The Medical Expansion, Life-Expectancy, and Endogenous Directed Technical Change*

Leon Huetsch[†] Dirk Krueger[‡] Alexander Ludwig[§]

March 20, 2025

Abstract

We build a unified quantitative theory of income growth and increasing life expectancy in the last two centuries, and the emergence and expansion of a modern health sector in the 20th century. We develop a two-sector overlapping generations model with endogenous and directed technical change in which income growth, life expectancy, the size of and technological progress in the health and the final goods sector are jointly determined in general equilibrium. The model interprets the historical record as three phases of a dynamic equilibrium in which households are initially poor and the quality-adjusted price of health goods is prohibitively high so that demand for them is non-existent, life is short and life expectancy stagnant. As technological progress starts to fuel income growth, households commence consuming basic health goods (such as a better diet and basic domestic hygiene), remaining life expectancy at age 20 starts to rise. A century later, directed technological progress eventually leads to the emergence (in ca. 1940) and then expansion of a modern health sector. In counterfactual analyses the model reveals that almost 50% of the life-expectancy gains between 1940 and 2020 are attributed to increased spending in the modern health sector. Furthermore, in the absence of public spending on health—in the form of subsidies to health R&D during WW II and subsequently, by the emergence of Medicare—, the kickoff of the modern health sector would have been delayed by roughly two decades in the U.S.

JEL Codes: E13, O41, I15

Keywords: Life Expectancy, Modern Health Production, Endogenous Technological Progress, Transition

*We thank Michael Haines, our discussant Minchul Yum, as well as seminar participants at various conferences and universities for helpful comments, and the National Science Foundation (under grant SES-S-175708) for financial support.

[†]University of Bonn

[‡]University of Pennsylvania, CEPR and NBER

[§]European University Institute, Goethe University Frankfurt, and CEPR.

1 Introduction

In 1820 remaining life expectancy of a twenty year old person living in the United States was approximately 40 years, real (in 2011 dollars) per capita income was $\$2,674$ and virtually none of that income was spent on health goods and services, abstracting from expenditures on basic food and hygiene.[1] In 2019, the year prior to the COVID-19 pandemic, remaining life expectancy at age twenty stood above 60 years, real income per capita had risen to $\$55,355$, and close to 20% of that income was spent on goods produced by a modern, high-tech health sector.[2] In this paper, we build a quantitative theory to explain these observations, and use the theory to investigate what role government health (care) policies have played in this transition from a life that was poor and short and without much medical care to one with high incomes, long lives and high-tech, expensive health care.

To do so we develop a two-sector overlapping generations model with endogenous and directed technical change in which income growth, life expectancy, technological progress in the health and the final goods sector, as well as the size of the health sector and the quality and price of the goods it produces are jointly determined in general equilibrium. The model interprets the facts as three phases of a dynamic equilibrium in which households are initially poor and the price of health goods is prohibitively high so that demand for health goods is zero, life is short and life expectancy stagnant. As income grows, fueled by technological progress, households start consuming basic health goods, life expectancy starts to rise, and directed technological progress eventually, with a delay of ca. 100 years, leads to the emergence and expansion of a modern health sector that commands a significant and growing share of labor demand, production and household expenditures.

We calibrate the model parameters (which includes the initial conditions for the state of technology in both sectors) to observations on income per capita and life expectancy in 1820 as well as to the timing of increase in life expectancy and the emergence of the modern health sector, and then use the model as a quantitative laboratory to answer two applied questions. After having verified that the model accounts well for the time series in per capita income, life expectancy and the size of the modern health sector, we ask what share of the increase in life expectancy from 1940 onward can be attributed to the modern health sector. In 2020, we find that share to be approximately one third. This suggests that both the modern health sector, but even more so income-growth induced increases in spending on basic health (e.g., better food and hygiene) both played a major role in the expansion of life expectancy, even into the 21th century. Second, we use the model to decompose the increase in the relative price of health goods and services

[1] Unless otherwise noted, we use *cohort life expectancy* as the relevant measure of life expectancy.

[2] Real income per capita is in constant 2011 US dollars, and is reported by the Maddison data project, Bolt and van Zanden (2024).

between 1940 and 2020 into two components: a term driven by the income-growth-induced rising household demand for health goods relative to final goods; and a term due to productivity growth in the modern health sector relative to the final goods sector, driven by endogenous technological progress. We show that between 1940 and 1980, both components contribute roughly half of the overall increase in the relative health price, but after 1980, technological progress in the modern health sector accelerates and becomes the dominant force.

Related Literature. Our model has three key building blocks, a two-period overlapping generations structure with production akin to Diamond (1965), endogenous investments into health and longevity by private households, as in Grossman (1972) and the endogenous evolution of technological change in the Schumpeterian growth tradition (see, e.g., Aghion and Howitt (1992) or Aghion and Howitt (1998) whose speed differs across sectors in the economy, akin to Acemoglu and Guerrieri (2008). It seeks to describe the path of economic and health stagnation, take-off during a transition period and, eventually, balanced growth as one dynamic equilibrium, as in the general literature on unified growth theory (see, e.g., Galor (2011) or Hansen and Prescott (2002)).

In trying to explain long run trends in life expectancy and connect it to technological progress our paper builds on the work by Cervellati and Sunde (2005) who develop a model of the take-off of life expectancy by modeling the feedback loop between income growth, human capital formation, increases in life-expectancy and the size of the population. In contrast to them, we seek to provide a unified theory not only of the take-off in life expectancy in the 19th century, but also the emergence of the modern health sector. That purpose is shared with Hejkal, Ravikumar, and Vandenbroucke (2022) but their focus is on explaining cross-country differences (and similarities) in the reduction of mortality as well as the evolution of the world population.

More broadly, in terms of model-building this paper contributes to the literature on health spending, R&D in the health sector and endogenous growth. Borger, Rutherford, and Won (2008) develop a model with endogenous technology adaptation in the health sector to predict future health spending shares and conclude that health spending will slow down. Ehrlich and Yin (2013) construct an endogenous growth model where human capital is the engine of growth. Both these elements are encompassed in the work of Kuhn and Prettner (2016) who model a final goods sector with an intermediate R&D sector and a labor-intensive health sector. They argue that an expansion of this sector may reduce growth by shifting resources away from R&D spending. Like us, Frankovic and Kuhn (2023) develop an overlapping generations model with endogenous health and two production sectors to evaluate the quantitative impact of the introduction of health insurance through Medicare and Medicaid on health spending trends, macroeconomic performance, and trends in life expectancy since 1960 when Medicaid was introduced. In their model, the growth rate of the final goods sector is exogenous and endogenous in the health sector

(as in the work by Böhm et al. (2018) who model the evolution of individual health through an accumulation of health deficits). In contrast, we model endogenous growth symmetrically in both sectors, but permit it to be unbalanced.[3]

Finally, a vibrant positive literature studies potential reasons behind the increase in the health expenditure share in U.S. postwar data and a normative literature explores what share of economic activity should be dedicated to health goods and services. The normative perspective includes the work by Hall and Jones (2007), Jones (2004) and Jones (2016). Hall and Jones (2007) model health spending as a superior good with an income elasticity larger than one. As a consequence, income growth leads to an expansion of the health expenditure share, as in our paper. We extend this framework to a two-sector model with a symmetric treatment of endogenous growth in both sectors, and where unbalanced growth emerges as an equilibrium outcome. In contrast to their paper our main purpose is positive, seeking to understand the expansion of the modern medical sector. Jones (2004) develops an endogenous growth model with R&D to explain the increasing health expenditure share. Our model shares the same broad narrative, but provides an explicit treatment of production in a two-sector model so that income growth spurs quality improvements and thus technological progress in both sectors. Jones (2016) also considers growth in two sectors by studying the optimal rate of consumption growth relative to growth of life-saving technologies.

On the positive side, Anderson, Reinhardt, and Hussey (2003) argue that the increase of the health spending share is primarily due to the relative price increase of medical goods,[4] which could either be due to an increase in market power of the supply side relative to the demand side in the market for health goods, as Anderson et al. (2003) suggest,[5] or perhaps due to imperfectly measured quality improvements as our model implies.[6] The quality improvements in turn are the result of costly technological progress. The Congressional Budget Office (2008) presents a review of the literature arguing that technological progress contributes to 40-60 percent of the growth in real per capita health care spending, with demographic change towards an older population

[3] An important mechanism in these models is typically a market size effect which triggers innovation spending. Empirical support for this mechanism in the health sector is provided by Acemoglu and Linn (2004).

[4] Fonseca, Langot, Michaud, and Sopraseuth (2023) estimate US health prices to be 33% higher than those of European countries, which explains 60% of differences in health expenditures.

[5] As suggested by Retzlaff-Roberts, Chang, and Rubin (2004), the U.S. health care system is plagued by larger inefficiencies than the systems in other industrialized countries. This is a central theme in a growing literature attributing increases in health spending in the US and cross-country differences across the OECD to health prices and inefficiencies in health care markets, see, e.g. Cooper, Craig, Gaynor, and Reenen (2019), Horenstein and Santos (2019), and Feldman and Pretnar (2023). Our monopolistically competitive pricing structure in the health sector may reflect such inefficiencies. Higher prices, and the associated higher average asset returns could also be a reflection of compensation for medical innovation risk, see e.g., Koijen, Philipson, and Uhlig (2016).

[6] In this regard, our paper relates to the literature on a lack of quality adjustments of health goods, e.g., Graboyes (1994), Lawver (2011) and Berndt, Cutler, Frank, Griliches, Newhouse, and Triplett (2000), who find that health price indices profoundly underestimate the quality improvements documented in, e.g., Cutler and McClellan (2001).

being another important contributor. Zhao (2014) argues that an introduction and expansion of social security added to increase of health expenditures. See also Fonseca, Michaud, Galama, and Kapteyn (2021) in this regard.[7]

The remainder of this paper proceeds as follows. Section 2 presents stylized facts on life-expectancy, aggregate health spending and prices of health goods used to motivate, calibrate and evaluate the model. Section 3 lays out the model and defines equilibrium. Section 4 contains a theoretical characterization of parts of the equilibrium and Section 5 presents the calibration of the model. Section 6 contains the quantitative evaluation of the model, both along dimensions targeted in the calibration and validates the model along dimensions not targeted in the calibration. Section 7 concludes the paper. Details of the construction of the empirical facts, the computational procedure, the calibration of the model as well as on technical-theoretical derivations are contained in Appendices A, and B, respectively.

2 Stylized Facts

In this section we document the main facts that motivate our model, starting with the timing of the increase in life expectancy in Subsection 2.1, then briefly reviewing the time path of income per capita, starting from the industrial revolution in Subsection 2.2, and finally documenting the emergence and evolution of the Modern Health Sector in Subsection 2.3.

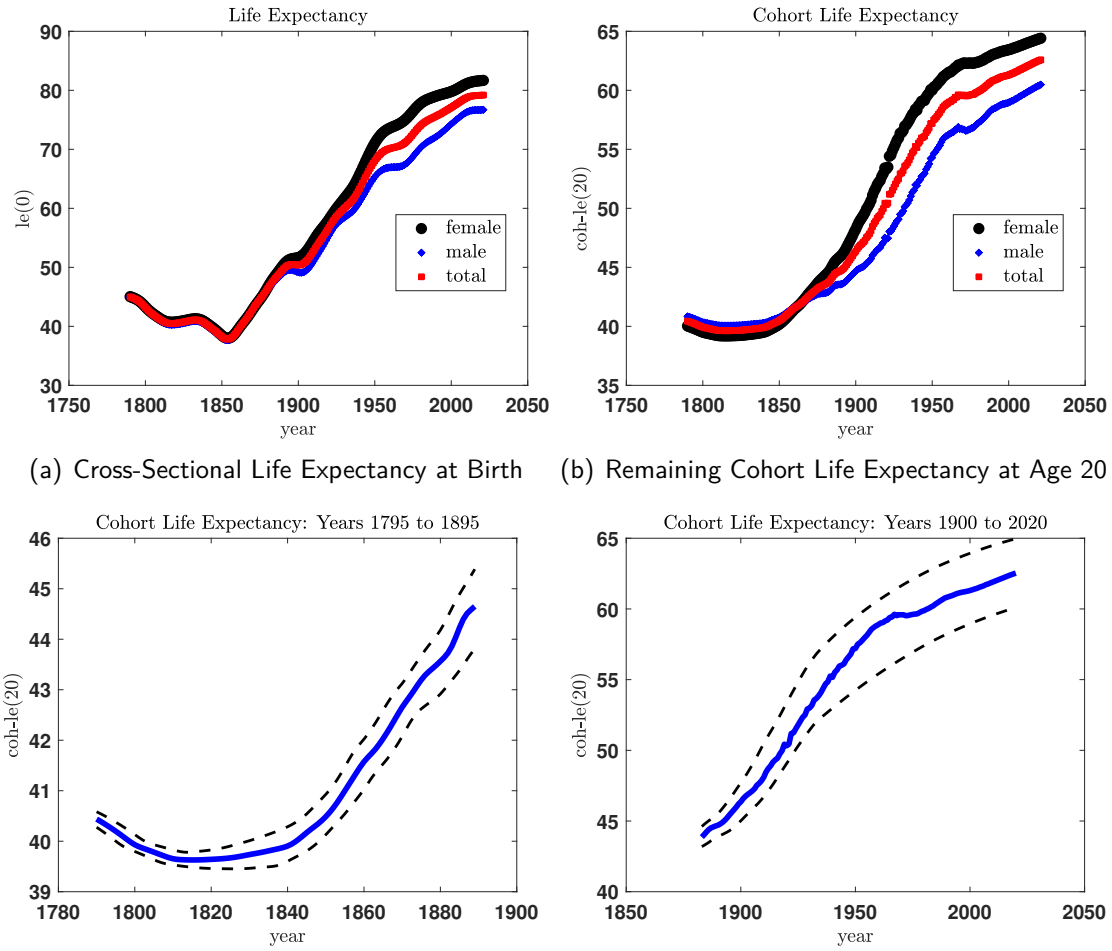
2.1 Life Expectancy

When documenting life expectancy for the last two centuries we face two crucial choices. First, in the early period a significant increase of life expectancy is due to a decline in child mortality, with later improvements mainly accruing to increased life expectancy conditional on survival to adulthood. Since our model focuses on the the second part, the improved longevity of adults, so will our discussion of the data. Second, life expectancy at a given point in time can be measured using purely cross-sectional survival rates or employ cohort survival rates. The first, cross-sectional concept only requires age-specific survival rates at a given point in time, but assumes that a current 20 year old adult will have the same survival rate at age 50 (that is, 30 years into the future) as a current 50 year old individual, thereby ignoring potential technological improvements in the health sector. Cohort life expectancy is more data-demanding since it requires future age-specific survival rates of the cohort under consideration. Since it fully captures the impact of medical innovations, a key aspect of our model, we prefer this concept for the purpose

[7] A parallel literature studies the sources of level differences in health spending shares across countries and attributes the higher share in the U.S. to the fact that the U.S. is the leading country for the invention of new costly health products, see Chandra and Skinner (2012).

of this paper. In conclusion, although we present various time series of life expectancy here, our main focus in the quantitative analysis will be on remaining cohort life expectancy at age 20.

Figure 1: Life-Expectancy in the US



(a) Cross-Sectional Life Expectancy at Birth

(b) Remaining Cohort Life Expectancy at Age 20

(c) Remaining Cohort Life Expectancy at Age 20: until 1900

(d) Remaining Cohort Life Expectancy at Age 20: after 1900

Notes: Cross-sectional life-expectancy in panel (a), cohort life expectancy in panels (b)–(e). 95% Bootstrapped confidence intervals are shown as dashed lines in panels (c)–(e). *Sources*: Hacker (2010), Human Life-Table Database, Human Mortality Database, own computations, see Appendix A for details.

Figure 1 shows in panel (a) life expectancy at birth according to the cross-sectional concept.[8] Moving from there to panel (b) of the figure we take three transformations of the data. First, we apply a Hodrick-Prescott filter on the age and time specific mortality rates to extract the age-specific trend components, second, we compute cohort life expectancy instead of cross-sectional life-expectancy and, third, we look at remaining (cohort) life expectancy at age 20 of

[8] A detailed description of data sources and our methods used to calculate life expectancy—according to both concepts—is contained in Appendix A.

a person in a given year. From this picture we make the following observations. First, before about 1840, remaining cohort life-expectancy in the US was basically flat and the average—taken for the years 1790 to 1840—was about 39.83 years. Second, since then life expectancy has been increasing so that now—i.e., in year 2020—remaining cohort life expectancy at age 20 stands at 66.22 years. Third, due to death related to pregnancy and child birth, the remaining cohort life expectancy of women was lower than that of men until 1864, when the familiar positive life expectancy gap between female and male life expectancy starts to emerge.[9] In panels (c) and (d) we focus on average life expectancy in the population—averaged across men and women—and we zoom in on the subperiods before and after 1900 and add bootstrapped confidence intervals. Acknowledging the uncertainty about the very precise timing of the take-off in adult life expectancy—which we illustrate here by displaying the bootstrapped 95% confidence intervals of the cohort life expectancy—, based on panel (c) we date this take-off at 1840 for the remainder of this paper. Panel (d) indicates diminishing gains to cohort life expectancy at age 20 in the past decades.

2.2 Income per Capita over the Last Two Centuries

Our theory ascribes crucial importance to the increase in income per capita in generating the takeoff, first in life expectancy and then in the emergence of a modern health sector, driven by rising demand for health goods from the household sector.[10] We therefore briefly review the main facts concerning income per capita (growth) in the U.S. over the long run (that is, abstracting from short-run business cycles).

Accordingly, Figure 2, which plots the natural logarithm of real income per capita for the U.S. for the last two centuries, as documented in the Maddison Project Data Base, displays the well-documented fact that income per capita started to rise significantly around 1820 and since then has grown at a roughly constant (in fact slightly accelerating) pace of approximately 2% per year.

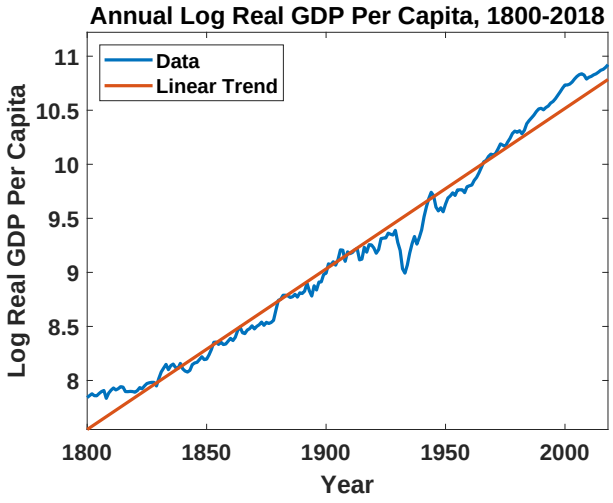
2.3 The Emergence and Evolution of the Modern Health Sector

The third set of facts that motivate our analysis is the emergence and growth of a modern health sector in the 20th century. Parker (2019) dates the start of the era of modern medicine at ca. 1920 in his book, pointing to landmark breakthroughs such as the discovery of Penicillin in 1928 and the start of its mass production towards the end of WWII, the discovery and analysis of blood

[9] On the contrary, the gap in remaining cohort life expectancy at age 40 between women and men is positive for all years.

[10] It is important to stress, though, that in our model income growth is not exogenously assumed, but emerges endogenously as part of the dynamic equilibrium transition path from an initial period of stagnation towards a balanced growth path with constant, positive growth rates of incomes and production in both sectors and constant expenditure shares.

Figure 2: Income per Capita in the U.S., 1800-2018



Source: Maddison Project Data Base 2020.

types and blood transfusions and the associated understanding of the causes of diabetes mellitus, as well as the emergence of cancer research and cures.

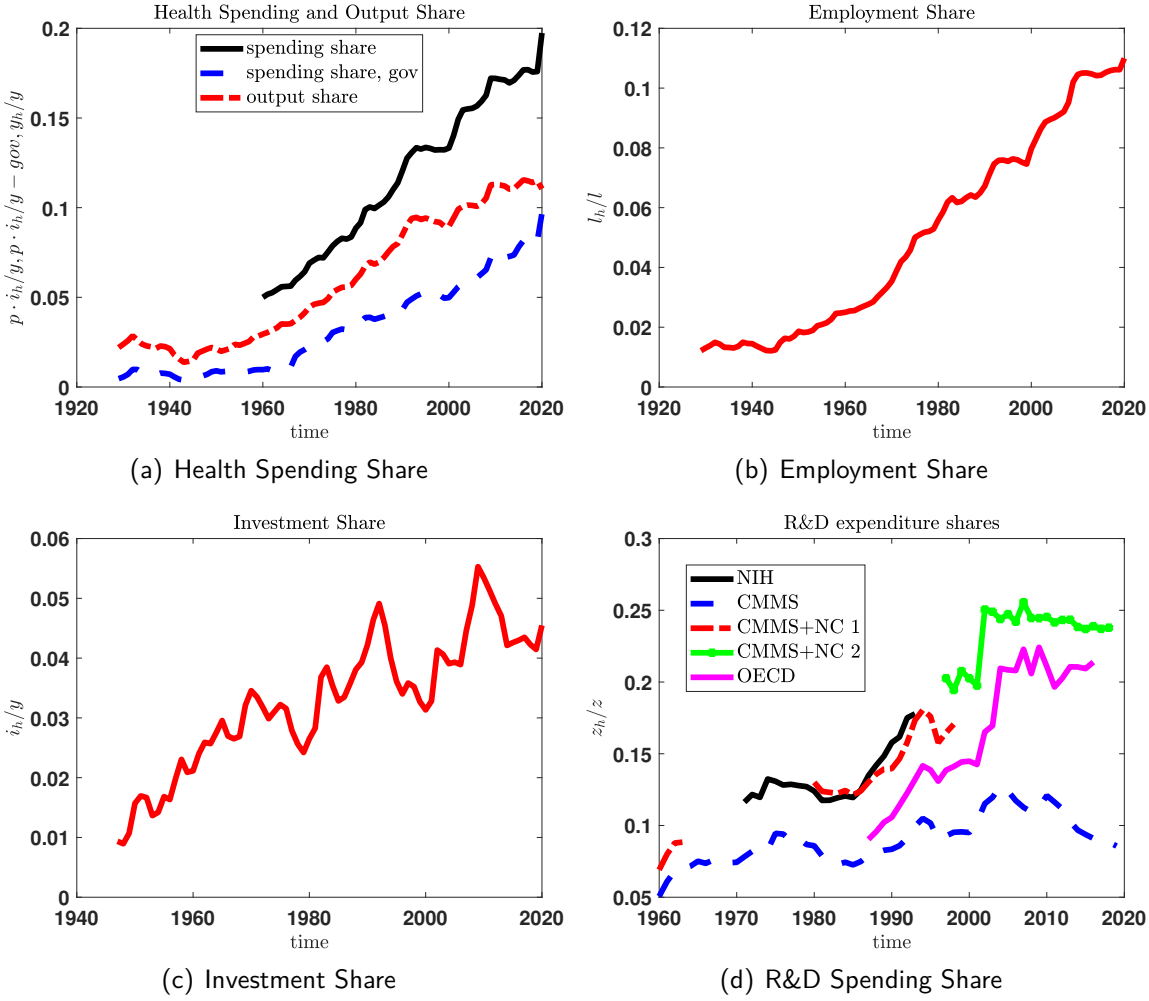
In Figure 3 we plot shares of various measures of output, investment and employment devoted to the health sector.[11] The range of the time series is dictated by data availability, but all measures point to the same broad observation: the share of economic activity contributed by the health sector was close to zero prior to World War II and since then has steadily been increasing over time, to more than 10% of total output (panel (a) and employment (panel (b)) and close to 20% of overall household spending and R&D investment. It is this emergence and continued expansion of the modern health sector we seek to explain as the endogenous equilibrium outcome driven by income growth on the demand side, and endogenous (temporally) unbalanced technological progress on the supply side of the model.

