The Future of Human Health, Longevity, and Health Costs

Sebastian Böhm†, Volker Grossmann‡ and Holger Strulik§

February 16, 2017

Abstract

Medical advances greatly enhance longevity but may threaten the sustainability of health care systems in its current form. In this paper, we investigate the future of human longevity, morbidity and health costs under different policy regimes. We propose a novel, multi-period overlapping generations model with endogenous medical R&D and endogenous survival that is closely associated with morbidity. We capture biologically founded ageing based on gerontology research in order to calibrate the model for the UK. Our analysis suggests substantial increases in human longevity under the current policy regime that go along with both reductions in morbidity and a rising health expenditure share in GDP. Following instead the goal of stabilizing the health expenditure share by extending health carerationing has sizable effects on morbidity and longevity in the longer run, associated with reduced medical R&D incentives. The implied welfare effects may be substantially negative particularly for future generations.

Key words: Longevity; Medical R&D; Morbidity; Health Care; Rationing.
JEL classification: H50; I10; C60.

Preliminary Version – Please do not cite
1 Introduction

In this paper we propose a new approach to study the interdependence of medical R&D, health expenditure, and the health status of an age-structured population. By developing a dynamic macroeconomic model that treats both medical progress and individual health of the citizens as endogenous we are able to make inferences about the future development of life expectancy and health expenditure, conditional on the extent of rationing in health care systems.

A salient feature of structural economic development over the last decades is the secular expansion of the health sector. In the U.S., health expenditure per capita grew by on average 4.1% annually since 1970 to a level of about 18% of GDP four decades later (Chernew and Newhouse, 2012; Gaynor et al., 2015). Starting at lower levels, other developed countries experienced similar rates of increase of the health sector such that, across the board, health expenditure increased faster than GDP. The resulting secular growth trends of health expenditure shares (Chernew and Newhouse, 2012) are predicted to continue in the future (e.g. European Union, 2010; OECD, 2016).

Scholars agree that most of the rise of health expenditure is caused by medical technological progress. Recent examples of cost-raising health innovations include computerized diagnostic tests (e.g. for medical imaging), personalized cancer therapy, and new treatments of virus infections like HIV or Hepatitis C. More generally, and supportive of our model, Lichtenberg (2007) shows that later vintages of pharmaceuticals are more powerful in the reduction of health deficits. Considering the evolution of 92 potentially lethal diseases he finds that conditions experiencing greater pharmaceutical innovation tend to have greater declines in mortality rates.

Most studies attribute between 50 and 80 percent of health expenditure growth to technological progress (Chernew and Newhouse, 2012). As argued convincingly by Chernew and Newhouse (2012), the persistent increase of health expenditure shares requires at least one other persistently growing explanatory variable (and thus rules out institutional changes like health care reforms and other only occasionally changing variables). Okunade and Murthy (2002) establish a long-run relationship between medical R&D expenditure and health care expenditure. There may be a role for demographics and income as a driver of increases health shares, although the latter is questioned by some recent studies refuting the luxury good hypothesis of health care by estimating an income elasticity of health expenditure below unity (Acemoglu, Finkelstein and Notowidigdo, 2013; Baltagi et al., 2016).

A promising example of a potentially powerful future technology is targeted genome editing like the clustered, regularly interspaced, short palindromic repeat (CRISPR) technology. It gives rise to the development of novel molecular therapeutics for human disease. Recent biotechnological R&D also
provision of salient health goods and services relies on highly regulated health insurance systems, the entailed increasing utilization of medical goods and services raised concerns about fiscal sustainability and, more generally, the overall desirability of these trends. It motivated the discussion of health care rationing as potential remedy to curb further rising expenditure shares (Aaron and Schwartz, 1990; Ham and Glenn, 2003; Singer, 2009). Indeed, health care rationing has become more and more visible in developed countries. For instance, the National Health Service (NHS) — managing tax-financed health care with guaranteed access in the UK — rations hip replacements and knee surgeries. In the mandatory German health system, if the amount of external costs attributable to a medical doctor exceeds a threshold per quarter, the doctor has to privately bear the costs above the cap. In addition, many health care systems (like the UK and Switzerland) have severely limited coverage of a novel (albeit expensive) drug that for the first time heals Hepatitis C.

In our setting, rationing care in order to constrain health expenditure growth has — aside from the obviously detrimental effects on health of the current population — severe “side-effects” on future health and longevity. According to the benchmark prediction of our model, that is calibrated to the UK, an individual who has reached age 20 in year 2080 could expect to die at age 100. The model highlights how such population ageing is associated with significant declines in morbidity. Nevertheless, as a result of population ageing, the health expenditure share in GDP would increase from 2010 by about two percentage points until 2080, along with an increasing employment share in the health sector. Stabilizing the health expenditure share, however, would reduce the increase of 

---

3 See Edwards, Crump and Dayan (2015). Contrary to many European health systems, US medicare (health insurance for the elderly) involves a co-insurance rate for pharmaceuticals of 25%. Co-insurance makes demand for pharmaceuticals price-elastic. In fact, in the US prices for pharmaceuticals are little regulated, compared to European health care systems. The fundamental issue of rationing health care provision is nevertheless present as well, of course, albeit in different form.

4 See http://www.hepatitisc.uw.edu/page/treatment/drugs/sofosbuvir-drug. NHS England has decided to provide treatment only to the 10,000 sickest persons of those being infected per year, a rather small fraction of the estimated 215,000 infected persons in the UK (https://www.theguardian.com/society/2016/jul/28/nhs-abandoning-thousands-by-rationing-hepatitis-c-drugs). In Switzerland, patients initially received Hepatitis C treatment only if they had severe liver damage (World Health Organization, 2016). Despite small modifications, rationing remains in place for the bulk of infected persons.
life-expectancy of a 20 year old person in year 2080 by almost 5 years compared to the baseline policy regime. In our model, the quality of health goods is endogenous and determines morbidity in interaction with access to health goods. Our results thus do not only reflect the detrimental health effects of increased health care rationing for a given health technology but also that it reduces the market size for new medical products, in turn suppressing medical R&D.

Deciding on the trade-off between promoting longevity and limiting increases in health costs is a fundamentally normative issue. We assume that instantaneous utility of surviving individuals depend on morbidity and material consumption. Marginal utility from consumption in each point in time negatively depends on morbidity. Our normative analysis suggests that particularly future generations would incur sizable welfare losses from extending health care rationing to a point that stabilizes the health expenditure share.5

Our main contribution is to highlight the interaction between endogenous medical technological progress and longevity as a function of health policy. Most empirical studies of the determinants of health expenditure estimate medical technological progress as a residual. The study by the European Union (2009), for example, regresses health expenditure against income, the population share above 65, and a time trend, and interprets the time trend (of on average 2 percent annually) as the rate of medical technological progress. Treating medical technological progress as a time trend, however, is problematic when predictions are made on long-run developments of population health and health expenditure under different policy scenarios. Implicitly these predictions assume that health policy does not affect medical progress. In our study we challenge this view by modeling endogenous medical innovation and endogenous population health for different policy scenarios. In particular, we show that limiting the rise in health expenditure has a detrimental effect on health R&D through a market size effect that is associated with the resulting reduction in health care utilization. We thus formalize an idea that

5We measure welfare changes by the factor we have to multiply material consumption levels under the baseline health policy scenario such that the ex ante life-time utility of a member of a given generation is the same as with health care reform (equivalent variation). We estimate that someone who is 20 years old in 2050 would experience a welfare loss of 8 percent when the health agency pursues the goal of stabilizing the health expenditure share.
goes back to Weisbrod (1991) who argues that the expansion of U.S. health care insurance has induced increasing health R&D and newly developed technologies that, in turn, have driven up health care utilization and costs. Testing this hypothesis, Acemoglu et al. (2006) could not show that the introduction of Medicare (the “Social Security Act of 1965” that covered hospital and doctor expenses) increased pharmaceutical demand and pharmaceutical R&D. This is perhaps not surprising since coverage of pharmaceuticals was not introduced until 2006. Extending the scope of analysis, however, Acemoglu and Linn (2004), showed large market size effects of the aging baby boomers on the development and market entry of new (age-specific) pharmaceuticals.

Related studies investigated the interaction of health R&D in “reduced form” by either assuming a direct utility gain from the consumption of pharmaceuticals (as in Garber, Jones and Romer, 2006) or by assuming a direct impact of health goods on the mortality rate of a representative individual (as in Jones, 2016). Garber et al. (2006) investigate the interaction between medical R&D and the generosity of the health care system, measured by the degree of coinsurance payment of individuals. New generations of pharmaceuticals are assumed to be directly utility increasing for individuals with the respective disease. Neither health nor longevity are explicitly modeled. Prices of pharmaceuticals are set in private markets (approximating the US health care system). In this setup, lower co-payments lead to higher demand and higher markups charged by drug producing firms. Consequently, profits of firms may exceed consumer surplus of patients such that, in this sense, there could be too much demand and too much medical R&D. In contrast to that paper, which highlights the problem of moral hazard when prices for pharmaceutical are set on markets, we assume that prices are regulated (approximating the British and German health care system, among others). We then focus on the interaction between health expenditure, medical R&D, morbidity and longevity of an age-structured population in a dynamic macroeconomic equilibrium model. For that purpose we employ the health deficit index developed by gerontologists (Mitnitski et al., 2000, 2002) and subsequently used in countless empirical studies in the natural sciences for measuring health status in an empirically meaningful way in order to calibrate our model. In contrast to health capital (a latent variable popular among economists; Grossman, 1972), health deficits are
observed and easily quantifiable. In our model, in line with the conceptualization of morbidity and physiological ageing in gerontology research, individuals accumulate health deficits over the life-cycle which in turn determine mortality rates at a given age. The individual accumulation process of health deficits depends on the interaction between the extent to which individuals are provided with health goods to treat their illnesses and the available quality of these health goods, where health good quality is endogenously determined by vertical R&D.

In another recent study, Jones (2016) proposes a macroeconomic model with deterministic, horizontal health innovations that affect longevity of a single cohort that privately buys health goods (similar to Grossman, 1972) with a trade-off to material consumption (featuring “love-of-variety” of consumers in both sectors). By investigating the optimal allocation of R&D effort directed towards innovations for health and non-health purposes, it is shown that under a mild condition non-health technological progress may optimally converge to zero growth such that the health expenditure optimally converges to 100 percent. The study makes an important, eye-opening contribution in the debate whether there is too much health care expenditure and it paves the way for our research. Our study, however, focusses on different research questions and proposes a multi-period, overlapping generations model with a much more detailed health care sector of the economy. In particular, we investigate the effects of health care (rationing) on health and longevity of an age-structured population with a public health care system and annuity markets. We consider uncertain, vertical R&D for health goods that are connected to potential illnesses, in turn determining the accumulation of health deficits.

In a development context, higher life expectancy may positively affect per capita income (e.g. Cervellati and Sunde, 2011). In fact, investments in human capital or entrepreneurship may be fostered because the gains of economic activity is spread on a longer time horizon. In advanced economies, however, longevity is enjoyed by retirees. Thus, publicly financed policy interventions to promote health good provision and health R&D do not necessarily raise per capita income and consumption levels. Rather there is a fundamental trade-off between longevity and material well-being that we focus on in this paper.
Finally, our paper is related to a strand of recent studies that utilized the health deficit approach to (re-)investigate the Preston curve (Dalgaard and Strulik, 2014), the role of adaptation for health behavior and health outcomes (Schuenemann, Strulik and Trimborn, 2015), the education gradient (Strulik, 2016), the historical evolution of retirement (Dalgaard and Strulik, 2017), and the optimal design of social welfare systems (Grossmann and Strulik, 2015).6

The reminder of the paper is organized as follows. Section 2 presents the model. Section 3 provides the positive analysis of health care rationing based on the calibrated model. The normative analysis is presented in section 4. The last section concludes.

