ill fit for most of the real world externalities referred to in the introduction. Second, the reality is of course made up of a multitude of externalities of particular types, relating to particular conditions and to particular age-groups. Thus, from a macro-perspective a host of positive and negative externalities are prone to overlap and offset each other with no empirical evidence as to the net effect. Finally, as highlighted in the introductory

section, we are unaware of any empirical evidence regarding the discrepancy between private and social incentives and can only guess at the overall magnitude of externalities - although some of the reading suggests that at some points in time and for some health care systems it may be very substantive.

Thus, it is clear that any of the following implications must be called tentative at best. Rather than addressing the whole list of externalities, in the following we focus on the two types of externalities we deem to be of greatest relevance in the context of ageing societies within developed (European) countries: These refer to the incentives for developing new (possibly age-specific) technologies depending on the overall (possibly age-specific) level of health expenditure; and to the presence of crowding effects within a health care system.

Murphy and Topel (2006) estimate at some 61 trillion USD the value of the longevity increase in the US between 1970 and 2000 net of the additional medical costs induced by the then-survivors. They also show, however, that the net gain has somewhat declined over time with two thirds of it falling on the 1970s and that it varies considerably with age and sex. While the availability of life saving technologies brings the greatest gains to the age-group 35-55 it tends to be associated with net social costs for the old and, in particular, for the female old. Murphy and Topel (2007) consider the supply-side incentives related to the value of longevity increases calculated in Murphy and Topel (2006). In particular, they show that the incentive for developing a new technology, and in turn the expected time to innovation, increases with the share of expenditure that can be appropriated by the innovating firm. What they develop within a stylized theoretical model is confirmed by Acemoglu and Linn (2004) who show with data from the US pharmaceutical industry that a larger (expected) market size tends to induce stronger entry of new drugs both, generic and on-patent drugs, including new molecular entities. More specifically, they show that the shift in the age profile towards older age-groups within the US population has been followed by a shift in the

targeting of drug innovation from the young to the middle aged and elderly. Thus, while one can be confident that innovation incentives shift with the ageing of the population there is an issue about whether these incentives are beneficial from a societal perspective.

The presence of positive spillovers suggests that the overall level of health expenditure and, thus, the investment in new life-saving technologies may be too small unless health expenditures are sufficiently subsidized. If individuals are required to pay out of pocket they do not include in their decision the expected benefit to society from the development of innovative technologies. As of now, this appears an unlikely problem for the European health care systems, where access to health care is predominantly free of charge with only modest co-payments in place. However, the dual pressure on public budgets stemming from unsustainable public debt, on the one side, and population ageing, on the other, leads to the prospect of health insurance coverage (or otherwise the availability of free public health services) being reduced to basic packages in the future, leaving many services to be paid for out of pocket. Such a development would lead to a direct reduction in the incentives to innovate, which after all may even be a desired effect. However, our analysis suggests that it may also compromise efficiency if the under-consumption of health care leads to an inefficiently low level of innovation and, therefore, to inefficiently high levels of mortality.

The finding by Murphy and Topel (2007)that life-saving innovation tends to be least beneficial from a social point of view for the oldest age groups may suggest it be efficient to ration access by these groups to certain forms of medical intervention. Disregarding ethical concerns about such a policy, it may be unwarranted even on efficiency grounds in the presence of externalities. Even if the highest age-groups do not stand to benefit much from the new technologies their access to these technologies and the resulting expenditure may trigger innovation incentives from which younger age-groups stand to benefit. In such a case, helping the survival of the oldest may generate

a positive externality on the survival of the middle aged.

Hospital crowding continues to be an issue in both the US and most European countries (Jayaprakesh et al., 2009; Pines et al., 2011). Medical and economic studies show that crowding is typically associated with poorer health outcomes even when risk factors among the patient population are controlled for. Clement et al. (2008) study the impact of crowding on mortality rates in a set of US hospitals in the wake of policy-mandated reimbursement constraints. While about a third of the hospitals in their sample were able to cope with the budget cuts without instances of crowding, hospital congestion led to a significant increase in mortality where it occurred. Furthermore, the share of elderly Medicare patients was weakly positively associated with crowding, which seems intuitive in the light of the greater demand for services by the elderly. While there is again no direct evidence from Europe it appears that a tightening of health care budgets, i.e. a reduction in capacity, together with increasing demand for services by an elderly population, is likely to foster congestion within the health care system and thereby provide the basis for negative externalities. If there is no way of relaxing the problem on the supply side then our results suggest that incentives should be provided to reduce the demand for services. Our results suggest that this would not lead to a strong increase in mortality but would then help to contain ineffective health care spending.

However, our analysis also shows that the implied tax on health care, implemented for instance in the form of co-payments, may be substantial and that it should increase with age up to the years with peak medical spending. The reduction in insurance coverage for those patients with the greatest demand for health care would clearly amount to a controversial policy, and is unlikely to gain political support. While this illustrates some of the difficulties in dealing with negative spillovers within the health care system, we should like to conclude this chapter by stressing the need for more parsimonious modeling to better understand particular types of externalities and the need for

empirical analysis to better quantify the externalities and their welfare implications.

8.4. References

Acemoglu, D., Linn, J., 2004, Market size in innovation: theory and evidence from the pharmaceutical industry, *Quarterly Journal of Economics* 119, 1049-1090.

Clement, J.P., Valdmanis, V.G., Bazzoli, G.J., Zhao, M., Chukmaitov, A., 2008, Is more better? An analysis of hospital outcomes and efficiency with a DEA model of output congestion, *Health Care Management Science* 11, 67-77.

Cohen, F.L., Tartasky, D., 1997, Microbial resistance to drug therapy: A review, *American Journal of Infection Control* 25, 51-64.

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

Easterlin, R.A., 1999, How beneficent is the market? A look at the modern history of mortality, *European Review of Economic History* 3, 257-294.

Ehrlich, I., 2000, Uncertain lifetime, life protection, and the value of life saving, *Journal of Health Economics* 19, 341-367.

Ehrlich, I., Chuma, H., 1990, A model of the demand for longevity and the value of life extension, *Journal of Political Economy* 98, 761-782.

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

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

Jayprakash, N., 2009, Crowding and delivery of health care in emergency departments: the European perspective, *Western Journal of Emergency Medicine* 10, 233-239.

Johansson, P.-O., 2002, On the definition and age-dependency of the value of a statistical life, *Journal of Risk and Uncertainty* 25, 251-263.

Kalbfleisch, J.D., Prentice, R.L., 1980, *The Statistical Analysis of Failure Time Data*, Wiley & Sons., New York.

Keyfitz, B.L., Keyfitz, N., 1997, The McKendrick partial differential equation and its uses in epidemiology and population study, *Mathematical Computation Modeling* 26, 1-9.

Kuhn, M., Wrzaczek, S, Prskawetz, A., and Feichtinger, G., 2011, Externalities in a life cycle model with endogenous survival. *Journal of Mathematical Economics* 47, 627-641.

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

Murphy K.M., Topel R.H., 2007, Social value and the speed of innovation, *American Economic Review* 97, 433-437.

Phillips, K.A., Luft, H., 1997, The policy implications of using hospital and physician volumes as 'indicators' of quality of care in a changing health care environment., *International Journal of Quality in Health Care* 9, 341-348.

Philipson, T., 2000, Economic epidemiology and infectious diseases, in: Culyer A.C., Newhouse J.P. (Eds.), *Handbook of Health Economics* (Vol. 1B). Elsevier, Amsterdam.

Philipson, T.J., Becker, G., Goldman, D., Murphy, K.M., 2010, Terminal care and the value of life near its end, NBER Working Paper 15649.

Pines, J.M. et al, 2011, International perspectives on emergency department crowding, *Academic Emergency Medicine* 18, 1358-1370.

Rosen, S., 1988, The value of changes in life expectancy, *Journal of Risk and Uncertainty* 11, 285-304.

Shepard, D.S., Zeckhauser, R.J., 1984, Survival versus consumption, *Management Science* 30, 423-439.

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

Zweifel, P., Breyer, F., Kifmann, M., 2009, *Health Economics*, Springer, Berlin, Heidelberg.

9. Immigration, Evolution of Skills and Social Security

by María Dolores Guilló, Alex Pérez-Laborda, and Fidel Perez-Sebastian

Abstract. Migration is sometimes seen as a panacea to circumvent the financial problems associated to ageing societies. Other authors, however, see it as a burden, because the arrival of migrant workers can depress the wage of non-educated natives and reduce the economy's average productivity. In this paper, we analyze the effect of migration on workers' productivity and the skill premium. We construct a life-cycle model with endogenous educational choices and health investment. Our results point out that, when human capital investment later in the life cycle is possible, migration does not universally increase the wage gap, nor depresses the economy's productivity; this depends on the economy's initial conditions.

9.1. Introduction

The challenges faced by the European economies will grow over the next decades as improvements in human longevity contribute to a significant increase in the proportion of aged population. According to Eurostat projections, the EU27 population is projected to continue to grow older, with the share of the population aged 65 years

and over rising from a 17.1% in 2008 to 30.0% in 2060 (Table 9.1). Moreover, the share of the population belonging to the oldest groups is also projected to increase considerably; those aged 80 and over are expected to rise from 4.4% to 12.1% over the same period.

Table 9.1.: Percentage of Older Population in EU27

	Percentage aged 65+			Percentage aged 80+			Old age dependency ratio (%)	
	2008	2035	2060	2008	2035	2060	2008	2060
EU27	**17.1**	**25.4**	**30.0**	**4.4**	**7.9**	**12.1**	**25.4**	**53.5**
Belgium	17.0	24.2	26.5	4.7	7.4	10.2	**25.8**	**45.8**
Bulgaria	17.3	24.7	34.2	3.6	7.1	12.8	**25.0**	**63.5**
Czech Republic	14.6	24.1	33.4	3.4	7.9	13.4	20.6	61.4
Denmark	15.6	24.1	25.0	4.1	7.7	10.0	23.6	42.7
Germany	20.1	30.2	32.5	4.7	8.9	13.2	30.3	59.1
Estonia	17.2	22.8	30.7	3.6	6.8	10.7	25.2	55.6
Ireland	11.2	17.6	25.2	2.8	5.0	9.6	16.3	43.6
Greece	18.6	26.3	31.7	4.1	7.9	13.5	27.8	57.1
Spain	16.6	24.8	32.3	4.6	7.2	14.5	24.2	59.1
France[3]	16.5	24.4	25.9	5.0	8.5	10.8	25.3	45.2
Italy	20.1	28.6	32.7	5.5	9.1	14.9	30.5	59.3
Cyprus	12.4	19.0	26.2	2.8	5.3	8.6	17.7	44.5
Latvia	17.3	23.7	34.4	3.6	6.7	11.9	25.0	64.5
Lithuania	15.8	24.3	34.7	3.3	6.4	12.0	23.0	65.7
Luxembourg	14.2	21.3	23.6	3.5	5.8	8.9	20.9	39.1
Hungary	16.2	23.1	31.9	3.7	7.6	12.6	23.5	57.6
Malta	13.8	24.8	32.4	3.2	8.3	11.8	19.8	59.1
Netherlands	14.7	25.9	27.3	3.8	8.0	10.9	21.8	47.2
Austria	17.2	26.1	29.0	4.6	7.2	11.4	25.4	50.6
Poland	13.5	24.2	36.2	3.0	7.7	13.1	19.0	69.0
Portugal	17.4	24.9	30.9	4.2	7.6	12.8	25.9	54.8
Romania	14.9	22.9	35.0	2.8	6.2	13.1	21.3	65.3
Slovenia	16.1	27.4	33.4	3.5	8.4	13.9	23.0	62.2
Slovakia	12.0	23.0	36.1	2.6	6.4	13.2	16.6	68.5
Finland	16.5	26.4	27.8	4.3	9.4	10.8	24.8	49.3
Sweden	17.5	23.6	26.6	5.3	8.1	10.0	26.7	46.7
United Kingdom	16.1	21.9	24.7	4.5	6.7	9.0	24.3	42.1
Norway	14.6	22.6	25.4	4.6	7.1	10.0	22.1	43.9
Switzerland	16.4	25.2	28.0	4.7	7.7	11.1	24.1	48.5

Source: eurostat (2011).

One of the main implications of this aging process is the negative effect that the increase of dependency ratios has on the viability of the current pay-as-you-go pension systems (that is, wether the programs begin to pay out more than are bringing in).[1]

[1] In addition, public pension schemes are severely affected by the current financial and economic crisis. Under proportional and progressive schemes, high unemployment and lower earnings is reducing the contribution to pensions systems exacerbating the problem due to aging.

In fact, the effect of ageing on pensions has been, and will continue to be, a subject of intense academic research. There is a consensus in the literature that the demographic transition will make the current pensions systems of most European countries unsustainable in the long-term. This fact has recently lead to the introduction of social security system reforms, mostly by delaying the retirement age.

Together with the ageing process, immigration flows towards European countries have been also increasing in recent decades. As longevity increases and fertility remains steady below replacement rates, positive net migration has become the only population growth factor in many EU countries. Consequently, migration has been lately regarded as positive by governments: the entry of young and middle age immigrants would increase working population which may help to mitigate the European problems associated to an ageing society.

Nevertheless, the net contribution of migration may be quite modest (Borjas (1995), Feldstein (2006)). First, the average productivity of a migrant is lower than the native as a consequence of less years of education on average, because of non-recognized educational credentials or just due to cultural distance. As the average immigrant has lower salary, the contributions she makes are less than those of an average native, especially under very progressive tax systems (as in EU). Second, migrants are also consumers of government benefits (especially health and education). Evidence of the US suggests that the immigrants are greater consumer of public assistance, since the immigrant household is greater and poorest on average than the native one, what makes them eligible for social benefits (Borjas and Hinton (1996)).

It is even more important that migration may only be a temporary solution, since the immigrant population will also become old and eventually beneficiary of the system as well.[2] Storesletten (2000) calibrates an OLG model for the US economy and concludes that the effect of immigrants in sustaining social security varies with their age and

[2]Although some may return to their country of origin before eligibility for social security benefits.

skills. According to his predictions, migration has a positive effect only if migration policy is directed to middle aged high skilled immigrants.

This is in contrast with the prior views of Smith and Edmonston (1996) and Lee and Miller (2000) who find that positive effects for the economy are higher if immigrants are younger. For the case of Europe, employing generational accounting methods Bonin et al. (2000) and Collado et al. (2004), find that increasing the number of immigrants reduces the financial burden for future natives in Germany and Spain, respectively.