Finally, in Figure 4 we plot the relative price of health goods, measured as the ratio of the price index of household expenditures on health services to the GDP price deflator. Health goods have become more expensive over time; it is an open question what share (potentially larger than 100%) of this increase can be attributed to improvement in the quality of modern health care. Our model will permit the interpretation of this observation since it will deliver a time series for the price of one unit of health care, and a series for the price of one unit of quality-adjusted health goods.

In the remainder of the paper we now construct a two-sector overlapping generations model with endogenous and directed technical change in which income growth, life expectancy, techno-

[11] Since we display shares, there is no need to deflate the measures by prices

Figure 3: Health Shares (Spending, Employment, Investment, R&D) in the U.S



(a) Health Spending Share

(b) Employment Share

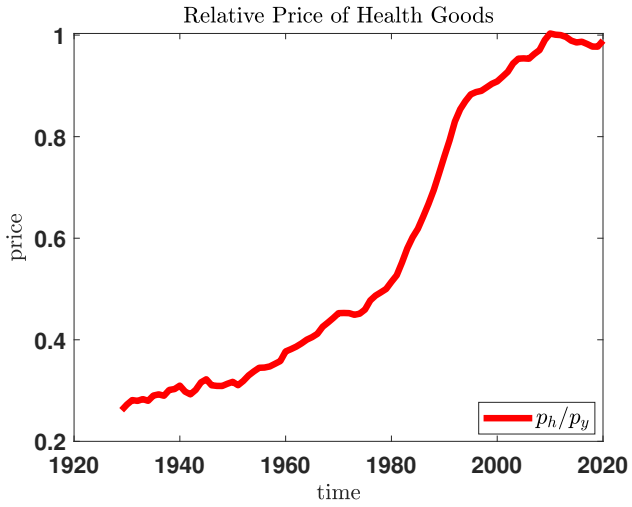
(c) Investment Share

(d) R&D Spending Share

Sources: Panel (a): The government spending share (blue dashed line) is taken from www.usgovernmentspending.com. The total health expenditure share (black solid line) is taken from the National Health Expenditures data of the Centers for Medicare & Medicaid Services (CMS). The output share (red dashed line) are household expenditures on health care taken from the Bureau of Economic Analysis (BEA), Table 1.5.5. All series are nominal and related to nominal GDP taken from the BEA data. Panel (b): The employment share is computed as full-time equivalent employees in the health sector relative to the total number of US full-time equivalent employees using data from the BEA, Tables 6.4 and 6.5. Panel (c): The investment share is computed as real investments in the health sector relative to total real investments in the US, with data taken from BEA, Tables 3.7 and 3.8. Panel (d): The real R&D expenditure share data are likewise defined as the ratio of real R&D expenditures in the health sector relative to total real R&D expenditures and compiled from various sources, the CMS, the National Science Foundation (NFS), and the OECD Stan R&D expenditure data.

logical progress in the health and the final goods sector, as well as the size of the health sector and the quality and price of the goods it produces are jointly determined in general equilibrium.

Figure 4: The Relative Price of Health Goods in the U.S.



Sources: The relative price of health is computed as the ratio of the price indices of household expenditures on health services to the GDP price index taken from BEA, Table 1.6.4.

3 The Model

We model a small open economy with overlapping generations, where in every period t a unit measure of identical young individuals is born. The number of old individuals is denoted by n_t^o and is endogenous in the model, as described below. The total population is denoted by $n_t = 1 + n_t^o$. Households work, earn income, spend resources on health and save in the first and consume in the second period of their lives. The other actors in the economy are firms in three sectors of the economy. First, output goods firms in the consumption good sector and the health sector that operate under perfect competition, where $j \in \{f, h\}$ denotes the final (consumption) and the health sector of the economy, respectively. Second, R&D sector firms that seek to invent intermediate goods of higher quality, and if successful, become the monopolistically competitive suppliers of intermediate goods of a certain variety that they sell to the final goods producers at a markup sufficient to recover the R&D costs needed to generate the new inventions. We now describe the household sector and then the several production sectors of the economy, before defining a competitive equilibrium.

Let $j \in \{f, h\}$ stand for the final and the health sector of the economy, respectively,

3.1 Households

Households derive utility from consumption in young age c_t^y, and old age c_{t+1}^o. Survival from the first to the second period of their life depends on investment i_t into health goods when young

which determines the mortality rate households face given by

$$\pi(i_t) = (1 + i_t)^{-\xi}. \tag{1}$$

The survival function, $\psi(i) = 1 - \pi(i)$, is increasing in health investment i, $\psi'(i) > 0$ with $\lim_{i\to\infty} \psi(i) = 1$ and $0 < \psi(0) < 1$. The marginal benefit of health investment is finite at $i = 0$, $\psi'(0) < \infty$, hence households survive with positive probability into the second period even absent any health investment.[12]

The utility of being dead is set to zero, and therefore expected lifetime utility is given by

$$(1 - \beta)u(c_t^y) + \beta\psi\left(i_t\right) u(c_{t+1}^o) \tag{2}$$

where the period utility function $u(c)$ is at least twice continuously differentiable with $u'(c) > 0$ and $u''(c) < 0$, and satisfies the lower Inada condition, thus $\lim_{c\to 0} u'(c) = \infty$.

Health investment i_t is the composite of two health goods, health good purchases i_{ft} from the final goods sector (standing in for expenditures on basic hygiene and nutritious food) and purchases i_{ht} of goods from a separate health production sector (such as modern hospital services and treatments as well as drugs). Thus,

$$i_t = f(i_{ft}, i_{ht}) \tag{3}$$

Young households are endowed with one unit of perfectly divisible time which they divide between working in the final goods sector (l_{ft}) and the health sector (l_{ht}). Their time constraint therefore reads as

$$l_{ft} + l_{ht} = 1. \tag{4}$$

We choose the final consumption good as the numeraire and normalize p_{ft} to 1. We then denote by $p_t = p_{ht}$ the relative price of goods produced by the modern health sector whenever it notationally is more convenient and there is no room for confusion. Wages in the two sectors of production are denoted by $w_{jt}, j \in \{f, h\}$. We envision the representative young household being composed of a large number of members of size 1, so that total labor income of the household

[12]The general form of our mortality function can be seen as $\pi(i_t) = (\kappa + i_t)^{-\xi}$. For $\kappa = 0$ it captures the specification of Jones (2016) which exhibits a constant elasticity of the mortality risk with respect to health investment equal to ξ. Choosing $\kappa > 0$ bounds the marginal benefit of health investment from below which is a necessary condition for the existence of a phase without any health investment. $\kappa = 1$ is therefore a calibration choice informed by the empirical finding that remaining life expectancy at age 20 is constant until about 1840. Note, in the long, as $i_t \Rightarrow \infty$, run we converge to the specification of Jones (2016) with a constant elasticity of the mortality risk with respect to health investment.

is given by $w_{ft}l_{ft} + w_{ht}l_{ht}$. Furthermore, households receive transfers T_t implied by accidental bequests from the share of the older generation $1 - \psi(i_{t-1})$ that do not survive until old age. Young households take these transfers as exogenously given. Lastly, government spending is financed by a lump-sum tax τ on young households that households also take as exogenous.[13] The maximization of the utility function (2) is then subject to the constraints:

$$c_t^y + i_{ft} + p_t i_{ht} + s_t = w_{ft}l_{ft} + w_{ht}l_{ht} + T_t - \tau \tag{5a}$$

$$i_t = f(i_{ft}, i_{ht}) \tag{5b}$$

$$1 = l_{ft} + l_{ht} \tag{5c}$$

$$c_{t+1}^o = Rs_t. \tag{5d}$$

We assume that the depreciation rate on capital is 1, so that the gross return on saving s_t is given by the world interest factor $R = 1 + r$ which we assume to be exogenous and constant. Since optimal saving is always strictly positive, potential borrowing constraints never bind and the period budget constraints can be consolidated to the lifetime budget constraint

$$c_t^y + i_{ft} + p_t i_{ht} + \frac{c_{t+1}^o}{R} = w_{ft}l_{ft} + w_{ht}l_{ht} + T_t - \tau \equiv x_t. \tag{6}$$

where x_t is disposable cash-on-hand of the household after taxes and transfers. In equilibrium, transfers to the generation born in period t due to accidental bequests from generation $t-1$ are given by:

$$T_t = Rs_{t-1}(1 - \psi(i_{t-1})). \tag{7}$$

Thus transfers are positive if and only if $\psi(i) < 1$ and households die with positive probability between young and old ages.

3.2 Firms, Production and R&D

3.2.1 Final Goods Producers

State the Profit max

In each sector a representative firm uses a continuum of intermediate inputs indexed by i and labor to produce sectorial output y_{jt} according to the production function

$$y_{jt} = \left(\int_0^1 q_{jit}^{1-\alpha} y_{jit}^{\alpha} di \right) l_{jt}^{1-\alpha}, \tag{8}$$

[13] See section 3.2.5 for details on government spending.

where $0 < \alpha < 1$ and y_{jit} is the quantity of intermediate input i used to produce the output good in sector j at date t and l_{jt} is the number of workers employed in sector j. The entity q_{jit} denotes the quality of intermediate input i at date t in sector j. Growth in this model results from innovations that increase the quality q_{jit} of intermediate inputs. Since the final good producer is competitive and takes factor input prices as given, she hires labor and intermediate inputs to equate marginal productivities to these input prices, taking as given their qualities q_{jit}. Let the wage rate in sector j be given by w_{jt}, the price of the final output good in sector j be p_{jt}, and denote the price of one unit of intermediate good i in sector j is p_{jit}. The representative final output firm in each sector solves the problem

$$\pi_{jt} = \max_{l_{jt},(y_{jit})_{i\in[0,1]}} p_{jt}\left(\int_0^1 q_{jit}^{1-\alpha} y_{jit}^{\alpha} di\right) l_{jt}^{1-\alpha} - w_{jt} l_{jt} - \int_0^1 p_{jit} y_{jit} di \tag{9}$$

with associated first order conditions for labor demand

$$p_{jt}\left(1-\alpha\right)\left(\int_0^1 q_{jit}^{1-\alpha} y_{jit}^{\alpha} di\right) l_{jt}^{-\alpha} = w_{jt} \tag{10}$$

and for the demand for intermediate goods

$$p_{jt}\alpha q_{jit}^{1-\alpha} y_{jit}^{\alpha-1} l_{jt}^{1-\alpha} = p_{jit} \tag{11}$$

given their quality q_{jit}.

3.2.2 Intermediate Goods Producers

Each intermediate good producer i is a monopolist that takes the demand function (11) as given and uses capital (which depreciates immediately after use) to produce the intermediate good according to:

$$y_{jit} = k_{jit}. \tag{12}$$

The gross rental rate of capital is given by R, so that each intermediate goods monopolist producer maximizes profits, taking as given the demand function of the final goods producer,

$$\pi_{jit} = \max_{y_{jit}} \left\{ \left[p_{jt}\alpha q_{jit}^{1-\alpha} y_{jit}^{\alpha-1} l_{jt}^{1-\alpha}\right] y_{jit} - R y_{jit} \right\},$$

with first order condition

$$y_{jit} = \left(\frac{p_{jt}\alpha^2}{R}\right)^{\frac{1}{1-\alpha}} q_{jit} l_{jt} \tag{13}$$

and profits

$$\pi_{jit} = \frac{1-\alpha}{\alpha} R y_{jit} > 0. \tag{14}$$

The monopolistic price follows from using (13) in (11) as

$$p_{jit} = \frac{1}{\alpha} R > R, \tag{15}$$

hence featuring the standard markup over marginal costs, R. It is the same across all intermediate input producers i.

Finally, observe from (13) that $\frac{y_{jit}}{q_{jit}}$ is constant across varieties i. Likewise the ratio of profits to quality $\frac{\pi_{jit}}{q_{jit}}$ is constant across varieties i, which we state for further reference using (13) in (14) as

$$\frac{\pi_{jit}}{q_{jit}} = \frac{1-\alpha}{\alpha} \left(\frac{p_{jt}\alpha^2}{R^\alpha} \right)^{\frac{1}{1-\alpha}} l_{jt}. \tag{16}$$

3.2.3 Aggregation of Production Sector

Because the ratios of variety-specific intermediate outputs to quality y_{jit}/q_{jit} and profits to output (or quality) π_{jit}/y_{jit} (π_{jit}/q_{jit}) are constant across varieties i we get immediate aggregation results for each sector.

For each production sector j we can determine aggregate capital input and production as

$$k_{jt} = \int_0^1 k_{jit} di = \int_0^1 y_{jit} di = \left(\frac{p_{jt}\alpha^2}{R} \right)^{\frac{1}{1-\alpha}} q_{jt} l_{jt} \tag{17}$$

where

$$q_{jt} = \int_0^1 q_{jit} di \tag{18}$$

is an aggregate quality index of intermediate inputs in sector j. Furthermore, exploiting (13) and (17) in (8) yields as aggregate production function for sector j

$$y_{jt} = k_{jt}^{\alpha} \left(q_{jt} l_{jt} \right)^{1-\alpha}. \tag{19}$$

Using equations (10) and (17) delivers as factor prices for labor inputs and capital inputs:

$$w_{jt} = (1-\alpha)\frac{p_{jt}y_{jt}}{l_{jt}} \tag{20a}$$

$$R = \alpha^2 \frac{p_{jt}y_{jt}}{k_{jt}}. \tag{20b}$$

Finally we can use (14) and (17) to determine aggregate profits in each sector j as

$$\pi_{jt} = \alpha\,(1-\alpha)\,p_{jt}y_{jt} \tag{21}$$

and thus in each sector j output exhausts factor input payments plus profits:

$$p_{jt}y_{jt} = \pi_{jt} + Rk_{jt} + w_{jt}l_{jt} \tag{22}$$

To summarize the aggregation result, in each of the two sectors output is produced with a Cobb-Douglas production function with capital and labor inputs in which the level of technology is given by q_{jt}. However, final goods producers cannot rent capital directly, but have to go through monopolistically competitive intermediaries. As a consequence owners of the capital (which will be the old households in equilibrium) command only a fraction α^2 of the value of output, with a fraction $\alpha(1-\alpha)$ accruing to the monopolist intermediaries.

3.2.4 Research and Development

An R&D developer that specializes in intermediate good i spends resources of the final consumption good z_{jit} on R&D to achieve innovation. If successful in innovation, the quality of the intermediate good increases from q_{jit-1} to

$$q_{jit} = \lambda q_{jit-1} \tag{23}$$

where $\lambda > 1$ is a parameter. The successful innovator immediately becomes the monopolist, and for one period enjoys monopoly profits π_{jit} associated with technology level $q_{jit} = \lambda q_{jit-1}$. In a product line i in which innovation is not successful a randomly chosen entrepreneur becomes the monopolist and produces at quality $q_{jit} = q_{jit-1}$ with associated profits.

We assume that the probability of innovating is related to R&D spending, z_{jit}, the quality reached when successfully innovating, given by λq_{jit-1} so that

$$\phi_j(z_{jit}; q_{jit-1}) = \left(\frac{(1+s_{jt})z_{jit}}{\lambda q_{jit-1}}\right)^{\gamma}, \tag{24}$$

with $\gamma \in (0,1)$. s_{jt} is a sector-specific R&D subsidy by the government that is exogenous to the entrepreneur.[14] The inverse relationship between the success probability and current quality q_{jit-1} reflects the fact that it becomes increasingly harder to innovate if already a level of quality is reached for variety i. Note that there is an upper bound on R&D spending, z_{jit}, which achieves an innovation probability of 1 and beyond which additional spending is unproductive. However, we check that in the benchmark calibration the innovation probability is always interior.

The R&D entrepreneur then spends resources z_{jit} and, if successful, collects profits π_{jit}. Hence the problem is

$$\max_{z_{jit}} \left\{ \pi_{jit} \phi_j \left(z_{jit}; q_{jit-1} \right) - \frac{z_{jit}}{\theta} \right\} \tag{25}$$

where θ captures the effectiveness of R&D spending, e.g. part of spending is on overhead cost which does not raise the probability of innovating. For interior solutions the first order condition is

$$\frac{(1+s_{jt})\pi_{jit}}{\lambda q_{jit-1}} \gamma \left(\frac{(1+s_{jt}) z_{jit}}{\lambda q_{jit-1}} \right)^{\gamma-1} = \frac{1}{\theta}, \tag{26}$$

which yields as solution for effective R&D spending, the ratio of R&D spending to potential period t technology $\frac{z_{jit}}{\lambda q_{jit-1}}$,

$$\begin{aligned} \frac{(1+s_{jt}) z_{jit}}{\lambda q_{jit-1}} &= \left[\theta \gamma \frac{(1+s_{jt})\pi_{jit}}{\lambda q_{jit-1}} \right]^{\frac{1}{1-\gamma}} \\ &= \left[\theta \gamma (1+s_{jt}) \frac{1-\alpha}{\alpha} \left(\frac{p_{jt} \alpha^2}{R^\alpha} \right)^{\frac{1}{1-\alpha}} l_{jt} \right]^{\frac{1}{1-\gamma}}, \end{aligned} \tag{27}$$

where in the second step we used the profit equation (16) and the fact that in case of success $q_{jit} = \lambda q_{jit-1}$. The innovation probability is then

$$\begin{aligned} \left(\frac{(1+s_{jt}) z_{jit}}{\lambda q_{jit-1}} \right)^{\gamma} &= \left[\theta \gamma \frac{(1+s_{jt})\pi_{jit}}{\lambda q_{jit-1}} \right]^{\frac{\gamma}{1-\gamma}} \\ &= \left[\theta \gamma (1+s_{jt}) \frac{1-\alpha}{\alpha} \left(\frac{p_{jt} \alpha^2}{R^\alpha} \right)^{\frac{1}{1-\alpha}} l_{jt} \right]^{\frac{\gamma}{1-\gamma}}. \end{aligned} \tag{28}$$

Notice that effective R&D spending and the probability of innovation are the same across all varieties i. Using the above back in (24) we observe that in the interior solution the share of

[14] See section 3.2.5 for a description of the government in the model, including the R&D subsidy.

varieties innovating is (due to the law of large numbers)

$$\mu_{jt} = \int \left(\frac{(1+s_{jt}) z_{jit}}{\lambda q_{jit-1}} \right)^{\gamma} di = \left[\theta\gamma(1+s_{jt}) \frac{1-\alpha}{\alpha} \left(\frac{p_{jt}\alpha^2}{R^{\alpha}} \right)^{\frac{1}{1-\alpha}} l_{jt} \right]^{\frac{\gamma}{1-\gamma}} \tag{29}$$

and is therefore independent of the distribution of qualities across varieties i. For future reference, also observe that private, public and total resources spend on R&D for variety i in sector j are, respectively,

$$z_{jit} = (1+s_{jt})^{\frac{\gamma}{1-\gamma}} \left[\theta\gamma \frac{1-\alpha}{\alpha} \left(\frac{p_{jt}\alpha^2}{R^{\alpha}} \right)^{\frac{1}{1-\alpha}} l_{jt} \right]^{\frac{1}{1-\gamma}} \lambda q_{jit-1}, \tag{30}$$

$$z_{jit}^{pub} = s_{jt} z_{jit}, \tag{31}$$

$$z_{jit}^{tot} = z_{jit} + z_{jit}^{pub} = (1+s_{jt}) z_{jit}. \tag{32}$$

Then respective resources devoted to R&D in sector j are equal to

$$z_{jt} = \int z_{jit} di = (1+s_{jt})^{\frac{\gamma}{1-\gamma}} \left[\theta\gamma \frac{1-\alpha}{\alpha} \left(\frac{p_{jt}\alpha^2}{R^{\alpha}} \right)^{\frac{1}{1-\alpha}} l_{jt} \right]^{\frac{1}{1-\gamma}} \lambda q_{jt-1}, \tag{33}$$

$$z_{jt}^{pub} = s_{jt} z_{jt}, \tag{34}$$

$$z_{jt}^{tot} = (1+s_{jt}) z_{jt}, \tag{35}$$

which are also independent of the distribution of qualities across varieties in sector j.

3.2.5 The Government

To capture the role of government spending for the emergence and growth of the modern health sector, we model two government spending programs: a social insurance program which corresponds to Medicare and parts of Medicaid, and a R&D subsidy.

First, we interpret the rising public spending share in modern health spending during the second half of the 20th century (see Figure 3) as being primarily driven by the introduction and expansion of social insurance programs, in particular Medicare and Medicaid (MC). To account for public health spending on the old as a driver of modern health demand, we assume the government covers old-age health spending i_{ht}^{MC} per old capita alive, financed by a lump-sum tax τ_t^{MC} on current young workers. We assume that old-age health spending is wasted from the perspective of households and therefore does not enter the household problem presented above, consistent with the view that individuals face substantial end-of-life health expenditures in the

data (e.g. intensive care unit treatments in the last months of life) which are associated with limited life-expectancy gains (Crimmins and Beltran-Sanchez (2010)).

