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

Leon Huetsch† Dirk Krueger‡ Alexander Ludwig§

November 2, 2023

Abstract

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. 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 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, 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.

JEL Codes: E13, O41, I15

Keywords: Life Expectancy, Modern Health Sector, Endogenous Technological Progress

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

†University of Pennsylvania

‡University of Pennsylvania, CEPR and NBER

§Goethe University Frankfurt, ICIR and CEPR.

1 Introduction

In 1820 the remaining life expectancy[1] of a twenty year old person living in the United States was approximately 40 years, per capita income was $\$2,674$ (in constant 2011 US dollars, according to the Maddison data base) and virtually none of that income was spent on health goods and services, abstracting from expenditures on basic food and hygiene. In 2019, the year prior to the COVID-19 pandemic, remaining life expectancy at age twenty stood above 60 years, income per capita rose to 55,355 (again in constant 2011 US dollars), and close to 20% of that income was spent on goods produced by a modern, high-tech health sector. 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 relative price for health goods, we ask what share of the increase in life expectancy from 1940 onward can be attributed to the modern health sector. We find that share to be approximately 30%, suggesting that the increased expenditures on basic health goods (e.g., better food and hygiene) played a major role in the expansion of life expectancy even in the 20th century, and even in the presence of a modern health sector that approaches an expenditure share of 20% of total GDP.

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

Second, we use the fact that the model accounts well for the increase in the relative price of health goods and services between 1940 and 2020 to decompose its increase 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 quality-adjusted 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 it builds on the work by Cervellati and Sunde (2005) who develop a model of the take-off of life expectancy by modeling a 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 adaption 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 (2018) 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.[2]

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. The normative perspective includes the work by Hall and Jones (2007), Jones (2004), 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, 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—, or due to perhaps imperfectly measured quality improvements as our model implies.[3] 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 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.[4]

[2]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).

[3]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 (), who find that health price indices profoundly underestimate the quality improvements documented in, e.g., Cutler and McClellan (2001).

[4]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

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.

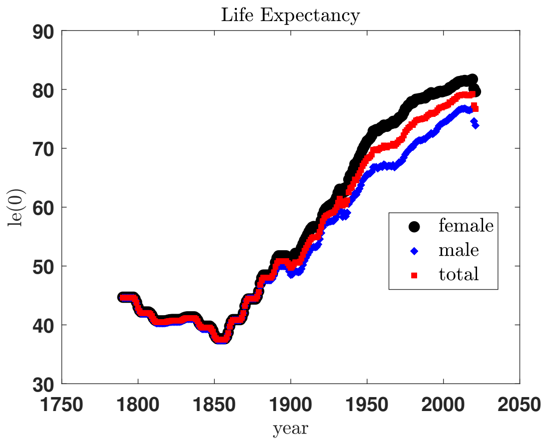
2 Stylized Facts

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.

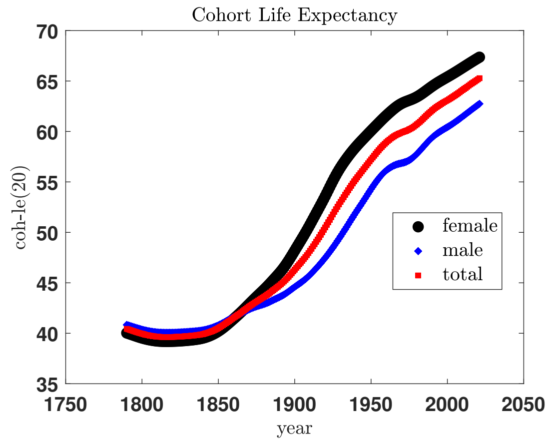
2.1 Life Expectancy

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, 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 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.