2 The Model

Consider the following multi-period overlapping generations model in discrete time, indexed by $t$, in which individuals age by accumulating bodily impairments (“health deficits”). In line with the evidence on human ageing, on average, individual health deficits correlate exponentially with age and are a highly relevant determinant of the probability of death (e.g. Mitnitski and Rockwood, 2002a, 2002b, 2005). Health goods are provided via a tax-financed health care system without co-insurance, like in the UK and Germany. Improved quality of provided health goods slows down the ageing process. Medical R&D rests on private firms and is competitive. Also the final good sector and factor markets are perfectly competitive, whereas health good providers charge mark-up prices. Mark-up factors can be thought of being determined by negotiations between health care representatives and health good suppliers, again, like in the UK and Germany. There exists a perfect private annuity market. For simplicity, we assume that there is an international capital market that fixes the real interest rate, $\bar{r}$.

---

6Grossmann and Strulik (2015) investigate the interaction between increasing health expenditure, which promotes longevity, and a publicly financed pension system that is challenged by (endogenously) changing demography. They do not incorporate health R&D or a health good sector, however. Moreover, their analysis is confined to two periods of life (with endogenous lengths).
2.1 Households

Each period a new cohort is born. Mortality is cohort-specific and determined by health status, that is measured by the fraction of the health deficits an individual suffers from out of a long list of potential impairments, ranging from mild deficits (reduced vision, incontinence) to near lethal ones (e.g. stroke). In section 4, we also capture that a higher health deficit index (i.e. a lower health status) reduces the marginal utility of consumption for a given survival probability, in line with empirical evidence (Finkelstein, Luttmer and Notowidigdo, 2013).

Formally, the probability \( m_{v,t} \) of a member of cohort \( v \) to die between period \( t \) and \( t+1 \), conditional on having reached age \( t-v \geq 0 \), is increasing and strictly convex as a function of the health deficit index at that age, \( d_{v,t} \in [0,1] \). There exists a threshold deficit state \( d_{\text{max}} \in (0,1) \) such that no sufficiently old individual survives beyond that state. Moreover, there is a maximum life span (irrespective of health deficits), \( T \). These properties are captured by the parsimonious specification

\[
m_{v,t} = \begin{cases} \frac{1-e^{-\left(d_{v,t}\right)^{\phi}}}{1-e^{-\left(d_{\text{max}}\right)^{\phi}}} = \tilde{m}(d_{v,t}) & \text{if } d_{v,t} < d_{\text{max}} \text{ and } t < v + T - 1 \\ 1 & \text{otherwise,} \end{cases}
\]

where \( \sigma \) and \( \phi \) affect the “curvature” of the mortality function as a function of age and \( \sigma > 1 \) and \( \phi \geq (1-1/\sigma)^{-1} \) are sufficient for \( \tilde{m}'' > 0 \). Note that \( \tilde{m}(0) = 0 \) and \( \tilde{m}(d_{\text{max}}) = 1 \).

As will become apparent, specification (1) enables us to capture empirically observed survival rates reasonably well with a small set of parameters. By definition, the survival rates and conditional mortality rates are related by

\[
S_{v,t} = S_{v,v} \prod_{u=v}^{t-1} (1 - m_{v,u}) \text{ for } t \geq v + 1,
\]

i.e., \( m_{v,t} = -(S_{v,t+1} - S_{v,t})/S_{v,t} \). The initial cohort size in period \( v \) is \( S_{v,v} \).

Each individual works for \( R \) periods and inelastically supply one unit of labor in working age (and zero afterwards). We thus implicitly assume that, conditional on

---

7 According to Rockwood and Mitnitski (2007) and Scarle et al. (2008), the exact choice of the set of potential deficits is not crucial, provided that the set is sufficiently large.

8 Allowing for a positive elasticity of labor supply with respect to net wages rather than assuming
survival, labor supply is independent of health status. The total units of labor supplied to the economy in period $t$ are given by $L_t = \sum_{u=t-R+1}^{t} S_{u,t}$.

Households have preferences over material consumption and health status. They optimize over their consumption path. Because the interest rate is fixed, saving decisions of households do not interact with R&D decisions of firms and physical capital input. We thus first analyze the supply side and introduce life-time utility later to analyze welfare implications of our model.

2.2 Firms

There is a standard numeraire goods sector, producing a standard final good, and a health sector.

2.2.1 Numeraire Good Sector

The final good is chosen as numeraire. It is produced under perfect competition according to

$$Y_t = \left( K_t^{Y} \right)^{\alpha} (A_t L_t^{Y})^{1-\alpha}, \quad (3)$$

$\alpha \in (0, 1)$, where $K_t$ denotes the physical capital input in period $t$, $L_t^{Y}$ is the amount of labor in the consumption goods sector, and $A_t$ is a measure of non-health knowledge. Its level is initially given by $A_0 > 0$ and exogenously grows over time with constant rate $g > 0$.

Physical capital depreciates at rate $\delta^K \geq 0$. Thus, the user cost per unit of capital is given by $\bar{\rho} + \delta^K$. It is equal to the marginal product of capital, $\bar{\rho} + \delta^K = \alpha (A_t L_t^{Y} / K_t)^{1-\alpha}$. The wage rate, $w_t$, equals the marginal product of labor, i.e. $w_t / A_t = (1 - \alpha) (A_t L_t^{Y} / K_t^{Y})^{-\alpha}$.

Thus,

$$\frac{w_t}{A_t} = (1 - \alpha) \left( \frac{\alpha}{\bar{\rho} + \delta^K} \right)^{\frac{1}{1-\alpha}} \equiv \omega. \quad (4)$$

---

exogenous labor supply would be conflict with the evidence that hours worked have declined over a longer time horizon in many growing economies (e.g. Lee, McCann and Messenger, 2007).

---

9 In fact, at the individual level, a decline in health status does not seem to have a large effect on labor supply (see e.g. Jaeckle and Himmler, 2010, as well as Hokayem, and Ziliak, 2014).
2.2.2 Health Good Producers

Health Technology and Health R&D There is a continuum of potential illnesses that are represented by the unit interval, indexed by $j \in [0,1]$. The health sector provides patentable health goods (and services) like pharmaceuticals to treat illnesses. For each illness, there is a competitive R&D sector aiming to advance the treatment quality. A successful innovator provides a quality level that is by an amount $\gamma > 0$ higher than the quality without innovation, being formally awarded an infinitely-lived patent. As will become apparent, however, patent holders will frequently be driven out of business by future innovators. The quality of the latest vintage of health good $j$ available in period $t$ is denoted by $q_t(j)$. The quality of all health goods (including older vintages) may deteriorate over time at rate $\delta^Q \in (0, \gamma)$. In the case of pharmaceuticals, depreciation of quality captures mutations of bacteria and viruses, with resistance of antibiotics being a prime example.

Denoting by $\mu_{t+1}(j)$ the probability of a successful innovation of a health good to treat illness $j$ that can be commercialized in $t+1$, the quality of health good $j$ thus evolves according to

$$q_{t+1}(j) = \begin{cases} (1 - \delta^Q)q_t(j) + \gamma & \text{with probability } \mu_{t+1}(j), \\ (1 - \delta^Q)q_t(j) & \text{otherwise}. \end{cases} \quad (5)$$

Hence, the expected quality of a health good targeted to illness $j$ in period $t+1$, $\mathbb{E}[q_{t+1}(j)]$, is given by

$$\mathbb{E}[q_{t+1}(j)] = \mu_{t+1}(j) [q_t(j)(1 - \delta^Q) + \gamma] + (1 - \mu_{t+1}(j))q_t(j)(1 - \delta^Q). \quad (6)$$

Let $l_t(j)$ denote the amount of labor devoted by a representative R&D firm in health sector $j$ at $t$ and assume that the perceived probability of a successful innovation is proportional to it:

$$\tilde{\mu}_{t+1}(j) = \tilde{\xi}_t l_t(j), \text{ with } \tilde{\xi}_t \equiv \xi \cdot (L_t^Q)^{-\vartheta}, \quad (7)$$

$\xi > 0$, $\vartheta \in (0,1)$, where $L_t^Q$ is the aggregate amount of health R&D labor in $t$. $\tilde{\xi}_t$ is taken as given in the decision of R&D firms and captures a negative R&D ("duplication")
externality: \( \theta > 0 \) implies a wedge between the private and social return to R&D that may arise because firms do not take into account that rivals may work on the same idea such that, from a social point of view, some of the R&D is duplicated.\(^{10}\) In a symmetric equilibrium, where \( l_t(j) = L_t^Q \) for all \( j \in [0,1] \), we obtain \( \tilde{\mu}_{t+1}(j) = \tilde{\mu}_{t+1} = \xi \cdot (L_t^Q)^{1-\theta} \) for all \( j \).

There also may be innovations that are not profit-driven (e.g. unintentional or commercialized directly by public research institutions).\(^{11}\) They become effective in \( t+1 \) with probability \( \tilde{\mu}_{t+1} \). Let \( Q_t \equiv \int_0^1 q_t(j) \, dj \) denote the average quality of the latest vintages of health goods ("stock of medical knowledge"). We assume that there is an intertemporal spillover from the stock of medical knowledge (reflected by \( Q \)) such as

\[
\tilde{\mu}_{t+1} = \mu^0 Q_t , \tag{8}
\]

\( \mu^0 \in [0, \delta^Q/\gamma] \). Thus, the total probability of medical progress in each sector, \( \mu_{t+1} \), given by

\[
\mu_{t+1} = 1 - (1 - \tilde{\mu}_{t+1})(1 - \tilde{\mu}_{t+1}) = \mu^0 Q_t + (1 - \mu^0 Q_t) \cdot \xi \cdot (L_t^Q)^{1-\theta} . \tag{9}
\]

By the law of large numbers, there is no aggregate risk. Thus, \( \int_0^1 \mathbb{E}[q_{t+1}(j)] \, dj \) is deterministic and equal to \( Q_{t+1} \). According to (6), it evolves as

\[
Q_{t+1} = \gamma \mu_{t+1} + (1-\delta^Q)Q_t , \tag{10}
\]

where initial level \( Q_0 > 0 \) is given. Substituting (9) into (10), we obtain the growth rate of \( Q \) as declining function of its level:

\[
\frac{Q_{t+1} - Q_t}{Q_t} = \frac{\gamma(1-\tilde{\mu}_{t+1})\tilde{\mu}_{t+1} - \delta^Q}{Q_t} = \frac{\gamma(1 - \mu^0 Q_t)\xi (L_t^Q)^{1-\theta}}{Q_t} - \delta^Q , \tag{11}
\]

\( \delta^Q \equiv \delta^Q - \gamma \mu^0 > 0 \).

\(^{10}\)For \( \theta \to 1 \), social returns to medical R&D investment approach zero. The argument is analogous to that in Jones and Williams (2000) for a non-health R&D context.