Second, government R&D spending during WWII and thereafter in the form of the NIH played an important role in the emergence and growth of the modern health sector in the US.[15] Therefore, we assume that the government implements an R&D subsidy in the modern health sector, s_{ht}, financed by a lump-sum tax τ_t^{RD} on current young workers. Total lump-sum taxes levied on the young are then given by $\tau_t = \tau_t^{MC} + \tau_t^{RD}$.

3.3 Definition of Equilibrium

In this section we define a competitive equilibrium for our economy. We immediately proceed to defining the equilibrium for the aggregate economy, thereby already exploiting the aggregation results developed in sections 3.2.3 and 3.2.4. Noticing that we can define either good as numeraire, we normalize $p_{ft} = 1$ and define all equilibrium conditions in terms of the price of health goods $p_t = \frac{p_{ht}}{p_{ft}}$. Recall that we assume a small open economy (SOE) facing the exogenous and constant interest rate factor R.

Definition 1. *Given an initial population, $1, n_t^o$, and initial conditions s_0, q_{f0}, q_{h0} and given exogenous return R, a competitive equilibrium is a sequence of household allocations $c_1^o, \{s_t, i_{ht}, i_{ft}, c_t^y, c_{t+1}^o\}_{t=1}^{\infty}$, a sequence of capital and labor inputs of goods producers $\{k_{jt}, l_{jt}\}_{t=1}^{\infty}$, foreign asset holdings $\{f_t\}_{t=1}^{\infty}$, a sequence of R&D expenditures, profits and consumption of R&D developers $\{z_{jt}, \pi_{jt}, c_{jt}\}_{t=1}^{\infty}$, a sequence of aggregate capital and technology $\{k_t, q_{ft}, q_{ht}\}_{t=1}^{\infty}$, prices $\{p_t, w_{ft}, w_{ht}\}_{t=1}^{\infty}$ and transfers $\{T_t\}_{t=1}^{\infty}$ and a law of motion of the old population n_t^o such that*

1. *Household maximization: for each $t \geq 1$, given prices and transfers $w_{ft}, w_{ht}, p_t, R, T_t$, the allocations $i_{ht}, i_{ft}, s_t, c_t^y, c_{t+1}^o$ maximize (2) subject to (5).*
2. *Transfers T_t satisfy equation (7).*
3. *Factor prices satisfy equations (20a) and (20b).*
4. *Optimal R&D spending z_{jt} in each sector is given by (33) and consumption of R&D entrepreneurs is determined as $c_{jt} = \pi_{jt} - z_{jt}$.*
5. *The equilibrium innovation intensity μ_{jt} is given by equation (29) and technology in each sector evolves according to*

$$q_{jt} = (1 - \mu_{jt}) q_{jt-1} + \mu_{jt} \lambda q_{jt-1} \tag{36}$$

[15] See e.g. Gross and Sampat

6. *Markets clear: for all* $t \geq 1$

(a) *Labor Market*

$$1 = g(l_{ft}, l_{ht}). \tag{37}$$

(b) *Capital Market*

$$\sum_j k_{jt} = k_t. \tag{38}$$

(c) *Final Goods Market*

$$s_t + c_t^y + c_t^o n_t^o + i_{ft} + \sum_j [c_{jt} + z_{jt}] = k_{ft}^{\alpha} \left(q_{ft} l_{ft}\right)^{1-\alpha} + Rf_t. \tag{39}$$

(d) *Health Goods Market*

$$i_{ht} + i_{ht}^{MC} n_t^o = k_{ht}^{\alpha} \left(q_{ht} l_{ht}\right)^{1-\alpha}. \tag{40}$$

(e) *International Capital Market*

$$f_t = s_{t-1} - k_t.$$

(f) *The population evolves according to*

$$n_t^o = \psi(i_{t-1}). \tag{41}$$

(g) *The budget of each government spending program is balanced.*

$$i_{ht}^{MC} p_t n_t^o = \tau_t^{MC}. \tag{42}$$

$$s_{ht} z_{ht} = \tau_t^{RD}. \tag{43}$$

4 Theoretical Characterization of Equilibrium

In this section we characterize the optimal solution to the household problem. Given prices and transfers, young households choose labor allocation l_{ft}, l_{ht}, health investment allocation i_{ft}, i_{ht} and thus total health investment i_t and the survival probability $\psi\left(i_t\right)$, and consumption in both periods of their life c_t^y, c_{t+1}^o, where they maximize (2) subject to (5). We think of i_{ft} as goods such

as expenditures on hygiene and fresh water that are beneficial for longevity but not measured as part of health expenditures. We would like to generate the following properties of the household problem, and will make appropriate functional form assumptions to insure that they are true in equilibrium.

Proposition 1. *Suppose that the sequence of prices and cash at hand* $\{p_t, x_t\}$ *satisfy*

$$\frac{x_{t+1}}{p_{t+1}} > \frac{x_t}{p_t}. \tag{44}$$

That is, there is real income growth along the transition. Then there exist threshold time periods $0 < T_1 < T_2 < \infty$ *such that*

1. *For all* $t < T_1$ *we have* $i_t = i_{ft} = i_{ht} = 0$ *and* $\psi(i_t) = \psi(0) \equiv \psi_0$ *and* $c^o_{t+1} = Rx_t$. *We call this phase 1.*

2. *For all* $t \in [T_1, T_2)$ *we have* $i_t = i_{ft} > 0$ *and* $i_{ht} = 0$ *as well as* $\psi(i_t) > \psi(0)$. *Life expectancy is increasing due to better basic hygiene and food intake, but the modern health sector remains inoperative. This is phase 2.*

3. *For all* $t \geq T_2$ *we have* $i_{ft} > 0$ *and* $i_{ht} > 0$ *as well as* $\psi(i_t) > \psi(0)$. *Life expectancy is further increasing fueled by increasing expenditures in the modern health sector. This is phase 3.*

4. *For* $t \to \infty$, *the economy converges to a balanced growth path with constant expenditure shares in cash-on-hand* x_t. *This is the balanced growth path.*

Health expenditures on final goods i_{ft} and in the health production sector i_{ht} are aggregated into effective health investment i_t according to the CES specification

$$i_t = \left[\chi (i_{ht} + \nu)^\rho + (1 - \chi)(i_{ft} + \nu)^\rho\right]^{\frac{1}{\rho}}. \tag{45}$$

Here $\chi \in (0, 1)$ is the share parameter, $\nu > 0$ allows for a non-homotheticity, and $\rho \leq 1$ governs the elasticity of substitution between basic and modern health goods, which is given by $\frac{1}{1-\rho}$. Health investment then enters the survival function satisfying the CDF of a type 2 Pareto distribution given by

$$\psi(i_t) = 1 - (1 + i_t)^{-\xi}. \tag{46}$$

for $\xi > 0$. Note that $\psi(\cdot)$ is strictly increasing in i_t with $\psi_0 \equiv \psi(i_{ft} = i_{ht} = 0) = 1 - [1+\nu]^{-\xi} > 0$ and $\lim_{i_t \to \infty} \psi(i_t) = 1$. The non-homotheticity parameter, ν, has two purposes. First, it gives

rise to a positive baseline survival probability for zero health investments, $\psi_0 > 0$. We use this property to pin down ν by matching baseline life expectancy before it began to grow around 1840. Second, $\nu > 0$ prevents the lower Inada condition to hold for both types of health goods investment such that $\lim_{i_{ft}\to 0} \frac{\partial \psi(i_t)}{\partial i_t}\frac{\partial i_t}{\partial i_{ft}} < \infty$ and $\lim_{i_{ht}\to 0} \frac{\partial \psi(i_t)}{\partial i_t}\frac{\partial i_t}{\partial i_{ht}} < \infty$ which is a necessary condition for the existence of phase 1 in which households optimally choose $i_{ft} = i_{ht} = 0$.

4.1 Labor Supply Allocation

The labor allocation problem is straight-forward and separated from the consumption-savings-health problem. As labor supply is perfectly substitutable, we have a corner solution whenever wages are not equalized:

$$l_{it} = \begin{cases} 1 & \text{if } \; w_{it} > w_{jt} \\ \in [0,1] & \text{if } \; w_{it} = w_{jt} \\ 0 & \text{if } \; w_{it} < w_{jt} \end{cases}$$

Thus, we get an interior solution in which both sectors are operative $l_{jt} \in (0,1), j \in \{f,h\}$ if and only if $w_{ht} = w_{ft}$.

4.2 Division of Health Expenditures

The consumption-health problem can be separated into two stages. In a first stage, for a given expenditure e_t on health and relative price of modern health goods p_t, households optimally allocate these expenditures across basic and modern health goods. The solution to this problem determines health investment $i_t(p_t, e_t)$, and in turn the survival probability $\psi(i_t)$, as a function of expenditures e_t and the relative price p_t. In a second stage households decide how to split their cash at hand x_t between consumption when young, c_t^y, savings for consumption when old $c_{t+1}^o = Rs_t$, and their expenditures on health goods e_t.

We first solve the first stage, the optimal split between final goods and health goods for a given amount of health expenditures e_t. Given e_t and the relative price p_t, households solve the following allocation problem:

$$\begin{aligned} i_t(p_t, e_t) &= \max_{i_{ft}, i_{ht} \geq 0} \left\{ \left[\chi \left(i_{ht} + \nu \right)^{\rho} + (1-\chi) \left(i_{ft} + \nu \right)^{\rho} \right]^{\frac{1}{\rho}} \right\} \\ s.t. \quad p_t i_{ht} + i_{ft} &= e_t \end{aligned}$$

Interior Solution. In the interior solution households spend on both types of health goods and the first-order conditions hold with equality. In that case, it is instructive to define "effective health spending" as $\tilde{e}_t = e_t + (1+p_t)\nu$ with households choosing "effective health investments"

$\tilde{i}_{ft} = i_{ft} + \nu$ and $\tilde{i}_{ht} = i_{ht} + \nu$. Then the solution is given by

$$\begin{aligned} \tilde{i}_{ft} &= \tilde{\vartheta}(p_t)\tilde{e}_t \\ p_t\tilde{i}_{ht} &= (1-\tilde{\vartheta}(p_t))\tilde{e}_t \end{aligned}$$

where $\tilde{\vartheta}(p_t) = \left(1 + \left(\frac{\chi}{1-\chi}\right)^{\frac{1}{1-\rho}} (p_t)^{\frac{1}{1-\frac{1}{\rho}}}\right)^{-1}$ is the effective health expenditure share spent on basic goods. Note that the effective share spent on basic goods $\tilde{\vartheta}(p_t)$ is strictly increasing in p_t if $\rho \in (0,1]$ and the two investments are relatively substitutable, is independent of p_t if $\rho = 0$ (the Cobb-Douglas case) and is strictly decreasing in p_t if $\rho < 0$. Thus, the level of health spending in both goods is given by:

$$\begin{aligned} i_{ft}(e_t, p_t) &= \tilde{\vartheta}(p_t)e_t + \nu(\tilde{\vartheta}(p_t)(1+p_t) - 1) \\ p_t i_{ht}(e_t, p_t) &= (1-\tilde{\vartheta}(p_t))e_t + \nu(1-\tilde{\vartheta}(p_t)(1+p_t)) \end{aligned} \tag{47}$$

Corner Solution for modern health goods: A corner solution for modern health spending, $i_{ht} = 0$, is the optimal choice of the household if the solution above yields $\tilde{i}_{ht} \leq \nu$ (that is, $i_{ht} \leq 0$) which, using equation (47) is the case if overall health expenditures e_t are sufficiently low relative to the price of modern health goods, that is

$$e_t \leq \nu \left(\left(\frac{1-\chi}{\chi} \right)^{\frac{1}{1-\rho}} (p_t)^{\frac{1}{1-\rho}} - 1 \right) = \underline{e}^h(p_t)$$

where $\underline{e}^h(p_t)$ is the health expenditure threshold below which the modern health sector remains inactive. The economy transitions from Phase 2 to Phase 3 when optimal health expenditures have risen sufficiently (due to secular income growth) that it becomes optimal to spend resources in the modern health sector. This expenditure threshold $\underline{e}^h(p_t)$ depends positively on the relative price of modern health goods, p_t, that is, optimal health expenditures have to be high enough relative to the price of modern health goods for this phase to commence. Of course, the threshold also depends on the size of the non-homotheticity ν, as well as the relative importance of basic and modern health goods governed by the elasticity of substitution ρ, and relative shares, χ.

Corner Solution for basic health goods: Symmetrically, the household does not spend on basic health goods if expenditures satisfy

$$e_t \leq \nu \left(\left(\frac{\chi}{1-\chi} \right)^{\frac{1}{1-\rho}} (p_t)^{\frac{\rho}{\rho-1}} - p_t \right) = \underline{e}^f(p_t)$$

The dependence of this threshold on the relative health price is more complex and depends on the elasticity of substitution. However, if the two health inputs are relatively substitutable (weakly more than Cobb-Douglas), $\rho \geq 0$ as in our calibration, then $\underline{e}^f(p_t)$ is strictly decreasing, and thus households will find it optimal to spend positive amounts on basic health goods only if the price p_t of modern health goods is sufficiently high and/or overall health expenditures are sufficiently high.

These results give rise to the following dynamics of health spending. Initially households are poor, find it optimal to not spend anything on health, and the (shadow) price of modern health goods is high. As incomes increase (but the price of modern health good remains relatively high), households optimally set $e_t > \underline{e}^f(p_t)$, the economy enters phase 2, prices remain too high so that $e_t < \underline{e}^h(p_t)$ and the modern health sector remains inactive. Finally, further income growth increases e_t and technological progress in the health sector, potentially sped up by public R&D subsidies leads to $e_t > \underline{e}^h(p_t)$ eventually, and the economy enters phase 3 in which households purchase health goods from both sectors.

Of course, this discussion presupposes a specific path for overall health expenditures $\{e_t\}$; we next characterize this path for a given sequence of relative health prices $\{p_t\}$.

Existence of Phase 2: Phase 2 corresponds to a specific case of the corner solution above with $i_{ht} = 0$ and $i_{ft} = e_t > 0$ in which households start spending on basic before modern health investment. It requires the corner solution to hold strictly for $e_t = 0$ which happens if and only if

$$p_t \geq \underline{\mathrm{p}}(e_t = 0) = \frac{\chi}{1-\chi}.$$

The existence of phase 2 requires the health price to be sufficiently large compared to the relative effectiveness of modern to basic health investment which is governed by the share parameter χ. In the benchmark calibration we set $\chi = 0.5$, meaning both health goods are equally important in deriving health, in which case the threshold price for health goods is equal to the price for basic health goods, $\underline{\mathrm{p}} = 1$. Then phase 2 exists if and only if there is a spell in which $p_t > 1$ and health goods (per unit) are more expensive than final goods.

Characterizing the Phases: For a given level of health expenditures e_t and health price p_t we can fully characterize the phases. Assuming $p_t > \underline{\mathrm{p}}$ for existence of phase 2, the phases are then characterized by

$$\text{Phase} = \begin{cases} 1, & \text{if} \quad e_t = 0 \\ 2, & \text{if} \quad e_t \in (0, \nu(\tilde{p}-1)] \\ 3, & \text{if} \quad e_t > \nu(\tilde{p}-1). \end{cases}$$

Conjecture 1. *Suppose that the sequence of health expenditures* $\{e_t\}$ *satisfies*

$$e_{t+1} > e_t \quad \forall\ t \geq T_1. \tag{48}$$

That is, there is growth in health expenditures along the transition from some period T_1 *onward. Then there exist a sequence of health prices* $\{p_t\}$ *such that*

1. *For all* $t \in [T_1, T_2)$ *we have* $i_{ft} > 0$ *and* $i_{ht} = 0$ *as well as* $\psi(i_t) > \psi(0)$*. Life expectancy is increasing due to better basic hygiene and food intake, but the modern health sector remains inoperative. This is phase 2.*

2. *For all* $t \geq T_2$ *we have* $i_{ft} > 0$ *and* $i_{ht} > 0$ *as well as* $\psi(i_t) > \psi(0)$*. Life expectancy is further increasing fueled by increasing expenditures in the modern health sector. This is phase 3.*

3. *For* $t \to \infty$*, the economy converges to a balanced growth path with constant expenditure shares in cash-on-hand* x_t*. This is the balanced growth path.*

Proof outline. Recall that the health investment production function above was assumed to be of the form

$$i_t = \left[\chi\left(\tilde{i}_{ht}\right)^{\rho} + (1-\chi)\left(\tilde{i}_{ft}\right)^{\rho}\right]^{\frac{1}{\rho}}$$

Then the lower Inada condition holds for both types of health goods investment such that $\lim_{i_{ft}\to 0} \frac{\partial \psi(i_t)}{\partial i_t}\frac{\partial i_t}{\partial i_{ft}} = \infty$ and $\lim_{i_{ht}\to 0} \frac{\partial \psi(i_t)}{\partial i_t}\frac{\partial i_t}{\partial i_{ht}} = \infty$, and the first stage solution yields that for all positive health expenditures $e_t > 0$ and finite price $p_t < \infty$, households spend on both types of health goods in the first stage: $\tilde{i}_{ft} > 0$ and $\tilde{i}_{ht} > 0$.

Now start with some health spending $e > 0$. The optimal level of basic health spending $\tilde{i}_{ft}$ is continuously and monotonically increasing in the health price $p_t \in (0, \infty)$, while the optimal level of modern health spending $\tilde{i}_{ht}$ is continuously and monotonically decreasing in the health price $p_t \in (0, \infty)$. The limit cases are given by $(\tilde{i}_{ft}, \tilde{i}_{ht}) \to (e_t, 0)$ for $p_t \to \infty$, and $(\tilde{i}_{ft}, \tilde{i}_{ht}) \to (0, \infty)$ for $p_t \to 0$. Then there exists a health price for the given level of health expenditures $\hat{p}(e)$ such that the optimal devision of health spending is given by $(\tilde{i}_{ft}, \tilde{i}_{ht}) = (e - \hat{p}(e)\nu, \nu)\ \ \forall\ \nu \in (0, \infty)$.

1. Then we can construct the threshold health price, $\hat{p}$, as described above for all combinations of health spending $e_t \in (0, \infty)$ and $\nu \in (0, \infty)$. Denote the price function by $\hat{p}(e, \nu)$ which is continuous in both arguments, continuously and monotonically increasing in e, and continuously and monotonically decreasing in ν.

2. Then there exists health spending $\hat{e}$ such that basic health spending at the threshold price is also equal to ν, that is, $\tilde{i}_{ft} = \hat{e} - \hat{p}(\hat{e}, \nu)\nu = \nu$, and, thus, $(\tilde{i}_{ft}, \tilde{i}_{ht}) = (\nu, \nu)$. Then $\tilde{i}_{ft} = e - \hat{p}(e, \nu)\nu > \nu \ \ \forall\ e > \hat{e}$. Now fix some $e > \hat{e}$. According to 1. and 2. there exists a price $\hat{p}(e, \nu)$ such that $\tilde{i}_{ft} > \nu$ and $\tilde{i}_{ht} = \nu$ which we call the unconstraint solution.

3. Assume that before solving for the first stage solution, the household is given ν units of non-tradable modern health units for free while her health spending is reduced by the market value of these units $\hat{p}(e, \nu)\nu$ such that $e_{new} = e - \hat{p}(e, \nu)\nu$. Then the household reoptimizes in this constraint case. Then spending all of the remaining health expenditures e_{new} on basic health goods yields the same basket of health goods $(\tilde{i}_{ft}, \tilde{i}_{ht})$ as in the unconstraint case which by defintion is optimal, that is, the household does not prefer to trade $(\tilde{i}_{ft}$ for $\tilde{i}_{ht})$ at the market price $\hat{p}(e, \nu)\nu$ in either direction. This is equivalent to a solution of the first stage problem formulated in effective units of health goods characterized by $i_{ft} > 0 \Rightarrow \tilde{i}_{ft} = i_{ft} + \nu > \nu$ and $i_{ht} = 0 \Rightarrow \tilde{i}_{ht} = i_{ht} + \nu = \nu$.