Recently, Gonzalez et al. (2009) also made use of an OLG model. In their study, they use demographic and immigration projections from the Instituto Nacional de Estadística (INE) to understand the impact of immigration on the Spanish pension system. They also simulate the full labor history of different workers (differing by age, gender, skill and nationality). Their findings show that, with current projections, immigration will not succeed to balance the Spanish pension system even under complete assimilation of immigrants (in terms of their revenues and contributions to social security). They may provide, however, additional time to undertake the necessary reforms.

Besides their roll as tax-payers, immigrants also have changed labor markets. Since migration represents an increasing share of the labor force, the mix of skills, education and productivity in the economy is increasingly determined by the skills, education and productivity of the foreign born population. A very large amount of the literature has focused in analyzing the impact of immigrants on the wages of natives.[3] Immigration will affect the wage inequality among natives if the skill distribution of immigrants differs from that of natives. The basic argument here is that since the average immigrant is less skilled than the native, migration would reduce a country's average productivity and also the wages of unskilled native workers (or in economies subject to rigidities, as in Europe, an increase the unskilled unemployment).

Although there is some empirical evidence that this is the case for the US (Borjas

[3]See for instance the surveys of Borjas (1994), (1999) Friedberg and Hunt (1995) or Card (2005))

(2003), empirical studies for Europe show little or no evidence of any impact of migration on wages.[4] Among the possible explanations of this fact, it has been argued that immigrant and native workers are not perfect substitutes, so that an increase of immigrant workers do not have to depress the wages of the native (Octaviano and Peri (2008)). Together with the wage premium, migration may also impact the accumulation of human capital. Since incentives to complete education are influenced by the wage structure, immigration may increase the price of human capital and, in particular, natives' high school educational attainment (Betts (1998), Hunt (2011)).

Since the benefit from migration for the receiving economy depends on the impact of migration on country's productivity, we believe that it is essential to understand jointly the roll of migration on both the labor market and human capital formation. Previous literature has mainly ignored the impact of migration on the incentives to acquire skills.[5] Although the assumptions of variable wages but fixed skill premium (Storesletten (2000)) introduce a link between migration and labor market outcomes, migration has the same effect on both skilled and unskilled workers. This is not satisfactory since migration is more likely to hit unskilled natives stronger.

We depart from the existent literature by considering the interaction between these two effects. In order to do so, we analyze migration in an overlapping generation model (OLG) with endogenous human capital formation (Ben-Porath (1969), Cunha and Heckman (2007)), where individuals can invest in human and physical capital in the different periods of their lives. The importance of jointly modeling labor supply and human capital decisions was already stressed by Becker (1967). Furthermore, by allowing individuals to undertake human capital investments at adult stages (health

[4]See Pischke and Velling (1997) for Germany, Zorlu and Hartog (2005) for the Netherlands and Norway, Dustmann et al. (2005) for the U.K. and Carrasco et al. (2008) for Spain.

[5]A noticeable exception may be found in Casarico and the Villanova (2003) who employ a theoretical model to analyze the effect of an unskilled migration shock when skill and unskilled workers are not perfect substitutes. The authors find that the unskilled migration shock modifies schooling premium and increases the incentives of natives to undertake more schooling. Thus, although in their setting migration helps to sustain social security, the arrival of migrants has strong redistributive effects.

investments and/or on-the job-training) they can compensate human capital depreciation when old.

This is of capital importance when one has in mind policy reforms affecting the retirement age. There is an agreement in the neuroscience literature that both health and mental training play an important role in diminishing the effects of physical and cognitive decline (Bosworth and Schaie (1997)), Adam et al. (2006), Scarmeas (2003)) or Stern (2002, 2003)). Furthermore, given that these investments may also be undertaken by the current immigrant population (even if assumed to be unskilled), a migration shock may increase the incentives to acquire human capital for the immigrant group as well.

The model allows us to shed some light over a broad set of important questions concerning migration, productivity and aging and the relation among them. More specifically, it allows looking at the skill and age composition of the labor force and the evolution of the skill premium; and at natives' education choice and human capital investments.

Our results suggest that, as expected, a massive entry of young unskilled migrant workers initially increases the wage premium and hence also the share of educated native workers. However, if human capital investment is possible at adult stages, it is no longer true that the final wage premium has necessarily to increase, or average productivity to decline as stated by the previous literature. The total impact of unskilled migration on the wage premium and the economy as a whole, ultimately depends on the initial average productivity of an educated versus an non-educated worker, and not only on the fraction of educated versus non-educated workers. This result is of importance for the policy maker, since it implies that the attitude towards migration (from a destination country viewpoint) depends on the particular country and should not be based on a general receipt. Immigration therefore may help to sustained the pension system in the short- and medium-run.

The remaining of the paper is organized as follows. Section 2 review some facts

concerning migration towards European countries. The model is presented in section 3. Equilibrium results may be found in section 4. Section 5 concludes.

9.2. Some Facts about EU Migration

From being traditional countries of emigration, many European countries have become a destination for international migration. According to Eurostat, in 2009 around 31.9 million people in EU27 are foreign citizens, representing a 6.4 percent of total EU27 population, being 4 percent citizens of countries outside EU27 (see Table 9.2).

Although foreign born residents represent a sizable part of total EU population, the numbers are still far from those of traditional migration countries: US, Canada and Australia have bigger shares of foreign born population (12 percent, 20 percent and 24 percent, respectively). Detailed statistics for each country may also be found in Table 9.2).. Germany accounts for the biggest number of foreign citizenship, followed by Spain, UK and France. But Luxembourg is the country with the largest fraction of immigrants with respect to total population, although most of them are EU citizens.

Concerning the nationalities of the EU immigrants, in 2009 around 40 percent of foreign EU27 residents are citizens of another EU27 state. Considering a wider Europe (including Turkey), more than a third of immigrants come from non-UE27 countries, followed by Africa, Asia and (Latin) America.

Migration flow data show that migration towards EU has decelerated in recent years, mostly as a consequence of the strong crisis hitting European countries. During 2009, about 3.0 million people immigrated into one of the EU Member States, while at least 1.9 million emigrants were reported to have left an EU Member State. UK, Spain and Italy reported the largest number of immigrants in 2009, but also the highest number of emigrants. Although most of EU Member States reported more immigration than emigration in 2009, Ireland, Malta and the three Baltic Member States already presented

Table 9.2.: Immigration by Main Citizenship Group, 2009

	Total foreign citizens		Citizens of another EU27 Member State		Citizens of countries outside the EU27	
	000s	% of total population	000s	% of total population	000s	% of total population
EU27*	**31 860.3**	**6.4**	**11 944.2**	**2.4**	**19 916.2**	**4.0**
Belgium**	971.4	9.1	659.4	6.2	312.0	2.9
Bulgaria	23.8	0.3	3.5	0.0	20.3	0.3
Czech Republic	407.5	3.9	145.8	1.4	261.7	2.5
Denmark	320.0	5.8	108.7	2.0	211.4	3.8
Germany	7 185.9	8.8	2 530.7	3.1	4 655.2	5.7
Estonia[3]	214.4	16.0	9.6	0.7	204.8	15.3
Ireland	504.1	11.3	364.8	8.2	139.2	3.1
Greece	929.5	8.3	161.6	1.4	767.9	6.8
Spain	5 651.0	12.3	2 274.2	5.0	3 376.8	7.4
France	3 737.5	5.8	1 302.4	2.0	2 435.2	3.8
Italy	3 891.3	6.5	1 131.8	1.9	2 759.5	4.6
Cyprus	128.2	16.1	78.2	9.8	50.0	6.3
Latvia[3]	404.0	17.9	9.4	0.4	394.6	17.5
Lithuania	41.5	1.2	2.5	0.1	39.0	1.2
Luxembourg	214.8	43.5	185.4	37.6	29.5	6.0
Hungary	186.4	1.9	109.8	1.1	76.6	0.8
Malta	18.1	4.4	8.2	2.0	9.9	2.4
Netherlands	637.1	3.9	290.4	1.8	346.7	2.1
Austria	864.4	10.3	317.0	3.8	547.4	6.6
Poland***	35.9	0.1	10.3	0.0	25.6	0.1
Portugal	443.1	4.2	84.7	0.8	358.4	3.4
Romania	31.4	0.1	6.0	0.0	25.3	0.1
Slovenia	70.6	3.5	4.2	0.2	66.4	3.3
Slovakia	52.5	1.0	32.7	0.6	19.8	0.4
Finland	142.3	2.7	51.9	1.0	90.4	1.7
Sweden	547.7	5.9	255.6	2.8	292.1	3.2
United Kingdom**	4 020.8	6.6	1 614.8	2.6	2 406.0	3.9
Iceland	24.4	7.6	19.4	6.1	5.0	1.6
Norway	302.9	6.3	165.4	3.4	137.6	2.9
Switzerland	1 669.7	21.7	1 033.6	13.4	636.1	8.3
Turkey	103.8	0.1	45.3	0.1	58.4	0.1

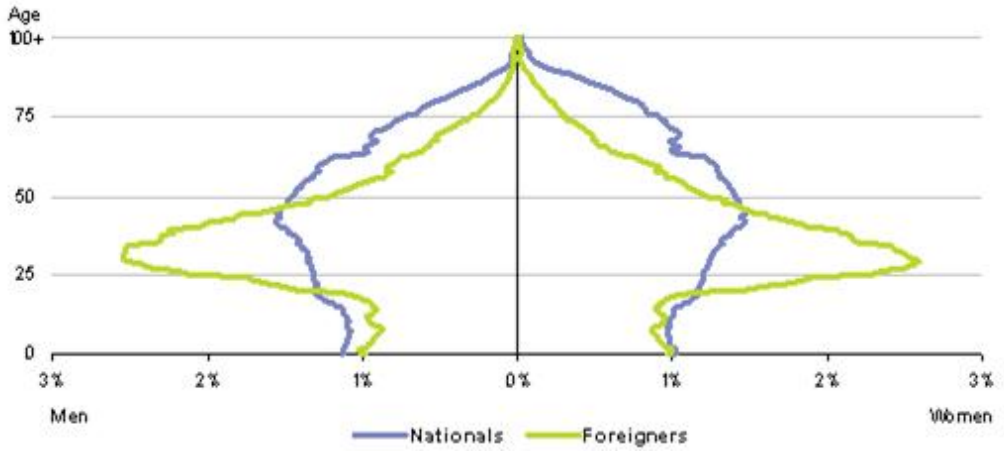
Source: eurostat (2011).

negative net migration balances.

The profile of the EU immigrant is a younger and less skilled individual than the average native. Foreign citizens living in the EU27 were significantly younger than the population of nationals with median age of 34.4 years compared with the 40.4 years of age for a native. Age differences between foreign and native populations are better understood by looking at population pyramids (Figure 9.1). As a consequence, immigration has slowed down population aging in European countries. Nevertheless,

medium and long term projections continue to forecast a considerable aging process.

Figure 9.1.: Age Structure of the National and Non-national Populations



, EU, 2010. Source: eurostat (2011).

No particular effort has been done by EU countries to attract high skilled people until recent dates. In 2000, 55 percent of foreign born population in the former EU15 was estimated to have less than secondary education. The average immigrant in EU has less years of education than the EU native. As Figure 9.2 shows, this number contrasts with the foreign born adults in North America (U.S. and Canada) who have, on average, more years of schooling. Table 9.3 presents the percentage of working age population by place of birth, level of education (Low, Medium and High) and country of residence, using data from the European Labor Force Survey (LFS), as it appears in Münz (2008). In the table, we can see that some countries are better at attracting high skill immigrants than others. For example, Ireland, Denmark and Estonia have a sizable proportion of high skilled adult foreign born population. In contrast, other countries as Portugal, France, Austria or Spain had mostly attracted low skilled workers.

Table 9.3.: Population Aged 25-65 by Place of Birth, Level of Education, and Country of Residence, 2005

	Born in country of residence			Born in an other EU27 country			Born in a country outside EU27		
	Low	Medium	High	Low	Medium	High	Low	Medium	High
EU27	**28.1**	**47.6**	**24.3**	**30.7**	**41.0**	**28.3**	**36.3**	**37.9**	**25.8**
Austria	16.5	65.8	17.7	14.0	57.7	28.3	45.6	41.5	12.9
Belgium	32.7	36.2	31.1	41.8	26.5	31.7	48.3	25.4	26.3
Cyprus	33.9	40.2	26.0	25.1	31.8	43.1	38.1	29.5	32.4
Czech Republic	9.9	77.2	13.0	23.6	62.2	14.3	15.9	54.2	29.9
Denmark	17.0	50.5	32.4	(10.6)	42.2	47.2	26.4	35.7	37.8
Estonia	11.0	56.2	32.8	:	:	:	10.5	52.5	37.0
Finland	20.8	44.6	34.6	20.5	47.0	32.5	28.3	44.8	26.9
France	31.3	43.5	25.2	51.0	28.7	20.3	47.6	27.9	24.5
Germany	12.4	62.2	25.4	:	:	:	:	:	:
Greece	40.4	38.9	20.8	25.3	51.3	23.4	44.4	40.5	15.0
Hungary	24.1	59.0	16.8	16.4	60.8	22.8	11.0	57.9	31.1
Ireland	37.0	35.9	27.2	25.5	35.5	39.0	13.1	27.9	59.0
Italy	50.0	38.1	11.9	:	:	:	:	:	:
Latvia	16.7	62.4	20.9	(33.7)	43.6	:	12.1	62.6	25.3
Lithuania	13.1	60.5	26.5	:	:	:	7.7	65.3	27.0
Malta	74.7	13.7	11.5	68.2	10.9	20.9	50.4	26.1	23.5
Netherlands	28.0	40.8	31.2	14.9	51.2	33.9	33.8	44.1	22.1
Poland	15.3	68.2	16.5	38.7	47.4	(13.9)	(19.9)	58.1	22.0
Portugal	75.7	12.5	11.8	45.3	27.9	26.8	50.5	25.9	23.6
Slovakia	12.3	73.9	13.8	(15.5)	63.9	20.6	:	:	:
Slovenia	18.4	60.7	20.8	(21.8)	(60.9)	(17.3)	30.3	57.5	12.2
Spain	52.8	19.1	28.2	32.2	33.0	34.8	43.9	30.0	26.1
Sweden	15.7	55.1	29.2	16.6	50.3	33.1	23.0	46.1	30.9
United Kingdom	14.4	56.2	29.5	14.8	56.7	28.6	20.0	50.0	30.0

Notes:
[1] Incomplete EU27 average: education levels of natives do not include data for Bulgaria, Luxembourg and Romania; education levels of immigrants (born in another EU27 country or outside EU27) do not include data for Bulgaria, Germany, Italy, Luxemburg and Romania.
Data in brackets are of limited reliability due to the small sample size.