\(^{11}\)The inventions of Penicillin and Viagra are prime examples of major breakthroughs that were not intended to treat the health problems they target today.
Marginal Costs of Health Goods and Price-Setting  Production of one dose of a health good in period $t$ requires $\chi > 0$ units of labor. That is, the capital requirement rises with the state of non-health technological progress. Thus, marginal production costs in period $t$ are $c_t \equiv \chi w_t$. The price mark-up of health goods can be thought of as an outcome of negotiations between the government as health insurer and (a representative body of) health good providers like pharmaceutical companies.\textsuperscript{12} For instance, in the UK, prices for pharmaceuticals are regulated and based on a non-contractual agreement between the UK Department of Health and the Association of the British Pharmaceutical Industry. Similarly, in Germany, health care suppliers negotiate with pharmaceutical companies the maximum price covered by the mandatory health insurance.

Suppose the government bids down suppliers of older vintages to their marginal costs and pays to the industry leader a mark up factor that is increasing in his quality advantage vis-à-vis previous vintages. Denoting by $q > 0$ the (absolute) quality advantage of the industry leader over the competitor with the second-highest quality product in the same market, suppose the mark up factor is given by $1 + f(q)$. Function $f$ captures the price setting power of health good providers. It is increasing and strictly concave with $f(0) = 0$. If the leading firm is one step ahead of the closest competitor (i.e. $q = \gamma$), it gets a profit per unit sold that is equal to $f(\gamma)c$, whereas the previous innovator receives zero profit. If the leading firm is two steps ahead of the closest competitor (i.e. $q = 2\gamma$), it gets a profit per unit equal to $f(2\gamma)c$. The profit increase for the industry leader by innovating, i.e. becoming two steps rather than one step ahead, is $[f(2\gamma) - f(\gamma)]c$. Since strict concavity of $f$ and $f(0) = 0$ imply $f(2\gamma) < 2f(\gamma)$, we have $[f(2\gamma) - f(\gamma)]c < f(\gamma)c$. Thus, it does not pay off for the leader to innovate. The incumbent firm would strictly prefer to invest in R&D in a second market rather than advancing its latest vintage.\textsuperscript{13} Consequently, the incumbent is driven out of business when there is an innovation in the market it leads and leader’s quality advantage to the closest competitor is exactly $q = \gamma$. Hence, the

\textsuperscript{12}Pharmaceutical companies may draw their negotiation power via lobbying and marketing that influences government negotiators and public opinion, respectively, on the merits of pharmaceuticals. For instance, interest groups representing the pharmaceutical sector strongly argue that they need to earn high profits enabling them to conduct R&D and therefore have to charge high prices that should be covered by health insurance.

\textsuperscript{13}See Grossman and Helpman (1991) for a similar argument in a context of Bertrand competition.
price \( p_t \) of each health good is given by

\[
p_t = \Gamma c_t = \Gamma \chi w_t = \Gamma \chi \omega A_t,
\]

where \( \Gamma \equiv 1 + f(\gamma) \) is the mark-up factor.

### 2.3 Health Care Provision

In many advanced countries, the bulk of individuals exclusively rely on a highly regulated health system with compulsory contributions (e.g. Germany and Switzerland) or is tax-financed like the National Health Service (NHS) in the UK. For simplicity, we assume that it is the only provider of health goods. There may be rationing in health care provision. As formally introduced and parameterized below, rationing is defined as falling short of the maximally health-effective provision of health goods for a given state of technology.

We do not consider the possibility of "out-of-pocket" health payments or co-insurance. Although the absence of these features are limitations of our analysis, we capture reasonably well the health system of the UK, to which we calibrate our model. In NHS, like in the mandatory German health system, co-insurance is absent. "Out-of-pocket" health expenditure as fraction of total health expenditure have been around 10 percent in the 2000s and also private health insurance coverage has been at a modest level (10.5 percent in the year 2014; see OECD, 2016). Many salient health goods (like surgeries treating orthopedic deficits or drugs for treating cancer and virus infections) may indeed be unaffordable for the bulk of individuals (presumably those who do not have private insurance in the UK either), if not covered by NHS.

### 2.4 Health Deficit Accumulation

We assume that physiological ageing starts when individuals become economically active, i.e. consume and supply labor.\textsuperscript{14} Modern gerontology describes ageing as an accumulation of health deficits (e.g. Mitnitski and Rockwood, 2002a, 2002b, 2005). The evidence suggests that individual health deficits grow exponentially with age in advanced countries

\textsuperscript{14} We will calibrate the model such that the initial period for each cohort member is at the age of 20.
(e.g. Mitnitski et al., 2002; Harttgen et al., 2013). Thus, we assume that the change in the deficit index of a member of cohort $\nu$ between period $t$ and $t+1$ is increasing in the deficit index accumulated until period $t$. The accumulation process is slowed down by receiving the publicly financed health input, $E_{\nu,t}$, conceptualized below. The health deficit index evolves according to

$$d_{\nu,t+1} - d_{\nu,t} = \begin{cases} \varrho d_{\nu,t} - \kappa E_{\nu,t} & \text{if } E_{\nu,t} < \frac{\varrho}{\kappa} d_{\nu,t}, \\ 0 & \text{otherwise}, \end{cases}$$

(13)

$\kappa > 0$, $\varrho > 0$, with initial value $d_{\text{min}} \equiv d_{\nu,0} > 0$. Parameter $\varrho$ is the growth rate of the health deficit index in absence of health interventions. It can be interpreted as the physiological “force of ageing”. Parameter $\kappa$ measures the effectiveness of health care.

We conceptualize health input, $E_{\nu,t}$, as total health good consumption to treat illnesses that are caused by existing health deficits, weighted by the quality of the consumed health goods. We thereby capture that health deficits derive from past, not fully cured or not fully curable, illnesses. For instance, consider the physical difficulty to move that is known to contribute to developing cardiovascular diseases. As another example, feeling lonely may cause major depressive disorder. Both health deficits are within the set of potential health deficits to calculate the deficit index in the empirical gerontology literature that motivates our modeling approach.

We normalize the maximally effective consumption per health good, per individual and per period, to unity. For instance, in the case of pharmaceuticals there is an optimal dose. We capture rationing in public health care provision by allowing the actual dose consumed per health good, per period and per individual, to be smaller than unity. Accordingly, rationing lowers the effective health input. Formally, an individual born in $v$ acquires a set $I_{v,t} \subset [0, 1]$ of illnesses in period $t \geq v$ with measure equal to the current deficit index, $|I_{v,t}| = d_{v,t}$. The “health care provision wedge” in $t$ is parameterized by $\varphi_t \in [0, 1]$. No rationing means $\varphi = 0$, whereas $\varphi = 1$ holds in absence of a health system. By the law of large numbers, suffering from a set of illnesses $I_{v,t}$ of measure $d_{v,t}$ in $t \geq v$,

---

15 Health deficit accumulation would cease if the effective health input became sufficiently high. Although this may not pure utopia but conceivable with further biotechnological advances (De Grey and Rae, 2007), we will not consider the case in this paper.
an individual born at \( v \) thus experiences the effective health input

\[
E_{v,t} = (1 - \varphi_t) \int_{j \in I_{v,t}} q_t(j) \, dj = (1 - \varphi_t) d_{v,t} Q_t. \tag{14}
\]

It depends on the contemporaneous health care provision wedge, \( \varphi_t \), as well as the interaction between the current deficit state and the average quality of health goods, \( d_{v,t} \cdot Q_t \). Substituting (14) into (13), the growth rate of the health deficit index is deterministic and independent of the deficit state. For \( t \geq v \) it is given by

\[
\frac{d_{v,t+1} - d_{v,t}}{d_{v,t}} = \begin{cases} 
\varrho - (1 - \varphi_t) \kappa Q_t & \text{if } Q_t < \frac{\varrho}{\kappa(1-\varphi_t)} \equiv \bar{Q}_t, \\
0 & \text{otherwise.} \end{cases} \tag{15}
\]

Each surviving member of cohort \( v \) in period \( t \) consumes \( h_{v,t} = (1 - \varphi_t) S_{v,t} d_{v,t} \) units per health good from the latest vintages (i.e. an average dose \( 1 - \varphi_t \) for any member of cohort \( v \) and for each illness \( j \in I_{v,t} \)), and nothing from the older vintages of the health good. Summing up over all cohorts already born, total demand in \( t \) for each selected vintage of a health good is \( H_t = \sum_{v=-\infty}^{t} h_{v,t} \).\(^\text{17}\) Using

\[
h_{v,t} = (1 - \varphi_t) S_{v,t} d_{v,t} \tag{16}
\]

and (2), we obtain

\[
H_t = (1 - \varphi_t) \sum_{v=-\infty}^{t} d_{v,t} S_{v,t} \prod_{u=v}^{t-1} (1 - m_{v,u}). \tag{17}
\]

Thus, more rationing in health care provision (a higher \( \varphi \)) saves health costs by reducing health good consumption, all other things being equal. However, a higher \( \varphi \) has two detrimental effects on morbidity and life expectancy. First, according to (15), it speeds up the evolution of health deficits for a given average quality of health goods, \( Q \). Second,\(^\text{16}\)

\[\text{If the economy reached a quality level of health goods } Q_t \geq \bar{Q}_t \text{ in finite time } \bar{t} \text{ and } \varphi_t \leq \varphi_{\bar{t}} \text{ for all } t \geq \bar{t}, \text{ all individuals born in } v \geq \bar{t} \text{ would remain with } d_{v,t} = d_{\text{min}} \text{ for all } t \geq \bar{t}. \text{ The case does not arise in our calibrated model, but would be implied by assuming zero depreciation of health good quality } (\delta^Q = 0).\]

\[\text{Recall that sufficiently old cohorts will not have any surviving members.}\]
according to (17), it lowers market size for health goods, \( H \). A lower market size, however, reduces innovation incentives to improve health good quality, \( Q \).

3 Positive Analysis

We first highlight some equilibrium conditions and then conduct a supply-side analysis of the calibrated model. For later use, as the number of retirees is \( O_t \equiv \sum_{u = t - T + 1}^{t - R} S_{u,t} \), the ratio of retirees to workers ("old-age dependency ratio", \( DPR \)) is given by

\[
DPR_t \equiv \frac{O_t}{L_t} = \frac{\sum_{u = t - T + 1}^{t - R} S_{u,t}}{\sum_{u = t - R + 1}^{t} S_{u,t}}. \tag{18}
\]

GDP in the economy is given by \( GDP_t \equiv Y_t + p_t H_t \). The ratio of health expenditure to GDP reads as

\[
s_t \equiv \frac{p_t H_t}{GDP_t} = \frac{p_t H_t}{Y_t + p_t H_t}. \tag{19}
\]

Finally, the inverse of the probability of an incumbent to be driven out of the market is the effective patent length (\( EPL \)), i.e. \( EPL_t \equiv \frac{1}{\mu_t} \).

3.1 Equilibrium Conditions

Denote by \( \pi_t \) the instantaneous profit of health good producers, which are all identical \( ex \ ante \) and thus also \( ex \ post \). Ruling out bubbles and arbitrage possibilities in the financial market and accounting for the probability \( \mu_u \) that health good producers are driven out of business in period \( u \geq t + 1 \), the value of a vertical innovation in the health sector created in \( t \) reads as

\[
V_t \equiv \pi_t + \sum_{u = t + 1}^{\infty} \frac{1}{(1 + \bar{\rho})^{u-t}} \prod_{s=t+1}^{u} (1 - \mu_s) \pi_u. \tag{20}
\]

Labor market clearing implies that

\[
L_t^Y + L_t^H + L_t^Q = L_t, \tag{21}
\]

where \( L_t^H \equiv \chi H_t \) denotes total employment in health goods production. For later use, denote employment shares by \( \ell_t^Y \equiv L_t^Y / L_t \), \( \ell_t^H \equiv L_t^H / L_t \) and \( \ell_t^Q \equiv L_t^Q / L_t \), i.e. in
equilibrium, \( t^e_t + t^H_t + t^Q_t = 1 \).