4. Lastly, need to argue that raising e by Δe will lead to interior solution with $i_h > 0$.

4.3 Level of Health Expenditures

Given the optimal division of health investment we now optimize over the allocation of cash-on-hand x_t into young consumption c_t^y, savings s_t and health expenditures e_t. That is, the household now solves

$$\max_{0 \leq c_t^y, s_t, e_t \leq x_t} (1-\beta)u(c_{t,y}) + \beta\psi\left(i_t(p_t, e_t)\right) u(Rs_t).$$
$$s.t. \quad s_t = x_t - e_t - c_{t,y}$$

Define the share of young consumption in cash-on-hand and the share of health expenditures in old-age spending, respectively, as

$$\vartheta_{t,c} = \frac{c_t^y}{x_t} \in [0,1],$$
$$\vartheta_{t,e} = \frac{e_t}{e_t + s_t} = \frac{e_t}{(1-\vartheta_{t,c})x_t} = \frac{p_t i_{ht} + i_{ft}}{(1-\vartheta_{t,c})x_t} \in [0,1].$$

Then, the maximization problem can be rewritten in terms of those two spending shares

$$\max_{0 \leq \vartheta_{t,c}, \vartheta_{t,e} \leq 1} (1-\beta)u(\vartheta_{t,c}x_t) + \beta\psi\left(i_t(p_t, (1-\vartheta_{t,c})\vartheta_{t,e}x_t)\right) u(Rx_t\left(1-\vartheta_{t,c}\right)\left(1-\vartheta_{t,e}\right)).$$

The interior solution with positive health expenditures, $\vartheta_{t,e} > 0$, corresponds to phases 2 and 3 in which the first-order conditions hold with equality and are given by

$$(1-\beta)u'_y = \beta\psi u'_o R(1-\vartheta_{t,e}) - \beta\psi' \frac{\partial i_t}{\partial \vartheta_{t,c}} \frac{u_o}{x}$$
$$\frac{\psi}{\psi'} x_t R(1-\vartheta_{t,c}) = \frac{u_o}{u'_o} \frac{\partial i_t}{\partial \vartheta_{t,e}}$$

Balanced Growth Consistent Functional Form Assumptions. Before further characterizing heath spending during the three phases, we derive functional form and parameter restrictions for which a balanced growth path with an interior solution for spending shares exists. We choose a functional form that is consistent with the existence of a balanced growth path by following Hall and Jones (2007) and others. We accordingly assume that the utility function takes the form

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma} + b,$$

where $\sigma \geq 0$ and $b \geq 0$ are parameters. Parameter b measures the value of being alive. Plugging in the functional forms, the first-order condition for the health expenditure share becomes

$$\frac{1}{\xi}(1+i_t)\left((1+i_t)^{\xi}-1\right) x_t R(1-\vartheta_c^*) = \left(\frac{1}{1-\sigma} + b(c^o_{t+1})^{\sigma-1}\right) c^o_{t+1} \frac{\partial i_t}{\partial \vartheta_t}. \tag{49}$$

To solve for the BGP solution, we take the limit case $x \to \infty$ in the long run for which the economy converges to the BGP. For a BGP with the properties $p_t = p^*$, and $i_t \to \frac{(1-\vartheta_c^*)\vartheta_e^* x_t}{p^*}$ so that i_t, c^y_t and c^o_{t+1} are all constant shares of cash-on-hand x_t to exist we therefore require $\xi = \sigma - 1$. Under this parametric restriction, solving for the limit case where $\psi(i_t) \to 1$, we can find the health expenditure share on the BGP

$$\vartheta_e^* = \left(1 + \left[\frac{(p^*R)^{1-\sigma}}{b\xi}\right]^{\frac{1}{\sigma}}\right)^{-1}.$$

Plugging ϑ_e^* back into the Euler equation yields the BGP share of young consumption in cash-on-hand

$$\vartheta_c^* = \left[1 + \left(\frac{\beta}{1-\beta}\psi\left[R(1-\vartheta_e^*)\right]^{1-\sigma}\right)^{\frac{1}{\sigma}}\right]^{-1}.$$

Corner Solution. The corner solution with $\vartheta_{t,e} = 0$ corresponds to phase 1 without any health expenditure during which the first-order condition for the consumption share, $\vartheta_{t,c}$, holds while the

first-order condition for the health spending share, $\vartheta_{t,e}$, does not hold with equality for $\vartheta_{t,e} = 0$:

$$(1-\beta)u_y' = \beta\psi u_o' R(1-\vartheta_{t,e}) - \beta\psi' \frac{\partial i_t}{\partial \vartheta_{t,c}} \frac{u_o}{x}$$
$$\frac{\psi}{\psi'} x_t R(1-\vartheta_{t,c}) > \frac{u_o}{u_o'} \frac{\partial i_t}{\partial \vartheta_{t,e}}$$

We can solve for split between young and old consumption analytically during phase 1 by solving the first-order condition at $\vartheta_{t,e} = 0$ which yields

$$\vartheta_{\mathsf{P1},c} = \left[1 + \left(\frac{\beta}{1-\beta}\psi_0 R^{1-\sigma}\right)^{\frac{1}{\sigma}}\right]^{-1}.$$

The first kickoff at which health expenditure become positive is then characterized by both equations holding with equality at $(\vartheta_{t,c}, \vartheta_{t,e}) = (\vartheta_{\mathsf{P1},c}, 0)$. For the specific case of $\sigma = 2$ and $\xi = 1$, we can solve the first-oder conditions analytically and obtain a solution for the level of cash-on-hand, x_{KO1}, and by extension for old-age consumption at the first kickoff given by

$$\begin{aligned} c^o_{\mathsf{KO1}} = R(1-\vartheta_{\mathsf{P1},c})x_{\mathsf{KO1}} &= \frac{1}{b}\left[\frac{1}{2}\left(1+\sqrt{1+4bRA_1}\right)\right] \\ &\equiv \frac{1}{b}\Delta(b,R,A_1), \end{aligned} \tag{50}$$

where

$$A_1 \equiv \frac{(1+\nu)\,\nu}{1-\chi}.$$

We express the first kickoff in terms of old-age consumption $c^o_{t+1} = (1-\vartheta_{\mathsf{P1},c})x_{\mathsf{KO1}}$ to gain further intuition. In the absence of health investment, $\vartheta_{t,e} = 0$, the lower bound on old-age consumption below which old-age utility becomes negative and households prefer to commit suicide is, for $\sigma = 2$, given by

$$c^o_{LB} = \frac{1}{b}.$$

Notice, that the non-homotheticity in health investment $\nu > 0$ ensures $A_1 > 0$ which in turn ensures $\Delta(b,R,A_1) > 1$ and $c^o_{\mathsf{KO1}} > c^o_{LB}$. Phase 1 is then precisely characterized by $\Delta(b,R,A_1)$, which captures the region of old-age consumption in which households prefer to live in the second

period but do not yet invest in health.

$$c^o_{t+1} \in [c^o_{LB}, c^o_{\text{KO1}}] = \frac{1}{b} \cdot [1\ ,\ \Delta(b, R, A_1)] \tag{51}$$

Interior Solution during the Transition: The interior solution with positive health expenditures, $\vartheta_{t,e} > 0$, corresponds to phases 2 and 3. We cannot solve for the shares analytically in the interior solution, instead we have a system of two equations from the first-order conditions given by

$$\begin{aligned}(1-\beta)u'_y &= \beta\psi u'_o R(1-\vartheta_{t,e}) - \beta\psi' \frac{\partial i_t}{\partial \vartheta_{t,c}} \frac{u_o}{x} \\ \frac{\psi}{\psi'} x_t R(1-\vartheta_{t,c}) &= \frac{u_o}{u'_o} \frac{\partial i_t}{\partial \vartheta_{t,e}}\end{aligned}$$

4.4 Transitional Dynamics

In equilibrium, the relative labor allocation $\frac{l_{ht}}{l_{ft}}$ adjust to clear the health goods market. Note that households are indifferent to the labor allocation in any interior equilibrium as long as which $w_{ft} = w_{ht}$. The health price is uniquely pinned down then to ensure that labor demand is consistent with $w_{ft} = w_{ht}$. To see this, combine the first-order condition for labor from final good producers with the intermediate good producer solution to solve for wages in each sector j

$$w_{jt} = p_{jt}^{\frac{1}{1-\alpha}} (1-\alpha) \left(\frac{\alpha^2}{R_t}\right)^{\frac{\alpha}{1-\alpha}} q_{jt}$$

Using the fact that labor supply requires wages to be equalized in the interior equilibrium, $\frac{w_{ht}}{w_{ft}} = 1$, we get

$$\begin{aligned}1 &= p_{ht}^{\frac{1}{1-\alpha}} \frac{1-\alpha}{1-\alpha} \alpha^{\frac{2\alpha}{1-\alpha}} \alpha^{\frac{2\alpha}{\alpha-1}} \frac{q_{ht}}{q_{ft}} \\ &= p_{ht}^{\frac{1}{1-\alpha}} \frac{q_{ht}}{q_{ft}} \\ &= p_{ht}^{\frac{1}{1-\alpha}} \frac{[1+(\lambda-1)\mu_{ht}(p_{ht})]\, q_{h,t-1}}{[1+(\lambda-1)\mu_{ft}(p_{ft})]\, q_{f,t-1}},\end{aligned}$$

which is the equation that determines the health price $p_t = p_{ht}\left(\frac{q_{h,t-1}}{q_{f,t-1}}\right)$ in any interior equilibrium as a function of the period $t-1$ quality ratio. This health price then determines demand for modern health goods by households. In turn, the labor allocation, $\{l_{ht}, l_{ft}\}$, adjusts to ensure such that supply of modern health goods meets that demand. Moreover, no interior equilibrium exists if the health price that ensure equalized wages is sufficiently high such that household

demand for modern health goods is zero. In that case, we are in phase 1 or 2 without an active modern health sector.

5 Calibration

In this section we exposit the calibration of the model. We first discuss our choice of initial conditions, and then the calibration of the remaining parameters.

We interpret each model period as 40 years. We assume that economic life starts at the age of 20, when adults make the health investment decisions which determine the probability to survive to the next period. Thus, the first period covers the biological age span 20-60. The second period is accordingly 60-100. Life-expectancy at biological age 20 in the model is therefore $20 + (1 + \psi(i)) \cdot 40$ years.

In our main experiment, we treat the years prior to 1940 as years prior to the onset of modern medical times. The data on investments share, employment share and output share in the health sector suggest that the modern medical time period starts in about 1940, cf. Figures 1 and 3. A number of additional salient facts support this interpretation. First, the widespread use of penicillin to treat infections started in the second world war and can thus be dated to about 1943. Second, while it is hard to obtain historical data on health spending shares, the data reported on https://www.usgovernmentspending.com/healthcare_spending support the interpretation of a start of modern medical times in the 1950s.

Also notice that the US medicare system was only introduced in 1965. Again, recall from Figures 1 and 3 that employment, investment and output shares in health start increasing prior to that year. This supports our interpretation of the data that emergence and growth of the health sector were essentially the consequence of research and development efforts, which in turn are triggered by economic developments, rather than interpreting the introduction of Medicare as the trigger as in Frankovic and Kuhn (2023), who develop their theory on the basis of a conjecture in Weisbrod (1991). This interpretation is also consistent with the emergence and importance of government spending on reserach and development of modern health treatments during WWII as documented by Sampat and Gross (2021) and others, which we account for by explicitely modeling R&D subsidies in the health sector.[16]

Given the assumed frequency of our model of 40 year periods, the years we look at immediately before the opening up of the modern health sector and thereafter are years 1940, 1980 and 2020. According to this interpretation of the data and the mechanics of our model, prior to 1940 any increase of life-expectancy through the lens of our model is attributed to health spending on the

[16] Here, we refer to the Office of Scientific Research and Development (OSRD) and particularly the Committee on Medical Research (CMR), which was subsumed by the NIH after the war. These programs had the goal to test and produce technologies and medical treatments to help the war effort, which included mass-produced penicillin and malaria treatment.

final consumption good, such as hygiene health and nutrition. Regarding the initial stage, we notice form Panel (b) of Figure 1 that remaining cohort life-expectancy at age 20 in the US was basically flat until cohorts that are of age 20 in year 1820 and starts to rise between 1820 and 1860. Thus, years 1780 and 1820 are the initial period before the kickoff in health spending where society is poor. Already in 1860 spending on the hygiene health good leads to increasing life-expectancy. Accordingly, years 1860 and 1900 correspond to stage 2. Table 1 summarizes the stages and the corresponding years.

Table 1: Stages and Calendar Years

1780	1820	1860	1900	1940	1980	2020	...
Stage 1		Stage 2		Stage 3			

Notes: Stage 1 is the initial stage where the economy is poor and $i_{ft} = i_{ht} = 0$. In stage 2 all investment in health takes place through spending on the hygiene health good, $i_{ft} > 0, i_{ht} = 0$. In stage 3 the modern health sector is also operative so that $i_{ft} > 0, i_{ht} > 0$.

We calibrate a subset of parameters exogenously either by reference to other studies or by fixing their values prior to calibration to reduce the degrees of freedom (first-stage parameters). The remaining parameters are then endogenously calibrated to match selected moments in the data.

5.1 External Calibration

We set the constant real interest rate to 1.5, which translates to an annual rate of 1%.

Preferences. We set the weight on second period utility relative to first period utility, β, to match an annualized discount factor of 0.94.

Technology. We set the intermediate good intensity in the final good sector, which translates into the capital share in the aggregate production function, to $\alpha_f = 0.33$ consistent with standard estimates for the US. Recall that intermediate markups are given by $1/\alpha_j$ in each sector j. To capture rising markups in the health sector in a reduced form way, consistent with the literature that emphasizes the role of rising markups and price wedges for rising health prices in the US, we let α_h decline from $\alpha_{h,1940} = 0.33$ in 1940 to $\alpha_{h,2020} = 0.2$, where $\alpha_{h,2020} = 0.2$ follows evidence for the health capital share in the US (Acemoglu and Guerrieri 2008).[17]

[17] See e.g. Horenstein and Santos (2019) and Fonseca et al. (2020).

Table 2: External Calibration

Parameter	Description	Value	Target
	Small open economy		
R-1	Real Rate	1.5	($\approx$ 1 % annually)
	Households		
β	Weight old utility	0.07	Annual discount factor of 0.94
χ	Shares in health function	0.957	Equal shares
	Production		
α_f	Intermediate intensity final sector	0.33	Capital share
$\alpha_{h,2020}$	Intermediate intensity health in 2020	0.2	Acemoglu and Guerrieri (2008)
$\alpha_{h,1940}$	Intermediate intensity health in 1940	0.33	Rising health markup

5.2 Internal Calibration

We calibrate the initial conditions and a subset of parameters such that the dynamic equilibrium has the desired properties in the initial phases (phases 1 and 2). We now derive the initial conditions and state the desired properties as well as the calibration strategy to satisfy those properties. While we calibrate these parameters jointly, we provide some intuition regarding which parameters are most informative for which moments.

Initial Conditions. The state variables of the model for any period $t < 1980$ are given by $(q_{f,t-1}, q_{h,t-1}, T_t)$. Recall that aggregation of the production sector and the fact that choices of entrepreneurs in the R&D sector are independent of intermediate qualities allows us to only keep track of the aggregate quality indices $q_{f,t}, q_{h,t}$ in each sector, rather than the full quality distribution over intermediate inputs. where aggregate quality index of intermediate inputs in sector j. Recall that we do not have to keep track of in the intermediate the distribution of qualities over intermediate goods q_{jit} Note, once the public insurance program of Medicare and Medicaid becomes active in 1980, the size of the current old population, n_t^o, becomes an additional state variable. We take the first period, $t = 1$, of the model to be 1820 and start by discussing the initial calibration of the state variables, $q_{f,0}, q_{h,0}, T_1$.

1. **Initial final sector quality.** We pin down the initial quality in the final goods sector, q_{f0}, by the no suicide condition. In particular, q_{f0} determines the initial income level of the

economy, x_0, which we calibrate such that households are exactly indifferent to suicide

$$0 = \beta\psi_0 u(R(1-\vartheta_{\mathsf{P1},c})x_0),$$
$$x_0 = \frac{[b(\sigma-1)]^{\frac{1}{1-\sigma}}}{(1-\vartheta_{\mathsf{P1},c})R},$$

where $(1-\vartheta_{\mathsf{P1},c})$ is the savings share in cash on hand during phase 1 when households do not spend on health.

initial quality in the final goods sector, q_{f0}, which governs the initial income level of the economy,

2. **Modern health share in health investment function χ.** We calibrate the relative weight of modern and basic health in the health investment function, χ, to match the timing and magnitude of the first kickoff. Recall from section 4.3 that for a given set of parameters on the household side, the first kickoff at which households begin purchasing basic health goods, i_{ft}, is determined by a threshold level for cash-on-hand, $x_{\mathsf{kickoff1}}(\chi)$, which depends on the importance of basic health χ.

   For $x_t \leq x_{\mathsf{kickoff1}}(\chi)$, we are in phase 1 without basic health investment, $i_{ft} = 0$, and for $x_t > x_{\mathsf{kickoff1}}$ households invest into basic health $i_{ft} > 0$. Moreover, let $\hat{i}_{f,1860}(\chi) > 0$ denote the level of basic health goods at which, in the absence of modern health investment $\hat{i}_{h,1860} = 0$, the corresponding remaining life-expectancy matches remaining cohort life-expectancy at age 20 in 1860 from data y_{1860}. Then for a given χ, we can back out from the household problem the corresponding level of income $\hat{x}_{1860}(\chi) > x_{\mathsf{kickoff1}}(\chi)$ at which households optimally choose $i_{ft} = \hat{i}_{f,1860}(\chi)$, assuming that the health price is sufficiently high such that $\hat{i}_{h,1860} = 0$. We then calibrate χ, thus the demand for basic health given a stream of income which is determined by the supply side, such that in 1860 we get for cash-on-hand

   $$x_{1860} = \hat{x}_{1860}(\chi) > x_{\mathsf{kickoff1}}(\chi),$$

   which ensures that remaining life-expectancy in 1860 is matched. We further validate that under the set of calibrated parameters this implies $x_{1820} < x_{\mathsf{kickoff1}}$ to match the timing of the first kickoff in 1860.

3. **Initial health sector quality.** We calibrate the initial quality in the health goods sector, q_{h0} to match the timing and magnitude of the second kickoff. Recall from section 4.4 that the period t health price in any interior equilibrium with an active health sector is pinned

down by the labor market clearing condition which yields an expression for the health price as a function of the period $t-1$ quality ratio $p_{ht}\left(\frac{q_{h,t-1}}{q_{f,t-1}}\right)$. The labor allocation, $\{l_{ht}, l_{ft}\}$, then adjusts to clear the health goods market at that price. We calibrate q_{h0} such that in 1940 the health price given by

$$p_{1940} = p_{h,1940}\left(\frac{q_{h,0}}{q_{f,1900}}\right)$$

is such that the corresponding health market clearing labor allocation matches the empirical health labor share in 1940, $ls_{h,1940}$, given by

$$s^l_{h,1940} = \frac{l_{h,1940}}{l_{h,1940} + l_{f,1940}}.$$

As for the first kickoff, we validate under the set of calibrated parameters that in all periods prior to 1940 the health price $p_{ht}\left(\frac{q_{h,0}}{q_{f,t-1}}\right)$ is sufficiently high such that households do not demand modern health goods, $i_{ht} = 0\ \forall t < 1940$, to match the kickoff timing of the modern health sector in 1940.

4. **Initial transfers.** Transfers in period 1 are determined by savings and the mortality rate of the previous period according to

$$T_1 = Rs_0(1 - n^o_1),$$

We assume that the small open economy has a zero net foreign asset position in period 1 so that $k_{f1} = s_0$. Final sector capital in period 1 is an equilibrium object given by

$$k_{f1} = \left(\frac{\alpha_f^2}{R}\right)^{\frac{1}{1-\alpha_f}} q_{f1}. \tag{52}$$

which depends on q_{f1} and, thus, the initial quality q_{f0}. Moreover, consistent with flat life-expectancy and no health investment prior to 1860, we set the initial old population to $n^o_1 = \psi(i_0 = 0)$. Note, since we calibrate the economy to ensure no health investment in 1820, $i_1 = 0$, we by extension ensure that $i_0 = 0$ is an equilibrium outcome, not an assumption, if we were to solve the model going back further.

Non-homotheticiy in survival function. We calibrate the non-homotheticity parameter ν to match initial life-expectancy in phase 1, that is in the absence of any health investment. Recall

that the survival probability is given by

$$\psi(i_t) = 1 - (1 + i_t)^{-\xi}$$

where health investment is given by

$$i_t = \left[\chi \left(i_{ht} + \nu\right)^{\rho} + (1 - \chi) \left(i_{ft} + \nu\right)^{\rho}\right]^{\frac{1}{\rho}}.$$

With zero health investments in both goods we get $i_t = \nu$ and, thus,

$$\psi_0 \equiv \psi(i_t = \nu) = 1 - (1 + \nu)^{-1}. \tag{53}$$

Remaining life-expectancy at economic birth (at age 20) in the model, y, is given by

$$y = (1 + \psi) \cdot 40. \tag{54}$$

Using (53) in (54) we get as value for ν given as

$$\nu = \frac{40}{80 - y_{\text{phase 1}}} - 1. \tag{55}$$

To calibrate ν in the above we take our estimate of remaining cohort life expectancy at age 20 in year 1790 (the beginning of our sample), of $y_{\text{phase 1}} = y_{1790} = 40.43$ years.

Step size of innovations. We calibrate the step size of innovations in productivity in both sectors, λ, targeting average GDP per capita growth between 1820 and 2020 of 2%.

Efficacy of R&D spending. The parameter θ, capturing the cost-effectiveness of R&D spending (e.g., costs such as overhead that are part of the R&D process), allows us to calibrate the size of R&D spending, z, separately from the probability of innovation, μ, which is necessary to separately target R&D spending as a share of GDP and GDP growth. We therefore calibrate θ targeting an average R&D expenditure to GDP ratio of 2.06%, which is the average ratio in the US between 1953 and 2020 according the National Science Foundation.[18]

Curvature innovation probability. The concavity of the innovation probability in R&D spending, γ, is closely linked to the relative quality growth in the health relative to the final good sector. We calibrate γ targeting the growth rate in the quality unadjusted price, $p_{ht}\frac{q_{ht}}{q_{ft}}$, between 1940 and 1980.

[18] See https://ncses.nsf.gov/pubs/nsf22320 for the original data.

Tail parameter of the survival function. We calibrate the parameter ξ to match evidence from Hall and Jones (2007) and Jones (2016) on the elasticity of the mortality rate with respect to health spending. To ensure the existence of a BGP, we then impose $\sigma = \xi + 1$, see 4.3 for the derivation of the BGP.

Elasticity of substitution between basic and modern health spending. We calibrate the elasticity of substitution between basic and modern health spending in the health investment function, ρ, to minimize the distance between the health employment share in the model and in the data between 1940 to 2020.

Value of life. We calibrate the value of life, b, to minimize the distance between life expectancy in the model and in the data from 1860 onwards.

Government. We set public health spending on the old in the model equal to the size of Medicare and the nursing home component of Medicaid up until 2020. In particular, we set the total tax levied on the young to finance public health spending on the old, τ_t^{MC}, in 1980 and 2020 to match estimates on Medicare spending as a share of GDP from the CBO.[19]

In order to construct τ_t^{MC} for future periods, we need to make an assumption on how Medicare spending will evolve going forward. Recall that the government budget constraint for public health spending on the old is given by

$$i_{ht}^{MC} p_t n_t^o = \tau_t^{MC}.$$

In the benchmark economy, we assume that nominal per-old-capita spending, $i_{ht}^{MC} p_t$, grows at a constant rate in the future equal to the growth rate of GDP. The size of public health spending is then endogenously determined by the growth rate of the old-age population, n_t^o, but will converge to a constant share of GDP over time as n_t^o converges to 1 on the BGP. Therefore, this calibration choice allows us to quantify the role of old-age population growth driven by endogenous life expectancy gains in different counterfactuals.