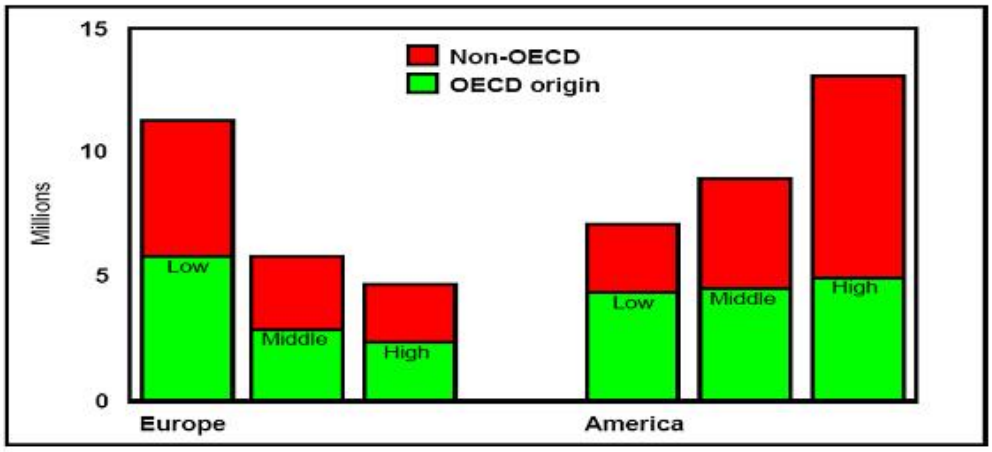
Source: European Labour Force Survey (LFS). From Münz (2008)

Source: eurostat (2011).

9.3. The Model

Let us consider first a closed developed economy to analyze the interplay between the educational choices and human capital investments and the endogenous determination of the skill premium in wages and productivity. Then we will study the effect of an unexpected inflow of workers on the skill premium and overall productivity.

Figure 9.2.: Adult Migrants in OECD Europe and North America, by Skill Level



Source: eurostat (2011).

The economy is populated by two-period overlapping generations of individuals with exogenous population growth.[6] Agents are endowed with a unit of time in each period, they don't value leisure and only care about consumption when old. During the first period of life a young individual has to decide whether she wants to go to college and become a high skill worker. If she does, a fraction of her time endowment when young, $e^H = e$, has to be allocated to education, where e is an idiosyncratic effort/ability level uniformly distributed on the interval $[0, 1]$. If she becomes a high skill worker, she will offer $(1 - e)\, h > 1$ units of efficiency labor during the first period, otherwise she will offer just 1 efficiency unit. Human capital of all types depreciates at a rate that will depend on the amount of health investment (this can also be interpreted as on the job training) during the first period of life. All markets are competitive.

9.3.1. Firms

The final goods production technology of the representative firm is

[6] We can think of young generations of workers aged between 15 and 55, and old generations aged between 55 and 95.

$$Y_t = AK_t^{\alpha} Q_t^{1-\alpha} \tag{9.1}$$

$$Q_t = (\mu_L L_t^{\rho} + \mu_H H_t^{\rho})^{1/\rho} \tag{9.2}$$

where Y_t is output, K_t is physical capital and Q_t is efficiency units of labor, where L_t and H_t represent low and high skill labor, respectively, both measured in units of efficiency. Let us call $k_t = K_t/Q_t$ and $x_t = H_t/L_t$, then assuming full depreciation of capital, the representative firm's optimality conditions are:

$$1 + r_t = \alpha A k_t^{\alpha-1} \tag{9.3}$$

$$w_t^L = (1-\alpha) A k_t^{\alpha} (\mu_L + \mu_H x_t^{\rho})^{-1+1/\rho} \mu_L, \tag{9.4}$$

$$w_t^H = (1-\alpha) A k_t^{\alpha} (\mu_L + \mu_H x_t^{\rho})^{-1+1/\rho} x_t^{\rho-1} \mu_H, \tag{9.5}$$

where r_t is the interest rate, w_t^i is the wage of i skill labor. Notice that the skill premium implied by this type of technology satisfies the following condition:

$$\frac{w_t^H}{w_t^L} = x_t^{\rho-1} \frac{\mu_H}{\mu_L}. \tag{9.6}$$

Finally, denoting by N_t the total population size at period t, we can express output per capita as follows:

$$y_t = A k_t^{\alpha} (\mu_L + \mu_H x_t^{\rho})^{1/\rho} L_t/N_t. \tag{9.7}$$

9.3.2. Households

All individuals have the same logarithmic preferences on future consumption and do not value leisure. Thus, first period labor income will be saved or invested in health.

In the first period an agent has to decide if she wants to go to college or not, this is a discrete choice of $e^i \in \left\{e^H = e, e^L = 0\right\}$. Working as a high skill worker requires the college diploma, which means acquiring $h > 1$ units of efficiency per unit of time supplied during the first period and supplying $(1-e)\,h$ efficiency units of high skill labor when young. Not going to college means supplying only 1 efficiency unit of low skill labor and not to be allowed to work as a high skill in the future. Moreover, human capital depreciates for all workers, but investing in health, I_t^i, can influence the future human capital stock as follows:

$$\begin{aligned} h_{2t+1}^i &= \left(1-\delta\left(I_t^i\right)\right)h_{1t}^i, \qquad \delta\left(I_t^i\right) = \left(D+I_t^i\right)^{-1}, \quad D>1, \qquad (9.8)\\ with \qquad h_{1t}^i &= h \quad if \quad i=H, \quad h_{1t}^i = 1 \quad if \quad i=L, \end{aligned}$$

where h_{jt}^i is the human capital stock in period t of an agent agent of age j with skill i.

The budget constraints when young and old of an agent born in period t are, respectively:

$$\begin{aligned} I_t^i + a_t^i &= \left(1-e^i\right)w_t^i h_{1t}^i,\\ c_{t+1}^i &= w_{t+1}^i h_{2t+1}^i + \left(1+r_{t+1}\right)a_t^i, \end{aligned}$$

where a_t^i and c_{t+1}^i are savings and consumption when old of an agent of type i.

Given the choice of e^i, the problem of an agent with skill i is equivalent to maximize:

$$\left(1-e^i\right)w_t^i h_{1t}^i + \frac{w_{t+1}^i h_{2t+1}^i}{1+r_{t+1}} - I_t^i,$$

which implies that the optimal amount of health investment is

$$I_t^i = \left(\frac{w_{t+1}^i h_{1t}^i}{1 + r_{t+1}} \right)^{1/2} - D. \tag{9.9}$$

Then we can compute the lifetime income left for consumption for each case, $e^H = e$ and $e^L = 0$, and conclude that an agent with effort/ability e will choose to go to college if and only if

$$(1-e)\, w_t^H h \geq w_t^L + \left[\frac{-\left(w_{t+1}^H h - w_{t+1}^L \right)}{1 + r_{t+1}} + 2 \frac{\left(\left(w_{t+1}^H h \right)^{1/2} - \left(w_{t+1}^L \right)^{1/2} \right)}{\left(1 + r_{t+1} \right)^{1/2}} \right] \tag{9.10}$$

The agent with effort ability e_t^* for which this condition holds with equality will determine the fraction of young workers in period t, N_{1t}, that decide to go to college. That is, the number of young $H - skill$ workers in period t will be $e_t^* N_{1t}$.

9.3.3. Equilibrium

In equilibrium the firms' optimality conditions imply that the factor proportions used will be determined by the relative factor prices, which will be a function of the economy's capital and labor ratios:

$$\frac{w_t^L}{1 + r_t} = \frac{(1-\alpha)}{\alpha} k_t \left(\mu_L + \mu_H x_t^{\rho} \right)^{-1+1/\rho} \mu_L \equiv \omega^L \left(k_t, x_t \right), \tag{9.11}$$

$$\frac{w_t^H}{1 + r_t} = \frac{(1-\alpha)}{\alpha} k_t \left(\mu_L + \mu_H x_t^{\rho} \right)^{-1+1/\rho} x_t^{\rho-1} \mu_H \equiv \omega^H \left(k_t, x_t \right). \tag{9.12}$$

The market clearing conditions imply that the employed levels of high skill labor and low skill labor must equal the corresponding labor supplies, so $x_t = H_t^s / L_t^s$. It follows from equations (9.9) and (9.10) that the supply of high skill labor and low skill labor measured in efficiency units are, respectively:

$$H_t^s = N_{1t}^H h \int_0^{e_t^*} (1-e)\, de + N_{1t-1}^H h_{2t}^{H*}, \tag{9.13}$$

$$L_t^s = N_{1t}^L + sl^* N_{2t}^L h_{2t}^{L*} = (1-e_t^*)\, N_{1t} + sl^* \left(1-e_{t-1}^*\right) N_{1t-1} h_{2t}^{L*}, \tag{9.14}$$

where N_{jt}^i is the amount of workers of type i with age j in period t, so $N_{1t}^H = e_t^* N_{1t}$ and $N_{1t}^L = (1-e_t^*)\, N_{1t}$. The term $h\int_0^{e_t^*}(1-e)\, de$ is the number of efficiency units supplied by a young high skill worker and h_{2t}^{i*} is the efficiency units supplied by an old worker of type i in period t, which follows from (9.8) and (9.9).

Dividing (9.13) by (9.14) and taking into account that population grows at the exogenous rate n, $N_{1t} = (1+n)\, N_{1t-1}$, it follows that in equilibrium the ratio of high skill to low skill labor is given by

$$x_t = \frac{(1-e_t^*/2)\, e_t^{*2}\,(1+n)\, h + e_{t-1}^* h_{2t}^{H*}}{(1-e_t^*)\,(1+n) + (1-e_{t-1}^*)\, h_{2t}^{L*}}. \tag{9.15}$$

Finally, the market clearing condition in the credit market implies that next period capital stock must be financed from total savings net of investment in health:

$$K_{t+1} = \left[w_t^L L_t + w_t^H H_t\right] - \left[I_t^L\,(1-e_t^*) + I_t^H e_t^*\right] N_{1t}. \tag{9.16}$$

Dividing both sides of this equation by L_t, and expressing everything in terms of k and x:

$$\begin{aligned}\left(\mu_L + \mu_H x_{t+1}^{\rho}\right)^{1/\rho} k_{t+1} \frac{L_{t+1}}{L_t} \;&=\; (1-\alpha)\, A k_t^{\alpha} \left(\mu_L + \mu_H x_{t+1}^{\rho}\right)^{1/\rho} + \\ &\quad - \frac{N_{1t}}{L_t}\left[I_t^{L^*}\,(1-e_t^*) + I_t^{H*} e_t^*\right].\end{aligned} \tag{9.17}$$

where $I_t^{i*} = I^i(k_{t+1}, x_{t+1})$ follows from (9.11) and (9.12), and the population ratios

$$\frac{L_{t+1}}{L_t} = \frac{\left(1-e_{t+1}^*\right)(1+n)+\left(1-e_t^*\right)\left(1-\delta\left(I_t^{L*}\right)\right)}{\left(1-e_t^*\right)(1+n)+\left(1-e_{t-1}^*\right)\left(1-\delta\left(I_{t-1}^{L*}\right)\right)}(1+n), \tag{9.18}$$

$$\frac{N_{1t}}{L_t} = \frac{1+n}{\left(1-e_t^*\right)(1+n)+\left(1-e_{t-1}^*\right)\left(1-\delta\left(I_{t-1}^{L*}\right)\right)}. \tag{9.19}$$

The equilibrium dynamic system is given by (9.10), (9.15) updated to $t+1$ and (9.17), where the term e_{t-1}^* is given by (9.15).

In a stationary equilibrium the fraction of young workers that enrolls in education remains constant, $e_t^* = e^*$, and so $L_{t+1}/L_t = 1+n$. It follows from (9.10), (9.15) and (9.17) that

$$x^* = \frac{e^*h}{(1-e^*)}\frac{\left(1-\delta\left(I^{H*}\right)\right)+e^*\left(1-e^*/2\right)(1+n)}{\left(1-\delta\left(I^{L*}\right)\right)+(1+n)} \tag{9.20}$$

$$(1-e^*) = \frac{\mu_L}{h\mu_H}x^{*1-\rho}+\frac{\omega^{L*}}{w^{H*}}\left[\left(\frac{h\mu_H}{\mu_L}x^{*\rho-1}\right)^{1/2}-1\right]\left[-1+\frac{2}{\left(\omega^{L*}\right)^{1/2}}\right] \tag{9.21}$$

$$k^{*1-\alpha} = \frac{(1-\alpha)A}{1+n}-\frac{\left[\left(1-e^*\right)I^{L*}+e^*I^{H*}\right]\left(\mu_L+\mu_H x^{*\rho}\right)^{-1/\rho}}{\left(1-e^*\right)\left(1+n+1-\delta\left(I^{L*}\right)\right)} \tag{9.22}$$

9.3.4. The Inflow of Workers

Suppose that the there is an exogenous and unexpected inflow of young and low skill workers in period t. Suppose also that they arrive in the middle of the period when education decisions by natives are already taken. So the amount of high skill workers in period t is fixed. Natives and immigrants can invest in human capital but this investment will depend on their expected future wages. Due to the inflow of low skill workers the ratio x_t will fall and so initially the wage gap will increase. Since the aggregate stock of capital at the time of the shock is already fixed, an inflow of workers will also imply a fall in k_t. As a result total labor income per efficiency unit of labor will decrease but the effect on aggregate human capital investment and savings is ambiguous.

It follows from (9.7) that initially productivity will fall unless the increase in the share of low skill labor over total population compensates for the fall in k_t and x_t. In period $t+1$ the fraction of young workers that will enroll in high education will rise (young natives and immigrants descendants) and this will have a small but positive effect on x_{t+1} and on human capital investment and savings. This process will be reinforced after the immigrant population is fully assimilated (school and human capital decisions) increasing the economy's long run productivity. How long this transition is will depend on the economy's initial conditions, the relative size of the immigrants inflow and their degree of assimilation. Future increases of productivity can in turn induce more human capital investment and, in an environment with endogenous retirement, longer working lifes.

9.4. Conclusion

Migration and aging are two main interconnected challenges that Europe is facing, and their effects on the economy are not well understood. Migration is sometimes seen as a panacea to circumvent the financial problems associated to ageing societies, other times is seen as a burden since the arrival of migrant workers can depress the wage of non-educated natives and reduce the average productivity of the economy as a whole. In this paper, we have focused on the effect of migration on workers'productivity and the wage gap between skill and unskilled workers.