As implied by the assumption that the interest rate is exogenous, consumer choices (introduced in section 4) do not play a role for the allocation of labor, health contribution rates, longevity and morbidity. The dynamical system, including the long run equilibrium, is summarized in Appendix A.

### 3.2 Calibration

We dynamically calibrate the model to endogenous observables in the UK whenever available; otherwise we use North American data. We first consider the output elasticity of labor, \( 1 - \alpha = wL^Y/Y \). According to Karabarbounis and Neiman (2014, “CLS KN merged”), the arithmetic average for the period 1987-2011 of the UK corporate labor share in total income has been 62 percent. Thus, we set \( \alpha = 0.38 \). For the real interest rate we choose the typical value \( \bar{\rho} = 0.05 \). For the depreciation rate of physical capital we follow Grossmann and Steger (2016) who argue that \( \delta^K = 0.07 \). The growth rate of wage rates is set equal to the long run growth rate of income per capita, \( g = 0.02 \).

We assume that individuals become economically active at age 20 and live for a maximum of 100 additional years (i.e. nobody reaches age 121); thus, \( T = 101 \). In fact, for modern times, 120 years seems to be the maximum life-span, irrespective of increasing life-expectancy in the last decades. The retirement age is reached after \( R = 43 \) working years (i.e. at age 63).\(^{18}\)

We set \( d_{\text{min}} \) equal to the average health deficit index for a 20 year old in recent times. Using Canadian data, Mitnitski et al. (2002a) suggest \( d_{\text{min}} = 0.03 \). Empirical evidence also suggests that the deficit state that leads to death for sure approximately is about two thirds (e.g. Harttgen et al., 2013); thus, \( d_{\text{max}} = 0.67 \).

The remaining parameters are the mortality rate curvature parameters \( (\sigma, \phi) \), the labor requirement per unit of health good \( (\chi) \), medical R&D technology parameters \( (\xi, \delta^Q, \vartheta) \), innovation step size \( (\gamma) \), health care effectiveness \( (\kappa) \), the strength of the intertemporal innovation spillover \( (\mu^0) \), the mark-up health deficit accumulation parameters \( (\varrho) \),

\(^{18}\)In the UK, the average age of withdrawl from the labor market is around 64 for males and slightly below 62 for females in the 2000s (Mitchell and Guled, 2010).
the initial quality index of health goods \((Q_0)\), the time path of the health care wedge, \(\{\varphi_t\}_{t=0}^{\infty}\), the vector of health deficits of all cohorts with living members in the initial period (denoted by vector \(d_0\)), and the time path of initial survival rates \(\{S_{v,t}\}_{t=0}^{\infty}\). They are chosen to simultaneously match (i) UK survival rates \(S_{v,t}\) for ages 20-100 and periods 1950, 1970, 1990 and 2010, (ii) the UK ratio of health expenditure to GDP \((s_t)\) between 1980-2010, (iii) the recent average rate of change of the health deficit index, \(\delta_{v,t}\), in the cross-section of Canadian cohorts, (iv) the UK employment share in the health sector \((\ell^H_t)\).19

To match UK survival rates from year 1950 onwards (www.mortality.org), with \(T = 101\), we need to specify initial conditions for the deficit index of all cohorts with living members in 1850, i.e. we choose year 1850 for \(t = 0\). Initial deficit states, \(d_0\), correspond to steady state values when the health system is non-existing (i.e. \(\varphi_0 = 1\)).20 According to (1) and (15), the evolution of survival functions is exclusively driven by \(d_{\text{min}}\), the exogenous time paths \(\{\varphi_t\}_{t=0}^{\infty}\) and \(\{S_{v,t}\}_{t=0}^{\infty}\), and the endogenous time path of the health good quality index \(\{Q_t\}_{t=0}^{\infty}\). Matching them turns out to require that \(\varphi_t\) is gradually declining and \(S_{v,t}\) is non-decreasing over time.21 The underlying time path of \(S_{v,t}\) is plausible as it reflects decreases over time in child mortality and fatal accidents for young individuals. The time path of \(\varphi_t\) is roughly consistent with the historical improvements in the British health care system (e.g. Stewart, 2015).22 For recent times, any health system is still characterized by some rationing in various forms. First, the density of physicians is much lower in rural areas than in urban areas, suggesting that access to health care is limited in rural regions (OECD, 2015, Fig. 7.10). Second, rationing takes place implicitly through waiting lists and often requires that a health problem is sufficiently severe (OECD, 2015, Fig. 7.11-7.13). Its implementation mitigates the improvements in access to health care over time. Nevertheless, although some rationing measures have recently

19 Initial labor productivity, \(A_0\), does not enter the dynamical system for the positive analysis (Appendix A).
20 Formally, denote by \(\delta_a\) the deficit index of a surviving individual of age \(a\) in absence of a health system. According to (15), the initial distribution of the deficit index is then given by \(\delta_a = d_{\text{min}}(1 + a)^a\), \(a \geq 0\).
21 See Figure A.1 in Online-Appendix.
22 For the past, innovations associated with health improvements may not exclusively be interpreted as being associated with the health sector but include better sanitation and better environmental conditions that more individuals received access to over time.
been introduced, we assume that urbanization and better information about treatment possibilities of patients acts as a counteracting force that overall has led and will continue to lead to an improving health care usage over time. Thus, also for the next decades, we assume for the baseline calibration that \( \varphi_t \) moderately decreases (from about 0.15 in 2010 to 0.05 in year 2080). We assume for the baseline scenario that such improvements in health care have been and will be anticipated by the economic subjects in our model. The initial quality index of health goods (in 1870), \( Q_0 \), is one percent of the steady state value of \( Q \) that results for \( \varphi = 0.05 \).23

The calibrated model fits the survival functions quite well, as shown in Figure 1. The most important deviation from the data (solid lines) is for the younger individuals in 1950 and 1970, which do, however, have fewer health deficits than older individuals and therefore are responsible for a small fraction of health costs.24 Second, the implied health expenditure share in GDP \( (\mathbf{s}) \) is 5.0 percent in 1980, 5.1 percent in 1990, 6.2 percent in 2000 and 8.0 percent in 2010, compared to the observed UK levels of 5.1, 5.1, 6.3 and 8.5 percent, respectively (OECD, 2016). Third, the rate of change of the health deficit index across cohorts implied by the calibration is 3.8 percent. According to Mitnitski et al. (2002a), the estimated rate of change of the health deficit index at a given year in the cross-section of Canadian cohorts is equal to 4.3 percent for men and 3.1 percent for women. Finally, we may approximate \( t^H_t \) with the employment share in human health activities as published by the OECD. For the UK, in 2010, it was 7.3 percent.25 Including additionally residential care and social work activities (that may not include other activities than health care provision) would suggest that \( t^H \) was 12.7 percent. Our calibrated model gives us a value in-between, equal to 10 percent in 2010.

---

23Our calibrated model leads to the case where steady state quality of health goods \( \hat{Q} \equiv \lim_{t \to \infty} Q_t < \bar{Q} \). We can verify that the steady state equilibrium of the calibrated model is saddle-point stable.

24Table A.1 in the Online-Appendix compares in detail the remaining life expectancy at a given age implied by the calibrated model with the actual ones in the UK. Remaining life expectancy is computed as follows. Recall that the number of persons surviving to age \( t - v \) is given by (2). We calculate the "person-years lived" between ages \( t - v \) and \( t - v + 1 \) when born in \( v \) as \( P_{v,t} = S_{v,t+1} + 0.5 \cdot S_{v,t} m_{v,t} \), where \( S_{v,t} m_{v,t} \) is the number of persons dying between age \( t - v \) and \( t - v + 1 \) for those born in \( v \). Now define \( \Sigma v,t := (S_{v,t})^{-1} \sum_{u=t-120} P_{u,t} \) as the total number of years lived after attaining age \( t - v \) divided by the \( S_{v,t} \) individuals reaching that age, where the maximum age is restricted to 120. Strictly speaking, although commonly reported, this is not "life expectancy" at a given age of some cohort member since medical technological progress is neglected.

Unfortunately, we do not have data for the UK employment share of medical R&D workers ($t^2_t$) and the effective patent length ($EPL_t$) in the medical sector. These outcomes are critically determined by the price mark up for medical goods ($\Gamma$). For $\Gamma = 1.25$, we obtain $t^2_t = 0.011$, $\bar{\mu}_t = 0.041$ and $EPL_t = 1/\mu_t = 11$ for the year 2010. The implied effective patent length is close to the 10 years assumed in Jones and Williams (2000).

3.3 Results

We now examine for alternative policy scenarios the future evolution (for periods $t \geq t_0$) of age-specific survival rates ($S_{\epsilon,t}$), age-specific morbidity ($d_{\epsilon,t}$), age-specific health care demand ($h_{\epsilon,t}$), the total health expenditure share ($s_t$), the employment structure ($t^H_t$, $t^Q_t$), and the old-age dependency ratio ($DPR_t$).

We start with the implications of the baseline scenario for the future, where the health care wedge ($\varphi_t$) continues to decrease slightly until 2080 (to 0.1). Panel (a) of Figure 2 displays survival curves for 2020 (solid black line), 2050 (dashed blue line) and 2080 (dotted green line), suggesting that they are considerably shifting upwards over time. An individual that has reached age 20 in year 2020 can expect to live until age 82.4. Analogous figures are 91.1 and 100.6 years when reaching age 20 in 2050 and 2080, respectively. Moreover, whereas someone having reached age 80 in 2020 can expect to live until age 90.5. Analogous figures are 96.6 and 105.7 years when reaching age 80 in 2050 and 2080, respectively.

Panel (b) suggests that rising survival rates are driven by improvements in morbidity. This is because age-specific mortality rates ($m_{\epsilon,t}$) decrease when health deficits ($d_{\epsilon,t}$) at a given age decline over time, according to (1). According to (15), as the future health...

---

26 See Appendix A for a discussion.
27 The implied share of workers in medical R&D occupations of about one percent seems high at the first glance, if a strict definition of this occupation is applied, like medical scientists and engineers. For instance, we may consider US data from the Bureau of Labor Statistics (2016) for data availability reasons. Adding up for the year 2015 the number of biological scientists (102,000 employees), medical scientists (110,000 employees) and biomedical engineers (21,000 employees) suggests a combined employment share of only 1.7 per mill. However, a more appropriate interpretation of employment related endogenous technical progress in our model requires to add managers (for strategic decisions and marketing) and other professionals (like patent lawyers) who organize and commercialize medical R&D.
28 See Table A.2 in Online-Appendix (left columns).
Figure 1: Survival curves for 1950, 1970, 1990 and 2010: Calibrated model vs. UK data. Notes. (1) Empirical series: solid lines, calibrated model: circles (2) Data source: www.mortality.org. (3) Time paths \( \{ \varphi_t \} \) and \( \{ S_{r,v} \} \) are displayed in Figure A.1 (Online-Appendix). (4) Initial quality index (in 1870) \( Q_0 = 0.01 \cdot \lim_{t \to \infty} Q_t \) for \( \lim_{t \to \infty} \varphi_t = 0.05 \). (5) Other parameters: \( \alpha = 0.38, \delta^K = 0.07, \sigma = 1.5, \phi = 2.65, \gamma = 0.9, \varphi = 0.04, \kappa = 0.05, \xi = 0.07, \mu^0 = 0.13, \delta^Q = 0.02, \vartheta = 0.6, \vartheta = 0.02, \bar{r} = 0.05, \delta_{\min} = 0.03, \delta_{\max} = 0.67, \gamma = 0.1, \Gamma = 1.25, T = 101, R = 43. \)
care wedge is time-invariant, the accumulation of health deficits is exclusively driven by a rising quality of health goods \((Q_t)\). For instance, the health deficit index (the fraction of actual health deficits out of a set of possible deficits) for someone having reached age 80 declines from 19.3 percent in 2020 to 14 percent in 2050 and 10 percent in 2080.