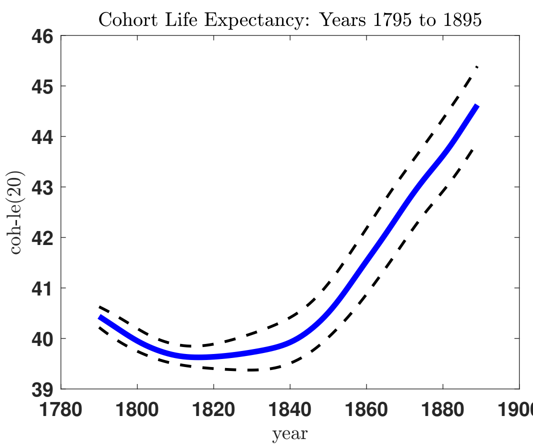
costly health products, see Chandra and Skinner (2012). This does not rule out the possibility that the U.S. health care system is plagued by larger inefficiencies than the systems in other industrialized countries, a view consistent with the cross-country evidence in Retzlaff-Roberts, Chang, and Rubin (2004). 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).



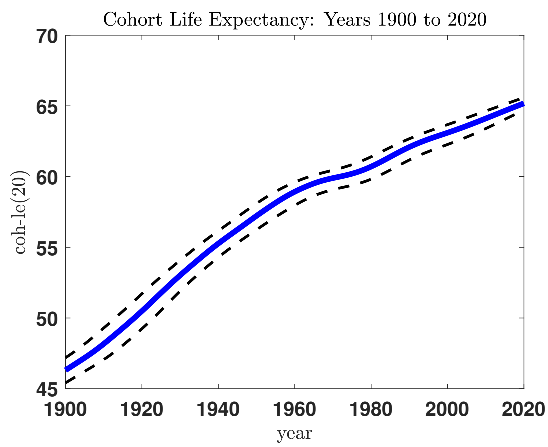
(a) Cross-Sectional Life Expectancy at Birth



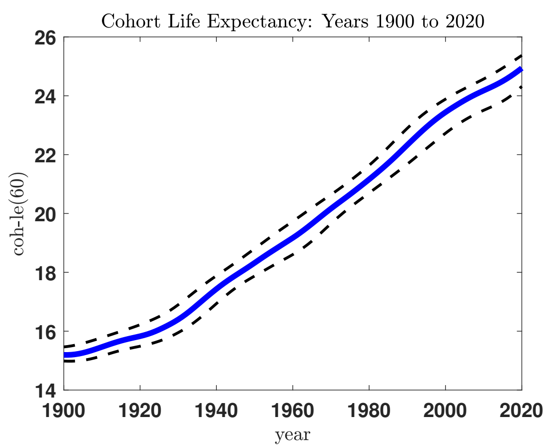
(b) Remaining Cohort Life Expectancy at Age 20



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



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



(e) Remaining Cohort Life Expectancy at Age 60: after 1900

Figure 1: Life-Expectancy in the US

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.[5] 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 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.[6] 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. In contrast to these diminishing gains, the path of cohort life expectancy at age 60—which we display in panel (e)—indicates increasing gains. This visualizes that in the past century, gains in life expectancy increasingly took place at older ages, which points to the importance of a modern health sector for shaping life expectancy.

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.[7] 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

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

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

[7]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.

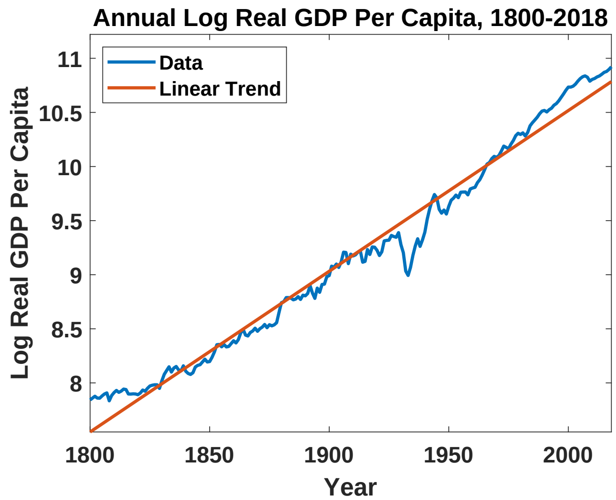


Figure 2: Income per Capita in the U.S., 1800-2018

Source: Maddison Project Data Base 2020.