In order to shed some light on the effect of migration on ageing receiving economies, we have constructed a two period OLG model where we have made explicit both educational choices and human capital investments. Endogeneity of educational choices is of key importance since the arrival of migrant young workers is more likely to depress the wage of non-educated natives – as they are more close substitutes – and hence to increase incentives for schooling. Also, as evidenced by the biological literature, indi-

vidual life cycles are influenced by health investment and maintenance practices.

The conclusion is that the evolution of human capital during the life cycle is badly described by existent theories that treat depreciation or, for our purposes, aging as purely exogenous. Our results point out that, when human capital investment over the life cycle is possible, migration does not universally increase the wage gap, nor necessarily deteriorates the economy's average productivity. This depends on the economy's initial conditions, and needs to be studied case by case. Therefore, our findings go against the main drawbacks emphasized by previous literature against immigration as a valuable tool to help the sustainability of the pension system.

Further work should carry out detailed calibration and simulation exercises to give quantitative predictions of the effects found. Introducing different pension systems such as the pay-as-you-go system, and considering general equilibrium effects on capital accumulation should also produce additional interesting insights. We leave these and other related issues to future research.

9.5. References

Adam S., Bay C., Bonsang E., Germain S., and Perelman, S., 2006, Occupational activities and cognitive reserve: A frontier approach applied to the Survey of Health, Ageing and Retirement in Europe (SHARE), CREPP WP 2006/05, University of Liege.

Becker, G.S., 1967, *Human capital and the personal distribution of income: an analytical approach*, Michigan: University of Michigan. Institute of Public Administration.

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

Betts, J.R., 1998, Educational crowding out: Do immigrants affect the educational attainment of American minorities?. In: D.S. Hamermesh and F.D. Bean (Eds.), *Help or Hindrance? The Economic Implications of Immigration for African-Americans*, New York: Russell Sage Foundation.

Bonin, H., B. Raffelhüschen and Walliser, J. , 2000, Can immigration alleviate the demographic burden? *FinanzArchiv*, 57.

Borjas, G.J., 1994, The economics of immigration, *Journal of Economic Literature* 37, 1667-1717.

Borjas, G.J., 1999, The economic analysis of immigration, *Handbook of Labor Economics*, Vol.3a.

Borjas, G.J, and Hilton, L., 1996, Immigration and the welfare state: Immigrant participation in means-tested entitlement programs, *Quarterly Journal of Economics* 111, 575-604.

Bosworth H.B., and Schaie, K.W., 1997, The relationship of social environment, social networks, and health outcomes in the Seattle Longitudinal Study: Two analytic approaches, *Journal of Gerontology: Psychological Sciences* 52B, 197-205.

Card, D., 2005, Is new immigration really so bad? *Economic Journa*, 115, 300-323.

Carrasco R., Jimeno J.F., and Ortega, C., 2008, The Effects of immiration on the labor market performance of native born workers: some evidence for Spain, *Journal of Population Economics* 21, 627-648.

Casarico, A., and Devillanova, C., 2003, Social security and migration with endogenous skill upgrading, *Journal of Public Economics* 87, 773-797.

Collado, M.D., Iturbe-Ormaetxe, I. and Valera, G., 2004, Quantifiying the impact of immigration on the Spanish welfare state, it International Tax and Public Finance 11, 335-353.

Cunha, F. and Heckman, J., 2007, The technology of skill formation, *American Economic Review* 97, 31-47.

Dustmann, C., Fabbri, F. and Preston, I., 2005, The impact of immigration on the British labor market, *Economic Journal* 115, 324-341.

Eurostat, 2011, Migration and migrant population statistics, `http://epp.eurostat.ec.europa.eu/statistics_explained/index.php`

Feldstein, M.S., 2006, The effects of the ageing European population on economic growth and budgets: Implication for immigration and other policies, NBER Working Paper No. 12736.

Friedberg, R.M, and Hunt, J., 1995, The impact of immigrants on host country wages, employment and growth, *Journal of Economic Perspectives* 9, 23-44.

Gonzalez, C.I, Conde-Ruiz, J.I. and Boldrin, M., 2009, Immigration and social security in Spain, FEDEA Working Paper No. 2009-26 .

Hunt, J., 2010, The Impact of immigration on the educational attainment of natives, Typescript, McGill University.

Lee, R. and Miller, T., 2000, Immigration, social security, and broader fiscal impacts, *American Economic Review P&P* 90, 350-354.

Octaviano, G, and Peri, G., 2008, Immigration and national wages: clarifying the theory and the empirics, NBER Working Paper No. 14188.

Pischke, J.S, and Velling, J., 1997, Employment effects of immigration on Germany: An analysis based on local labor markets, *Quarterly Journal of Economics*, 594-604.

Scarmeas, N., and Stern, Y., 2003, Cognitive reserve and lifestyle, *Journal of Clinical & Experimental Neuropsychology* 25, 625-663.

Smith, J. and Edmonston, B. (eds.),. 1997, *The New Americans: Economic, Demographic, and Fiscal Effects of Immigration*, Washington DC, National Research Council, National Academy Press.

Stern, Y., 2002, What is cognitive reserve? Theory and research application of the reserve concept, *Journal of the International Neuropsychological Society* 8, 448-460.

Storesletten, K., 2000, Sustaining fiscal policy through immigration, *Journal of Political Economy* 108, 300-323.

Zorlu, A., and Hartog, J., 2005, The effects of immigration on wages in three European countries, *Journal of Population Economics* 18, 113-151.

10. An Alternative to Retirement: Part-time Work

by Alfonsa Denia Cuesta and Maria Dolores Guilló

Abstract. Part time work can facilitate participation in the labor market and smooth the transition to retirement. Part-time employment, however, represents for the most part an involuntary choice. The aim of this chapter is to conduct an empirical investigation of the determinants of part-time work. Using Spanish labor market data, we find that part-time work becomes a more desired employment alternative as people age, and that education and children's age have opposite effects on women and men's probabilities of voluntary part-time employment. Interestingly, most part-time work among women occurs in low-skill occupations, whereas part-time work among men are mainly concentrated in high-skill jobs.

10.1. Introduction

Part time work can offer a variety of benefits for the problems faced by ageing societies. First, it can be used as an instrument to facilitate participation in the labor market and smooth the transition to retirement. Second, it can be complementary tool to policies that want to extend the legal retirement age. Third, it can be key to rise government's

revenues. Within this context, a main issue that the policymaker faces is to create attractive conditions in part-time employment to make it a desired option to retirees; the existence of disadvantageous labor contract conditions often associated to part-time jobs limits the extent of part-time jobs as a voluntary choice.

Part-time work as a form of partial retirement is already a possibility in many nations. In countries like, for example, the United States about 18% of the cohort of workers born between 1931 and 1941 were in phased or partial retirement in 1998 and 2000 (Scott 2004). In Europe, part-time work programs for retirees have also been implemented; in the Netherlands, for example, about one-third of employees said their last employer offered the possibility of phased retirement (Van Soest *et al.* 2006). In Spain, about 25.5% of men employed that are 65 years old and over work part-time; among women the percentage is 38.2.

The goal of this chapter is to conduct an empirical exploration of the determinants of part-time work, and the extent to which it represents an individual's preferred labor market situation. More specifically, we explore the relative importance of individual and family variables on the probabilities of part-time and full-time employment. Given that pat-time work occurs more often among women, we differentiate between genders to introduce additional variation into the analyses. This is also important because men and women may have different reasons to choose part-time work.[1] The data used come from the Spanish labor force survey EPA (for the Spanish initials) and the focus is on individuals that live with a partner.

[1] Recent studies, for example, point out that men and women not only allocate their time between market and home duties in very different ways but also that there exist important differences across countries in time use and values that can be related to the existence of social norms or gender role attitudes (Burda, Hammermesh and Weil, 2007, and Fortin 2005). For instance, in a sample of 14 EU countries Jaumotte (2004) documents the preferences of couples with small children over three working options, non-work, part-time and full-time, and compares these preferences with their actual patterns of employment. Only ten percent of the couples prefer the traditional male breadwinner model (the man works full-time and the woman does not work), but 38 percent of them have it actually, and in all countries the rate of couples preferring the man working full-time and the woman working part-time is higher than the rate of couples with this employment arrangement, which is indicative of the potential rise in participation that can arise from more and better jobs on a part-time basis.

Our main empirical findings are the following. First, part-time work becomes a more desired employment alternative as people age, independently of gender; although old people's preference for PT work is weaker among women than among men. Second, education and family characteristics tend to have opposite effects on women and men's probabilities of voluntary part-time employment. Third, having grown-up children or a temporary contract increases significantly the probability of being involuntary part-time employed. Finally, after exploring the implications of different definitions of involuntary employment for the employment probabilities, we conclude that the EPA definition, which is the same as the one adopted by EUROSTAT, understates the extent of involuntary part-time employment.

The fact that part-time employment in developed countries is mainly concentrated on low educated workers and low skilled occupations in the service sector is well documented, for example, by Manning and Petrongolo (2004). They also find that there exists an important wage penalty associated to part-time employment which can be explained to a large extent by the high degree of occupational segregation, being Spain one of the countries with the largest occupational segregation. Fernandez-Kranz and Rodriguez-Planas (2009) show that the part-time pay penalty in Spain is larger and more persistent in the case of women in fixed-term contracts, whereas O'Dorchai, Plasman and Rycx (2007) show that the part-time wage penalty of men in Spain is negligible. This and the positive association we find between involuntary part-time employment and fixed term contracts in Spain suggest that the part-time wage penalty can be responsible of a large fraction of Spanish involuntary employment.[2]

The rest of the chapter is organized as follows. Section 2 presents some data about Spanish part-time employment with a focus on individuals that live with a partner, trying to identify the main (individual and family) factors that shape their labor supply.

[2] These issues are analyzed for the case of Britain in Connolly and Gregory (2008), Paull (2008) and Booth and van Ours (2008), the case of Australia is analyzed in Booth and van Ours (2009), and the case of a developing country (Honduras) in López-Bóo, Madrigal and Pagés (2009).

Section 3 contains the empirical investigation. Section 4 concludes.

10.2. Part-time Employment in Spain

There are mainly two types of definitions for part-time (PT) employment: objective – a PT worker usually works less hours than those of a comparable full-time (FT) worker; and subjective – the employee's spontaneous answer to 'what type of employment do you have, FT or PT?'. In general, it is not possible to establish a precise distinction between PT and FT since the standard workweek can vary from one country to another or from one activity to another. In this Section we follow the subjective definition used in the Spanish labor force survey, but we explore the extension of this definition in several directions in the appendix. These extensions try to capture, on the one side, the large heterogeneity and dispersion of part-time employment relative to full-time employment and, on the other, the voluntary or involuntary character of the labor situation. In this Section we provide some data of part-time employment in Spain trying to illustrate these special features. Although EPA is a rotating panel, in some cases we use data on two different years only for purposes of comparative statics.

Table 10.1 clearly shows that PT employment is an important alternative within the old population. More specifically, within the 65+ age group, 25.5 and 38.2% of employees have PT jobs among men and women, respectively; those same numbers for the total population are 4.1 and 22.7%.

Table 10.2, in turn, establishes that this is mainly a voluntary decision. In particular, only 0.9% of old workers among men, and 9.2% among women declare that they are PT because they do not find a FT job; whereas 30.8 and 26.3% of men and women, respectively, say that they are PT because they want to. This is the opposite to what we find for the total population: 27.6% of men are PT because they do not find a FT job, and only 10.0% because they prefer that alternative; among women, 31.2% of PT

Table 10.1.: Percentage of Part-time Workers by Age and Gender

Age	Men	Women
16-54	3.9	22.6
55-64	3.5	22.9
65+	25.5	38.2
Total	4.1	22.7

Source: EPA and own calculation.

Table 10.2.: Percentage of Part-time Workers (Not) Liking a Full-time (FT) Job

Age	Do not find a FT job		Do not want a FT job	
	Men	Women	Men	Women
16-54	31.0	32.5	7.5	13.0
55-64	19.5	22.8	15.5	22.7
65+	0.9	9.2	30.8	26.3
Total	27.6	31.2	10.0	14.2

Source: EPA and own calculation.

employers would like to have a FT job, and only 14.2% are happy with their current employment status. Interestingly, the table also suggests that the preference for PT work in the 65+ category is weaker among women than among men: the percentage of PT employed men that would not like to have a FT job is 30.6%, whereas the same figure for women is 26.3%; for the total population these numbers display again the opposite message.

To account for the heterogeneity of part-time employment some authors distinguish between 'substantial' PT and 'marginal' PT, depending on the number of working hours per week (e.g. it can be considered marginal up to 19 hours and substantial from 20 up to 34 hours). Table 10.3 illustrates this heterogeneity in the Spanish case, where we can also observe that men work more hours than women in all types of employment; one possible explanation to this fact is that women hold part-time jobs for very different reasons than men do, what somehow conditions the type of labor contract they have. The reasons of having a PT job given in the survey are reported in

Table 10.3.: Total Worked Hours per Week, Household's Reference person or Spouse.

	Men		Women	
	2000	2008	2000	2008
FT				
mean	43.14	43.9	40.47	38.89
sd	7.92	8.09	6.73	6.77
PT				
mean	19.96	19.27	17.87	19.14
sd	6.06	7.07	6.72	7.10
Substantial PT				
mean	22.73	23.24	22.33	23.22
sd	3.20	3.55	3.07	3.66
Marginal PT				
mean	11.88	11.14	11.01	11.26
sd	4.05	4.74	4.33	4.70

Source: EPA and own calculation.

Table 10.4 and correspond to all possible answers to the question *'why do you have a part-time employment?'*[3]

In each year box of Table 10.4, the second and fourth columns report the gender distribution of a given answer (row) and the third and fifth show how often each answer is given by men and women (col.), respectively. For example, 47.4 percent of workers who have a PT job because they are undertaking some education or training program are men in 2000 and this share has fallen to 42.8 in 2008; whereas 13.7 percent of men and only 4.2 percent of women in 2000 have it for that reason. This table also reveals that having a PT job due to family obligations is mainly a women's motive. The fact that most of PT workers that do not want a FT job are also women (around 82 percent in 2000 and 80 per cent in 2008) goes probably in the same direction, since these data usually correspond to women being in households where men hold FT jobs. Moreover, the majority of workers that have a PT job because they have not found a FT one are also women.

[3] Since 2004 the 'type of activity' developed is not listed as a possible reason of having a part-time job, and 'family obligations' is split into two different reasons, 'taking care of children and other dependent adults' and 'other family reasons'.