The evolution of health deficits \((d_{v,t})\) also determines, in interaction with survival rates \((S_{v,t})\), the evolution of age-specific health care demand \((h_{v,t})\), according to (16). As displayed in panel (c), total age-specific health care demand is inverted U-shaped as a function of age. Over time it decreases for younger individuals, whereas it increases for older ones. The result reflects that, for younger individuals, improvements in the quality of health goods have little effect on survival rates, whereas the opposite holds for older individuals. In fact, survival rates of younger individuals are high to begin with and their deficit index is low. For them, both also change little over time, compared to older individuals. For an older age-group, total health care demand is rising over time since health deficits grow exponentially (holding the quality of health goods constant) and quality improvements in health care help reducing age-specific mortality rates substantially.

Since older individuals have higher levels of health care demand, population ageing may result in increasing health expenditure shares \((s_t)\). As shown in panel (d), despite declining health deficits at given age over time, \(s_t\) indeed increases from 8.2 percent in 2020 to 9 percent in 2050 and 10 percent in 2080. This holds despite the feature of our model that prices of health goods do not grow at higher rate than the source of health care finance (i.e. wages) over time and despite declining morbidity at any age. Panel (e) shows that rising health care demand is, quite intuitively, associated with a rising employment share in the production of health goods (see \(t^H_t\) at the right axis). Importantly, it also raises incentives for health innovations through increased market size. This implies that the medical R&D labor share \((t^Q_t)\) is rising over time as well (see \(t^Q_t\) at the left axis).

As is well known, demographic change induced by human ageing leads to a rising old-age dependency ratio \((DPR_t)\) over time. The interesting question is by how much. Projections in the literature that do not account for the endogeneity of health care quality are not very informative in this respect. Panel (f) shows the evolution of the ratio of population size aged 63+ (retirement age) to the population size aged 20-62 (working
age). It suggests that $DPR_t$ rises from $x$ percent in 2020 to $x$ percent in 2050 and $x$ percent in 2080.

In sum, our model gives rise to an important insight that has yet not been clearly worked out in the literature: population ageing that is associated with health improvements at any age may inevitably be associated with rising health expenditure shares. In this sense, rising health costs are good news. As we will argue in the welfare analysis next section, therefore, measures to raise health care rationing may not be desirable.

Before doing so, we analyze the consequences of a health care rationing scheme that stabilizes the health expenditure share for $t \geq t_0$, starting in 2020. It requires to substantially increase in the health care provision wedge ($\varphi_t$) over time, from 13 percent in 2020 to 19 percent in year 2050 and 29 percent in year 2080. ²⁹

The thin lines in Figure 3 repeat the results for the baseline scenario shown in Figure 2, whereas the thick lines correspond to the reform scenario where health care rationing is extended. Panels (a) and (b) suggest that age-specific survival rates ($S_{i,t}$) and morbidity ($d_{e,t}$) do not improve much over time, respectively, in contrast to the baseline scenario. The differences across policy regimes are particularly visible for 2080, whereas differences are small for 2050. Panel (c) shows that age-specific health care demand ($h_{e,t}$) is reduced compared to the baseline scenario, particularly for older age-groups. It reflects that survival rates of older cohorts do not improve much anymore over time. This comes at the benefit of a time-invariant health expenditure share ($s_t$) over time, as displayed in panel (d). Relatedly, panel (e) shows that the employment share in the production of health goods ($\ell_t^H$) is basically time-invariant. This reflects a policy-driven stabilization of market size, which in turn is associated with a medical R&D labor share ($\ell_t^Q$) that is quite stable over time as well. That is, compared to the baseline scenario, medical R&D effort is reduced. This dynamic incentive effect of health care rationing adds to the static reduction in health care usage, jointly slowing down both the process of population ageing and health improvements in society. According to panel (f), it also materializes in a slowdown of the increase over time in the old-age dependency ratio ($DPR_t$), now rising only slightly from $x$ percent in 2020 to $x$ percent in 2050 and $x$ percent in 2080.

²⁹The scheme is displayed in Figure A.2 of the Online-Appendix.
Figure 2: The future of human health, longevity, and health costs for the baseline policy scenario. Notes: Panels (a)-(c): Solid black line for 2020, dashed (blue) line for 2050, dotted (green) line for 2080. Parameters as for Figure 1.
For the year 2050, with the considered health care reform, life expectancy increases by 0.8 years in 2050 and 4.8 years less when reaching age 20 in year 2050 and 2080, respectively, than in the baseline scenario. The respective increases in life expectancy are lowered by 0.6 and 3.6 years for someone having reached age 80 in 2050 and 2080, respectively. In sum, shorter run effects from implementing the cost-saving health care reform are smaller than longer run effects. Over time, by raising morbidity, the reform induces sizable losses in life expectancy.

4 Normative Analysis

In this section, we examine the welfare implications of the switch in health policy from the baseline scenario (Figure 2) to the one of extended rationing from year 2020 onwards (Figure 3). For concreteness, we assume that the policy regime switch is not anticipated by living members of generations born before the shock.

4.1 Welfare Behind the Veil of Ignorance

We first need to define an appropriate welfare criterion. Facing uncertain death, rational individuals calculate (under rational expectations) the expected utility from life-time consumption by multiplying the instantaneous utility experienced in a given period with the probability to survive beyond that period. In this section we consider welfare behind the veil of ignorance. We assume that instantaneous utility depends on current consumption levels of the numeraire and health status as measured by the health deficit index.

Formally, with maximum life span $T$, a member of cohort $v$ has preferences that are represented by the intertemporal utility function

$$U_v = \sum_{t=v}^{v+T-1} \beta^{t-v} S_{v,t} u(c_{v,t}, d_{v,t}), \quad (22)$$

---

30 Again, see Table A.2 in Online-Appendix (right columns).
Figure 3: Effects of extending health care rationing from year 2020 onwards for stabilizing the health expenditure share (reform scenario). Notes: (1) Panels (a)-(c): Solid black line for 2020, dashed (blue) line for 2050, dotted (green) line for 2080. (2) Thin lines repeat the baseline scenario, thick lines show the new scenario. (3) Time path \( \{ \varphi_t \} \) that is displayed in Figure A.2 (Online-Appendix). Parameters as for Figure 1.
with instantaneous utility with

\[ u(c_{v,t}, d_{v,t}) = \frac{\log c_{v,t}}{(1 + d_{v,t})^\zeta}, \]  

(23)

where \( c_{v,t} \) is consumption of an individual born in \( v \) in period \( t \geq v \). Recall that \( S_{v,t} \) is the probability to survive beyond age \( t - v \) and the cohort size of generation \( v \) in \( t \). \( \beta \geq 0 \) is the discount factor. \( \zeta > 0 \) measures to which extent a higher deficit state reduces the marginal utility of consumption. For an individual without health deficits (\( d_{v,t} = 0 \)) or in the case where \( \zeta = 0 \), we are back to a standard instantaneous utility function. With log-utility, the intertemporal elasticity of substitution is unity, as supported by Chetty (2006), among others.

We assume that the health care system is financed by a constant contribution rate out of wage income, denoted by \( \tau_t \) in period \( t \).\(^{31}\) The health care budget is balanced in each point in time; that is, revenue, \( \tau_t w_t L_t \), equals expenses, \( p_t H_t \). Recalling (12), the health contribution rate is proportional to the share of labor (\( \ell^H_t \)) allocated for producing health goods and services:

\[ \tau_t = \Gamma \ell^H_t. \]  

(24)

Denote asset holding (“wealth”) of a member of cohort \( v \) in \( t \) by \( a_{v,t} \). Initial asset holding is \( a_{v,v} = 0 \) since there is no bequest motive and the annuity market is perfect. We assume fair insurance within a cohort in the annuity market, where zero-profit insurance companies keep the individuals’ wealth after death. The law of motion for an individual of cohort \( v \), wealth at \( t \geq v \) can be written as

\[ a_{v,t+1} = (1 - \tau_t) w_t + (1 + r_{v,t}) a_{v,t} - c_{v,t}, \]  

(25)

\(^{31}\)Assuming that health insurance is paid by workers and enjoyed by retirees greatly simplifies the analysis. If health insurance were also be financed by capital income, we would have to keep track of aggregate asset holdings in the economy. Recall that these are unrelated to investments in our model. In our model, tax-financing health care no distortions because we abstract from a labor supply choice.
where the cohort-specific interest factor between date $t$ and $t+1$ is given by\footnote{See e.g. Heijdra, Mierau and Trimborn (2016).} \[1 + r_{v,t} = \frac{1 + \bar{r}}{1 - m_{v,t-1}}. \tag{26}\]

Individuals of each generation $v$ choose their consumption paths $\{c_{v,t}\}_{t \geq v}$ to maximize utility $U_v$ s.t. (25) and non-negativity constraint $a_{v,v+T} \geq 0$. They have perfect foresight about the health contribution rate and health deficit states (including their implications on mortality risks as given by 1)) that would result in the baseline policy regime and take these as given when optimizing.

In period $t_0$ when an unanticipated policy shock hits (again, in year 2020), those living members of generations $v < t_0$ (i.e. those already born) reoptimize, correctly assuming that the new policy regime prevails in the aftermath. Members of future generations $v \geq t_0$ optimize under perfect foresight of the outcomes resulting from the new policy regime. The optimization problems are solved in Appendix B. Welfare behind the veil of ignorance equals expected life-time utility from the resulting consumption paths.

Welfare effects of policy reforms are evaluated as follows. Let superscript 0 on consumption levels, deficit states and survival rates denote the values of these variables in the baseline policy regime and superscript 1 the values in the policy reform regime. Moreover, let

\[ U_v^i(\psi) \equiv \sum_{t=v}^{v+T-1} \beta^{t-v} S_{v,t}^i \log(\psi c_{v,t}^i) \left(1 + d_{v,t}^i \right)^{-1}, \tag{27}\]

denote the life-time utility of cohort $v$ when consumption levels in scenario $i \in \{0, 1\}$ are multiplied with factor $\psi > 0$. By definition (27), life-time utility is $U_v^1(1)$ in the reform scenario. We report cohort-specific factors $\psi_v$ that solve

\[ U_v^0(\psi_v) = U_v^1(1). \tag{28}\]

Thus, $\psi_v$ is the equivalent variation (EV) welfare measure in the baseline scenario such that cohort $v$ gets the same utility than in the reform scenario.\footnote{See Jones and Klenow (2016) for a similar way to measure welfare differences of randomly chosen individuals in a cross-country context rather than across policy regimes.}
4.2 Calibration

We choose a typical value for the subjective discount rate, $\beta$, such that $\beta(1 + \bar{r}) > 1$; setting $\beta = 0.98$ (recall $\bar{r} = 0.05$). Next, we calibrate $\zeta$, which determines the loss in marginal utility from consumption caused by health deficits. Finkelstein et al. (2013) find that, starting at the mean, a one-standard deviation increase of chronic diseases is associated with a decline in the marginal utility of consumption, denoted by $LOSS$, of 11.2 percent. The 95% confidence interval ranges from 2.7% to 16.8%. Marginal consumption utility reads as $(1 + \delta_{v,t})^{-\zeta}/c_{v,t}$. Evaluated at the mean deficit index, $E(d)$, and denoting the standard deviation by $STD(d)$, the estimate of Finkelstein et al. (2013) then suggests that $\zeta$ is given by

$$\frac{[1 + E(d) + STD(d)]^{-\zeta}}{[1 + E(d)]^{-\zeta}} = 1 - LOSS. \quad (29)$$

According to Mitnitski et al. (2002), the mean deficit index in the population is $E(d) = 0.054$ and the standard deviation is $STD(d) = 0.024$. Hence, $\zeta = -44.42 \cdot \log(1 - LOSS)$. The point estimate of $LOSS = 0.112$ thus suggests $\zeta = 5.1$; the bounds of the 95% confidence interval imply $\zeta \in [1.22, 8.17]$.