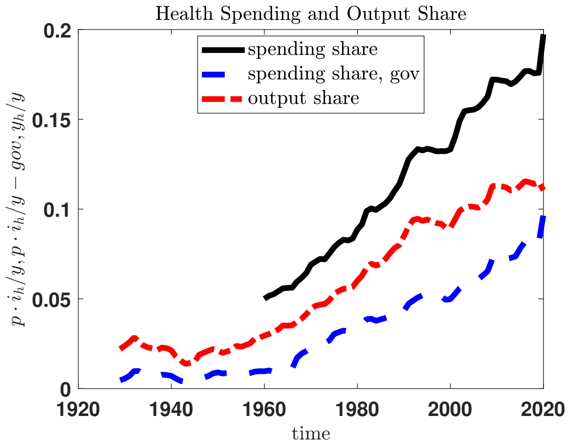
well-documented fact that income per capita started to rise significantly around 1820 and since then has grown at a roughly constant (in fact slightly accelerating) pace of approximately 2% per year.

2.3 The Emergence and Evolution of the Modern Health Sector

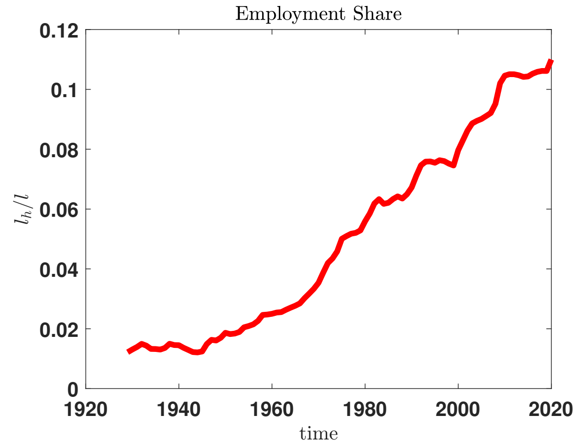
The third set of facts that motivate our analysis is the emergence and growth of a modern health sector in the 20th century. Parker (2019) dates the start of the era of modern medicine at ca. 1920 in his book, pointing to landmark breakthroughs such as the discovery of Penicillin in 1928 and the start of its mass production towards the end if WWII, the discovery and analysis of blood 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.[8] 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.

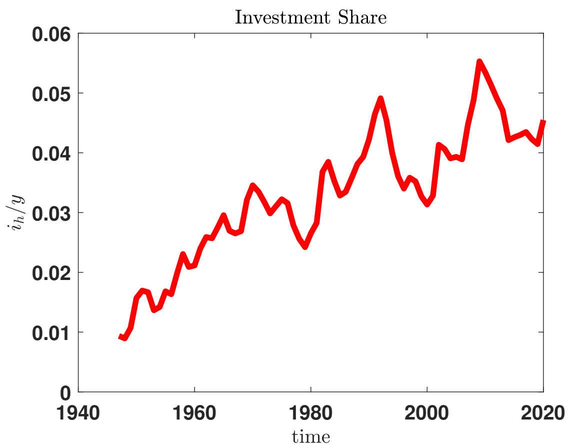
[8]Since we display shares, there is no need to deflate the measures by prices



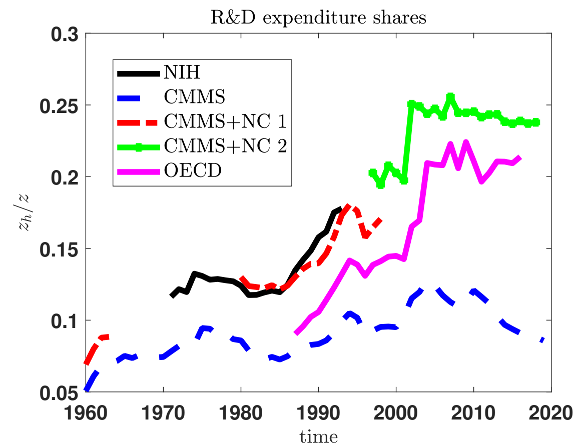
(a) Health Spending Share



(b) Employment Share



(c) Investment Share



(d) R&D Spending Share

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

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.

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.