Table 10.4.: PT Employment Reasons Across Genders

2000	Men		Women	
	Row	Col.	Row	Col.
Education, training	47.4	13.7	52.6	4.2
Illness	60.0	3.0	40.0	0.6
Family obligations	1.2	0.6	98.8	13.7
FT not found	22.0	22.5	78.0	21.9
FT not wanted	16.1	4.1	83.9	5.9
Type of activity	20.9	37.0	79.1	38.2
Other reasons	26.3	18.1	73.7	13.9
Unknown reason	13.1	0.6	86.9	1.7

2008	Men		Women	
	Row	Col.	Row	Col.
Education, training	42.8	26.2	57.2	8.5
Illness	43.8	4.5	56.2	1.4
Family obligations	5.4	3.4	94.6	14.3
Children care	1.4	1.2	98.3	17.5
FT not found	17.7	27.6	82.3	31.2
FT not wanted	14.7	10.0	85.3	14.2
Other reasons	34.0	26.3	66.0	12.4
Unknown reason	26.9	0.9	73.1	0.6

Population: household's reference person or spouse.
Source: EPA and own calculation.

Table 10.5.: Unemployed Searching Options Across Genders, Percentage

2000	Men		Women	
	Row	Col.	Row	Col.
FT only	53.9	20.8	46.1	12.6
FT, PT	38.5	27.6	61.5	31.2
PT, FT	15.7	0.9	84.3	3.4
PT only	17.6	1.5	82.4	5.1
Any type	42.2	49.2	57.8	47.7

2008	Men		Women	
	Row	Col.	Row	Col.
FT only	52.2	57.0	47.8	47.0
FT, PT	38.5	6.9	61.5	7.7
PT, FT	15.6	0.6	84.4	1.7
PT only	19.9	6.1	80.1	13.3
Any type	41.4	29.3	58.6	30.4

Population: household's reference person or spouse.
Source: EPA and own calculation.

Table 10.6.: Individuals Wanting to Work More Hours by Hours Worked, Age, and Gender

	+35		20-34		1-19	
Age	Men	Women	Men	Women	Men	Women
16-54	7.4	4.6	26.9	27.4	40.4	48.0
55-64	2.8	1.7	7.9	11.9	28.6	31.3
65+	2.0	0.0	0.0	8.3	10.0	18.2
Total	6.8	4.3	24.2	26.0	38.0	45.6

Source: EPA and own calculation

Finally, if we are interested in the profile of a PT worker, we cannot ignore the unemployed. Unemployed workers have preferences about the type of workweek they want. Table 10.5 reports the frequency of different workweek types searched for by unemployed workers. We can see that the distribution of each searching option across genders has remained practically the same but the distribution of all options within each gender has experienced important changes. For example, searching for *FT only* and searching for *FT as the first option* (*FT, PT*) have become more and less important over the years, respectively. The same thing happens for the categories *PT only* and *PT as the first option* (*PT, FT*). In other words, preferences about workweek types have become more polarized. Thus, if we want to properly account for the determinants of the (voluntary) PT labor supply we have to account also for those unemployed that seek for a PT job. So, first of all, we have to specify what do we mean by a (voluntary) PT worker (unemployed or employed).

With respect to the unemployed we could just say that she is a PT seeker if she is searching for PT only or for PT as the first option, but with respect to the employed there is not such a clear cut between voluntary and involuntary. For instance, we could say that an involuntary PT employed worker is a worker with a PT job who wants to work FT or more hours, whereas a voluntary PT employed worker does not want a FT job. But then we were left with a large number of PT employed workers who are neither voluntary nor involuntary.

Table 10.6 illustrates the importance of the distinction between a free choice (voluntary) and a constrained (involuntary) employment situation. It shows that the percentage of workers who prefer to work more hours is quite high for people with 'marginal' PT and that the majority of workers who want to work more hours, both substantial and marginal PT, are in the 16-54 year group. In the empirical analyses of Section 4 we try to solve this ambiguity combining the reasons of having a PT employment with a control variable for hours.

10.3. Model and Results

The aim in this Section is to find the factors that shape the profile of a (voluntary) part-time worker. As discussed above we cannot ignore unemployed individuals searching for a part-time job, so we try to explore the determinants of different labor situations including full-time employed and full-time seekers, voluntary part-time employed and part-time seekers, and involuntary part-time employed. The first problem we face is that the voluntary or involuntary nature of the part-time labor status that comes from employed and unemployed individuals responds to very different circumstances and factors. In the theoretical framework individuals take wages and technologies as given and choose the employment supply option that maximizes their utility subject to their budget constraint. In the data, however, individuals have a labor status that can be different from their first-best option, depending for example of how long they have been searching for a part-time work or the availability of full-time jobs.

10.3.1. Measurement

There are some peculiarities in the employment categories that we find in the data that are worth mentioning. On the one side, in the estimation of employment probabilities it is a common practice to include inactive as well as unemployed individuals in the same non-work category (e.g., Bardasi and Gornick, 2003), but unemployment is usually an involuntary status. On the other side, part-time employment can be also involuntary. Hence, to explore the implications of accounting for involuntary situations, we propose (i) to include unemployed workers either in the part-time or in the full-time categories depending on their job searching options, and (ii) to distinguish between voluntary and involuntary part-time employed workers depending on the reasons of having a part-time job and the willingness to work more hours. That is, we need to address two questions: how to classify the unemployed individuals into part-

Table 10.7.: Job Searching Distributions by Unemployment Spells 2008

	FT only		FT/PT		PT/FT		PT only		Any type	
Months	Men	Wom	Men	Wom	Men	Wom	Men	Wom	Men	Wom
$m < 1$	64.7	49.9	3.9	5.4	1.0	2.7	6.9	19.8	21.6	20.2
$1 \leq m < 3$	59.5	40.1	7.6	8.4	0.2	1.9	6.9	20.8	24.4	27.8
$3 \leq m < 6$	57.8	40.0	8.2	10.2	0.2	3.3	6.1	17.5	26.3	28.1
$6 \leq m < 12$	54.6	36.1	7.3	8.8	1.5	2.5	5.9	18.1	29.6	32.9
$12 \leq m < 18$	50.9	37.2	8.1	6.6	0.9	2.6	5.4	19.1	32.6	32.8
$18 \leq m < 24$	47.1	36.0	5.8	9.5	0.5	3.1	8.3	17.3	35.9	32.4
$24 \leq m < 48$	46.3	31.9	8.7	9.4	0.3	2.3	2.8	17.5	39.4	37.5
$48 \leq m$	48.8	36.2	5.6	7.8	0.8	2.8	3.6	14.9	40.5	36.6

time and full-time workers and how to classify the part-time employed individuals into the voluntary and involuntary employment situations.

To address the classification of unemployed workers between part-time and full-time workers the only available information we have comes from the job searching categories reported in Table 10.5. This classification could be done trying to order the searching options categories according with the workers' preferences, but then the question is whether these preferences follow really an order or not, see for instance Baslevent (2002). From a worker's perspective, the five searching options can be ordered in different alternative ways but the order itself can be influenced by the length of the unemployment spell, as Table 10.7 illustrates. This Table shows the distribution of the five searching options for each of the unemployment spells in EPA 2008. For longer unemployment spells the options full-time only or part-time only become less frequent and the option searching for any type more frequent, whereas there is not a clear pattern for the rest of searching alternatives.

To explore further the influence of unemployment duration on the part-time searching option we have estimated different multinomial specifications and obtained similar results in all of them. In Table 10.8 we report the marginal effects of a multinomial logit with three possible states, searching full-time only, searching part-time only and

Table 10.8.: Searching Options, Multinomial logit Marginal Effects

	Women		Men	
	PT only	Either	PT only	Either
Age	0.055	−0.010	0.011	−0.015
	(0.007)	(0.008)	(0.003)	(0.009)
Age2	−0.0001	0.0001	−0.0001	0.00009
	(0.00008)	(0.0001)	(0.00003)	(0.0001)
Edu2	0.062*	−0.070*	0.032**	−0.078*
	(0.023)	(0.025)	(0.017)	(0.027)
Edu3	0.006	−0.076*	0.064	−0.120*
	(0.024)	(0.027)	(0.026)	(0.028)
Edu4	0.00003	−0.161*	0.092*	−0.137*
	(0.029)	(0.030)	(0.044)	(0.034)
Married	0.0425*	0.0450**	−0.008	−0.076*
	(0.0216)	(0.027)	(0.009)	(0.033)
Em.par.	0.0435**	0.0142	−0.013**	0.020
	(0.026)	(0.032)	(0.008)	(0.026)
Un.par.	−0.024	0.0853*	−0.023*	0.011
	(0.035)	(0.042)	(0.007)	(0.034)
Child1	0.131*	−0.077*	0.006	−0.017
	(0.024)	(0.025)	(0.008)	(0.03)
Child2	0.091*	−0.048*	0.010	−0.049**
	(0.021)	(0.023)	(0.008)	(0.027)
Child3	0.015	−0.012	−0.002	−0.019
	(0.089)	(0.023)	(0.007)	(0.027)
Child4	−0.051*	0.017	0.0006	0.028
	(0.020)	(0.026)	(0.009)	(0.031)
Child5	−0.039	0.066*	−0.005	−0.058
	(0.025)	(0.032)	(0.010)	(0.037)
Em.other	0.024	−0.045	0.010	0.034
	(0.028)	(0.031)	(0.012)	(0.040)
Adult65	0.006	−0.013	−0.016	0.044
	(0.045)	(0.054)	(0.001)	(0.064)
Search	−0.006	0.020*	0.002**	0.027*
	(0.003)	(0.0042)	(0.001)	(0.0053)
Obs.Total		3131		1939
Log likel.		−3183.768		−1400
Ps.R^2		0.051		0.075

(*) and (**) stand for significance at the 5 and 10 percent levels, respectively; s. e. in parenthesis.

The dependent var is 0 if looking for a FT only, 1 if PT only, 2 if Either (FT/PT or PT/FTor Any type)

searching either, which includes the other 3 searching options. The sample consists of unemployed individuals aged between 16 and 64 that live with a partner and are either the household's head or the partner of the household's head. The individual characteristics are age (and square age, as a control variable) and education (four different levels, primary, secondary first level, secondary second level and university), which are not only proxies for experience and productivity, they also can influence the age and number of children. The family variables are marital status (non-married includes single, divorced and widowed), the partner's employment, the partner's unemployment, age of children (six school-age intervals), the presence of other employed adults living in the household, and the presence of other adults older than 64.[4] The variable 'search' is an ordinal variable that accounts for the different lengths of the unemployment spells. Note that the marginal effect of search duration is irrelevant for the 'part-time only' option and significant for the 'either' option in the case of women, whereas both effects are significant and positive in the case of men. Other noticeable gender asymmetries are that being married or having small children have positive and significant effects for women, and negative and not always significant for men. So these two variables arise as possible determinants (positive effects) of the (voluntary) part-time labor supply of women and full-time labor supply of men.

Moreover, since there is not a clear mapping from the unemployed searching categories to the categories of full-time and part-time workers, in the estimations that follow we have decided to classify as part-time workers all the unemployed individuals that search for 'part-time only' or for part-time as the first option, and to consider the rest of the unemployed as full-time workers (except in Definition 1, where seekers that would be willing to work part-time are excluded), see bottom panel of Table 10.9 .

Next, we turn to issues related to the classification of part-time employment between

[4] All the model specifications also include dummy variables for the Spanish regions which are not reported.We have also estimated all the models including instead the regional unemployment rates and found very similar results.

Table 10.9.: Alternative Definitions for Employment Categories

Voluntary (PT) and Involuntary (IPT) Part-Time employment	
Definition 1	VPT: FT not wanted IPT: FT not found
Definition 2	VPT: FT not wanted IPT: other than FT not wanted
Definition 3	VPT: FT not wanted+ other reason if not want more hours IPT: FT not found + other reason if want more hours
Def. EPA	VPT: other than FT not found IPT: FT not found
Part-Time (PTU) and Full-Time (FTU) Unemployment	
For all Def.	PTU: searching PT only + PT as first option
All but Def. 1	FTU: searching FT only + FT as first option + any type
Only Def. 1	FTU: searching FT only

voluntary and involuntary. This can be done combining at least two criteria: (i) the hours criterium, which accounts whether the worker prefers working more hours or not and (ii) the reasons criterium, which accounts for the reasons of having a part-time work (Table 10.4). A selection of alternative ways for the classification of these employment categories is given in the top panel of Table 10.9. The official statistics classification (Def. EPA) identifies the involuntary part-time employment with the share of part-time workers that have not found a full-time job, so in this case the voluntary part-time employment share will include the rest of part-time employment categories described in Table 10.4. One way to solve this ambiguity using the reasons criterium is to define the voluntary part-time employment as the number of workers that do not want a full-time job and compute the involuntary part-time employment as the residual (Definition 2) or as those that have not found a full-time job (Definition 1). Notice that Definition 1 is the less ambiguous of all, but the problem with this definition is that we loose a lot of observations and part-time employment is very heterogenous. In the next section we explore the implications of these alternative employment definitions, and those of unemployment described above, for the labor status probabilities of all individuals.