We model a health R&D subsidy in 1940 to capture the role of government spending on reserach and development in medical and health-related fields during WWII as documented by Sampat and Gross (2021) and others. We target total federal health R&D spending from two sources between 1941 and 1945. First, as part of the Office of Scientific Research and Development (OSRD), established in 1941, the Committee on Medical Research (CMR) was specifically tasked with overseeing health-related R&D focused on medical research projects such as the development of penicillin, antimalarial drugs, and blood plasma techniques. The budget of CMR was $25 million between 1941 and 1945. Second, the U.S. military branches, including the Army

[19] We take these estimates from historical budget data by the CBO which can be found here.

and Navy, conducted their own medical R&D programs focusing on advancements in battlefield medicine, including improvements in trauma care, surgical techniques, and disease prevention with approximately $75 million allocated to these military medical programs (Harden, 1986; Rosenberg, 1997). The two programs combined then account for approximately $100 million federal health R&D spending between 1941 and 1945 which translates to $20 million annually or 0.02% of GDP. We calibrate the health R&D subsidy in 1940, $s_{h,1940}$, to target total federal health R&D spending, $z^{pub}_{h,1940}$, of 0.02% of GDP.

Table 3: Internal Calibration

Parameter	Description	Value	Target
	Households		
ν	Non-homotheticity	0.071	LE20 of 40.43 years in 1820
b	Value of Life b	2	Avg LE20 over time
ρ	Subst. elasticity health goods	0.65	Avg health labor share over time
ξ	Tail parameter, survival function	0.3	Mortality elasticity 0.18 1940-2020
$1/\sigma$	IES	0.769	BGP condition: $\sigma = \xi - 1$
	Production and R&D		
λ	Innovation step size	6.8	2% avg GDPpc growth
θ	R&D spending efficacy	0.59	2% avg R&D-GDP ratio
γ	Curvature innovation prob	0.28	Growth unadj. health price 1940-1980
$q_{f,0}$	Initial quality final sector	1400	LE20 of 42.47 years in 1860
$q_{h,0}$	Initial quality health sector	0.53	Health labor share 1940
	Government		
τ^{MC}_{1980}	1980 Medicare	190	1.57% of GDP in 1980
τ^{MC}_{2020}	2020 Medicare	1000	4.32% of GDP in 2020
$s_{h,1940}$	1940 health R&D subsidy	1	0.02% of GDP in 1940

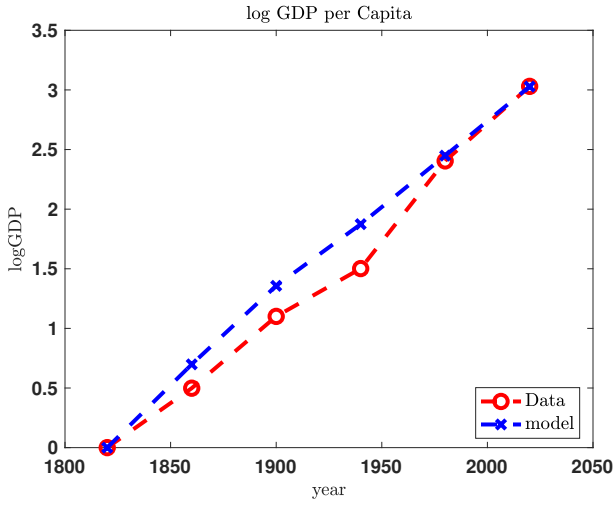
6 Results

In this section we contrast the positive predictions of the model concerning the time paths of income, life expectancy, the share of economic activity devoted to the health sector and the relative price of health with the empirical facts documented in Section 2 of the paper. We will also use the model as a measurement tool to quantify a) the contribution of the modern health sector to the overall increase in life expectancy and b) how much in the increase in the observed relative price of health services in the data should be attributed, from the perspective of the model, to changes in the relative quality q_{ht}/q_{ft}.

6.1 How the Model Makes Sense of the Facts?

In Figure 5 we display the time series of income per capita (real GDP per capita) in the model and in the data, both plotted in (natural) log-scale. Per capita income growth in the model is endogenous and driven by innovation and the associated growth in quality (q_{ft}, q_{ht}), initially only in the final goods sector (since the health sector is inoperative), and after 1940 also driven by innovation and thus quality/productivity growth in the modern health sector. We observe that the model matches the roughly linear income growth in the data well.

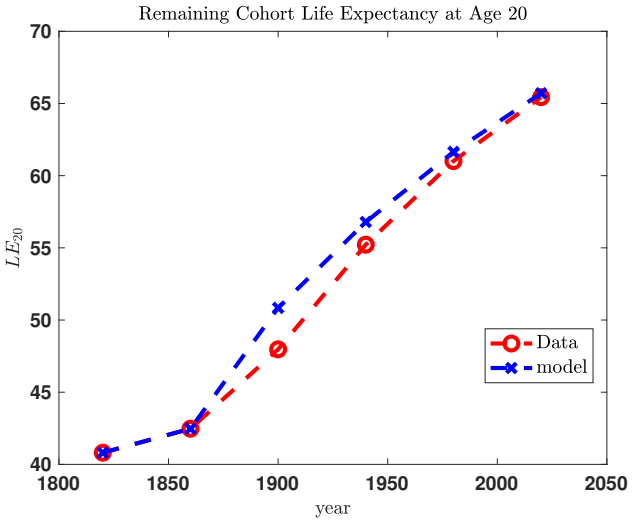
Figure 5: GDP per Capita [Logarithmic Scale]



Notes: Natural logarithm of real GDP per capita in model and data. *Source:* Data sources as in Figure 2, own calculations.

Figure 6 displays the time series of cohort life expectancy at age 20 from the data (as estimated for the U.S. in Section 2) and from the model. We observe that the model, on the account of endogenous income growth and unbalanced endogenous growth between the final goods sector and the emergence of the modern health sector in the 1940s, implies that life expectancy grows continually throughout the last 2 centuries and matches the data on cohort life expectancy well.

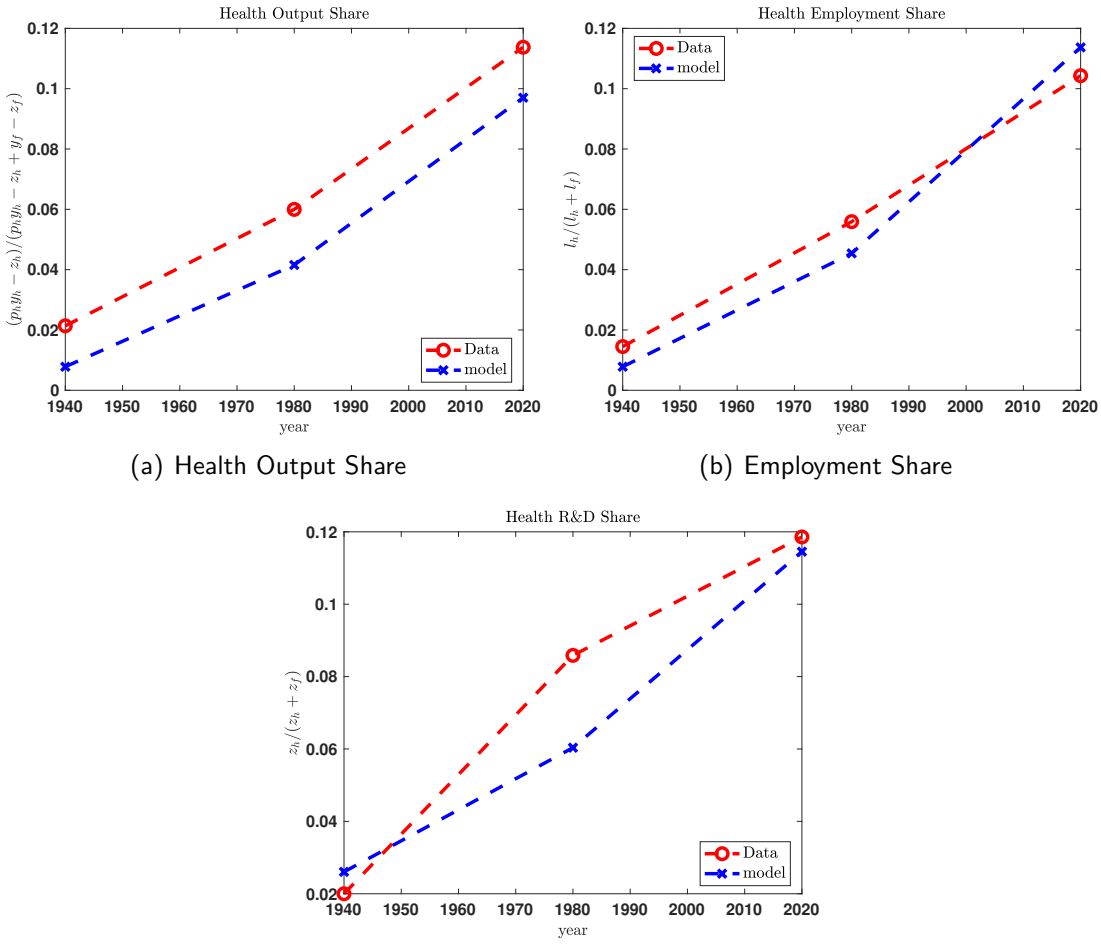
Figure 6: Remaining Cohort Life Expectancy in the U.S.



Sources: Data sources as in Figure 1, own calculations.

The health-related output-, labor- and R&D shares underlying the emergence of the modern health sector are displayed in Figure 7. Note that we plot these shares only from 1940 onward as they are not available empirically, and are equal to zero in the model, beforehand. The figure shows that the model quantitatively reproduces the increasing shares of labor, R&D and consequently output accruing to the modern health sector. Note that among these shares, the employment share in 1940 and the average growth rate in the employment share over time was targeted in the calibration.

Figure 7: Health Shares (Spending, Employment, R&D) in the U.S.



(a) Health Output Share

(b) Employment Share

(c) R&D Spending Share

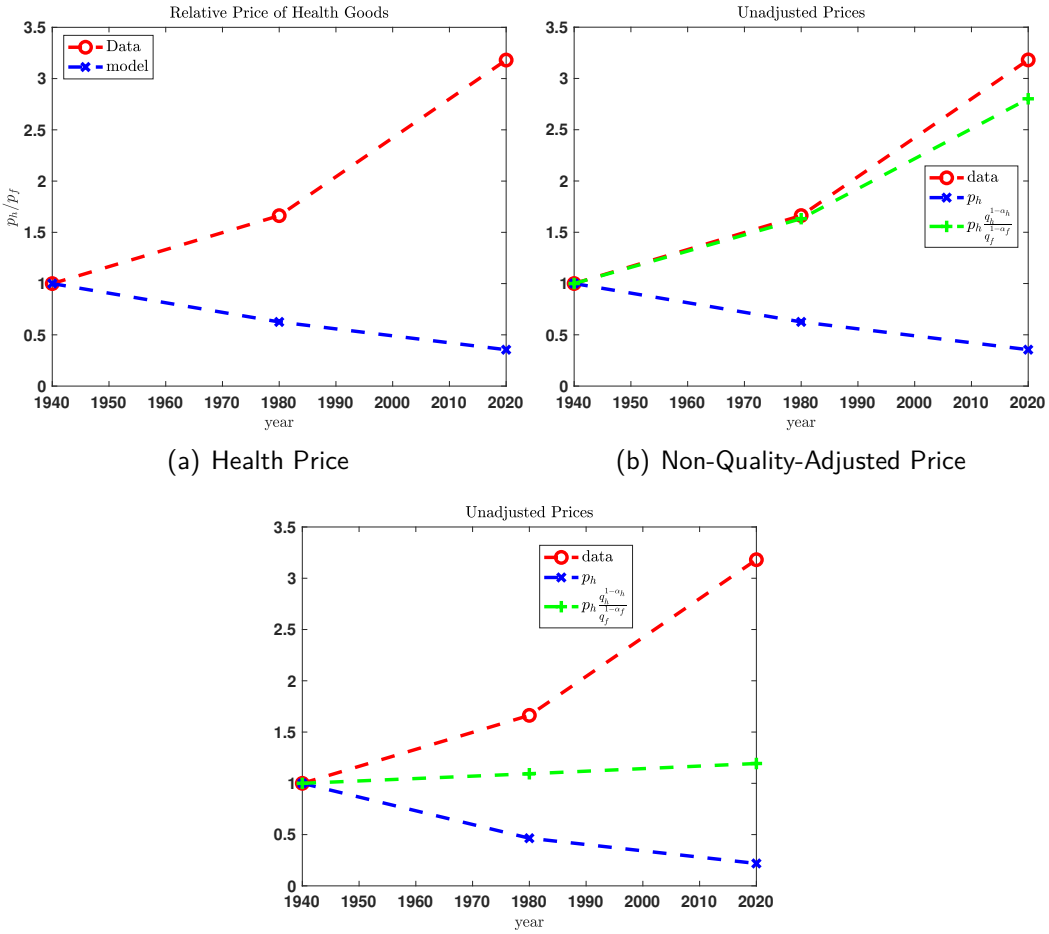
Sources: Data sources as in Figure 3, own calculations.

Finally, Figure 8 plots the relative price of health goods from Section 2 of the data, combined with the relative price times the growth component, $p_t \frac{q_{ht}^{\frac{1-\alpha_h}{1-\alpha_f}}}{q_{ft}}$, which is the model analogue price that would be observed in the data if there were no quality adjustments at all in the empirically measured relative price of health. We refer to $p_t \frac{q_{ht}^{\frac{1-\alpha_h}{1-\alpha_f}}}{q_{ft}}$ as the non-quality-adjusted price. According to the model, the price of one raw unit of health goods is falling slowly over time, accounting for the improvement in the quality of health care the non-quality-adjusted price of health goods increases at a rate close to the data.[20] Panel (c) shows the same price measures from a model with a constant intermediate good intensity in the health sector, $\alpha_h = 0.2$, akin to constant markups. In that case, the unadjusted price flattens relative to the data, consistent with the

[20] Thus, this figure implicitly assumes that the empirically observed relative price for health goods has not been appropriately quality adjusted.

literature that attributes the increase in relative health prices in the U.S. since 1980 to markups and price wedges.

Figure 8: Health Price Index



(a) Health Price

(b) Non-Quality-Adjusted Price

(c) Non-Quality-Adjusted Price in the Benchmark with constant α_h

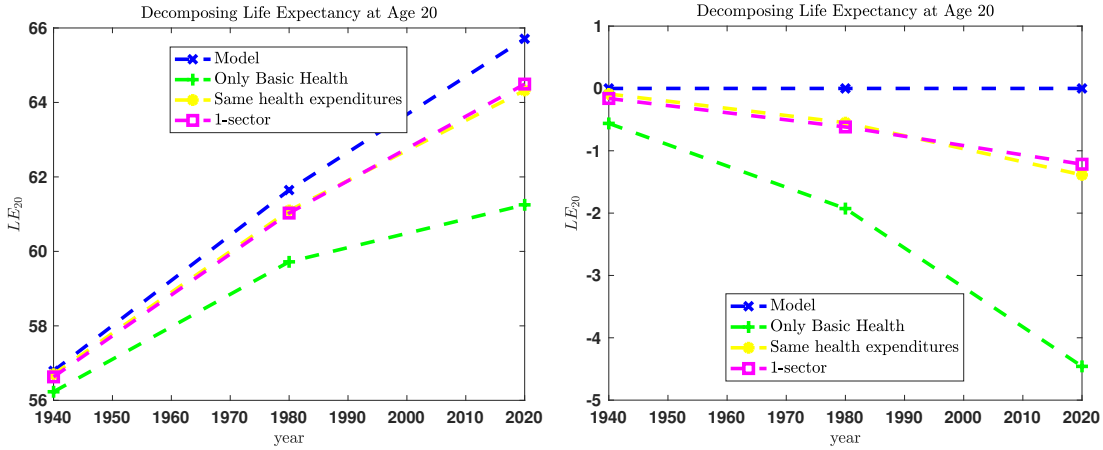
Sources: Data sources as in Figure 4, own calculations.

6.2 Decomposing Life Expectancy

We next use of our model to quantify the importance of the emergence of the modern health sector for the improvement in life expectancy in the last 80 years.

to decomposing improvements in life expectancy into spending on traditional health goods, i_{ft}, and on modern high-tech health goods, i_{ht}.

Figure 9: Decomposing Remaining Cohort Life Expectancy



(a) Remaining life expectancy 1940-2020

(b) Difference life expectancy relative to benchmark

Notes: The figure shows what LE would be if we simply set spending on modern health goods to zero (green), and if we reallocate modern health spending $p_h i_h$ towards basic health spending i_f, thus keeping total health spending constant (yellow).

In Figure 9 we plot remaining life expectancy at age 20, for the benchmark model, and in a sequence of counterfactual thought experiment to quantify the importance of the modern health sector for the increase in life expectancy from 1940 onwards. The left panel shows the level of life expectancy and the right panel displays deviations from the benchmark model. The upshot of the figure is that the modern health goods are a quantitatively important driver of the 10 year rise in life expectancy between 1940 and 2020, but that income growth and the associated increase in basic health spending (such as a better and more balanced diet and hygiene) would have driven up life expectancy significantly even in the absence of a modern health sector.

To make this point, the most extreme version of abstracting from the importance of the modern health sector is to render expenditures in this sector completely ineffective. This is the green line in the figure, which assumes that modern health spending is effectively wasteful, leaving the income-growth induced expansion of basic health spending as the only source of the mortality reduction. As the figure shows, the impact of this expansion is still quite potent, resulting in an improvement in residual life expectancy of almost 7 years between 1940 and 2020. However, the improvement is roughly 4.5 years below that in the benchmark model (and in the data), leaving a gap of 4 years by 2020. Put differently, between 1940 and 2020 slightly more than 4.5 years of expanded life expectancy are accounted for by the modern health sector (see again the difference between the green and the blue line in both sub-panels), or more than 50% of all life expectancy gains since 1940.

Assuming that all expenditures in the modern health sector are completely wasteful to assess the importance of this sector might be too extreme. The figure presents two alternatives, by assuming either that all modern health spending in a given period, $p_t i_{ht}$, is reallocated to basic health spending, i_{ft} (the yellow line), or by considering a one-sector economy without modern health (which differs from the yellow line) thereby permitting households to re-optimize the level of health expenditures over time (the magenta line). When resolving the one-sector economy, we retain the benchmark calibration of the model. Figure 9 shows that slightly more than half of the loss in life expectancy of four years by 2020 is offset when expenditures are transferred to basic health expenditures, leaving a gap of close to 1 year. When households are able to re-optimize their basic health spending in an one-sector economy, given that the marginal benefit of health expenditures falls, they choose to spend less on health, and the implied life expectancy loss from the absence of modern health initially rises above the case in which spending is reallocation (yellow). Note, the lower health spending share in the one-sector economy partially is offset by the absence of Medicare taxes which leaves households with more cash on hand. By 2020, the implied life expectancy loss in the one-sector economy is slightly more than 1 year of life expectancy.

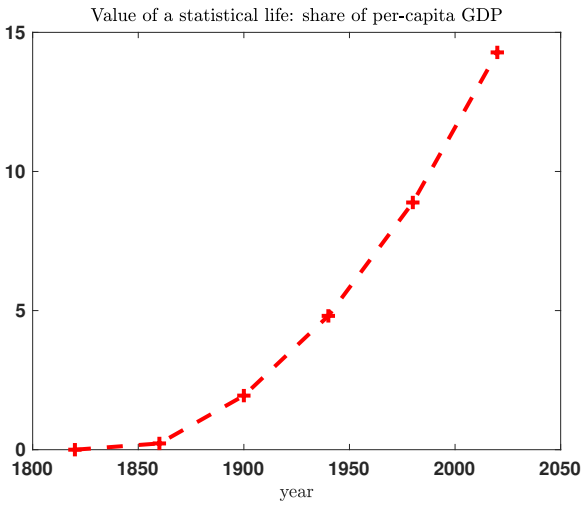
6.3 The Value of a Statistical Life

We next use the model to quantify the value of a statistical life over time, and validate the model output against existing empirical estimates. The value of a statistical life at a given age is usually defined, and empirically measured, as the willingness to pay of an individual for an intervention that reduces their mortality risk at their current age. Given our model of mortality risk in old-age from age 60 onward, we focus on the value of a statistical life at age 60 which we measure as follows

$$VSL_t = \psi_t \frac{u(c_{t+1})}{u'(c_t)}.$$

The term $\psi_t u(c_{t+1})$ is the utility value from reaching old-age and living until the age of life-expectancy, dividing by the marginal utility of consumption $u'(c_t)$ converts the value of a statistical life into units of the numeraire. Consistent with the literature, here VSL_t measures an individual's willingness to pay at age 60, right before reaching old-age, for an intervention that reduces their current mortality risk. Below we plot the the value of a statistical life as a share of GDP per capita over time.

Figure 10: Value of a statistical life

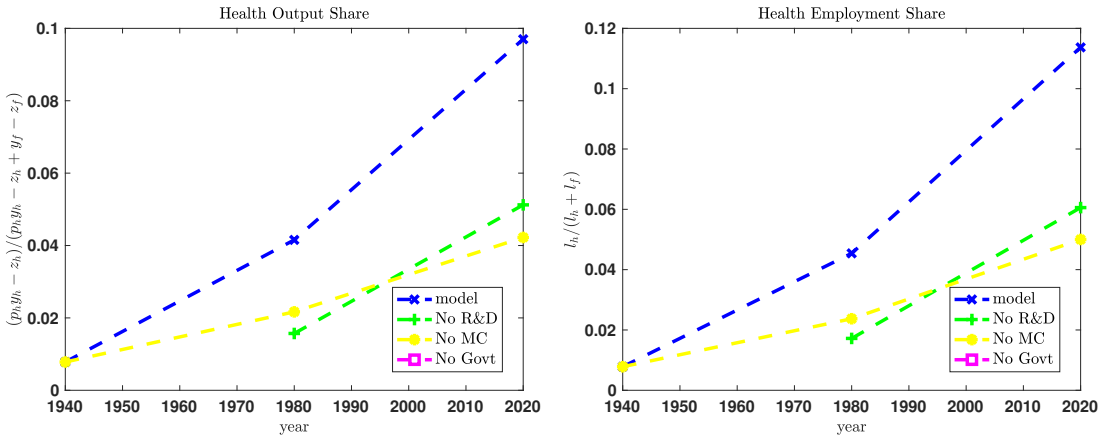


The value of a statistical life is rising over time relative to the level income or consumption that workers have. The reason is two-fold. First, rising life-expectancy raises the value of an intervention that saves a life at age 60 as workers expect to live longer after the intervention. Second, higher consumption levels further raise the value of a statistical life as workers enjoy higher utility while being alive. Translating the value of a statistical life into current dollars yields $1.1 million for the value of a statistical life in 2020, broadly in line with existing empirical estimates for the value of a statistical life (see e.g. Viscusi and Aldy (2003), Cordoba and Ripoll (2017)).