Finally, we have to calibrate the initial general state of technology, $A_0$. This is because the time path of productivity and wage income potentially affects welfare changes. We do this by targeting a certain ratio of the value of life to GDP per worker, $y \equiv GDP/L$. Denote the value of life of an individual born in $v$ by $\Omega_v$ and assume it is given by expected (indirect) life-time utility in the baseline scenario normalized by the marginal instantaneous (indirect) utility in the initial period of life:

$$W_v \equiv \frac{U^0_v(1)}{u_c(c_{v,v}, d_{\min})}. \quad (30)$$

We set $A_0$ (still assuming that $A_t$ grows annually at rate $g = 0.02$) such that $W_v/y_v = 20$ for the year 2010.\footnote{If we assumed $\beta(1 + \bar{r}) = 1$, then individual consumption would monotonically decrease with age, which is inconsistent with the evidence.} \footnote{This is an explicit normative judgement. In fact, the value of life is an inherently normative concept. Any attempt to compute it in the literature has necessarily been based on strong assumptions. For}
4.3 Results

Figure 4 displays the cohort-specific welfare effects (EV) of extending health care rationing from year 2020 onwards for stabilizing the health expenditure share, compared to the baseline scenario without policy change. For older cohorts (born in the 1960s), the policy reform is almost neutral for welfare. On the one hand, those close to retirement age do not save much health care contributions (that we assumed to be entirely paid by workers) from the reform. On the other hand, the detrimental effects from the reform on longevity and morbidity are small in earlier periods. For later cohorts, however, the welfare change becomes substantial and the higher, the later the year of birth. This is remarkable since younger cohorts save health contributions over a long working period. Those born after the reform year 2020 benefit from reduced contributions even for the entire working life, whereas reductions in survival rates in response to the reform are minor. However, reduced survival rates during retirement and reduced instantaneous utility from higher health deficits by far outweigh the utility increases from higher disposable income.

We estimate that someone who is 20 years old in 2020 would experience a welfare loss of 3 percent when pursuing the goal of stabilizing the health expenditure share. Someone who is 20 years old in 2050 would experience a corresponding welfare loss of 8 percent.

5 Conclusion

In this paper we have taken up the important debate on the sustainability of health care systems that results from medical advances and associated gains on longevity. We have proposed a novel, multi-period overlapping generations model with endogenous medical R&D and endogenous survival that is closely associated with morbidity. Considering different policy regimes, our main focus has been on the trade-off between promoting human longevity by reducing morbidity and preventing increases in health costs. A instance, Hall and Jones (2007) assume that the value of an additional year lived equals the health costs to increase life expectancy by an additional year. This implicitly assumes that the US health system is optimal in the sense of equating marginal benefits and marginal costs of saving lives.
Figure 4: Cohort-specific welfare effects (EV) of extending health care rationing for stabilizing the health expenditure share. Notes: (1) $\beta = 0.98$, $\zeta = 5.1$. Time path $\{\varphi_t\}$ in baseline scenario 0 and reform scenario 1 are displayed in Figure A.1 and Figure A.2, respectively (Online-Appendix). Other parameters as for Figure 1. (2) For instance, the displayed change in welfare evaluated at year 2020 corresponds to the EV of the cost-saving reform for someone who is 20 years old in 2020.
salient feature that enables us to perform such analysis is to capture biologically founded ageing, based on gerontology research, in order to calibrate the model.

Our calibrated model implies substantial future increases in human longevity under the current policy regime that go along with both reductions in morbidity and a rising health expenditure share in GDP. Following instead the goal of stabilizing the health expenditure share by extending health care rationing has sizable effects on morbidity and longevity in the longer run, associated with reduced medical R&D incentives. The implied welfare effects of extending health care rationing may be substantially negative particularly for future generations.

Our results may be understood as warning for policy makers to overlook the detrimental welfare costs from reduced longevity and increased morbidity of health care rationing for the purpose of addressing budgetary problems in health care finance. Our analysis suggests, in fact, that population ageing and rising health costs go hand in hand. In this sense, rising health expenditure shares are not a problem but symptom of a blessing for human health and longevity.

In future research we aim to allow for the possibility of private purchases of health goods and services in a health care system with rationing. Its consideration would naturally refocus the debate on health inequality issues, for instance, when purchases of life-saving drugs may be available only for richer individuals. Such policy regime could give rise to major distributional conflicts that are different to those in societies without significant health care rationing. The associated challenges for modern societies appear profound and discomforting.

Appendix

A. Positive Analysis

- **Dynamical System:** Recall that $V_{t+1}(j)$ is the value of an innovation in health sector $j$ resulting from R&D effort in $t$. A representative R&D firm searching for a
vertical innovation to treat illness $j$ solves

$$\max_{l_t(j)} \{\mu_{t+1}(j) V_{t+1}(j) - w_t l_t(j)\} = \left(\xi_t V_{t+1}(j) - w_t\right) l_t(j),$$  \hspace{1cm} (31)$$

according to (7). Thus, $\xi_t V_{t+1}(j) = w_t$ for all $j$. Thus, in equilibrium, R&D firms do not earn profits. Moreover, $l_t(j) = L_t^Q$ and $V_{t+1}(j) = V_{t+1}$ are the same for all $j \in [0, 1]$. Using $\tilde{\xi}_t = \xi \cdot (L_t^Q)^{-\theta}$, the zero-profit condition for R&D firms reads as

$$V_{t+1} \xi (L_t^Q)^{-\theta} = w_t.$$  \hspace{1cm} (32)$$

Given that there is a unit mass of health sectors, the total and per firm amount of labor allocated to the production of health goods is given by $\chi H_t$. Thus, the profit per health good producer is

$$\pi_t = (p_t - c_t) H_t = (\Gamma - 1) \chi w_t H_t = (\Gamma - 1) w_t L_t^H,$$  \hspace{1cm} (33)$$

according to (12). According to (20),

$$V_t = \pi_t + \frac{1 - \mu_{t+1}}{1 + \bar{r}} \pi_{t+1} + \frac{(1 - \mu_{t+1})(1 - \mu_{t+2})}{(1 + \bar{r})^2} \pi_{t+2} + \frac{(1 - \mu_{t+1})(1 - \mu_{t+2})(1 - \mu_{t+3})}{(1 + \bar{r})^3} \pi_{t+3} + \ldots,$$  \hspace{1cm} (34)$$

$$V_{t+1} = \pi_{t+1} + \frac{1 - \mu_{t+2}}{1 + \bar{r}} \pi_{t+2} + \frac{(1 - \mu_{t+2})(1 - \mu_{t+3})}{(1 + \bar{r})^2} \pi_{t+3} + \ldots = \frac{1 + \bar{r}}{1 - \mu_{t+1}} (V_t - \pi_t).$$  \hspace{1cm} (35)$$

Using (33) in (35), we get the following no-arbitrage condition in the market that finances health R&D:

$$\frac{1 - \mu_{t+1}}{1 + \bar{r}} \frac{V_{t+1}}{V_t} + \frac{(\Gamma - 1) w_t L_t^H}{V_t} = 1.$$  \hspace{1cm} (36)$$

Now let us define $V_t = V_t/A_t$. Denote by $\vartheta_{a,t}$ the health deficit index of a surviving individual of age $a$ in period $t$, define $\bar{a}_t$ as the largest integer in $t$ such that $\vartheta_{a,t} \leq d_{\text{max}}$ and define $\overline{a}_t \equiv \min(\bar{a}_t, T)$ as the age at which an individual dies for sure.
Neglecting the household side (which is relevant for the welfare analysis only), the dynamical system can be summarized as follows:

\[ d_{1,t+1} = [1 + g - (1 - \varphi)\kappa Q_t] d_{\min}, \]  
\[ d_{2,t+1} = [1 + g - (1 - \varphi)\kappa Q_t] d_{1,t}, \]  
\[ d_{3,t+1} = [1 + g - (1 - \varphi)\kappa Q_t] d_{2,t}, \]  
\[ \vdots \]

\[ \mu_{t+1} = \mu^0 Q_t + (1 - \mu^0 Q_t) \cdot \xi \cdot (L_t^Q)^{1-\delta}, \]  
\[ Q_{t+1} - Q_t = \gamma (1 - \mu^0 Q_t) \xi (L_t^\theta)^{1-\delta} - (\delta^Q - \gamma \mu^0) Q_t, \]  
\[ \frac{1 - \mu_{t+1}}{1 + \bar{r}} \nu_{t+1} (1 + g) + (\Gamma - 1) w_t L_t^H = \nu_t, \]  
\[ \nu_{t+1} (1 + g) \xi \cdot (L_t^Q)^{-\delta} = \omega, \]

\[ H_t = (1 - \varphi_t) S_{t,t} d_{\min} + (1 - \varphi_t) (1 - \tilde{m}(d_{\min})) \times \]
\[ \{ S_{t-1,t-1} d_{1,t} + d_{2,t} S_{t-2,t-2} (1 - \tilde{m}(d_{1,t-1})) + \]
\[ d_{3,t} S_{t-3,t-3} (1 - \tilde{m}(d_{2,t-1})) (1 - \tilde{m}(d_{1,t-2})) + \]
\[ \vdots \]
\[ d_{\pi,t} S_{t-\pi,t-\pi} (1 - \tilde{m}(d_{\pi-1,t-1})) (1 - \tilde{m}(d_{\pi-2,t-2})) \times \cdots \times (1 - \tilde{m}(d_{1,t-\pi+1})) \]
\[ L_t^Y + L_t^H + L_t^Q = L_t, \]

according to (15), (9), (11), (36), (32), (17) and (21), respectively, for a given \( Q_0 > 0 \) and a given vector of current deficit states of the cohorts living in period 0, \( d_0 \equiv (d_{1,0}, d_{2,0}, d_{3,0}, \ldots, d_{\pi,0}) \).