10.3.2. Results

We analyze the influence of individual and family characteristics on the labor status of women and men, exploring the implications of the different employment definitions discussed above. The sample consists of all individuals aged between 16 and 64 that live with a partner and are either the household's head or the partner of the household's head. The set of explanatory variables is the same as that in Table 10.8 except for the 'search' variable. Tables 10.10 and 10.11 report the marginal effects of the ordered and multinomial models for women and men, respectively, under Definition 2. In the statistical appendix we show that the significance of grown up children and adults over

Table 10.10.: Women's Labor Status under Definition 2.,Mmarginal Effects

	ordered model			multinomial model	
	NW	PT	FT	PT	FT
Age	-0.031^{*} (0.003)	-0.001^{*} (0.0001)	0.033^{*} (0.003)	0.002^{*} (0.001)	0.030^{*} (0.012)
Age2	0.0005^{*} (0.000031)	-0.00002^{*} (0.00001)	-0.0006^{*} (0.00003)	-0.00003^{*} (0.00001)	-0.0005^{*} (0.00003)
Edu2	-0.0715^{*}	-0.0034^{*} (0.0004)	0.075^{*} (0.008)	0.0055^{*} (0.0034)	0.068^{*} (0.008)
Edu3	-0.197^{*} (0.0072)	-0.012^{*} (0.0006)	0.209^{*} (0.008)	-0.007^{*} (0.0032)	0.204^{*} (0.008)
Edu4	-0.352^{*} (0.0057)	-0.030^{*} (0.0011)	0.382^{*} (0.006)	-0.018^{*} (0.0029)	0.375^{*} (0.006)
Married	0.116^{*} (0.010)	0.008^{*} (0.0009)	-0.123^{*} (0.011)	0.003 (0.004)	-0.122^{*} (0.011)
PartnerE	-0.099^{*} (0.009)	-0.003^{*} (0.0002)	0.102^{*} (0.010)	0.006 (0.004)	0.095^{*} (0.010)
PartnerU	-0.169^{*} (0.0013)	-0.0013^{*} (0.0014)	0.182^{*} (0.014)	0.011 (0.171)	0.166^{*} (0.015)
Child1	0.175^{*} (0.010)	0.003^{*} (0.0002)	-0.178^{*} (0.010)	0.001 (0.004)	-0.181^{*} (0.010)
Child2	0.083^{*} (0.008)	0.002^{*} (0.0002)	-0.085^{*} (0.009)	0.006^{*} (0.032)	-0.085^{*} (0.009)
Child3	0.023^{*} (0.008)	0.0009^{*} (0.0003)	-0.024^{*} (0.008)	-0.0006 (0.003)	-0.024^{*} (0.008)
Child4	0.011 (0.008)	0.0005 (0.0031)	-0.012^{*} (0.009)	-0.009 (0.003)	-0.013^{*} (0.009)
Child5	0.004 (0.010)	0.0002 (0.0004)	-0.004 (0.010)	-0.009 (0.004)	-0.007 (0.010)
OtherE	0.0003 (0.009)	0.00001 (0.0004)	0.0003 (0.010)	0.005 (0.004)	0.002 (0.010)
Adult65	0.028^{*} (0.013)	0.0010^{*} (0.0004)	-0.029^{*} (0.014)	0.009^{**} (0.005)	-0.026^{**} (0.014)
LIMIT_1	1.453^{*} (0.255)				
LIMIT_2	1.661^{*} (0.255)				
Obs.Total	33427			33427	
Log likel.	-23170			-23086	
Ps.R^2	0.153			0.156	

Table 10.11.: Men's Labor Status under Definition 2, Marginal Effects

	ordered model			multinomial model	
	NW	PT	FT	PT	FT
Age	-0.016^{*} (0.0013)	-0.0004^{*} (0.00006)	0.016^{*} (0.0014)	-0.00008 (0.0002)	0.016^{*} (0.0013)
Age2	0.0003^{*} (0.00001)	-0.000007^{*} (0.000001)	-0.00026^{*} (0.00002)	0.000002 (0.000002)	-0.00025^{*} (0.00001)
Edu2	-0.015^{*} (0.0030)	-0.0004^{*} (0.00009)	0.015^{*} (0.0030)	0.00013 (0.00082)	0.014^{*} (0.003)
Edu3	-0.013^{*} (0.0031)	-0.0003^{*} (0.00009)	0.014^{*} (0.003)	0.0013 (0.0010)	0.012^{*} (0.003)
Edu4	-0.035^{*} (0.0029)	-0.0009^{*} (0.00013)	0.036^{*} (0.0029)	0.004^{*} (0.0016)	0.032^{*} (0.003)
Married	-0.019^{*} (0.007)	-0.0005^{*} (0.0002)	0.0200^{*} (0.007)	0.000004 (0.0007)	0.020^{*} (0.007)
PartnerE	-0.031^{*} (0.0028)	-0.0008^{*} (0.0001)	0.032^{*} (0.003)	-0.0002 (0.0005)	0.031^{*} (0.003)
PartnerU	-0.013^{*} (0.0045)	-0.0004^{*} (0.00013)	0.014^{*} (0.0049)	-0.0016^{*} (0.0006)	0.014^{*} (0.005)
Child1	-0.024^{*} (0.005)	-0.0006^{*} (0.0002)	0.0242^{*} (0.0052)	-0.0004 (0.0006)	0.025^{*} (0.005)
Child2	-0.008^{**} (0.005)	-0.0002^{**} (0.0001)	0.0079^{**} (0.0048)	-0.0006 (0.0006)	0.008^{**} (0.005)
Child3	-0.013^{*} (0.004)	-0.00034^{*} (0.0001)	0.014^{*} (0.0037)	-0.0005 (0.0006)	0.013^{*} (0.004)
Child4	-0.012^{*} (0.0033)	-0.0003^{*} (0.00009)	0.012^{*} (0.0034)	-0.0009^{**} (0.0005)	0.012^{*} (0.003)
Child5	-0.008^{*} (0.004)	-0.0002^{*} (0.0001)	0.0079^{*} (0.0037)	-0.0012^{*} (0.0006)	0.008^{*} (0.004)
OtherE	-0.0099^{*} (0.0035)	-0.0003^{*} (0.0001)	0.0101^{*} (0.0036)	0.00006 (0.0008)	0.0097 (0.0035)
Adult65	-0.0034 (0.0044)	-0.00009 (0.0001)	0.0035 (0.005)	0.0001 (0.001)	0.032 (0.005)
LIMIT_1	1.263^{*} (0.534)				
LIMIT_2	1.293^{*} (0.533)				
Obs.Total	31245			31245	
Log likel.	-9044			-8996	
Ps.R^2	0.258			0.262	

64 depends on the definition used, but in general all definitions yield similar results. For women, the probability of choosing full-time work decreases with the presence of children of any age, although the marginal effects become less negative for older children, whereas the marginal effects are positive (but small) if the alternative is part-time work. As we can observe, this relationship between children and labor status is just the opposite in the case of men. A similar conclusion arises for the marital status, being married decreases the probability of full-time by 12 per cent and increases slightly the probability of part-time in the case of women, whereas the signs of the marginal effects are again the opposite in the case of men. Moreover, in the case of women, the presence of adults older than 64 in the household has a positive and significant effect on the probability of part-time and a negative and significant effect on the probability of full-time, but none of the coefficients are significant in the case of men. Finally, it is worth noting that, regardless of gender, the higher is the education level the higher is the probability of having a full-time job, whereas in the case of part-time employment things are very different across genders: more education decreases the probability of part-time employment for a woman but it only matters the university degree and with a positive effect for the employment probability of a man.

10.3.3. Robustness

To check the robustness of the results, we now restrict the sample to salaried individuals and add a set of employment related variables as controls. These variables include four categories for the type of activity and occupation that are defined from the EPA socioeconomic classification, primary sector, blue-collar, white-collar/professional and service sector (base category); the 'contract' variable takes the value 1 if the employment contract is permanent and 0 if it is temporary; the 'private' variable takes the value 1 if the individual is employed in the private sector and 0 if employed in the

Table 10.12.: Salaried Workers, Marginal Effects

	Women			Men		
	(1)	(2)	(3)	(1)	(2)	(3)
Age	−0.028	−0.062*	−0.005	0.003*	0.0026*	0.002*
	(0.033)	(0.003)	(0.003)	(0.0006)	(0.0005)	(0.0004)
Age2	0.00003	0.00003	0.000002	−0.00004*	−0.00003*	−0.00003*
	(0.00004)	(0.00004)	(0.0004)	(0.0001)	(0.0001)	(0.00000)
Edu2	0.061*	0.046*	0.042*	−0.00008	−0.0005	0.00004
	(0.010)	(0.010)	(0.010)	(0.003)	(0.002)	(0.002)
Edu3	0.149*	0.113*	0.103*	−0.003	−0.0014	−0.0005
	(0.009)	(0.010)	(0.010)	(0.003)	(0.002)	(0.002)
Edu4	0.246*	0.155*	0.140*	−0.012*	−0.005	−0.004**
	(0.009)	(0.011)	(0.012)	(0.004)	(0.003)	(0.002)
Married	−0.025*	−0.038*	−0.062*	0.010*	0.005*	0.003**
	(0.011)	(0.010)	(0.010)	(0.0032)	(0.0024)	(0.018)
Epartner	0.021	0.0074	−0.002	−0.012	−0.0059	−0.001
	(0.014)	(0.013)	(0.013)	(0.002)	(0.0015)	(0.001)
Upartner	0.023	0.030	0.054*	−0.008*	−0.0058**	−0.001
	(0.021)	(0.019)	(0.018)	(0.004)	(0.003)	(0.002)
Child1	−0.150*	−0.153*	−0.173*	−0.003	−0.0027	−0.002
	(0.012)	(0.012)	(0.013)	(0.003)	(0.002)	(0.002)
Child2	−0.077*	−0.083*	−0.088*	−0.0006	−0.0006	−0.0001
	(0.010)	(0.010)	(0.011)	(0.0022)	(0.002)	(0.001)
Child3	−0.022*	−0.025*	−0.020*	0.00009	0.00006	0.0008
	(0.093)	(0.009)	(0.009)	(0.002)	(0.002)	(0.001)
Child4	−0.010	−0.009	0.005	0.0021	0.0017	0.001
	(0.010)	(0.010)	(0.010)	(0.002)	(0.002)	(0.001)
Child5	0.013	0.012	0.012	0.006*	0.004*	0.003**
	(0.012)	(0.012)	(0.012)	(0.002)	(0.002)	(0.0016)
Eother	−0.021**	−0.010	−0.005	−0.006	−0.0045	−0.003
	(0.013)	(0.012)	(0.013)	(0.004)	(0.003)	(0.002)
Adult65	0.015	0.009	0.016	0.004	0.003	0.003
	(0.019)	(0.018)	(0.019)	(0.003)	(0.003)	(0.002)
Primary		0.190*	0.181*		0.005*	0.004*
		(0.0051)	(0.005)		(0.002)	(0.0015)
Blue col.		0.163*	0.152*		0.013*	0.010*
		(0.006)	(0.006)		(0.002)	(0.0015)
Profess.		0.032*	0.025*		−0.002	−0.003**
		(0.010)	(0.011)		(0.002)	(0.0016)
Contract		0.164*	0.111*		0.022*	0.011*
		(0.010)	(0.010)		(0.003)	(0.002)
Private		−0.210*	−0.197*		−0.005*	−0.004*
		(0.007)	(0.007)		(0.001)	(0.001)
Hours			−0.528*			−0.064*
			(0.016)			(0.008)
Obs.	14576	14576	14576	19470	19470	19470
R^2	0.068	0.154	0.239	0.036	0.075	0.151

(*) and (**) stand for significance at the 5 and 10 percent levels, respectively; s.e. in parenthesis.

The dependent var is 1 if FT, 0 if PT.

Table 10.13.: Gender Asymmetries of Salaried Workers, Binary Estimations

	Family	Personal	Market
Men	Married (+) Children (not)	High education (-) Rest education levels (not)	Professional (-)
Women	Married (-) Children (-)	↑ Education ⇒↑ Coef All education levels (+)	Professional (+)

public sector; finally, the variable 'hours' takes the value 1 if the individual wants to work more hours and 0 otherwise.[5]

The purpose is to analyze the relative impact of individual, family and job related variables on the probabilities of three possible employment situations: voluntary part-time, involuntary part-time and full-time. Due to sample limitations this kind of multivariate analysis cannot be done for the male population, so we propose to make gender comparisons through binary estimations and restrict that multinomial estimation to the female population.[6]

The binary choice model we propose takes the self-reported labor situations in the labor force survey (full-time or part-time) but tries to control for involuntary employment using the 'hours' criterium. To do that, the set of explanatory variables related to the job characteristics includes the variable 'hours' capturing whether the worker wants to work more hours or not. The estimation results are reported in Table 10.12. Clearly, the variable 'hours' improves considerably the explanatory power of the model, which can be indicative of the relevance of involuntary part-time employment. The negative sign and significance of this coefficient reflect that the involuntary employment situation is positively related to part-time employment. The decrease of

[5] We have also included four firm size categories in an extended model and find that these variables are very significant, being more likely to have PT with small firms than with large firms. These results are available only for the year 2000 and can be obtained from the authors upon request.

[6] The few observations for men with a PT job do not allow to split the sample into voluntary and involuntary PT workers.

Table 10.14.: Salaried Women's Labor Status, Marginal Effects

	Vol PT	Inv PT	FT
Age	0.00313	0.036	−0.0068*
	(0.0024)	(0.018)	(0.003)
Age2	0.00000	−0.0004**	0.00004
	(0.00003)	(0.0002)	(0.00004)
Edu2	−0.02208*	−0.021*	0.043*
	(0.0075)	(0.005)	(0.009)
Edu3	−0.0541*	−0.053*	0.107*
	(0.0074)	(0.005)	(0.009)
Edu4	−0.0794*	−0.069*	0.148*
	(0.0086)	(0.006)	(0.010)
Married	0.0447*	−0.005	−0.040*
	(0.007)	(0.006)	(0.009)
Epartner	−0.0069	−0.0004	0.007
	(0.0099)	(0.008)	(0.013)
Upartner	−0.0384	0.008	0.031**
	(0.012)	(0.013)	(0.018)
Child1	0.1439*	0.012*	−0.156*
	(0.010)	(0.006)	(0.012)
Child2	0.0705*	0.010**	−0.080*
	(0.008)	(0.006)	(0.010)
Child3	0.0147*	0.009**	−0.024*
	(0.006)	(0.005)	(0.009)
Child4	−0.0045	0.012*	−0.007
	(0.010)	(0.006)	(0.009)
Child5	−0.0101	−0.0007	0.011
	(0.008)	(0.007)	(0.011)
Eother	−0.0178**	0.028	−0.010
	(0.012)	(0.007)	(0.012)
Adult65	0.007	0.008	0.010
	(0.009)	(0.011)	(0.017)
Primary	−0.1072*	−0.070*	0.178*
	(0.004)	(0.003)	(0.0027)
Blue col.	−0.0877*	−0.066*	0.153*
	(0.004)	(0.003)	(0.006)
Profess.	−0.0195*	−0.011**	0.031*
	(0.007)	(0.006	(0.010)
Private	0.1288*	0.072*	−0.200*
	(0.005)	(0.004)	(0.007)
Contract	−0.0379*	−0.118*	0.157*
	(0.007)	(0.007)	(0.009)
Obs.Total			14579
Log likel.			−9052
Ps.R^2			0.142

(*) and (**) significance at the 5 and 10 percent, respectively; s. e. in parenthesis.

The dependent var is 0 if Voluntary PT, 1 if Involuntary PT, 2 if FT.

the contract type coefficient (all coefficients of the market variables decrease) suggests that the desire of working more hours is mainly associated to temporary contracts regardless of gender.