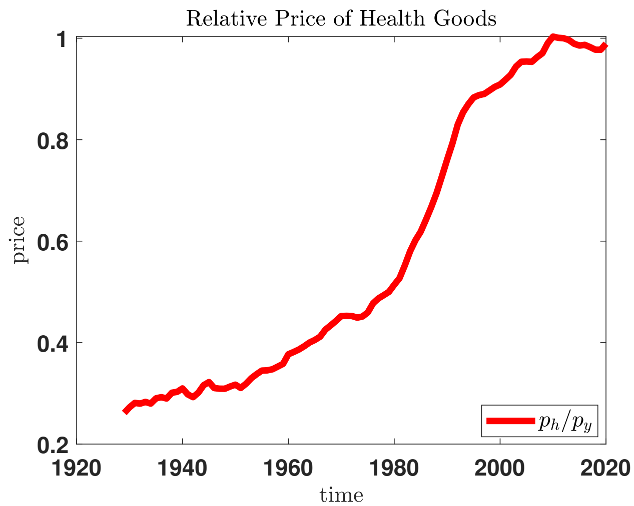


Figure 4: The Relative Price of Health Goods in the U.S.

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.

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, 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.

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, final goods firms in the consumption good sector and the health sector that operate under perfect competition, and 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.

3.1 Households

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 that depends on their investment i_t into health goods when young. The utility of being dead is set to zero, and therefore expected lifetime utility is given by[9]

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

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 $\lim_{c\to 0} u'(c) = \infty$. The survival function $\psi(i)$ is increasing in health investment i, that is, $\psi'(i) > 0$. We further assume that $\lim_{i\to\infty} \psi(i) = 1$ and $0 < \psi(0) < 1$ and $\psi'(0) < \infty$ that is, households survive with positive probability into the second period even absent any health investment, and the marginal benefit of health investment is finite at $i = 0$.

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,

$$i_t = f(i_{ft}, i_{ht}) \tag{2}$$

Young households supply labor to both sectors of the economy, with l_{ft} denoting labor supply to the final goods sector and l_{ht} denoting the corresponding supply to the health sector. To model, in a reduced form, the frictions associated with butchers (workers in the consumption sector) becoming surgeons (workers in the health sector) we assume the following effective constraint on the labor supplied by the unit mass of households of the form:

$$1 = g(l_{ft}, l_{ht}). \tag{3}$$

A constraint of the form of (3) is sometimes used in multi-sector models to model, in a simple manner, imperfect labor mobility across sectors.[10] It also implies that wages can potentially differ across the two sectors.

We choose the final consumption good as the numeraire and denote by p_t the price of goods produced by the modern health sector, and by $w_{jt}, j \in \{f, h\}$ wages in the two sectors of

[9] Since households survive the first period of life for sure, and since we assume that they only value consumption in the second period, the level of utility from being alive in the first period is immaterial.

[10] E.g., Giagheddu and Papetti (2019) calibrate function $g(.)$ referencing evidence by Cardi and Restout (2015).

production. 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 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 exogenous. The maximization of the utility function (1) is then subject to the constraints:

$$c_t^y + i_{ft} + p_t i_{ht} + s_t = w_{ft}l_{ft} + w_{ht}l_{ht} + T_t \tag{4a}$$

$$i_t = f(i_{ft}, i_{ht}) \tag{4b}$$

$$1 = g(l_{ft}, l_{ht}) \tag{4c}$$

$$c_{t+1}^o = Rs_t. \tag{4d}$$

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

$$c_t^y + i_{ft} + p_t i_{ht} + \frac{c_{t+1}^o}{R} = w_{ft}l_{ft} + w_{ht}l_{ht} + T_t \equiv x_t. \tag{5}$$

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

$$T_t = Rs_{t-1}(1 - \psi(i_{t-1})). \tag{6}$$

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

Let $j \in \{f, h\}$ stand for the final and the health sector of the economy, respectively, and p_{jt} for the price of the output of each of the two sectors. We normalize p_{ft} to 1 and simply let $p_t = p_{ht}$ denote the relative price of health goods whenever it notationally is more convenient and there is no room for confusion. 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

$$y_{jt} = \left(\int_0^1 q_{jit}^{1-\alpha_j} y_{jit}^{\alpha_j} di \right) l_{jt}^{1-\alpha_j}, \tag{7}$$

where $0 < \alpha_j < 1$ 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} and the price of one unit of intermediate good i in sector j is p_{jit}. The first order conditions are

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

for labor demand and

$$p_{jt} \alpha_j q_{jit}^{1-\alpha_j} y_{jit}^{\alpha_j - 1} l_{jt}^{1-\alpha_j} = p_{jit} \tag{9}$$

for the demand for intermediate goods, given their quality q_{jit}.