- **Long Run Equilibrium:** We next derive the long run equilibrium (focussing on the case where \( Q_{t+1} = Q_t \) holds for \( t \to \infty \) only). Setting \( Q_{t+1} = Q_t = Q \) in (10)
and omitting the time index, we obtain
\begin{equation}
L^Q = \left( \frac{\bar{Q} - \bar{\mu}}{(1 - \bar{\mu})\xi} \right) \frac{1}{1 - \bar{\mu}}. \tag{46}
\end{equation}

Using \( \mathcal{V}_{t+1} = \mathcal{V}_t = \mathcal{V} \) in (42) implies
\begin{equation}
\mathcal{V} = \frac{(\Gamma - 1)(1 + \bar{r})\omega L_t^H}{\bar{r} - g + \mu (1 + g)}. \tag{47}
\end{equation}

Moreover, according to (43) and (46),
\begin{equation}
\mathcal{V} = \frac{\omega (L^Q)^\alpha}{(1 + g)\xi}. \tag{48}
\end{equation}

Combining (47) and (48) implies
\begin{equation}
\frac{(L^Q)^\alpha}{\xi} = \frac{(\Gamma - 1)(1 + \bar{r})L_t^H}{\bar{r} - g + \mu}. \tag{49}
\end{equation}

Let \( \delta_a \) denote the long run health deficit index of someone of age \( a \geq 0 \), associated with the steady state quality index \( \bar{Q} < \bar{Q} \). Moreover, let \( \hat{\varphi} \equiv \lim_{t \to \infty} \varphi_t \) and suppose that \( \hat{S} \equiv \lim_{v \to \infty} S_{v,v} = 1 \). According to (15),
\begin{equation}
\delta_{a+1} = \left[ 1 + g - (1 - \hat{\varphi})\hat{Q} \right] \delta_a, \tag{50}
\end{equation}

with initial condition \( \delta_0 = d_{\text{min}} > 0 \). The solution of difference equation (50) gives us the steady state age-path of the health deficit index conditional on \( \varphi \) and \( \hat{Q} \), denoted by \( \mathcal{D}(a, \varphi, \hat{Q}) \), \( a \geq 0 \). Function \( \mathcal{D}(a, \varphi, Q) \) is increasing in age, \( a \), increasing in \( \varphi \), and decreasing in quality index, \( Q \). Let \( a_\infty \) denote the largest age \( a \) such that \( \mathcal{D}(a, \varphi, Q) \leq d_{\text{max}} \) and define function
\begin{equation}
\hat{H}(\varphi, Q) \equiv (1 - \varphi) \left[ \sum_{n=0}^{\infty} \mathcal{D}(a, \varphi, Q) \prod_{n=0}^{a} \left[ 1 - \tilde{m}(\mathcal{D}(u, \varphi, Q)) \right] \right], \tag{51}
\end{equation}

34
where \( \bar{a}_\infty \equiv \min(\bar{a}_\infty, T) \). Substituting (46) into (40), we have

\[
\mu = \frac{\delta^Q}{\gamma} Q. \tag{52}
\]

Substituting (52) into (49) and using \( \bar{\mu} = \mu^0 Q, L^H = \chi \tilde{H}(\varphi, Q) \) and (46) we obtain

\[
\frac{\delta^Q}{\gamma} Q = \left( \frac{\frac{1}{Q} - \mu^0}{\frac{\delta^Q}{\gamma} - \mu^0} \right)^{\frac{\sigma}{1 - \sigma}} \xi^{\frac{1}{1 - \sigma}} (\Gamma - 1)(1 + \bar{r}) \chi \tilde{H}(\varphi, Q) - \frac{\bar{r} - g}{1 + g}, \tag{53}
\]

which implicitly defines \( \hat{Q} \). We see that \( \hat{Q} \) is unique when \( \tilde{H}(\varphi, Q) \) is non-increasing in \( Q \). The other long run values follow.

- **Calibration:** A steady state analysis is instructive to understand the relationship between endogenous observables and helps us to calibrate the model.

1. First, setting \( Q_{t+1} = Q_t = Q \) in (41) and using both \( \bar{\mu} = \mu^0 Q \) and \( \bar{\mu} = \xi (L^Q)^{1-\sigma} \), we obtain

\[
[\bar{\mu} + (1 - \bar{\mu}) \bar{\mu}] = \mu = \frac{\delta^Q}{\gamma \mu^0} \bar{\mu} \tag{54}
\]

(recall that \( \mu^0 < \frac{\delta^Q}{\gamma} \)). Thus, in the long run, the total innovation probability \( \mu \) is proportional to \( \bar{\mu} \).

2. Second, according to (40),

\[
\frac{(L^Q)^{\sigma}}{\xi} = \frac{(1 - \bar{\mu})L^Q}{\mu - \bar{\mu}}. \tag{55}
\]

Combining (55) with (49) and using both \( L^H = \chi H \) and (54) implies that

\[
\ell^Q = \frac{\frac{\delta^Q}{\gamma \mu^0} - 1}{\frac{\bar{\mu} - 1}{1 + g} + \frac{\delta^Q}{\gamma \mu^0}} (\Gamma - 1)(1 + \bar{r})L^H \tag{56}
\]

holds in the long run.

3. Third, according to (19), the long run health expenditure share can be written
as
\[ s = \frac{pH}{Y + pH} = \frac{1}{V + 1} = \frac{1}{\frac{LY}{\omega L}} + 1 = \frac{1}{\Gamma(1 - \alpha)L + 1}. \]  
(57)

where we used (3), \( \omega = (1 - \alpha)(AL^Y/K^Y)^{-\alpha} \) and (12).

4. Fourth, according to (1), (2) and (15), the time paths of the health care wedge, \( \{\varphi_t\}_{t=0}^{\infty} \), and initial cohort sizes, \( \{S_{v,t}\}_{v=0}^{\infty} \), drive (along with the endogenous time path of the quality of health goods, \( \{Q_t\}_{v=0}^{\infty} \)) the evolution of survival functions over time.

B. Normative Analysis: Consumption Paths

- **Anticipated Health Policy:** Let us start with the case without unanticipated policy shocks. Using \( S_{v,t} = S_{v,v} \prod_{u=v}^{t-1} (1 - m_{v,u}) \) in (22), the Lagrangian \( \mathcal{L}_v \) associated with maximizing \( U_v \) subject to (25) and \( a_{v,v+T} \geq 0 \) is

\[
\mathcal{L}_v = \ldots + \beta^{-v} S_{v,v} \prod_{u=v}^{t-1} (1 - m_{v,u}) \frac{\log c_{v,t}}{(1 + d_{v,t})^\xi} + \\
\beta^{t+1-v} S_{v,v} \prod_{u=v}^{t} (1 - m_{v,u}) \frac{\log c_{v,t+1}}{(1 + d_{v,t+1})^\xi} + \ldots +
\lambda_{v,t} [(1 - \tau_{t})w_t + (1 + r_{v,t})a_{v,t} - c_{v,t} - a_{v,t+1}] +
\lambda_{v,t+1} [(1 - \tau_{t+1})w_{t+1} + (1 + r_{v,t+1})a_{v,t+1} - c_{v,t+1} - a_{v,t+2}] + \ldots
\]  
(58)

where \( \lambda_{v,t}, \lambda_{v,t+1}, \text{etc.} \) denote the multipliers for period \( t, t+1, \text{etc.} \). The first-order conditions \( \partial \mathcal{L}_v / \partial c_{v,t} = \partial \mathcal{L}_v / \partial c_{v,t+1} = \partial \mathcal{L}_v / \partial a_{v,t+1} = 0 \) can be written as

\[
\frac{\beta^{-v} S_{v,v} \prod_{u=v}^{t-1} (1 - m_{v,u})}{(1 + d_{v,t})^\xi c_{v,t}} = \lambda_{v,t},
\]  
(59)

\[
\frac{\beta^{t+1-v} S_{v,v} \prod_{u=v}^{t} (1 - m_{v,u})}{(1 + d_{v,t+1})^\xi c_{v,t+1}} = \lambda_{v,t+1},
\]  
(60)

\[
\lambda_{v,t} = \lambda_{v,t+1}(1 + r_{v,t+1}).
\]  
(61)

Combining (59)-(61) leads to

\[
\frac{(1 + d_{v,t+1})^\xi c_{v,t+1}}{(1 + d_{v,t})^\xi c_{v,t}} = \beta (1 - m_{v,t})(1 + r_{v,t+1}).
\]  
(62)
Using (26) in (62) implies
\[ c_{v,t+1} = \left( \frac{1 + d_{v,t}}{1 + d_{v,t+1}} \right)^\varsigma \beta (1 + \bar{r}) c_{v,t}. \]  

(63)

Iterating and using \( d_{v,v} = d_{\text{min}} \), we obtain
\[ c_{v,t} = \left( \frac{1 + d_{\text{min}}}{1 + d_{v,t}} \right)^\varsigma \beta^{t-v} (1 + \bar{r})^{t-v} c_{v,v}. \]  

(64)

From (25), (26), \( a_{v,v} = 0 \) and \( a_{v,v+T} = 0 \) (reflecting that it is optimal not to hold wealth after certain death), we find that the intertemporal budget constraint of a member of cohort \( v \) is given by
\[ c_{v} + \sum_{t=v+1}^{v+T-1} \left( \frac{c_{v,t}}{\prod_{u=v+1}^{t}(1 + r_{v,u})} \right) = (1 - \tau_v) w_v + \sum_{t=v+1}^{v+R-1} \left( \frac{(1 - \tau_t) w_t}{\prod_{u=v+1}^{t}(1 + r_{v,u})} \right). \]  

(65)

Using (26) and (64), we obtain for the left-hand side of (65) that
\[ c_{v,v} + \sum_{t=v+1}^{v+T-1} \left( \frac{c_{v,t}}{\prod_{u=v+1}^{t}(1 + r_{v,u})} \right) = c_{v,v} \left( 1 + \sum_{t=v+1}^{v+T-1} \beta^{t-v} \left( \frac{1 + d_{\text{min}}}{1 + d_{v,t}} \right)^\varsigma \prod_{u=v}^{t-1} (1 - m_{v,u}) \right). \]  

(66)

Equating the right-hand sides of (65) and (66), and using (26), \( w_t = \omega A_t \) with \( \omega \) given by (4), (2) with \( S_{v,v} = 1 \) and \( A_t = A_v (1 + g)^{t-v} \), implies that the initial consumption level, \( c_{v,v} \), is given by
\[ c_{v,v} = \omega A_v \left( 1 - \tau_v + \sum_{t=v+1}^{v+R-1} (1 - \tau_t) \frac{(1 + g)^{t-v} S_{v,t}}{S_{v,v}} \right) \frac{1}{1 + \sum_{t=v+1}^{v+T-1} \beta^{t-v} \left( \frac{1 + d_{\text{min}}}{1 + d_{v,t}} \right)^\varsigma \frac{S_{v,t}}{S_{v,v}}}. \]  

(67)

**Unanticipated Health Policy Shock:** We now turn to the case where some individuals experience an unanticipated policy shock in period \( t_0 \). That is, for \( t < t_0 \) they follow the same consumption path as computed in the previous case and then they reoptimize in \( t_0 \). According to (63), knowing \( c_{v,t_0} \), the path of consumption of
any living member of generation \( v \) for future dates for \( t \geq t_0 \) evolves as

\[
c_{v,t} = \left( \frac{1 + d_{v,t}}{1 + d_{v,t}} \right)^\zeta \beta^{t-t_0} (1 + \bar{r})^{t-t_0} c_{v,t_0}.
\] (68)

Using (25) and \( a_{v,v} = 0 \), for \( t_0 < v + R \) we have

\[
a_{v,t_0} \prod_{u=v+1}^{t_0} (1 + r_{v,u}) = (1 - \tau_v) w_v - c_{v,v} + \sum_{t=v+1}^{t_0-1} \prod_{u=v+1}^{t} (1 + r_{v,u}) (1 - \tau_t) w_t - c_{v,t}
\] (69)

Using (26), (64) and (2), we obtain

\[
c_{v,v} + \sum_{t=v+1}^{t_0-1} \left( \frac{c_{v,t}}{\prod_{u=v+1}^{t} (1 + r_{v,u})} \right) = c_{v,v} \left( 1 + \sum_{t=v+1}^{t_0-1} \beta^{t-v} \frac{1 + d_{\min}}{1 + d_{v,t}} \frac{S_{v,t}}{S_{v,v}} \right).
\] (70)