It is clear from Table 10.12 that the inclusion of the market variables improves considerably the explanatory power of the models and, at the same time, implies similar values and significance levels for most individual and family variable coefficients. One remarkable feature is the change of the 'married' coefficient relative to the non-market case, it increases (decreases) by more than fifty percent in the case of women (men). That is, the marital status tends to be more (less) relevant for women (men) when we consider job characteristics. Moreover, some of the gender asymmetries that showed up in the multinomial estimations in the non-market framework of the previous section (all individuals) also arise here in the binomial estimations (only employed individuals). Being married and having children decrease the probability of full-time employment for a woman, whereas for men being married increases the probability of full-time and having children is in general not significant; furthermore, education is in general not significative in the case of men, only matters at the highest level and has a negative influence on full-time employment, whereas in the case of women all education categories are significant and more education increases the probability of full-time employment. A related feature is that working as a professional has opposite effects across genders, it increases (decreases) the probability of having a part-time employment for a man (woman). A summary of these important gender asymmetries is reported in Table 10.13, although they should be interpreted with caution due to the small number of male observations.

Finally, we estimate a multinomial model for women where we distinguish between voluntary and involuntary part-time employment. The aim of this final stage is to confirm the importance of involuntary employment and to identify the profile of a voluntary (involuntary) part-time worker. In this case, the dependent variable takes

the value 0 if the status is voluntary part-time, the value 1 if the status is involuntary part-time and the value 2 if it is full-time. Table 10.14 reports the marginal effects on the probabilities of the different labor situations when we use Definition 3, which combines the 'reasons' of having a part-time job with the willingness to work more hours (i.e., the 'hours' criterium). Using Definition 2 instead, as we did in the non-market framework, yields similar results. Note that now with this choice of definition the explanatory variable 'hours' is not necessary, since whether the worker wants to work more hours or not is already accounted for in the definitions used (see Table 10.9).

As above, the inclusion of the market variables adds substantial explanatory power to the model and both the signs and significance levels of most variables here are similar to those in the non-market setup. For concreteness Table 10.14 only reports the extended model with the market variables. The results complement those of Table 10.12 in that they confirm the importance of accounting for involuntary employment in the study of part-time situations and throw some light on the profile of a voluntary PT female worker. It is clear now that more education decreases the probability of involuntary and voluntary part-time jobs and that being 'married' and having small children are very important determinants of voluntary part-time employment. Indeed whereas the sign of the marginal effect of being 'married' is negative on both the involuntary part-time and the full-time probabilities, it is positive on the probability of voluntary part-time. Moreover, having an unemployed husband decreases the probability of a voluntary part-time job. With respect to children, it is clear that their age is crucial for the workweek type of a woman living with a partner, children aged 11 or less (child1 to child3) increase the probability of having a voluntary part-time job and decrease the probability of having a full-time job by 14 per cent and 15 per cent, respectively. These probability effects and the difference between the voluntary and involuntary part-time effects become weaker with the age of children. Furthermore,

having children aged above 10 is not significative for voluntary employment but it increases the probability of involuntary part-time employment by 1.2 percent if children are aged below 16.

Finally, the influence of job characteristics on the employment probabilities confirms the effects found in the binary estimation of Table 10.12.

10.4. Conclusion

Part-time employment is a recent phenomenon. It surged in the mid nineties as a possible mechanism to conciliate family responsibilities and market work, for both men and women. Now it is also seen as a mechanism to transit to retirement.

Our empirical study has focused on the features of part-time employment in Spain, and also analyzed the extent of part-time employment as a voluntary choice for employed and non-employed individuals. It has been developed in two stages. In the first stage we have omitted market variables and consider both non-employed and employed individuals. The unemployed population was split between part-time and full-time workers according with their workweek searching options, and the population employed part-time was classified into part-time or full-time workers according with their voluntary (do not want a full-time job) or involuntary employment situation. In the second stage, we have included some market variables as controls and restricted the analysis to employed individuals. The inclusion of the market variables improved considerably the explanatory power of the model and, at the same time, kept similar values and significance levels for most individual and family coefficients.

Results pointed out to a positive association between temporary contracts and the presence of involuntary part-time employment, independently of gender. In contrast, being married and having children decreased the probability of full-time employment for a woman, but not for a man. For men, education only mattered at the highest level

and, in contrast to the case of women, it favored part-time employment. Furthermore, we estimated the profile of a part-time employed woman using the extended definitions of voluntary and involuntary part-time employment; the results clearly showed the importance of education in reducing the woman's probability of involuntary employment, as well as the highly significant positive effect of being married and having small children on the probability of voluntary part-time employment.

We have also found that part-time work becomes a more desired employment alternative as people age. Interestingly, old people's preference for PT work is weaker among women than among men. These findings put part-time employment as a main policy tool to smooth the transition to retirement, especially among men, and mitigate the financial problems associated with the ageing society.

Appendix: Descriptive Statistics

Variable	Women		Men	
	Mean	S.D	Mean	S.D
Married	0.908	0.289	0.903	0.295
Age	43.21	10.61	44.54	10.16
Age2	1979.91	913.09	2086.77	896.71
Edu1	0.265	0.441	0.237	0.426
Edu2	0.292	0.456	0.307	0.461
Edu3	0.260	0.438	0.289	0.454
Edu4	0.184	0.387	0.166	0.371
Child1	0.173	0.378	0.185	0.388
Child2	0.184	0.387	0.196	0.397
Child3	0.217	0.412	0.232	0.422
Child4	0.165	0.374	0.175	0.380
Child5	0.315	0.464	0.308	0.462
Epartner	0.784	0.411	0.551	0.497
Upartner	0.036	0.185	0.065	0.247
Ipartner	0.180	0.384	0.384	0.486
Eother	0.251	0.433	0.245	0.430
Adult65	0.051	0.219	0.050	0.218
Primary sector	0.040	0.196	0.0527	0.224
Service sector	0.528	0.499	0.181	0.385
Private sector	0.307	0.461	0.501	0.500
Blue collar	0.074	0.263	0.369	0.483
Professional	0.348	0.476	0.386	0.486
Contract type	0.323	0.467	0.514	0.499
Hours	0.050	0.217	0.051	0.22

Source: EPA 2008

Appendix: Alternative Classifications of Employment

The next table describes the 2008 female results of the ordered model for the alternative definitions listed in Table 10.9. The dependent variable takes on 0 if the status is non-work, on 1 if the status is voluntary part-time employment or unemployment searching for a part-time job, and on 2 if the status is full-time employment, involuntary part-time employment or unemployment searching for a full-time job.

Women ordered model, alternative labor status definitions

	Def 1	Def 2	Def 3	EPA Def
Age	0.081*	0.081*	0.076*	0.081*
	(0.007)	(0.006)	(0.006)	(0.006)
Age2	−0.001*	−0.001*	−0.001*	−0.001*
	(0.00008)	(0.00008)	(0.00007)	(0.00007)
Edu2	0.235*	0.193*	0.203*	0.222*
	(0.022)	(0.020)	(0.020)	(0.019)
Edu3	0.632*	0.555*	0.544*	0.573*
	(0.023)	(0.022)	(0.021)	(0.021)
Edu4	1.285*	1.141*	1.136*	1.169*
	(0.027)	(0.026)	(0.024)	(0.024)
Married	−0.329*	−0.306*	−0.316*	−0.279*
	(0.031)	(0.030)	(0.028)	(0.027)
Epartner	0.249*	0.252*	0.246*	0.253*
	(0.024)	(0.023)	(0.022)	(0.021)
Upartner	0.428*	0.497*	0.523*	0.418*
	(0.047)	(0.044)	(0.042)	(0.042)
Child1	−0.516*	−0.427*	−0.542*	−0.491*
	(0.025)	(0.024)	(0.023)	(0.022)
Child2	−0.250*	−0.199*	−0.258*	−0.242*
	(0.022)	(0.021)	(0.020)	(0.019)
Child3	−0.079*	−0.058*	−0.081*	−0.077*
	(0.021)	(0.020)	(0.019)	(0.018)
Child4	−0.036**	−0.029	−0.024	−0.027
	(0.021)	(0.021)	(0.020)	(0.019)
Child5	−0.023	−0.010	−0.013	−0.014
	(0.023)	(0.024)	(0.023)	(0.022)
Eother	0.022	−0.001	−0.002	0.010
	(0.025)	(0.024)	(0.023)	(0.023)
Adult65	−0.045	−0.067*	−0.056**	−0.045
	(0.034)	(0.033)	(0.032)	(0.031)
LIMIT_1	1.075*	0.917*	0.723*	0.982*
	(0.158)	(0.149)	(0.143)	(0.139)
LIMIT_2	1.220*	1.041*	0.999*	1.423*
	(0.158)	(0.149)	(0.143)	(0.139)
Obs.total	30131	33448	33448	33448
Log likel.	−21272	−23202	−27239	−29254
Ps.R^2	0.165	0.152	0.133	0.129

(*) and (**) stand for significance at the 5 and 10 percent levels, respectively; s.e. in parenthesis.

10.5. References

Bardasi, E., Gornick, J., 2003, Women's part-time employment across countries: Workers' choices and wage penalties, *Women in the labor Market in changing economies: demographic Issues. OUP,* 209-243.

Baslevent, C., 2002, Preference of the unemployed between full-time and part-time jobs, *Applied Economics Letters* 9, 855-857

Booth, A.L., van Ours, J.C., 2008, Job satisfaction and family happiness: The part-time work puzzle, *The Economic Journal* 118, 77-99

Booth, A.L., van Ours, J.C., 2009, Hours of work and gender identity: Does part-time work make the family happier?, *Economica* 76, 176-196.

Burda, M., Hamermesh, D. S., Weil, P., 2007, Total work, gender and social norms, *NBER Working Paper* 1300, 2007.

Connolly, S., Gregory, M., 2008, Moving down: women's part-time work and occupational change in Britain, 1991-2001, *The Economic Journal* 118, 52-76.

Fortin, N.M., 2005, Gender role attitudes and the labor market outcomes of women across OECD countries, *Oxford Review of Economic Policy* 21, 416-438.

Fernandez-Kranz, D., Rodriguez-Planas, N., 2009, The part-time pay penalty in a segmented labor market, *IZA Working Paper* 4342.

Jaumotte, F., 2004, labor force participation of women: Empirical evidence of the role of Policy and other determinants in OECD countries, *OECD Economic Studies* 37, 2.

López-Bóo, F., Madrigal, L., Pagés, C., 2009, Part-time work, gender and job satisfaction: Evidence from developing countries, *IZA Working Paper* 3994.

Manning, A., Petrongolo, B., 2008, The part-time pay penalty for women in Britain, *Economic Journal* 118, 28-51.

O'Dorchai, S., Plasma, R., Rycx, F., 2007, The part-time wage penalty in European countries: How large is for Men?, *International Journal of Manpower* 28, 571-603.

Paull , G., 2008, Children and women's hours of work, *The Economic Journal* 118, 8-27.

Scott, J., 2004, Is phased retirement a state of mind? The role of individual preferences in retirement outcomes, *mimeo*, Cornell University.

Van Soest, A., Kapteyn, A., Zissimopoulos, J., 2006, Using stated preferences data to analyze preferences for full and partial retirement, *DNB working paper* 081, Netherlands Central Bank, Research Department.

11. A Model of Health Investment, Medicine Consumption, and Income

by Fidel Perez-Sebastian

Abstract. The human population, all over the planet, is enjoying an increasing average life span that is expected to extend another 10 years by 2050. The rising number of older adults generates an unprecedented demand on the public health system. This chapter presents a growth framework of health prevention investment and medical technology consumption in which the use of medical techniques, and diseases endogenously determine the levels of income and health. In the model, there are two forms of medical technology: one is embodied in medicines that try to cure sick individuals, whereas the other is disembodied knowledge that tries to prevent health deterioration. I find that larger saving rates are not only a consequence of a larger productivity of the consumption-goods technology; more productive pharmaceutical products and higher levels of preventive-health knowledge also increase the incentive to saving, and allow the economy to achieve larger levels of income, mitigating the adverse effects of ageing.

11.1. Introduction

The human population, all over the planet, is enjoying an increasing average life span that is expected to extend another 10 years by 2050. The rising number of older adults generates an unprecedented demand on the public health system and on medical and social services. On the one hand, old individuals are more susceptible to infectious diseases. On the other, chronic illness affect older adults disproportionately. All this contributes to disability, diminishes quality of life, and increases health and long-term care costs within the old-agent population. Health programs must respond to this challenge.

This is important not only from a humanitarian perspective, but also from an economic perspective. Health, along with education, is one of the main components of an economy's human capital. It is then clear that understanding the forces that lead to improvements in health levels is a key economic policy question in any economy.

To understand the evolution of health and mortality, we need first to measure it. We can do that through indicators such as life expectancy and mortality rates. Simple casual observation teaches us that the factors that have contributed to reduce mortality and raise life expectancy in the last two centuries include the increase in income levels and the creation and diffusion of medical technology in advanced nations. In addition, several studies find statistical evidence that all countries benefit from the diffusion of medical technology, like pharmaceutical products, and medical equipment. For example, evidence that technical progress in medicine has permitted huge improvements in health even among the less developed nations is provided by Kremer (2002) and Papageorgiou, Savvides and Zachariadis (2007).

Pharmaceutical-products invention and production worldwide is, however, heavily concentrated. As WHO (2004) highlights, trends from 1985 to 1999 indicate that the value of medicine production has grown four times more rapidly than the world's income. Medicine production is highly concentrated in the industrialized countries,

where just five countries – the USA, Japan, Germany, France and the UK – account for two-thirds of the value of all medicines produced. A small number of transnational companies dominate the global production, trade and sales of medicines. Ten of these companies now account for almost half of all sales; concentration that has increased considerably since 1987. Therefore, for most nations, medical treatments and products are more a consumption problem than a production problem.

Motivated by the above evidence, this chapter presents a growth framework of medical technology consumption in which the spread of new medical techniques and the increasing influence of diseases endogenously determine the levels of income and health. In the model, there are two forms of medical technology: one is embodied in products, whereas the other is disembodied knowledge. Disembodied knowledge supplies adequate nutritional and sanitary information. For example, education about how to have healthier children and habits. In a wider sense, it could also proxy for investment with important externalities such as construction of a national health system, and infrastructure to have access to clean water supplies and trash disposal. Embodied knowledge, on the other hand, comes in health capital equipment and pharmaceutical products.

The two forms of technology also differ in terms of their adoption process. The acquisition of disembodied knowledge needs time investment, that is, education; whereas the acquisition of embodied knowledge requires income to buy pharmaceutical products. Another difference is that, in the model, disembodied technology prevents the acquisition of diseases, whereas embodied technology fights illnesses once the person becomes affected by the disease.[1]

We find that investment in prevention becomes more appealing as its productivity increases and as getting sick becomes more likely. For example, a decrease in the model in the surviving probability causes an increase in prevention investment. A con-

[1] In reality, there are exceptions. Nutrition and vaccines require income, but are useful prevention means.

sequence of this is that investment in prevention and the consumption of medicines behave as substitutes. The model also help to characterize the relationship between health technology and income. We show that larger saving rates are not only a consequence of a larger productivity in the consumption-goods technology, more productive pharmaceutical products and higher levels of healthy behavior increase as well the incentive to saving, and allow the economy to achieve larger levels of income.