6.4 The Importance of Public Spending for Modern Health

How important has public spending been for the emergence and growth of the modern health sector, and life expectancy since 1940? We study that question by shutting down the two public spending programs and then resolving those counterfactual economies, keeping the benchmark calibration. We first shut down each program separately, we then shut them down jointly. The plots show results for the benchmark as well as the three counterfactual economies.

Figure 11: Health Shares (Spending, Employment, R&D) in the U.S.



(a) Health Output Share

(b) Employment Share

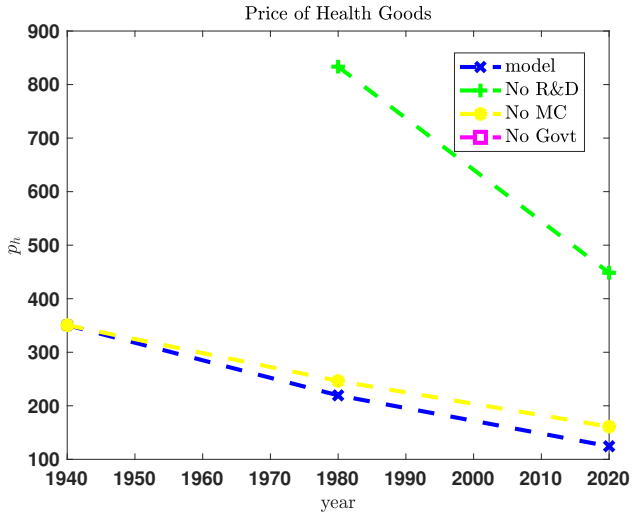
Sources: Data sources as in Figure 3, own calculations.

In figure 11 we plot the health employment and output share for all four economies. Through the lens of the model, government spending on reserach and development in medical and health-related fields during WWII has been crucial to the emergence of the modern health sector in the 1940s. Without public R&D subsidies its emergence would have been delayed, in the model by one period which is 40 years. To better approximate the counterfactual kickoff timing, we interpolate the state variables between 1940 and 1980 and resolve the economy on an interpolated grid of the state variables between 1940 and 1980. We find that without R&D subsidies, the kickoff timing would have been around 1957.

The government starts spending on the health of the old in the form of Medicare and nursing home expenditures through Medicaid in 1980. Through the lens of the model, public spending has been an important driver of the modern health growth by raising demand for health goods. We find that public old-age health spending through the two programs accounts for almost half of the output and employment growth in the modern health sector relative to the rest of the economy. Consistent with the effects of each individual program, we find that an economy without either public spending programs would have seen a modern health sector that emerges later and, by 2020, is less than half the size in terms of its output and employment share to what we observe in the data.

Figure 12 makes this point in terms of the health price. It reveals that particularly R&D subsidies during WWII, by raising R&D and therefore productivity in the health sector, q_h, drove down health prices.

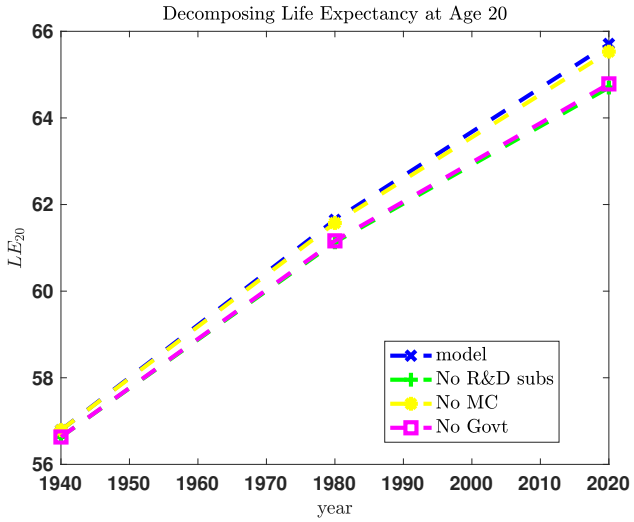
Figure 12: Health Price Index



Sources: Data sources as in Figure 1, own calculations.

We now turn to public spending effects on life expectancy, shown in Figure 13. We find that life expectancy is much less affected by public spending than the size of the sector, driven by the fact that households endogenously adjust their choices, and by a relatively low elasticity of substituting between consumption and health. Without R&D subsidies, households respond to higher prices by substituting away from modern health goods towards basic health goods which reduces life expectancy by roughly half a year in 1980 and 2020. The effect of health spending of the old is reversed. Removing Medicare and Medicaid raises cash on hand for young households who pay for these programs in the benchmark economy, while also removing the upward pressure on health prices from public demand. Both effects mean that young households spend more on modern health, and less on basic health, resulting in slightly higher life expectancy. Note, this also explains why the health price in 12 rises when public health spending is shut down. Thus, the low substitutability between consumption and health means that in both cases households achieve a similar level of life expectancy as in the benchmark economy by shifting their spending in response to the removal of public expenditures and a changing health price.

Figure 13: Remaining Cohort Life Expectancy in the U.S.



Sources: Data sources as in Figure 1, own calculations.

7 Conclusion

In this paper we build a quantitative theory of income growth, the increase in life expectancy in the last two centuries, and the emergence and expansion of a modern health sector in the 20th century. Our two-sector overlapping generations model with endogenous and directed technical change endogenously determines income growth, life expectancy, and technological progress in the health sector and the final goods sector, as well as the size of the health sector and the quality and price of the goods in general equilibrium. We show that it can generate an economic path in which households are initially poor and the quality-adjusted price of health goods is prohibitively high so that demand for health goods is zero, life is short and life expectancy stagnant. As income grows, fuelled by technological progress, households start consuming basic health goods, life expectancy starts to rise, and directed technological progress eventually, with a delay of ca. 100 years, leads to the emergence and expansion of a modern health sector.

We employ the model to quantify the role of the modern health sector for the evolution of life-expectancy and find that it contributes to about 35% of the total life expectancy gains between 1940 and 2020. We further investigate the role of government intervention through R&D subsidies and medicare insurance. We find that without these public spending programs the kickoff of the modern medical sector would have appeared about 17 years later.

References

Acemoglu, D. and V. Guerrieri (2008). Capital Deepening and Nonbalanced Economic Growth. *Journal of Political Economy 116*(3), 467–498.

Acemoglu, D. and J. Linn (2004). Market size in innovation: Theory and evidence from the pharmaceutical industry*. *The Quarterly Journal of Economics 119*, 1049–1090.

Aghion, P. and P. Howitt (1992). A Model of Growth through Creative Desctruction. *Econometrica 60*, 323–351.

Aghion, P. and P. Howitt (1998). *Endogenous Growth Theory*. Cambridge, MA: MIT Press.

Anderson, G. F., U. E. Reinhardt, and P. S. Hussey (2003). It's The Prices, Stupid : Why The United States Is So Different From Other Countries. *Health Affairs 22*(3), 89–105.

Berndt, E. R., D. M. Cutler, R. G. Frank, Z. Griliches, J. P. Newhouse, and J. E. Triplett (2000). Medical Care Prices and Output. In A. J. Cuyler and J. P. Newhouse (Eds.), *Handbook of Health Economics*, Volume 1, Chapter 3. Elsevier Science B.V.

Böhm, S., V. Grossmann, and H. Strulik (2018). R&D-Driven Medical Progress, Health Care Costs, and the Future of Human Longevity. Working Paper.

Bolt, J. and J. L. van Zanden (2024). Mpd version 2023: Maddison style estimates of the evolution of the world economy: A new 2023 update. *Journal of Economic Surveys*, 1–41.

Borger, C., T. F. Rutherford, and G. Y. Won (2008). Projecting long term medical spending growth. *Journal of Health Economics 27*, 69–88.

Cervellati, M. and U. Sunde (2005). Human Capital Formation, Life Expectancy, and the Process of Development. *The American Economic Review 95*(5), 1653–1672.

Chandra, A. and J. Skinner (2012). Technology Growth and Expenditure Growth in Health Care. *Journal of Economic Literature 50*(3), 645–680.

Congressional Budget Office (2008). Technological Change and the Growth of Health Care Spending. *359*(1), 50–60.

Cooper, Z., S. V. Craig, M. Gaynor, and J. V. Reenen (2019, 2). The price ain't right? hospital prices and health spending on the privately insured. *Quarterly Journal of Economics 134*, 51–107.

Cutler, D. M. and M. McClellan (2001). Is technological change in medicine worth it? *Health Affairs 20*(5), 11–29.

Diamond, P. A. (1965). National Debt in a Neoclassical Growth Model. *American Economic Review 55*, 1126–1150.

Ehrlich, I. and Y. Yin (2013). Equilibrium Health Spending and Population Aging in a Model of Endogenous Growth: Will the GDP Share of Health Spending Keep Rising? *Journal of Human Capital 7*(4), 411–447.

Feldman, M. and N. Pretnar (2023). The causal factors driving the rise in health services prices.

Fonseca, R., F. Langot, P. C. Michaud, and T. Sopraseuth (2023, 8). Understanding cross-country differences in health status and expenditures: Health prices matter. *Journal of Political Economy 131*, 1949–1993.

Fonseca, R., P.-C. Michaud, T. Galama, and A. Kapteyn (2021). Accounting for the rise of health spending and longevity. *Journal of the European Economic Association 19*, 536–579.

Frankovic, I. and M. Kuhn (2023). Health Insurance, Endogenous Medical Progress, and Health Expenditure Growth. *Journal of Health Economics 87*, doi: 10.1016/j.jhealeco.2022.102717.

Galor, O. (2011). *Unified Growth Theory*. Princeton: Princeton University Press.

Graboyes, R. F. (1994). Medical Care Price Indexes. *FRB Richmond Economic Quarterly 80*(4), 69–89.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy 80(2)*, 223–255.

Hacker, J. D. (2010, 4). Decennial Life Tables for the White Population of the United States, 1790-1900. *Historical Methods 43*, 45–79.

Haines, M. R. (1994). Estimated Life Tables for the United States, 1850-1900.

Hall, R. E. and C. I. Jones (2007). The Value of Life and the Rise in Health Spending. *Quarterly Journal of Economics 122*(1), 39–72.

Hansen, G. and E. C. Prescott (2002). Malthus to Solow. *American Economic Review 92 (4)*, 1205–1217.

Hejkal, J., B. Ravikumar, and G. Vandenbroucke (2022). Technology adoption, mortality, and population dynamics.

Horenstein, A. R. and M. S. Santos (2019, 2). Understanding growth patterns in us health care expenditures. *Journal of the European Economic Association 17*, 284–326.

Jones (2004). Why Have Health Expenditures as a Share of GDP Risen So Much? Working Paper.

Jones, C. (2016). Life and Growth. *Journal of Political Economy 124*(2), 1–36.

Koijen, R. S. J., T. J. Philipson, and H. Uhlig (2016). Financial Health Economics. *Econometrica 84*(1), 195–242.

Kuhn, M. and K. Prettner (2016). Growth and Welfare Effects of Health Care in Knowledge-based Economies. *Journal of Health Economics 46*, 100–119.

Lawver, D. (2011). Measuring Quality Increases in the Medical Sector. Working Paper.

Lee, R. D. and L. Carter (1992). Modeling and Forecasting U.S. Mortality. *Journal of the American Statistical Association 87 (419)*, 659–671.

Parker, S. (2019). *A Short History of Medicine*. London: Penguin Random House.

Retzlaff-Roberts, D., C. F. Chang, and R. M. Rubin (2004). Technical efficiency in the use of health care resources: A comparison of OECD countries. *Health Policy 69*(1), 55–72.

Weisbrod, B. (1991). The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care, and Cost Containment. *Journal of Economic Literature 29*, 523–552.

Zhao, K. (2014). Social Security and the Rise in Health Spending : A Macroeconomic Analysis. *Journal of Monetary Economics 64*, 21–37.

A Data Appendix

A.1 Life-Expectancy

Our data to construct cohort life-expectancy comes from three different sources. First, we use gender specific historical mortality rates for the US that were originally collected and imputed by Haines (1994), which were updated by Hacker (2010). This data covers the time period 1790 to 1899 and comes at a decennial frequency for the age groups $\{0, 1-4, 5-9, \ldots, 80\}$, where the authors report mortality rates of 1 for age 80 onward. We carry out two key transformations to this data, whereby we suppress the gender index for ease of presentation[21]: (i) we hold mortality rates constant within age group and compute age specific mortality rates in each age group i, m^i_j such that m^i is the respective geometric average within this age group, i.e., $m^i_j = 1-(1-m^i)^{1/5}$; (ii) to obtain estimates of mortality rates above (and including) age 80, we estimate per period a Gompertz–Makeham mortality model on the (constructed) age specific mortality rates and set the mortality rates for all ages 80 and older to the predicted values. Second, for the years 1900 to 1932 we use data from the Human Life-Table Database, which have age specific mortality data for ages $0, 1, \ldots, 105$. We append those to the historical mortality rates and smooth the resulting mortality rates over years 1790 to 1932 and ages 5 and older with a kernel density smoother and a bandwidth parameter of 5. Third, for the years 1933 to 2021 we use data from the Human Mortality Database with age and time specific mortality rates up to age 110.

To predict future mortality rates—required for the computation of cohort life expectancy—we estimate future trends in mortality by a Lee-Carter (Lee and Carter 1992) method decomposing the age and time specific weighted average (across genders) mortality rates into the age-specific components a_j, b_j and a single time index k_t as

$$\ln(m_{jt}) = a_j + b_j k_t.$$

We assume a stochastic trend of the single index,

$$k_t = \alpha + \rho k_{t-1} + \epsilon_t.$$

and estimate the parameters a_j, b_j, k_t in a first stage and α, ρ in a second stage using data for the period 1933 to 2021.[22]

[21] To compute the average (across genders) mortality rates and their prediction, which we use in the paper, in additional to female and male mortality rates, all of the following data operations are performed on the population weighted average mortality rates.

[22] To smooth out the Corona mortality crises induced reduction of life expectancy, we set the mortality rates in 2020 and 2021 to their year 2019 values. Our estimation procedure decomposes a_j, b_j and k_t by a standard Lee-Carter method. We additionally scale the determined k_t-sequence to exactly match cross-sectional life-expectancy.

The baseline path of predicted mortality rates is computed using the estimates values $\hat{a}_j, \hat{b}_j, \hat{\alpha}, \hat{\rho}$ and setting to 0 the innovations ϵ_t for all years $t \geq 2021$ until our forecast horizon 2110.[23] To also obtain the remaining cohort life of a year 2022 20 year which is our forecast horihon. The bootstrapped confidence intervals shown in Panels (c) and (d) in Figure 1 are computed by bootstrapping along the time dimension of the cross-sectional mortality rates for the historical data before 1933.[24] For the period 1933 to 2021, as well as for the prediction until 2131, we bootstrap by drawing from the residuals ϵ_t.[25]

A.2 R&D Spending

Summary overall R&D.

- **Overall R&D:** Overall R&D was small until the 1950s, at most 1% of GDP. In the post war period R&D spending rose, today it is roughly 3.4% of GDP.
- **Industry vs Federal/public R&D:** Prior to WW2, federal R&D was basically non-existent, the little R&D that existed was mostly by industry and rising slowly. Both industry and federal R&D jumped up / grew faster during and after WW2 as it became more institutionalized.
- **Composition of R&D spending:** While overall R&D spending has been rising since WW2, the composition has moved from federal to industry. In the 1950s and 1960s the federal share of total R&D spending peaked at almost 60%, now it is down to roughly 20%. This is mostly driven by rising industry R&D spending as a share of GDP, though federal R&D spending as a share of GDP has also been falling slowly since 1970.

Summary health R&D.

- **Overall health R&D:** Similar to overall R&D, data before WW2 is sparse. But the existing estimates suggests there was almost no federal or industry health R&D prior to WW2. Health R&D grew over time relative to overall R&D. Toole (2012) argues health R&D grew from <5% of total R&D (pre-1950) to around 25% today.
- **Federal health R&D during WW2:** Total Federal Health R&D between 1941 and 1945 was around $100 million, or 0.02% of GDP.

[23] For calibration purposes, we need a time sequence of remaining cohort life expectancy of 20 year old persons up to year 2020 (i.e. for a person of age 20 in year 2020). Human Mortality Database range to age 110, a person of age 20 in year 2020 lives at most until year 2110.

[24] We implement the bootstrap procedure as a block bootstrap and set the width of the blocks according to the standard rule of thumb $bw = T^{1/3}$. Since the historical data span $T = 143$ years, this gives $bw = T^{1/3} = 5.23$, of which we take the ceil.

[25] These residuals are not autocorrelated, thus there is no need to apply a block bootstrap approach.

A.2.1 Details on overall R&D spending over time

Comprehensive national R&D data is available from 1953 onwards from the National Science Foundation. Pre-1953 estimates rely on corporate archives, patent records, and federal budgets. Sectoral breakdowns (e.g., health vs. defense) are rare.

- 19th century: R&D was minimal and largely informal (e.g., Edison's labs, agricultural stations). Mowery and Rosenberg (1998) estimate R&D spending at $<$0.1% of GDP in the 19th century.
- Early 20th century: Industrial R&D grew with firms like DuPont, GE, and Bell Labs. Kuznets (1962) and Mowery (1995) suggest R&D spending rose to 0.2-0.4% of GDP by the 1930s. World War II: Federal R&D surged (Manhattan Project, radar, etc.), reaching 0.7% of GDP by 1945.
- Bolt and van Zanden (2020) (Our World in Data) compile GDP-adjusted proxies and argue that R&D rose from 0.1% to 1% of GDP between 1870 and 1950.

A.2.2 Details on federal R&D spending over time

- Similarly, overall federal R&D spending was also tiny before world war 2.
- Federal R&D rose during and after WW2, and also became more institutionalized. Industry R&D spending started to rise as well.
- While overall R&D spending has been rising slowly since WW2, the composition has moved from federal to industry. In the 1950s and 1960s the federal share of total R&D spending peaked at almost 60%, now it is down to roughly 20%. This is mostly driven by rising industry R&D spending as a share of GDP, though federal R&D spending as a share of GDP has also been falling slowly since 1970.

A.2.3 Details on overall health R&D spending over time

For the health sector particular, R&D tracking is sparse before the 1960s. While there are some key milestones that we outline below, the available data is consistent with our narrative that overall R&D as well as federal R&D were minimal prior to WW2.

- Early 20th century: 1) Pharmaceuticals: Companies like Pfizer and Merck invested in antibiotics and vaccines. 2) Public Funding: NIH's predecessor (Hygienic Laboratory, 1887) had minimal budgets until the 1930s.
- Post-1930s NIH Budgets. 1938: roughly $3 million ($60 million today); 1950: roughly $52 million ($580 million today); Post-1950: Explosive growth (e.g., War on Cancer, 1971).

- PhRMA Data: Industry R&D rose from $1 Billion (1960s) to more than $100 Billion (2020s).
- Health R&D also grew over time relative to overall R&D. Toole (2012): Analyzes health R&D trends, showing it grew from $<$5% of total R&D (pre-1950) to around 25% today.

A.2.4 Details on federal health R&D spending over time

The NIH's predecessor, the Hygienic Laboratory, was established in 1887. It became the National Institute of Health (NIH) in 1930 and expanded into multiple institutes post-WWII. We reconstruct a time series of federal R&D spending through these programs from different sources below. The Hygienic Laboratory was basically non-existent in terms of spending. In 1901 its budget was roughly $35,000, or around $1.5 \cdot 10^4$% of GDP. In 1930, the Hygienic Laboratory was renamed into the National Institute of Health with a budget of roughly $750,000, or $8 \cdot 10^4$% of GDP.

Federal health R&D spending during WW2.

1. The CMR program (1941-1945) had a total budget: $25 million (nominal) and focused on penicillin mass production, malaria treatments (e.g., Atabrine), blood plasma preservation, and combat medicine.
2. Additional Military Medical R&D (Army/Navy) had an estimated size of $75 million (nominal) and focused trauma surgery, vaccines (e.g., typhus), battlefield triage, and psychological health.

Total Federal Health R&D between 1941 and 1945 was then around $100 million, and $20 million annually. Given nominal GDP of $100 billion in 1940, total federal health R&D during WW2 translates into 0.02% of GDP.

B Model Appendix

B.1 Analytical Solution to the Household Problem

Households derive utility from consumption in young age c_t^y,, and old age c_{t+1}^o,, they survive from the first to the second period of their life with probability ψ which depends on their investment i_t into health goods when young. Expected lifetime utility is given by

$$(1-\beta)u(c_t^y) + \beta\psi\left(i_t\right)u(c_{t+1}^o).$$

The maximization of the utility function is subject to the constraints:

$$c_t^y + i_{ft} + p_t i_{ht} + s_t = x_t \equiv w_{ft}l_{ft} + w_{ht}l_{ht} + T_t \tag{56a}$$

$$i_t = f(i_{ft}, i_{ht}) \tag{56b}$$

$$1 = g(l_{ft}, l_{ht}) \tag{56c}$$

$$c_{t+1}^o = Rs_t. \tag{56d}$$

Since optimal saving is always strictly positive, potential borrowing constraints never bind and the period budget constraints can be consolidated to the lifetime budget constraint

$$c_t^y + i_{ft} + p_t i_{ht} + \frac{c_{t+1}^o}{R} = w_{ft}l_{ft} + w_{ht}l_{ht} + T_t \equiv x_t. \tag{57}$$

B.1.1 Division of Health Investment

We first derive the optimal split between final goods and health goods for a given amount of health expenditures e_t.

Corner Solution with $i_{ht} = 0$, $i_{ft} = e_t$: In the corner solution the first-order conditions do not hold with equality. From the budget constraint we get $i_{ht} = 0$, $i_{ft} = e_t$. Then health investment i_t is given by

$$i_t = f(i_{ht}, i_{ft}) = i_{ht} + (\nu + i_{ft})^\zeta = (\nu + e_t)^\zeta$$

Note, for $e_t = 0$ the corner solution corresponds to stage 1 (no health expenditures), for $e_t > 0$ it corresponds to stage 2 (positive health expenditures that are fully invested into final goods).