3.2.2 Intermediate Goods Producers

Each intermediate good producer i is a monopolist that takes the demand function (9) as given and uses capital (which depreciates immediately after use) to produce the intermediate good according to:

$$y_{jit} = k_{jit}. \tag{10}$$

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,

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

with first order condition

$$y_{jit} = \left(\frac{p_{jt} \alpha_j^2}{R} \right)^{\frac{1}{1-\alpha_j}} q_{jit} l_{jt} \tag{11}$$

and profits

$$\pi_{jit} = \frac{1-\alpha_j}{\alpha_j} R y_{jit} > 0. \tag{12}$$

The monopolistic price follows from using (11) in (9) as

$$p_{jit} = \frac{1}{\alpha_j} R > R, \tag{13}$$

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

Finally, observe from (11) 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 (11) in (12) as

$$\frac{\pi_{jit}}{q_{jit}} = \frac{1-\alpha_j}{\alpha_j} \left(\frac{p_{jt}\alpha_j^2}{R^{\alpha_j}} \right)^{\frac{1}{1-\alpha_j}} l_{jt}. \tag{14}$$

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

$$k_{jt} = \int_0^1 k_{jit} di = \int_0^1 y_{jit} di = \left(\frac{p_{jt}\alpha_j^2}{R} \right)^{\frac{1}{1-\alpha_j}} q_{jt} l_{jt} \tag{15}$$

where

$$q_{jt} = \int_0^1 q_{jit} di \tag{16}$$

is an aggregate quality index of intermediate inputs in sector j. Furthermore, exploiting (11) and (15) in (7) yields as aggregate production function for sector j

$$y_{jt} = k_{jt}^{\alpha_j} \left(q_{jt} l_{jt} \right)^{1-\alpha_j}. \tag{17}$$

Using equations (8) and (15) delivers as factor prices for labor inputs and capital inputs:

$$w_{jt} = (1 - \alpha_j) \frac{p_{jt} y_{jt}}{l_{jt}} \tag{18a}$$

$$R = \alpha_j^2 \frac{p_{jt} y_{jt}}{k_{jt}}. \tag{18b}$$

Finally we can use (12) and (15) to determine aggregate profits in each sector j as

$$\pi_{jt} = \alpha_j (1 - \alpha_j) p_{jt} y_{jt} \tag{19}$$

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

$$p_{jt} y_{jt} = \pi_{jt} + R k_{jt} + w_{jt} l_{jt} \tag{20}$$

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

$$q_{jit} = \lambda_j q_{jit-1} \tag{21}$$

where $\lambda_j > 1$ is a parameter. The successful innovator immediately becomes the monopolist, and for one period enjoys monopoly profits π_{jit} associated with technology level $q_{jit} = \lambda_j q_{jit-1}$. 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 the quality reached when successfully innovating given by λq_{jit-1} as well as the size of the corresponding final production sector given by the employment share l_{jt} so that

$$\phi_j(z_{jit}; l_{jt}, q_{jit-1}) = \min \left[\varphi_j \left(\frac{z_{jit}}{\lambda_j q_{jit-1}} \right)^{\gamma_j} \cdot l_{jt}^{-1}, 1 \right], \tag{22}$$

with $\gamma_j \in (0,1)$ and $\varphi_j > 0$. First, an increase in the scale of the final production sector measured as the employment share l_{jt} dilutes the effects of research outlays, z_{jit}. Conditional on having a new product, successful innovation requires supplying the intermediate good to the respective final production sector. The negative dependence on l_{jt} captures that a larger final production sector benefits the incumbent monopolist (e.g. due to existing supply chain networks, contracts and relationships with hospitals/doctors, etc.), thus lowering the probability of successful innovation. Second, 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 the probability of innovating is bounded between 0 and 1. As a result, 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. The upper bound is more likely to become binding in the early stages of a final production sector in which l_{jt} is small.