Previous literature have already studied the effect of health on income levels and growth rates. Regarding the empirical evidence, Acemoglu et al. (2003) find that diseases and geography affect development through institutions. Gallup and Sachs (2001) argue that malaria causes poverty. McCarthy et al. (2000) estimate a sizeable effect of malaria on growth. Lorentzen et al. (2008) find evidence that adult mortality explains almost all of Africa's growth tragedy in the past 40 years.

Theoretical models of health and growth include Birchenall (2007), Momota, Tabata and Futagami (2005), and Chakraborty et al. (2010). The first of these papers argue that agricultural productivity improvements due to health care improvements triggered the escape from the Malthusian trap in North America and Western Europe. Momota et al. (2005) and Chakraborty et al. (2005) present OLG models of preventive behavior and infectious-disease spread. None of them consider the adoption of medical knowledge. Grossman (1972) and, more recently, Murphy and Topel (2006) and Kuhn et al. (2011), for example, also study the effect of health care in a life-cycle model; they do not discriminate between different forms of health investment.

The rest of the chapter is organized as follows. Section 2 presents the environment. The main results are presented in section 3. Section 4 concludes.

11.2. The Model

I consider an economy composed of households and firms. Household members are two-period-life agents. The model has an overlapping-generations structure. The size of new generations is constant. Agents value consumption at the end of each period of life, and have no preference over leisure. Young individuals are endowed with one unit of time that they can supply as labor to obtain income. Old individuals only enjoy capital rents from their savings and invest in their health through medicine consumption.

The idea that we want to formalize is that health conditions depend on habits and health investment during the whole life time. Put differently, young individuals that take good care of their health have a lower probability of suffering from diseases when old. We introduce this notion in a simple way by assuming that the disease condition is obtained in the first period of life and carried out to the second period.

During the first period of life, there is a the probability of getting a disease (i_t). If the person becomes ill, the probability of surviving and enjoying consumption in the second period of life is $q_t < 1$. Agents can reduce both the probability of getting the disease and the probability of dying if affected. Young individuals invest an amount of time l_{xt} in health education at the beginning of the period to increase their knowledge on diseases and prevent contagion. In turn, sick old individuals can fight the illness consuming the available pharmaceutical products. We suppose that the discovery of new medicines and treatments occurs outside the model.

The problem that a new-born faces is the following:

$$\max_{c_{yt},s_t,l_{xt},m_{jt+1}} \left\{ u(c_{yt}) + \beta \left[(1 - i_{t+1}) u(c^u_{ot+1}) + i_{t+1} q_{t+1} u(c^i_{ot+1}) \right] \right\} \tag{11.1}$$

subject to

$$
\begin{aligned}
c_{yt} &= w_t(1 - l_{xt}) - s_t && (11.2)\\
c^u_{ot+1} &= (1 + r_{t+1})s_t, && (11.3)\\
c^i_{ot+1} &= (1 + r_{t+1})s_t - \sum_{j=1}^{h_{t+1}} p_{jt+1}m_{jt+1}; && (11.4)
\end{aligned}
$$

where w_t represents the salary at date t; c_{yt}, c^u_{ot} and c^i_{ot} are consumption levels for a young, a healthy old and a sick old, respectively; s_t is saving; r_t represents the interest rate; and h_t is the amount of different types of pharmaceutical products available in period t; finally, p_{jt} and m_{jt} are the price and quantity consumed of medicine i.

As suggested above, the infection rate among adults is endogenous, and depends on the amount of time invested in accumulating disease-prevention skills such as right exercising, healthy food intake, and non-risky behavior. More specifically, the motion equation for the prevalence rate is:

$$
i_{t+1} = i_t \left(\frac{d_t}{x_t} \right)^{\theta}; \tag{11.5}
$$

where x_t is the prevention health-knowledge accumulated; and d_t is an index of disease virulence in the economy. The latter index is related to the incidence of infectious diseases, pollution levels, and other environmental factors that affect human health conditions.

The evolution of the last two variables is given by

$$
x_{t+1} - x_t = \mu x_t^{1-\rho} h_t^{\rho} l_{xt}^{\lambda}, \quad \text{with } x_t \leq h_t, \mu > 0, \tag{11.6}
$$

and

$$
(d_{t+1} - d_t)/d_t = g_d. \tag{11.7}
$$

Expression (11.6) establishes that agents invest an amount of time l_{xt} in education to accumulate knowledge that helps to reduce the probability of acquiring an ill condition. In this health-education technology, there is an intertemporal knowledge-spillover effect. In particular, as the economy accumulates health skills, x_t, the return from education investment rises. Investment productivity also increases with the world's health knowledge, captured by variable h_t.

Equality (11.7), in turn, simply assumes that disease virulence grows exogenously. Here we want to capture ideas in Deaton (2004) who points out that much of the literature in public health sees globalization as a threat to international health. For example, the spread of HIV/AIDS accelerated by the ease and volume of modern travel. It could be argued as well that the influence of chronic diseased has increased, not only because of the ageing of society, but also because of environmental degradation.

The probability of surviving to the illness depends on the consumption of available health products. In particular,

$$q_t = \sum_{j=1}^{h_t} \frac{m_{jt}^{\gamma}}{d_t}; \tag{11.8}$$

where $\gamma \in (0, 1)$. Different types of medicines are, therefore, complementary. As will be clear later, their complementarity generates a positive external effect on the surviving probability. Health technology is assumed to come from the rest of the world. Its index h_t evolves according to

$$(h_{t+1} - h_t)/h_t = g_h. \tag{11.9}$$

Finally, the price of medicines $p_{jt} = p_t$, for all j, is determined in international markets and is taken as given.

Consumption goods are supplied to the market by perfectly competitive firms. Their production technology uses capital and labor under constant returns to scale. It takes a standard Cobb-Douglas form:

$$Y_t = A_t K_t^{\alpha} L_t^{1-\alpha};$$

where Y_t is the amount of the consumption good produced; and K_t and L_t are the amounts of capital and labor employed by firms. From the production technology, assuming that all markets are perfectly competitive, we obtain the following input prices:

$$w_t = (1-\alpha)\frac{Y_t}{L_t},$$

and

$$r_t = \alpha\frac{Y_t}{K_t}.$$

11.3. Optimal Health Investment

To solve the problem, we work backwards. In his second period of life, the problem of an sick agent is choosing the amount of each type of medicine (or treatment) that he wants to consume to reduce the probability of dying. From the above equations, we can write this problem as

$$\max_{\{m_{jt}\}_{j=1}^{h_t}} \{q_t u(c_{ot}) + (1-q_t) * 0\}$$

subject to (11.4) and (11.8).

The first order condition for this problem is

$$\left(\frac{\gamma}{d_t}\right) m_{jt}^{\gamma-1} u(c_{ot}) - q_t u'(c_{ot}) p_t = 0, \text{ for all } j = 1, ..., h_t. \tag{11.10}$$

We concentrate on symmetric equilibrium, in which consumers buy the same amount of each medicine, that is, $m_{jt} = m_t$ for all j. Hence, (11.10) implies that

$$m_t = \left(\frac{\gamma}{h_t p_t}\right) \frac{u(c_{ot})}{u'(c_{ot})}. \tag{11.11}$$

Let us give an specific form to the utility function. We choose the CES specification

$$u(c_t) = \frac{c_t^{1-\sigma}}{1-\sigma};$$

where σ is the inverse of the intertemporal elasticity of substitution between present and future consumption. Substituting it into (11.11), we get

$$m_t = \left[\frac{1+r_t}{1+(1-\sigma)/\gamma}\right]\left(\frac{s_{t-1}}{h_t p_t}\right). \tag{11.12}$$

Optimality implies that the amount bought of each medicine increases with the agent's income $(1+r_t)s_{t-1}$, and decreases with the price. A larger number of varieties, h_t, also diminishes the optimal consumption of each medicine because of the complementarity effect. This also captures a taste for variety: the old agent prefers to diversify, benefiting from different treatments and medicines to improve their health status. The elasticity of consumption substitution also affect m_t. A larger elasticity reduces the incentive to invest in pharmaceutical products.

Once we know the optimal investment in pharmaceutical products when agents are ill, we can solve for the optimal education, saving and consumption decisions in period 1. From the above equations, the problem faced by an individual in the first period of life can be written as:

$$\max_{c_{yt},s_t,l_{xt}} \left\{u(c_{yt}) + \beta\left[(1-i_{t+1})u\left[(1+r_{t+1})s_t\right] + i_{t+1}q_{t+1}u\left(\frac{(1+r_{t+1})s_t(1-\sigma)/\gamma}{1+(1-\sigma)/\gamma}\right)\right]\right\}$$

subject to (11.2), (11.5), (11.6), (11.7), (11.9), and

$$q_{t+1} = \frac{h_{t+1}}{d_{t+1}}\left[\left(\frac{1+r_t}{1+\frac{1-\sigma}{\gamma}}\right)\frac{s_{t-1}}{h_t p_t}\right]^{\gamma}.$$

Now the first order conditions with respect to l_{xt} and s_t are, respectively,

$$u'(c_{yt})w_t = \beta\theta i_{t+1}\lambda\left(\frac{x_{t+1}-x_t}{x_{t+1}l_{xt}}\right)\left[u(c^u_{ot+1}) - q_{t+1}u(c^i_{ot+1})\right], \tag{11.13}$$

and

$$\begin{aligned} u'(c_{yt}) &= \beta(1+r_{t+1})(1-i_{t+1})u'(c^u_{ot+1}) \\ &\quad +i_{t+1}q_{t+1}\left[\frac{\gamma u(c^i_{ot+1})\left(1+\frac{1-\sigma}{\gamma}\right)^{-1}}{h_{t+1}p_{t+1}m_{t+1}} + \frac{(1-\sigma)\,u'(c^i_{ot+1})}{\gamma+(1-\sigma)}\right]. \end{aligned} \tag{11.14}$$

The interpretation of these two Euler equations is straightforward. They imply that at the optimum the marginal cost (LHS in both expressions) must be equalized to the marginal benefit (RHS's).

Let us focus first on Equation (11.13). The amount invested in preventive education decreases with the return to the alternative activity, that is, with the wage. Investment in prevention, on the other hand, becomes more appealing as its productivity increases and when getting infected becomes more severe. We can see, for example, that a decrease in the surviving probability must come with an increase in l_{xt} so as to restore the equality. Because of this, investment in education and the consumption of medicines behave as substitutes.

Equation (11.14), in turn, implies that the marginal benefit of saving one additional unit of income increases with the number of available health treatments and products, because, as explained above, total expenditure in health ($h_{t+1}p_{t+1}m_{t+1}$) falls and the probability of surviving (q_{t+1}) rises. The marginal benefit of saving also goes up with investment in prevention, since prevention reduces disease prevalence (i_{t+1}). As a consequence of all this, saving and health investment move together in the same direction; furthermore, saving rises with both prevention and the productivity of medicine consumption and, as a consequence, so does income.

11.4. Conclusion

This chapter has presented a growth framework of medical technology adoption that is relevant to study the implications of increasing disease incidence and health-product consumption in ageing societies. We have considered an overlapping-generation structure in which people live for two periods. Its main features are the following. There are two forms of medical technology: one is embodied in products, whereas the other is disembodied knowledge. Disembodied knowledge supplies adequate information that helps to avoid getting ill. Embodied knowledge, on the other hand, comes in pharmaceutical products and treatments that fight the illness after the agent has fallen into the ill condition. The acquisition of disembodied knowledge needs time investment in education, whereas the acquisition of embodied knowledge requires income to buy medicines.

The model has helped to characterize the relationship between health technology and income. Income levels depend on capital accumulation. Capital, in turn, is the result of saving. In the model, larger saving rates are not only a consequence of a larger productivity of the consumption-goods technology, more productive pharmaceutical products and higher levels of preventive-health knowledge increase the incentive to saving, allowing the economy to achieve larger levels of income.

The key policy implication of the model is that additional investment in prevention and medical technology can have a sizable effect on income levels, mitigating the adverse effects of ageing. However, whether the effect of additional investment in prevention has a higher payoff than the one of treatment still remains an open question that we leave to future research

The model could also be extended to analyze another important issue pointed out, for example, by WHO (2004). It is estimated that half of all medicines are inappropriately prescribed, dispensed or sold, and that half of all patients fail to take their medicine properly. This inappropriate use is not only widespread, it is costly and ex-

tremely harmful both to the individual and the population as a whole. Incorporating these aspects into the model could provide relevant information for the effective monitoring and regulation of medicine spending and consumption.

11.5. References

Acemoglu, D., Johnson, S., Robinson, J., 2003, Disease and development in historical perspective, *Journal of the European Economic Association* 1, 397–405.

Birchenall, J., 2007, Escaping high mortality, *Journal of Economic Growth* 12, 351-387.

Chakraborty, S., Papageorgiou, C., Perez-Sebastian, F., 2010, Diseases, infection dynamics, and development, *Journal of Monetary Economics* 57, 859-872.

Deaton, A., 2004, Health in an age of globalization, *NBER Working Paper* 10669.

Gallup, J.L., Sachs, J.D., 2001, The economic burden of malaria, *The American Journal of Tropical Medicine and Hygiene* 64, 85-96

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

Kuhn, M., Wrzaczek, S., Prskawetz, A., Feichtinger, G., 2011, Externalities in a life-cycle model with endogenous survival, *Journal of Mathematical Economics* 47, 627-641.

Kremer, M., 2002, Pharmaceuticals and the developing world, *Journal of Economic Perspectives* 16, 67-90.

Lorentzen, P., McMillan, J., Wacziarg, R., 2008, Death and development", *Journal of Economic Growth* 13, 81-124.

McCarthy, D., Wolf, H., Wu, Y., 2000, The growth costs of malaria, *NBER working paper* 7541.

Momota, A., Tabata, K., Futagami, K., 2005, Infectious disease and preventive behavior in an overlapping generations model, *Journal of Economic Dynamics and Control* 29, 1673-1700.

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

Papageorgiou, C., Savvides, A., Zachariadis, M., 2007, International medical technology diffusion, *Journal of International Economics* 72, 409-427.

WHO, 2004, *The World Medicines Situation*, World Health Organization, Geneve.