Interior Solution: In the interior solution the first-order conditions hold with equality, they are given by

$$\begin{aligned} \zeta \left(\nu + i_{ft}\right)^{\zeta-1} &= \lambda \\ 1 &= \lambda p_t \end{aligned}$$

Combining yields

$$\begin{aligned} \zeta \left(\nu + i_{ft}\right)^{\zeta-1} &= \frac{1}{p_t} \\ \nu + i_{ft} &= \left(\zeta p_t\right)^{\frac{1}{1-\zeta}} \\ i_{ft} &= \left(\zeta p_t\right)^{\frac{1}{1-\zeta}} - \nu \end{aligned}$$

Define $\tilde{\lambda}_t \equiv \left(\zeta p_t\right)^{\frac{1}{1-\zeta}}$. Then

$$i_{ft} = \tilde{\lambda}_t - \nu$$

The quasi-linear specification means that in the interior solution there are no wealth effects for the final health good i_{ft}. Thus, $i_{ft} = \tilde{\lambda}_t - \nu$ is independent of e_t. Using the budget constraint, we get

$$i_{ht} = \frac{e_t - \left(\tilde{\lambda}_t - \nu\right)}{p_t}$$

For $\zeta \to 0$ we get $\tilde{\lambda}_t \to 0$ and, thus, taking into account the non-negativity constraints $(i_{ft}, i_{ht}) \geq 0$

$$\begin{aligned} i_{ft} &\to 0 \\ i_{ht} &\to \frac{e_t}{p_t} \\ i_t &\to \frac{e_t}{p_t} + \nu^{\zeta} \end{aligned}$$

Thus, ζ needs to be sufficiently large for stage 2 to exist, otherwise the marginal benefit from final goods investment is too small.

Existence of Stage 2: We want to ensure that initial health investment will be allocated towards i_{ft}, that is, that the corner solution with $i_{ht} = 0$ and $i_{ft} = e_t$ characterized above exists. We get the undesired corner solution with $i_{ht} > 0$ and $i_{ft} = 0$ (stage 2 is skipped with initial health investment directly being allocated towards i_{ht}) if the marginal cost of investing into i_{ft}

exceeds the marginal benefit, evaluated at $i_{ft} = 0$:

$$\begin{aligned} \zeta \left(\nu + i_{ft}\right)^{\zeta - 1} &\leq \tfrac{1}{p_t} \\ \zeta \nu^{\zeta - 1} &\leq \tfrac{1}{p_t} \\ \left(\zeta p_t\right)^{\frac{1}{1-\zeta}} &\leq \nu \\ \tilde{\lambda}_t &\leq \nu \end{aligned}$$

Then stage 2 exists as long as the optimal interior solution computed above yields $i_{ft} = \tilde{\lambda}_t - \nu > 0$ which requires the non-homotheticity factor ν to be sufficiently small relative to health sector price p_t and ζ.

Characterizing the Stages: For a given level of health expenditures e_t we can fully characterize the stages now. Assuming $\tilde{\lambda}_t > \nu$ for existence of stage 2, the stages are then characterized by

$$\text{Stage} = \begin{cases} 1, & \text{if} \quad e_t = 0 \\ 2, & \text{if} \quad e_t \in (0, \tilde{\lambda}_t - \nu] \\ 3, & \text{if} \quad e_t > \tilde{\lambda}_t - \nu \end{cases}$$

B.1.2 Level of Health Expenditures

Given the optimal division of health investment we now optimize over the allocation of cash x_t into consumption in young age, savings s_t, and health expenditures e_t. The household solves

$$\max_{0 \leq c_t^y, e_t \leq x_t} (1-\beta) u(c_{t,y}) + \beta \psi \left(i_t(p_t, e_t)\right) u(R \left[x_t - e_t - c_{t,y}\right]$$

We can rewrite the household problem in terms of expenditure shares. Define the share of consumption when young and the share of health expenditures in old age spending, respectively, as

$$\begin{aligned} \vartheta_{t,c} &= \frac{c_t^y}{x_t} \in [0,1], \\ \vartheta_{t,e} &= \frac{e_t}{e_t + s_t} = \frac{e_t}{(1-\vartheta_{t,c}) x_t} = \frac{p_t i_{ht} + i_{ft}}{(1-\vartheta_{t,c}) x_t} \in [0,1], \end{aligned}$$

Then the maximization problem can be rewritten in terms of those two spending shares

$$\max_{0 \leq \vartheta_{t,c}, \vartheta_{t,e} \leq x_t} (1-\beta) u(\vartheta_{t,c} x_t) + \beta \psi \left(i_t(p_t, (1-\vartheta_{t,c}) \vartheta_{t,e} x_t)\right) u(R x_t \left(1-\vartheta_{t,c}\right) \left(1-\vartheta_{t,e}\right))$$

The first-order conditions are given by

$$[\vartheta_{t,c}] \;\; 0 = (1-\beta)u_y' x_t - \beta\psi u_o' R x_t (1-\vartheta_{t,e}) + \beta\psi' \frac{\partial i_t}{\partial \vartheta_{t,c}} u_o$$

$$[\vartheta_{t,e}] \;\; 0 = \psi' \frac{\partial i_t}{\partial \vartheta_{t,e}} u_o - \psi u_o' R x_t (1-\vartheta_{t,c})$$

Note, $\frac{\partial i_t}{\partial \vartheta_{t,e}}$ depends on the optimal division of health investment and, thus, the level of $\vartheta_{t,e}$. Rearranging yields

$$(1-\beta)u_y' = \beta\psi u_o' R(1-\vartheta_{t,e}) - \beta\psi' \frac{\partial i_t}{\partial \vartheta_{t,c}} \frac{u_o}{x}$$

$$\frac{\psi}{\psi'} x_t R(1-\vartheta_{t,c}) = \frac{u_o}{u_o'} \frac{\partial i_t}{\partial \vartheta_{t,e}}$$

The second optimality condition describes the tradeoff between health investment and consumption in old age.

B.1.3 Corner Solution of Optimal Health Investment

In stage 1 all resources are spent on consumption without any health investment, $\vartheta_{t,e} = 0$. Moreover, $\frac{\partial i_t}{\partial \vartheta_{t,e}}$ depends on the optimal division of health investment, that is, it depends on how the first marginal unit of health investment e_t is allocated between i_{ht} and i_{ft}. We assume $\tilde{\lambda}_t \geq \nu$ which we later verify in equilibirum. As shown above, this ensures that phase 2 exists and initial health investment is allocated towards the final good sector i_{ft}. Then health investment and its derivatives are given by

$$i_0 = \nu^\zeta$$

$$\frac{\partial i_t}{\partial \vartheta_{t,e}}|_{\vartheta_{t,e}=0} = \frac{\partial}{\partial \vartheta_{t,e}} \left(\nu + (1-\vartheta_{t,c})\vartheta_{t,e} x\right)^\zeta |_{\vartheta_{t,e}=0} = \zeta(1-\vartheta_{t,c}) x_t \nu^{\zeta-1}$$

$$\frac{\partial i_t}{\partial \vartheta_{t,c}} = -\zeta \nu^{\zeta-1} \vartheta_{t,e} x |_{\vartheta_{t,e}=0} = 0$$

Consumption in young and old age are given by

$$c_t^y = \vartheta_{t,c} x_t$$

$$c_{t+1}^o = R(1-\vartheta_{t,c}) x_t$$

The intertemporal Euler equation can be solved analytically for the optimal share of consumption in young age during phase 1. It is given by

$$
\begin{aligned}
(1-\beta)u_y' &= \beta\psi_0 u_o' R(1-\vartheta_{t,e}) - \beta\psi_0' \frac{\partial i_t}{\partial \vartheta_{t,c}} \frac{u_o}{x} \\
(1-\beta)u_y' &= \beta\psi_0 u_o' R \\
(\vartheta_{t,c}x)^{-\sigma} &= \frac{\beta}{1-\beta}\psi_0 \left(R(1-\vartheta_{t,c})x\right)^{-\sigma} R \\
\left(\frac{1-\vartheta_{t,c}}{\vartheta_{t,c}}\right)^{\sigma} &= \frac{\beta}{1-\beta}\psi_0 R^{1-\sigma} \\
\vartheta_{t,c} &= \left[1+\left(\frac{\beta}{1-\beta}\psi_0 R^{1-\sigma}\right)^{\frac{1}{\sigma}}\right]^{-1} \equiv \vartheta_{\text{phase 1},c}
\end{aligned}
$$

where $\psi_0 = 1 - \left(1+\nu^{\zeta}\right)^{-\xi}$ is the base survival probability during phase 1. We can further use the 2. FOC characterizing the tradeoff between health investment and consumption in old age to find the threshold cash level at which phase 2 begins. During stage 1 without any health expenditure, the 2. FOC does not hold with equality, evaluated at $i_t(p_t, e_t = 0)$, such that

$$
\begin{aligned}
\frac{\psi}{\psi'} x_t R(1-\vartheta_{t,c}) &\geq \frac{u_o}{u_o'}\frac{\partial i_t}{\partial \vartheta_{t,e}} \\
\frac{1}{\xi}(1+i_t)\left((1+i_t)^{\xi}-1\right) R x_t (1-\vartheta_{t,c}) &\geq \left(\frac{1}{1-\sigma} + b(c_{t+1}^o)^{\sigma-1}\right) c_{t+1}^o \frac{\partial i_t}{\partial \vartheta_{t,e}}
\end{aligned}
$$

Plugging $i_t(p_t, e_t = 0)$ and $\frac{\partial i_t}{\partial \vartheta_{t,e}}$ in yields

$$
\begin{aligned}
\frac{1}{\xi}\left(1+\nu^{\zeta}\right)\left(\left(1+\nu^{\zeta}\right)^{\xi}-1\right) R x_t (1-\vartheta_{t,c}) &\geq \left(\frac{1}{1-\sigma} + b(c_{t+1}^o)^{\sigma-1}\right) c_{t+1}^o \zeta (1-\vartheta_{t,c}) x_t \nu^{\zeta-1} \\
\frac{1}{\xi}\left(1+\nu^{\zeta}\right)\left(\left(1+\nu^{\zeta}\right)^{\xi}-1\right) R &\geq \left(\frac{1}{1-\sigma} + b(c_{t+1}^o)^{\sigma-1}\right) c_{t+1}^o \zeta \nu^{\zeta-1}
\end{aligned}
$$

Rewriting consumption in old age in terms of overall cash in the first stage ($\vartheta_{t,e} = 0$) yields

$$
c_{t+1}^o = R s_t = R(1-\vartheta_{t,e})(1-\vartheta_{\text{phase 1},c}) x_t = R(1-\vartheta_{\text{phase 1},c}) x_t = R\tilde{x}_t
$$

where

$$
\tilde{x}_t = (1-\vartheta_{\text{phase 1},c}) x_t = \left(1 - \left[1+\left(\frac{\beta}{1-\beta}\psi_0 R^{1-\sigma}\right)^{\frac{1}{\sigma}}\right]^{-1}\right) x_t
$$

is total cash spent on health investment and old age consumption during phase 1. Plugging back in then gives

$$\frac{1}{\xi}\left(1+\nu^{\zeta}\right)\left(\left(1+\nu^{\zeta}\right)^{\xi}-1\right)R \geq \left(\frac{1}{1-\sigma}+b\left(R\tilde{x}_t\right)^{\sigma-1}\right)R\tilde{x}_t\zeta\nu^{\zeta-1} \tag{58}$$

$$\frac{1}{\xi}\left(1+\nu^{\zeta}\right)\left(\left(1+\nu^{\zeta}\right)^{\xi}-1\right) \geq \left(\frac{1}{1-\sigma}+b\left(R\tilde{x}_t\right)^{\sigma-1}\right)\tilde{x}_t\zeta\nu^{\zeta-1} \tag{59}$$

This equation characterizes the region of cash $\tilde{x}_t$ for which health expenditures are zero, $e_t = 0$, and the economy is in phase 1. For convenience define

$$A_1 \equiv \frac{1}{\xi}\frac{\left(1+\nu^{\zeta}\right)\left(\left(1+\nu^{\zeta}\right)^{\xi}-1\right)}{\zeta\nu^{\zeta-1}} \begin{cases} = 0, & \text{if } \nu = 0 \\ > 0, & \text{if } \nu > 0 \end{cases}$$

Thus, $A_1 \geq 0$ and in our case we have $A_1 > 0$ because $\nu > 0$ is necessary for the existence of stage 2. Plugging A_1 back into the condition for the corner solution yields

$$A_1 \geq \left(\frac{1}{1-\sigma}+b\left(R\tilde{x}_t\right)^{\sigma-1}\right)\tilde{x}_t$$

To make progress analytically we need to make assumptions on the parameters. Assume $\sigma = 2$, then

$$\begin{aligned} A_1 &\geq -\tilde{x}_t + bR\tilde{x}_t^2 \\ 0 &\geq \tilde{x}_t^2 - \frac{1}{bR}\tilde{x}_t - \frac{A_1}{bR} \end{aligned}$$

The two solutions to this quadratic equation are given by

$$\tilde{x}_{1/2} = \frac{1}{2bR}\left(1 \pm \sqrt{1+4bRA_1}\right)$$

$A_1 \geq 0$ means that $\sqrt{1+4bRA_1} \geq 1$. Thus, $\tilde{x}_2 \leq 0$, then $\tilde{x}_1$ is the only economically relevant solution. The first kick-off separating stage 1 and stage 2 is then given by

$$\begin{aligned} \tilde{x}_{\text{kickoff1}} &= \frac{1}{2bR}\left(1+\sqrt{1+4bRA_1}\right) \\ &= \frac{1}{bR}\left(\frac{1}{2}\left(1+\sqrt{1+4bRA_1}\right)\right) \\ &\equiv \frac{1}{bR}\Delta(b,R,A_1) \end{aligned} \tag{60}$$

Notice that $\nu > 0$ ensures $A_1 > 0$ which in turn ensures $\Delta(b, R, A_1) > 1$. Further, $\tilde{x}_{\text{lowerbound}} = \frac{1}{b}$ is the lower bound on cash such that there is no suicide. Then the interval

$$\tilde{x} \in [\tilde{x}_{\text{lowerbound}}, \tilde{x}_{\text{kickoff1}}] \tag{61}$$

is non-empty for $\nu > 0$ and characterizes the cash region for stage 1 without suicide. Finally, we can map this characterization of phase 1 in terms of cash spent in old age $\tilde{x}$ back into overall cash x.

$$x \in [x_{\text{lowerbound}}, x_{\text{kickoff1}}] = \left[\frac{\tilde{x}_{\text{lowerbound}}}{(1-\vartheta_{\text{phase 1},c})}, \frac{\tilde{x}_{\text{kickoff1}}}{(1-\vartheta_{\text{phase 1},c})}\right] \tag{62}$$

B.1.4 Interior Solution of Optimal Health Investment

In the interior solution we cannot solve for the shares analytically. Instead we have a system of two equations given by

$$\begin{aligned}(1-\beta)u_y' &= \beta\psi u_o' R(1-\vartheta_{t,e}) - \beta\psi' \frac{\partial i_t}{\partial \vartheta_{t,c}} \frac{u_o}{x} \\ \frac{\psi}{\psi'} x_t R(1-\vartheta_{t,c}) &= \frac{u_o}{u_o'} \frac{\partial i_t}{\partial \vartheta_{t,e}}\end{aligned}$$

Where health investment i_t and its derivatives $(\frac{\partial i_t}{\partial \vartheta_{t,c}}, \frac{\partial i_t}{\partial \vartheta_{t,e}})$ both depend on whether we are in stage 2 or 3 (both stages are interior solutions with respect to the optimal $\vartheta_{t,e}$). We know that we move from stage 2 to stage 3 when e_t is sufficiently large, formally when $e_t = (1-\vartheta_{t,c})\vartheta_{t,e}(x_t)x_t > \tilde{\lambda}_t - \nu$. For a given x_t we can solve the FOCs and find $\vec{\vartheta}(x_t) = (\vartheta_{t,c}(x_t), \vartheta_{t,e}(x_t))$ with the following steps

1. Guess $\hat{\vec{\vartheta}}$
2. Determine the current stage through $\hat{e} = (1-\hat{\vartheta}_{t,c})\hat{\vartheta}_{t,e}x$. Specifically,

$$\text{Stage} = \begin{cases} 2, & \text{if} \quad \hat{e} \in (0, \tilde{\lambda}_t - \nu] \\ 3, & \text{if} \quad \hat{e} > \tilde{\lambda}_t - \nu \end{cases}$$

3. Compute $\hat{c}_t^y$, $\hat{c}_{t+1}^o$, $\hat{i}$, $\frac{\partial \hat{i}}{\partial \hat{\vartheta}_{t,e}}$ and $\frac{\partial \hat{i}}{\partial \hat{\vartheta}_{t,c}}$ based on the stage and check if the FOCs hold

$$\hat{i}_f = \begin{cases} \hat{e} = (1-\hat{\vartheta}_{t,c})\hat{\vartheta}_{t,e}x, & \text{if Stage} = 2 \\ \tilde{\lambda} - \nu, & \text{if Stage} = 3 \end{cases}$$

$$\hat{i}_h = \begin{cases} 0, & \text{if Stage} = 2 \\ \frac{\hat{e}-\hat{i}_f}{p_h} = \frac{(1-\hat{\vartheta}_{t,c})\hat{\vartheta}_{t,e}x-\hat{i}_f}{p_h}, & \text{if Stage} = 3 \end{cases}$$

$$\frac{\partial \hat{i}}{\partial \hat{\vartheta}_{t,e}} = \begin{cases} \zeta(\nu+\hat{i}_f)^{\zeta-1}(1-\hat{\vartheta}_{t,c})x, & \text{if Stage} = 2 \\ \eta_h \frac{(1-\hat{\vartheta}_{t,c})x}{p_h}, & \text{if Stage} = 3 \end{cases}$$

$$\frac{\partial \hat{i}}{\partial \hat{\vartheta}_{t,c}} = \begin{cases} -\zeta(\nu+\hat{i}_f)^{\zeta-1}\hat{\vartheta}_{t,e}x, & \text{if Stage} = 2 \\ -\eta_h \frac{\hat{\vartheta}_{t,e}x}{p_h}, & \text{if Stage} = 3 \end{cases}$$

4. Repeat until $\hat{\vec{\vartheta}}$ is found that solves the FOCs

B.1.5 Interior Solution on the BGP

We solve for the BGP using $x_t \to \infty$ and $\psi(i_t) = 1$. In the interior solution of the BGP we have

$$\frac{\partial i_t}{\partial \vartheta_{t,e}} = \frac{\partial}{\partial \vartheta_{t,e}} \frac{1}{p_t} \left(\vartheta_{t,e}(1-\vartheta_{t,c})x_t + \nu_t + \frac{1-\zeta}{\zeta}\tilde{\lambda}_t \right) = \frac{(1-\vartheta_{t,c})x_t}{p_t}$$

Then the 2. FOC becomes

$$\frac{1}{\xi}(1+i_t)\left((1+i_t)^{\xi}-1\right)x_t R(1-\vartheta_{t,c}) = \left(\frac{1}{1-\sigma} + bc_{t+1}^{\sigma-1}\right)c_{t+1}\frac{(1-\vartheta_{t,c})x_t}{p_t}$$

$p_t = p$ is constant, x_t converges to infinity and $i_t \to \eta_h i_{ht} \to \frac{e_t}{p_t} = \frac{(1-\vartheta_{t,c})\vartheta_e x}{p_t}$. Thus i_t and c_{t+1} are both constant shares of cash x_t on the BGP. For a BGP to exist we therefore require $\xi = \sigma - 1$. Plugging in

$$i = \frac{(1-\vartheta_c)\vartheta_e x}{p}$$
$$c = R(1-\vartheta_e)(1-\vartheta_c)x$$

Solving for the limit case where $x \to \infty$ we can find the health expenditure share on the BGP

$$\vartheta_e^* = \left(1 + \left[\frac{(pR)^{1-\sigma}}{b\xi}\right]^{\frac{1}{\sigma}}\right)^{-1}$$

Plugging ϑ_e^* into the Euler equation yields the BGP share of young consumption in cash

$$\vartheta_c^* = \left[1 + \left(\frac{\beta}{1-\beta}\psi\left[R(1-\vartheta_e^*)\right]^{1-\sigma}\right)^{\frac{1}{\sigma}}\right]^{-1}$$

B.2 Model without Young Consumption: Characterization of Health Investment

If we abstract from consumption at young age, $\beta = 1$, we can characterize the full household solution analytically along the transition, including the cash levels at which both kickoffs happen. We again first solve for the optimal split between final goods and health goods for a given amount of health expenditures e_t. Then we solve for the optimal amount of health expenditures e_t. Thus, we start by solving

$$\begin{aligned} i_t &= i_t(p_t, e_t) = \max_{i_{ft}, i_{ht}} f(i_{ft}, i_{ht}) \\ s.t.\ p_t i_{ht} + i_{ft} &= e_t \\ i_{ft}, i_{ht} &\geq 0 \\ f(i_{ft}, i_{ht}) &= i_{ht} + (\nu + i_{ft})^{\zeta} \end{aligned}$$

Corner Solution with $i_{ht} = 0$, $i_{ft} = e_t$: In the corner solution the first-order conditions do not hold with equality. From the budget constraint we get $i_{ht} = 0$, $i_{ft} = e_t$. Then health investment i_t is given by

$$i_t = f(i_{ht}, i_{ft}) = i_{ht} + (\nu + i_{ft})^{\zeta} = (\nu + e_t)^{\zeta}$$

Note, for $e_t = 0$ the corner solution corresponds to stage 1 (no health expenditures), for $e_t > 0$ it corresponds to stage 2 (positive health expenditures but fully invested into final goods).