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

$$\max_{z_{jit}} \left\{ \pi_{jit} \phi_j \left(z_{jit}; l_{jt}, q_{jit-1} \right) - z_{jit} \right\} \tag{23}$$

For interior solutions the first order condition is

$$\frac{\pi_{jit}}{\lambda_j q_{jit-1} l_{jt}} \varphi_j \gamma_j \left(\frac{z_{jit}}{\lambda_j q_{jit-1}} \right)^{\gamma_j - 1} = 1, \tag{24}$$

which yields as solution a ratio of R&D spending to potential period t technology $\frac{z_{jit}}{\lambda_j q_{jit-1}}$

$$\frac{z_{jit}}{\lambda_j q_{jit-1}} = \left[\varphi_j \gamma_j \frac{\pi_{jit}}{\lambda_j q_{jit-1} l_{jt}} \right]^{\frac{1}{1-\gamma_j}}. \tag{25}$$

In the interior solution the innovation probability is then

$$\varphi_j \left(\frac{z_{jit}}{\lambda_j q_{jit-1}} \right)^{\gamma_j} l_{jt}^{-1} = \varphi_j \left[\varphi_j \gamma_j \frac{\pi_{jit}}{\lambda_j q_{jit-1} l_{jt}} \right]^{\frac{\gamma_j}{1-\gamma_j}} l_{jt}^{-1}. \tag{26}$$

The condition for the upper bound on the innovation probability to bind is then

$$1 \geq \varphi_j \left[\varphi_j \gamma_j \frac{\pi_{jit}}{\lambda_j q_{jit-1} l_{jt}} \right]^{\frac{\gamma_j}{1-\gamma_j}} l_{jt}^{-1} \tag{27}$$

Noticing that in case of success $q_{jit} = \lambda_j q_{jit-1}$, we can now use the profit equation (14) in the above to find effective R&D spending in the binding and non-binding case

$$\frac{z_{jit}}{\lambda_j q_{jit-1}} = \begin{cases} \left[\frac{l_{jt}}{\varphi_j}\right]^{\frac{1}{\gamma_j}} & \text{if } \phi_j\left(z_{jit}; l_{jt}, q_{jit-1}\right) = 1 \\ \left[\frac{1-\alpha_j}{\alpha_j}\varphi_j\gamma_j\left(\frac{p_{jt}\alpha_j^2}{R^{\alpha_j}}\right)^{\frac{1}{1-\alpha_j}}\right]^{\frac{1}{1-\gamma_j}} & \text{if } \phi_j\left(z_{jit}; l_{jt}, q_{jit-1}\right) \in (0,1). \end{cases} \tag{28}$$

Notice that effective R&D spending and the probability of innovation are the same across all varieties i. Using the above back in (22) we observe that in the interior solution the share of varieties innovating is (due to the law of large numbers)

$$\mu_{jt} = \int \varphi_j\left(\frac{z_{jit}}{\lambda_j q_{jit-1}}\right)^{\gamma_j} l_{jt}^{-1} di = \varphi_j^{\frac{1}{1-\gamma_j}}\left[\gamma_j\frac{1-\alpha_j}{\alpha_j}\left(\frac{p_{jt}\alpha_j^2}{R^{\alpha_j}}\right)^{\frac{1}{1-\alpha_j}}\right]^{\frac{\gamma_j}{1-\gamma_j}} l_{jt}^{-1} \tag{29}$$

and is therefore independent of the distribution of qualities across varieties i. For future reference, also observe that resources spend by entrepreneur i are

$$z_{jit} = \left[\frac{1-\alpha_j}{\alpha_j}\varphi_j\gamma_j\left(\frac{p_{jt}\alpha_j^2}{R^{\alpha_j}}\right)^{\frac{1}{1-\alpha_j}}\right]^{\frac{1}{1-\gamma_j}} \lambda_j q_{jit-1} \tag{30}$$

so that total resources devoted to R&D in sector j are equal to

$$z_{jt} = \int z_{jit} di = \left[\frac{1-\alpha_j}{\alpha_j}\varphi_j\gamma_j\left(\frac{p_{jt}\alpha_j^2}{R^{\alpha_j}}\right)^{\frac{1}{1-\alpha_j}}\right]^{\frac{1}{1-\gamma_j}} \lambda_j q_{jt-1}, \tag{31}$$