Using (26) and (2), we also get

\[
\frac{1}{\prod_{u=v+1}^{t_0-1} (1 + r_{v,u})} = \frac{S_{v,t_0-1}}{S_{v,v}(1 + \bar{r})^{t_0-1-v}}.
\] (71)

Substituting (70), (71), \( w_t = \omega A_t \) and \( A_t = A_v (1 + g)^{t-v} \) into (69), the wealth holding of a member of generation \( v \) in \( t_0 < v + R \) is given by

\[
a_{v,t_0} = (1 + \bar{r})^{t_0-v-1} \frac{S_{v,v}}{S_{v,t_0-1}} \left[ A_{v,\omega} \left( 1 + \sum_{t=v+1}^{t_0-1} (1 - \tau_t) \left( 1 + g \frac{1 + \bar{r}}{1 + \bar{r}} \right)^{t-v} \frac{S_{v,t}}{S_{v,v}} \right) - c_{v,v} \sum_{t=v+1}^{t_0-1} \left( 1 + d_{\min} \right)^\zeta \beta^{t-v} \frac{S_{v,t}}{S_{v,v}} - c_{v,v} \right].
\] (72)

Analogously, for \( t_0 \geq v + R \), we have

\[
a_{v,t_0} = (1 + \bar{r})^{t_0-v-1} \frac{S_{v,v}}{S_{v,t_0-1}} \left[ A_{v,\omega} \left( 1 + \sum_{t=v+1}^{v+R-1} (1 - \tau_t) \left( 1 + g \frac{1 + \bar{r}}{1 + \bar{r}} \right)^{t-v} \frac{S_{v,t}}{S_{v,v}} \right) - c_{v,v} \sum_{t=v+1}^{t_0-1} \left( 1 + d_{\min} \right)^\zeta \beta^{t-v} \frac{S_{v,t}}{S_{v,v}} - c_{v,v} \right].
\] (73)

Recall that \( c_{v,v} \) is the initial consumption level chosen before the unanticipated
shock occurs. Next, use (25) and \(a_{v,v+T} = 0\) to obtain

\[
c_{v,t_0} + \sum_{t=t_0+1}^{v+T-1} \frac{c_{v,t}}{\prod_{u=t_0+1}^{t} (1 + r_{v,u})} = (1 + r_{v,t_0})a_{v,t_0} + (1 - \tau_{t_0}) w_{t_0} + \sum_{t=t_0+1}^{v+R-1} \frac{(1 - \tau_t) w_t}{\prod_{u=t_0+1}^{t} (1 + r_{v,u})},
\]

Using (68) implies

\[
c_{v,t_0} + \sum_{t=t_0+1}^{v+T-1} \frac{c_{v,t}}{\prod_{u=t_0+1}^{t} (1 + r_{v,u})} = c_{v,t_0} \left( 1 + \sum_{t=t_0+1}^{v+T-1} \left( \frac{1 + d_{v,t_0}}{1 + d_{v,t}} \right) \beta^{t-t_0} \prod_{u=t_0}^{t-1} (1 - m_{v,u}) \right) \quad (74)
\]

Equating the right-hand sides of (74) and (75) and using (26), \(w_t = \omega A_t\) and \(A_t = A_{t_0} (1 + g)^{t-t_0}\) implies, for \(t_0 < v + R\), the consumption level:

\[
c_{v,t_0} = \frac{1 + \frac{1}{1 - m_{v,t_0}} a_{v,t_0}}{1 + \sum_{t=t_0+1}^{v+T-1} \left( \frac{1 + d_{v,t_0}}{1 + d_{v,t}} \right) \beta^{t-t_0} \prod_{u=t_0}^{t-1} (1 - m_{v,u})} \left( 1 - \tau_{t_0} + \sum_{t=t_0+1}^{v+R-1} (1 - \tau_t) \left( \frac{1 + g}{1 + g} \right)^{t-t_0} \prod_{u=t_0}^{t-1} (1 - m_{v,u}) \right)
\]

with \(a_{v,t_0}\) given by (72) and \(A_{t_0} = A_v (1 + g)^{t_0-v}\). Analogously, for \(t_0 \geq v + R\) (i.e. the individual is retired when the shock hits), we have

\[
c_{v,t_0} = \frac{1 + \frac{1}{1 - m_{v,t_0}} a_{v,t_0}}{1 + \sum_{t=t_0+1}^{v+T-1} \left( \frac{1 + d_{v,t_0}}{1 + d_{v,t}} \right) \beta^{t-t_0} \prod_{u=t_0}^{t-1} (1 - m_{v,u})} \left( 1 - \tau_{t_0} + \sum_{t=t_0+1}^{v+R-1} (1 - \tau_t) \left( \frac{1 + g}{1 + g} \right)^{t-t_0} \prod_{u=t_0}^{t-1} (1 - m_{v,u}) \right)
\]

with \(a_{v,t_0}\) given by (73).

References


Online-Appendix:

Additional Figures and Tables

Figure A.1. Calibration of the time paths of the health care wedge and initial cohort sizes in the baseline scenario.

Figure A.2. Calibration of the time paths of the health care wedge in the reform scenario.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>51.8</td>
<td>52.8</td>
<td>53.8</td>
<td>54.1</td>
<td>56.7</td>
<td>56.4</td>
<td>61.0</td>
<td>59.8</td>
</tr>
<tr>
<td>25</td>
<td>47.0</td>
<td>47.9</td>
<td>49.0</td>
<td>49.2</td>
<td>51.9</td>
<td>51.5</td>
<td>56.1</td>
<td>54.9</td>
</tr>
<tr>
<td>30</td>
<td>42.4</td>
<td>43.1</td>
<td>44.2</td>
<td>44.4</td>
<td>47.1</td>
<td>46.6</td>
<td>51.2</td>
<td>50.0</td>
</tr>
<tr>
<td>35</td>
<td>37.8</td>
<td>38.3</td>
<td>39.4</td>
<td>39.5</td>
<td>42.3</td>
<td>41.8</td>
<td>46.4</td>
<td>45.2</td>
</tr>
<tr>
<td>40</td>
<td>33.2</td>
<td>33.6</td>
<td>34.7</td>
<td>34.8</td>
<td>37.5</td>
<td>37.1</td>
<td>41.6</td>
<td>40.4</td>
</tr>
<tr>
<td>45</td>
<td>28.7</td>
<td>29.0</td>
<td>30.1</td>
<td>30.2</td>
<td>32.8</td>
<td>32.4</td>
<td>36.9</td>
<td>35.7</td>
</tr>
<tr>
<td>50</td>
<td>24.4</td>
<td>24.6</td>
<td>25.7</td>
<td>25.8</td>
<td>28.2</td>
<td>27.9</td>
<td>32.3</td>
<td>31.2</td>
</tr>
<tr>
<td>55</td>
<td>20.3</td>
<td>20.4</td>
<td>21.6</td>
<td>21.5</td>
<td>23.9</td>
<td>23.6</td>
<td>27.8</td>
<td>26.8</td>
</tr>
<tr>
<td>60</td>
<td>16.6</td>
<td>16.5</td>
<td>17.7</td>
<td>17.5</td>
<td>19.8</td>
<td>19.5</td>
<td>23.5</td>
<td>22.6</td>
</tr>
<tr>
<td>65</td>
<td>13.2</td>
<td>12.9</td>
<td>14.3</td>
<td>13.9</td>
<td>16.1</td>
<td>15.7</td>
<td>19.4</td>
<td>18.6</td>
</tr>
<tr>
<td>70</td>
<td>10.2</td>
<td>9.8</td>
<td>11.3</td>
<td>10.7</td>
<td>12.9</td>
<td>12.3</td>
<td>15.5</td>
<td>14.9</td>
</tr>
<tr>
<td>75</td>
<td>7.6</td>
<td>7.1</td>
<td>8.7</td>
<td>7.9</td>
<td>10.0</td>
<td>9.3</td>
<td>12.0</td>
<td>11.7</td>
</tr>
<tr>
<td>80</td>
<td>5.6</td>
<td>5.0</td>
<td>6.5</td>
<td>5.6</td>
<td>7.5</td>
<td>6.8</td>
<td>8.9</td>
<td>8.8</td>
</tr>
<tr>
<td>85</td>
<td>4.1</td>
<td>3.3</td>
<td>4.8</td>
<td>3.8</td>
<td>5.5</td>
<td>4.7</td>
<td>6.4</td>
<td>6.4</td>
</tr>
<tr>
<td>90</td>
<td>3.0</td>
<td>2.1</td>
<td>3.5</td>
<td>2.4</td>
<td>4.0</td>
<td>3.1</td>
<td>4.4</td>
<td>4.4</td>
</tr>
<tr>
<td>95</td>
<td>2.2</td>
<td>1.2</td>
<td>2.6</td>
<td>1.5</td>
<td>2.9</td>
<td>2.0</td>
<td>3.1</td>
<td>2.9</td>
</tr>
<tr>
<td>100</td>
<td>1.8</td>
<td>0.6</td>
<td>2.0</td>
<td>0.8</td>
<td>2.1</td>
<td>1.1</td>
<td>2.2</td>
<td>1.8</td>
</tr>
<tr>
<td>Age</td>
<td>2020 Baseline</td>
<td>2020 Baseline</td>
<td>2050 Experiment</td>
<td>2080 Baseline</td>
<td>2080 Experiment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>--------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>-----------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>62.4</td>
<td>71.1</td>
<td>70.3</td>
<td>80.6</td>
<td>75.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>57.5</td>
<td>66.2</td>
<td>65.4</td>
<td>75.7</td>
<td>70.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>52.6</td>
<td>61.4</td>
<td>60.6</td>
<td>70.9</td>
<td>66.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>47.8</td>
<td>56.6</td>
<td>55.7</td>
<td>66.0</td>
<td>61.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>43.0</td>
<td>51.8</td>
<td>51.0</td>
<td>61.3</td>
<td>56.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>38.3</td>
<td>47.1</td>
<td>46.3</td>
<td>56.5</td>
<td>51.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>33.7</td>
<td>42.4</td>
<td>41.6</td>
<td>51.8</td>
<td>47.2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>29.2</td>
<td>37.8</td>
<td>37.1</td>
<td>47.2</td>
<td>42.7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>25.0</td>
<td>33.3</td>
<td>32.6</td>
<td>42.7</td>
<td>38.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>20.9</td>
<td>29.0</td>
<td>28.3</td>
<td>38.2</td>
<td>34.0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>17.1</td>
<td>24.8</td>
<td>24.2</td>
<td>33.9</td>
<td>29.8</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>13.6</td>
<td>20.8</td>
<td>20.2</td>
<td>29.7</td>
<td>25.9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>10.5</td>
<td>17.2</td>
<td>16.6</td>
<td>25.7</td>
<td>22.1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>7.8</td>
<td>13.8</td>
<td>13.3</td>
<td>21.8</td>
<td>18.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>5.6</td>
<td>10.8</td>
<td>10.3</td>
<td>18.2</td>
<td>15.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>3.8</td>
<td>8.2</td>
<td>7.8</td>
<td>14.8</td>
<td>12.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>2.5</td>
<td>6.0</td>
<td>5.7</td>
<td>11.7</td>
<td>9.6</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table A.2.** Implied remaining life expectancies according to age: baseline vs. reform scenario for years 2020, 2050, 2080.