Interior Solution: In the interior solution the first-order conditions hold with equality, they are given by

$$\begin{aligned} \zeta\left(\nu + i_{ft}\right)^{\zeta - 1} &= \lambda \\ 1 &= \lambda p_t \end{aligned}$$

Combining yields

$$\begin{aligned}\zeta\left(\nu+i_{ft}\right)^{\zeta-1} &= \frac{1}{p_t}\\ \nu+i_{ft} &= \left(\zeta p_t\right)^{\frac{1}{1-\zeta}}\\ i_{ft} &= \left(\zeta p_t\right)^{\frac{1}{1-\zeta}}-\nu\end{aligned}$$

Define $\tilde{\lambda}_t \equiv \left(\zeta p_t\right)^{\frac{1}{1-\zeta}}$. Then

$$i_{ft} = \tilde{\lambda}_t - \nu$$

The quasi-linear specification means that in the interior solution there are no wealth effects for the concave good i_{ft}. Thus, $i_{ft} = \tilde{\lambda}_t - \nu$ is constant and independent of e_t. Using the budget constraint, we get

$$i_{ht} = \frac{e_t - (\tilde{\lambda}_t - \nu)}{p_t}$$

For $\zeta \to 0$ we get $\tilde{\lambda}_t \to 0$ and, thus, taking into account the non-negativity constraints $(i_{ft}, i_{ht}) \geq 0$

$$\begin{aligned}i_{ft} &\to 0\\ i_{ht} &\to \frac{e_t}{p_t}\\ i_t &\to \frac{e_t}{p_t}+\nu^{\zeta}\end{aligned}$$

Thus, ζ needs to be sufficiently large for stage 2 to exist, otherwise the marginal benefit from final goods investment is too small.

Existence of Stage 2: We want to ensure that initial health investment will be allocated towards i_{ft}, that is, that the corner solution with $i_{ht} = 0$ and $i_{ft} = e_t$ characterized above exists. We get the undesired the corner solution with $i_{ht} > 0$ and $i_{ft} = 0$ (stage 2 is skipped with initial health investment directly being allocated towards i_{ht}) if the marginal cost of investing into i_{ft} exceeds marginal benefit, evaluated at $i_{ft} = 0$:

$$\begin{aligned}\zeta\left(\nu+i_{ft}\right)^{\zeta-1} &\leq \tfrac{1}{p_t}\\ \zeta\left(\nu\right)^{\zeta-1} &\leq \tfrac{1}{p_t}\\ \left(\zeta p_t\right)^{\frac{1}{1-\zeta}} &\leq \nu\\ \tilde{\lambda}_t &\leq \nu\end{aligned}$$

Then stage 2 exists as long as the optimal interior solution computed above yields $i_{ft} = \tilde{\lambda}_t - \nu > 0$ which requires the non-homotheticity factor ν to be sufficiently small relative to health sector price p_t and ζ.

Characterizing the Stages: For a given level of health expenditures e_t we can fully characterize the stages now. Assuming $\tilde{\lambda}_t > \nu$ for existence of stage 2, the stages are then characterized by

$$\text{Stage} = \begin{cases} 1, & \text{if} \quad e_t = 0 \\ 2, & \text{if} \quad e_t \in (0, \tilde{\lambda}_t - \nu] \\ 3, & \text{if} \quad e_t > \tilde{\lambda}_t - \nu \end{cases}$$

B.2.1 Level of Health Expenditures

Given the optimal division of health investment we now optimize over the allocation of cash x_t into savings s_t and health expenditures e_t. That is, the household now solves

$$\max_{0 \leq e_t \leq x_t} \psi\left(i_t(p_t, e_t)\right) u(r_{t+1}\left[x_t - e_t\right]).$$

Define the share of health expenditures in total spending as

$$\vartheta_t = \frac{e_t}{x_t} = \frac{p_t i_{ht} + i_{ft}}{x_t} \in [0, 1],$$

Then we can rewrite the above problem in terms of ϑ_t

$$\max_{\vartheta_t \in [0,1]} \psi\left(i_t(p_t, \vartheta_t x_t)\right) u\left(R_{t+1} x_t \left(1 - \vartheta_t\right)\right).$$

The first-order condition is given by

$$0 = \psi' \frac{\partial i_t}{\partial \vartheta_t} u - \psi u' R_{t+1} x_t$$

Note, the first-order condition depends on $\frac{\partial i_t}{\partial \vartheta_t}$ which in turn depends on the optimal division of health investment and, thus, the level of ϑ_t. Rearranging yields

$$\frac{\psi}{\psi'} x_t R_{t+1} = \frac{u}{u'} \frac{\partial i_t}{\partial \vartheta_t} \tag{63}$$

B.2.2 Corner Solution of Optimal Health Investment

We get the corner solution with $\vartheta = 0$, corresponding to stage 1 without any health expenditure, when the FOC does not hold with equality evaluated at $i_t(p_t, e_t = 0)$ such that

$$\frac{\psi}{\psi'} x_t R_{t+1} \geq \frac{u}{u'} \frac{\partial i_t}{\partial \vartheta_t}$$

$$\text{or} \quad \frac{1}{\xi}(1+i_t)\left((1+i_t)^{\xi}-1\right) x_t R_{t+1} \geq \left(\frac{1}{1-\sigma} + b c_{t+1}^{\sigma-1}\right) c_{t+1} \frac{\partial i_t}{\partial \vartheta_t}$$

Thus, we get the corner solution when the marginal benefit from saving one more unit is larger than the marginal benefit from the first unit of health investment. To further characterize this condition we can plug in $i_t(p_t, e_t = 0) = \mu_t \nu^{\zeta}$. However, $\frac{\partial i_t}{\partial \vartheta_t}$ depends on the optimal division of health investment, that is, it depends on how the first marginal unit of health investment e_t is allocated between i_{ht} and i_{ft}. We assume $\tilde{\lambda}_t \geq \nu$ which, as shown above, ensures that initial health investment is allocated towards the final good sector i_{ft}. Then the derivative evaluated at $\vartheta_t = 0$ is given by

$$\frac{\partial i_t}{\partial \vartheta_t} = \frac{\partial}{\partial \vartheta_t}(\nu + e_t)^{\zeta} = \frac{\partial}{\partial \vartheta_t}(\nu + \vartheta_t x_t)^{\zeta} = \zeta x_t \nu^{\zeta-1}$$

Plugging $i_t(p_t, e_t = 0)$ and $\frac{\partial i_t}{\partial \vartheta_t}$ back into the corner solution condition yields

$$\frac{1}{\xi}\left(1+\nu^{\zeta}\right)\left(\left(1+\nu^{\zeta}\right)^{\xi}-1\right) R_{t+1} \geq \left(\frac{1}{1-\sigma} + b c_{t+1}^{\sigma-1}\right) c_{t+1} \zeta \nu^{\zeta-1}$$

Rewriting consumption in terms of overall cash in the first stage ($\vartheta_t = 0$) yields

$$c_{t+1} = R_{t+1} s_t = R_{t+1}(1-\vartheta_t) x_t = R_{t+1} x_t$$

Plugging back in then gives

$$\frac{1}{\xi}\left(1+\nu^{\zeta}\right)\left(\left(1+\nu^{\zeta}\right)^{\xi}-1\right) \geq \left(\frac{1}{1-\sigma} + b\left(R_{t+1} x_t\right)^{\sigma-1}\right) x_t \zeta \nu^{\zeta-1} \tag{64}$$

Thus, this equation characterizes the first stage: The region of cash x_t for which health expenditures are zero, $e_t = 0$. For convenience define

$$A_1 \equiv \frac{1}{\xi} \frac{\left(1+\nu^{\zeta}\right)\left(\left(1+\nu^{\zeta}\right)^{\xi}-1\right)}{\zeta \nu^{\zeta-1}} \begin{cases} = 0, & \text{if } \nu = 0 \\ > 0, & \text{if } \nu > 0 \end{cases}$$

Thus, $A_1 \geq 0$ and in our case we have $A_1 > 0$ because $\nu > 0$ is necessary for the existence of stage 2. Plugging A_1 back into the condition for the corner solution yields

$$A_1 \geq \left(\frac{1}{1-\sigma} + b \left(R_{t+1} x_t \right)^{\sigma-1} \right) x_t$$

To make progress analytically we need to make assumptions on the parameters. Assume $\sigma = 2$, then

$$\begin{aligned} A_1 &\geq -x_t + bR_{t+1} x_t^2 \\ 0 &\geq x_t^2 - \frac{1}{bR_{t+1}} x_t - \frac{A_1}{bR_{t+1}} \end{aligned}$$

The two solutions to this quadratic equation are given by

$$x_{1/2} = \frac{1}{2bR_{t+1}} \left(1 \pm \sqrt{1 + 4bR_{t+1}A_1} \right)$$

$A_1 \geq 0$ means that $\sqrt{1 + 4bR_{t+1}A_1} \geq 1$. Thus, $x_2 \leq 0$, then x_1 is the only economically relevant solution. The first kick-off separating stage 1 and stage 2 is then given by

$$\begin{aligned} x_{\text{kickoff1}} &= \frac{1}{2bR_{t+1}} \left(1 + \sqrt{1 + 4bR_{t+1}A_1} \right) \\ &= \frac{1}{bR_{t+1}} \left(\frac{1}{2} \left(1 + \sqrt{1 + 4bR_{t+1}A_1} \right) \right) \\ &\equiv \frac{1}{bR_{t+1}} \Delta(b, R_{t+1}, A_1) \end{aligned} \tag{65}$$

Notice that $\nu > 0$ ensures $A_1 > 0$ which in turn ensures $\Delta(b, R_{t+1}, A_1) > 1$. Further, $x_{\text{lowerbound}} = \frac{1}{bR_{t+1}}$ is the lower bound on cash such that there is no suicide. Then the interval

$$x \in [x_{\text{lowerbound}}, x_{\text{kickoff1}}] \tag{66}$$

is non-empty for $\nu > 0$ and characterizes the cash region for stage 1 without suicide.

B.2.3 Interior Solution of Optimal Health Investment

In the interior solution with positive health expenditures, $\vartheta_t > 0$, the FOC holds with equality

$$\frac{1}{\xi} (1 + i_t) \left((1 + i_t)^{\xi} - 1 \right) x_t R_{t+1} = \left(\frac{1}{1-\sigma} + b c_{t+1}^{\sigma-1} \right) c_{t+1} \frac{\partial i_t}{\partial \vartheta_t}$$

Plugging in $c_{t+1} = R_{t+1}s_t = R_{t+1}(1-\vartheta_t)x_t$ yields

$$\frac{1}{\xi}(1+i_t)\left((1+i_t)^{\xi}-1\right) = \left(\frac{1}{1-\sigma} + b\left(R_{t+1}(1-\vartheta_t)x_t\right)^{\sigma-1}\right)(1-\vartheta_t)\frac{\partial i_t}{\partial \vartheta_t}$$

As before, in order to solve this equation for the share of health expenditure as a function of cash $\vartheta(x_t)$ we need to plug in i_t and $\frac{\partial i_t}{\partial \vartheta_t}$ which now both depend on whether we are in stage 2 or 3 (both stages are interior solutions with respect to the optimal ϑ_t). We know that we move from stage 2 to stage 3 when e_t is sufficiently large, formally when $e_t = \vartheta_t(x_t)x_t > \tilde{\lambda}_t - \nu$. For a given $x_t = \hat{x}$ we can solve the FOC and find $\vartheta(\hat{x})$ with the following steps

1. Guess $\hat{\vartheta}(\hat{x})$
2. Determine the current stage through $\hat{e} = \hat{\vartheta}(\hat{x})\hat{x}$. Specifically,

$$\text{Stage} = \begin{cases} 2, & \text{if} \quad \hat{e} \in (0, \tilde{\lambda}_t - \nu] \\ 3, & \text{if} \quad \hat{e} > \tilde{\lambda}_t - \nu \end{cases}$$

3. Compute $\hat{i}$ and $\frac{\partial \hat{i}}{\partial \hat{\vartheta}}$ based on the stage and check if the FOC holds
4. Repeat until $\hat{\vartheta}(\hat{x})$ that solves the FOC is found

In order to further characterize the level of cash x_{kickoff2} at which the second kick-off happens, we can utilitze that at the second kick-off total health expenditures e_t exactly equal the interior level of final good investment

$$\tilde{\lambda}_t - \nu = \vartheta_t(x_{\text{kickoff2}})x_{\text{kickoff2}}$$

Moreover, at the second kick-off health investment and its derivative w.r.t. the expenditure share are given by

$$\begin{aligned} i_t &= \lambda_t^{\zeta} \\ \frac{\partial i_t}{\partial \vartheta_t} &= \frac{\lambda_t - \nu}{p_t \vartheta_t} \end{aligned}$$

Plugging those two into the FOC yields

$$\frac{1}{\xi}\left(1+\lambda_t^{\zeta}\right)\left(\left(1+\lambda_t^{\zeta}\right)^{\xi}-1\right) = \left(\frac{1}{1-\sigma} + b\left(R_{t+1}(1-\vartheta_t)x_t\right)^{\sigma-1}\right)(1-\vartheta_t)\frac{\lambda_t - \nu}{p_t\vartheta_t}$$

Now using the fact that at the second kick-off $\tilde{\lambda}_t - \nu = \vartheta_t(x)x$ we get

$$\frac{1}{\xi}\left(1+\lambda_t^{\zeta}\right)\left(\left(1+\lambda_t^{\zeta}\right)^{\xi}-1\right) = \left(\frac{1}{1-\sigma} + b\left(R_{t+1}x - R_{t+1}(\lambda_t-\nu)\right)^{\sigma-1}\right)(1-\vartheta_t)\frac{\lambda_t-\nu}{p_t\vartheta_t}$$

As before assume $\sigma = 2$, then

$$\frac{1}{\xi}\left(1+\lambda_t^{\zeta}\right)\left(\left(1+\lambda_t^{\zeta}\right)^{\xi}-1\right) = \left(-1 + b\left(R_{t+1}x - R_{t+1}(\lambda_t-\nu)\right)\right)\frac{(1-\vartheta_t)}{\vartheta_t}\frac{\lambda_t-\nu}{p_t}$$

We can again use $\tilde{\lambda}_t - \nu = \vartheta_t(x)x$ to derive

$$\frac{(1-\vartheta_t)}{\vartheta_t} = \frac{x-(\lambda_t-\nu)}{(\lambda_t-\nu)} = \frac{x}{(\lambda_t-\nu)} - 1$$

and thereby eliminate the remaining ϑ_t to get

$$\frac{1}{\xi}\left(1+\lambda_t^{\zeta}\right)\left(\left(1+\lambda_t^{\zeta}\right)^{\xi}-1\right) = \left(-1 + b\left(R_{t+1}x - R_{t+1}(\lambda_t-\nu)\right)\right)\left(\frac{x}{(\lambda_t-\nu)} - 1\right)\frac{\lambda_t-\nu}{p_t}$$

Similar to before when finding the first kick-off, again define for convenience

$$A_2 \equiv \frac{p_t}{\xi}\left(1+\lambda_t^{\zeta}\right)\left(\left(1+\lambda_t^{\zeta}\right)^{\xi}-1\right)\begin{cases} = 0, & \text{if } \nu = 0 \\ > 0, & \text{if } \nu > 0 \end{cases}$$

Then the FOC at the second kick-off reduces to

$$A_2 = \left(x - (\lambda_t-\nu)\right)\left(bR_{t+1}x - bR_{t+1}(\lambda_t-\nu) - 1\right)$$

This is again a quadratic equation in x, simplifying further yiels

$$0 = x^2 - x\frac{2(\lambda_t-\nu)bR_{t+1}+1}{bR_{t+1}} + \frac{(\lambda_t-\nu)^2bR_{t+1} + (\lambda_t-\nu) - A_2}{bR_{t+1}}$$

The roots of this equation are given by

$$x_{1/2} = \frac{2(\lambda_t-\nu)bR_{t+1}+1}{2bR_{t+1}} \pm \sqrt{\left(\frac{2(\lambda_t-\nu)bR_{t+1}+1}{2bR_{t+1}}\right)^2 - \frac{(\lambda_t-\nu)^2bR_{t+1} + (\lambda_t-\nu) - A_2}{bR_{t+1}}}$$

which simplifies to

$$\begin{aligned} x_{1/2} &= \frac{1}{2bR_{t+1}}\left(2(\lambda_t-\nu)bR_{t+1}+1\pm\sqrt{1+4A_2bR_{t+1}}\right) \\ &= (\lambda_t-\nu)+\frac{1}{2bR_{t+1}}\left(1\pm\sqrt{1+4A_2bR_{t+1}}\right) \end{aligned}$$

This looks very similar to the solution for the first kick-off. The new term in the beginning $(\lambda_t-\nu)$ is exactly the lower bound for the second kick-off because by definition it has to be that $x_{\text{kickoff2}} \geq i_{ft} = (\lambda_t-\nu)$. Since as before $4A_2bR_{t+1} > 0$ if $\nu > 0$ the negative solution violates that condition and can be ruled out. Then the second kick-off is given by

$$\begin{aligned} x_{\text{kickoff2}} &= (\lambda_t-\nu)+\frac{1}{2bR_{t+1}}\left(1+\sqrt{1+4bR_{t+1}A_2}\right) \\ &= (\lambda_t-\nu)+\frac{1}{bR_{t+1}}\left(\frac{1}{2}\left(1+\sqrt{1+4bR_{t+1}A_2}\right)\right) \\ &\equiv (\lambda_t-\nu)+\frac{1}{bR_{t+1}}\Delta(b,R_{t+1},A_2) \qquad (67) \\ &= i_{ft}+\frac{1}{bR_{t+1}}\Delta(b,R_{t+1},A_2) \end{aligned}$$

Notice that the wedge $\Delta(b,R_{t+1},A)$ appears exactly the same way in both kick-offs. The only difference is the value of A plugged into the wedge function. To compare the relative size of those wedges first note that the wedge is increasing in A and further notice that

$$\begin{aligned} A_2 &= \frac{p_t}{\xi}\left(1+\lambda_t^{\zeta}\right)\left(\left(1+\lambda_t^{\zeta}\right)^{\xi}-1\right) \\ &> \frac{p_t}{\xi}\left(1+\nu^{\zeta}\right)\left(\left(1+\nu^{\zeta}\right)^{\xi}-1\right) \\ &> \frac{1}{\xi}\frac{\left(1+\nu^{\zeta}\right)\left(\left(1+\nu^{\zeta}\right)^{\xi}-1\right)}{\zeta\nu^{\zeta-1}} \\ &= A_1 \end{aligned}$$

where the last inequality comes from the FOC of the optimal division of health investment $\zeta\left(\nu+i_{ft}\right)^{\zeta-1} = \frac{1}{p_t}$ and, thus, $p_t > \zeta\left(\nu\right)^{\zeta-1}$ for $i_{ft} > 0$. We can directly link the wedges to savings at the kick-offs. Then the amount by which savings grow between the two kick-offs is given by

$$\Delta s = s_{\text{kickoff2}} - s_{\text{kickoff1}} = \frac{1}{bR_{t+1}}\left(\Delta(b,R_{t+1},A_2)-\Delta(b,R_{t+1},A_1)\right)$$

Thus, it makes sense that we get $\Delta(b, R_{t+1}, A_2) > \Delta(b, R_{t+1}, A_1)$. Also note that if $\lambda_t \leq \nu$ (stage 2 does not exist), we get $A_2 = A_1$ and the two kick-offs derived above collapse to the same cash value. (not yet sure expressing quantities in terms of these wedges is particularly useful but it is a good sanity check).

B.2.4 Interior Solution on the BGP

In the interior solution of the BGP we have

$$\frac{\partial i_t}{\partial \vartheta_t} = \frac{\partial}{\partial \vartheta_t} \frac{1}{p_t} \left(\vartheta_t x_t + \nu_t + \frac{1-\zeta}{\zeta} \tilde{\lambda}_t \right) = \frac{x_t}{p_t}$$

Then the FOC for ϑ_t becomes

$$\frac{1}{\xi} (1 + i_t) \left((1 + i_t)^{\xi} - 1 \right) p_t R_{t+1} = \left(\frac{1}{1-\sigma} + b c_{t+1}^{\sigma - 1} \right) c_{t+1}$$

p_t is constant, x_t converges to infinity and $i_t \to i_{ht} \to \frac{e_t}{p_t}$. Thus i_t and c_{t+1} are both constant shares of cash x_t on the BGP. For a BGP to exist we therefore require $\xi = \sigma - 1$. Plugging in

$$\begin{aligned} i &= \frac{\vartheta x}{p} \\ c &= R(1 - \vartheta) x \end{aligned}$$

and solving for the limit case where $x \to \infty$ we can find the health expenditure share on the BGP

$$\vartheta^* = \left(1 + \left[\frac{(pR)^{1-\sigma}}{b\xi} \right]^{\frac{1}{\sigma}} \right)^{-1}$$

B.2.5 Summarizing the Phases

To summarize, the three phases are characterized by three separate region of cash at hand. The thresholds are given by

$$\begin{aligned} x_{\text{LB}} &= \frac{1}{b R_{t+1}} = \frac{1}{b R_2} \\ x_{\text{KO1}} &= \frac{1}{b R_{T_1+1}} \Delta(b, R_{T_1+1}, A_1) \\ x_{\text{KO2}} &= (\lambda_t - \nu) + \frac{1}{b R_{T_2+1}} \Delta(b, R_{T_2+1}, A_2) \end{aligned}$$

where T_1 and T_2 are the points in time at which the two kick-offs occur. The second equality in the first equation comes from the assumption that cash starts at its lower bound in period 1

$x_1 = \frac{1}{bR_2}$. Recall that

$$
\begin{aligned}
A_1 &= \frac{1}{\xi} \frac{\left(1+\nu^{\zeta}\right)\left(\left(1+\nu^{\zeta}\right)^{\xi}-1\right)}{\zeta \nu^{\zeta-1}} \\
A_2 &= \frac{p_t}{\xi}\left(1+\lambda_t^{\zeta}\right)\left(\left(1+\lambda_t^{\zeta}\right)^{\xi}-1\right) \\
\Delta(b, R_{t+1}, A) &= \frac{1}{2}\left(1+\sqrt{1+4bR_{t+1}A}\right)
\end{aligned}
$$

Note that A_1 only depends on time-invariant parameters. Then the three regions are given by

$$
\text{Phase} = \begin{cases} 1, & \text{if} \quad x_t \in [x_{\text{LB}}, x_{\text{KO1}}) \\ 2, & \text{if} \quad x_t \in [x_{\text{KO1}}, x_{\text{KO2}}) \\ 3, & \text{if} \quad x_t \in [x_{\text{KO2}}, \infty) \end{cases}
$$