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

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 hold-*

ings $\{f_t\}_{t=1}^{\infty}$*, 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 (1) subject to (4).*

2. *Transfers* T_t *satisfy equation (6).*

3. *Factor prices satisfy equations (18a) and (18b).*

4. *Optimal R&D spending* z_{jt} *in each sector is given by (31) 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*

$$q_{jt} = (1 - \mu_{jt})q_{jt-1} + \mu_{jt}\lambda q_{jt-1} \tag{32}$$

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

   (a) *Labor Market*

$$1 = g(l_{ft}, l_{ht}). \tag{33}$$

   (b) *Capital Market*

$$\sum_j k_{jt} = k_t. \tag{34}$$

   (c) *Final Goods Market*

$$s_t + c_t^y + c_t^o n_t^o + i_{ft} + \sum_j [c_{jt} + z_{jt}] = k_{ft}^{\alpha} (q_{ft} l_{ft})^{1-\alpha} + R f_t. \tag{35}$$

   (d) *Health Goods Market*

$$i_{ht} = k_{ht}^{\alpha} (q_{ht} l_{ht})^{1-\alpha}. \tag{36}$$

(e) *International Capital Market*

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

(f) *The population evolves according to*

$$n_t^o = \psi(i_{t-1}). \tag{37}$$

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 $\psi(i_t)$, and consumption in both periods of their life c_t^y, c_{t+1}^o, where they maximize (1) subject to (4). 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*

$$\frac{x_{t+1}}{p_{t+1}} > \frac{x_t}{p_t}. \tag{38}$$

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 $\psi(i_t) = \psi(0)$ and $c_{t+1}^o = Rx_t$. 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.*

4.1 The Division of Labor Across the Two Sectors

Note that the labor allocation problem is straight-forward and separated from the health spending consumption problem. Thus, we can solve the household model sequentially, first by solving for the labor allocation that maximizes income and second, taking transfers T_t as given, allocating cash-on-hand x_t optimally to consumption, saving and health investment which boils down to a two-dimensional maximization problem.

We specify the function $g(l_{ft}, l_{ht})$ assuming a constant elasticity of ϵ as

$$1 = g(l_{ft}, l_{ht}) = \left(\sum_j l_{jt}^{1+\frac{1}{\epsilon}}\right)^{\frac{1}{1+\frac{1}{\epsilon}}}$$

Notice that $\epsilon = \infty$ this implies perfect labor mobility across the two sectors, and $\epsilon = 0$ it implies no mobility at all.[11] For interpretational purposes we think of a by age cohort representative household who optimally decides to allocate its labor across the two sectors.

In the case of perfect substitution ($\epsilon = \infty$) we have a corner solution whenever wages are not equalized given by

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

Thus, we get an interior solution in which both sectors are operative $l_{jt} \in (0,1), j \in \{f, h\}$ iff $w_{ht} = w_{ft}$. In the interior solution the labor supply to each sector is not determined.

In the general case of imperfect substitution ($\epsilon < \infty$), which we take as the benchmark, we directly obtain from (3) the transformation function

$$l_{ht} = \left(1 - l_{ft}^{1+\frac{1}{\epsilon}}\right)^{\frac{1}{1+\frac{1}{\epsilon}}}. \tag{39}$$

The first-order condition w.r.t l_{ft} gives

$$w_{ht} = w_{ft}\left(1 - l_{ft}^{1+\frac{1}{\epsilon}}\right)^{\frac{\frac{1}{\epsilon}}{1+\frac{1}{\epsilon}}} l_{ft}^{-\frac{1}{\epsilon}},$$

[11] Standard estimates of ϵ range between $\epsilon = 0.6$ and $\epsilon = 1.8$ (Giagheddu and Papetti 2019).

which simplifies to

$$\frac{l_{ht}}{l_{ft}} = \left(\frac{w_{ht}}{w_{ft}}\right)^{\epsilon} \tag{40}$$

and thus determines the relative labor allocation across the two sectors in dependence of relative wages.

4.2 Quasi-Linear Health Investment Function and the Division of Health Investment

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 quasi-linear specification

$$i_t = i_{ht} + (\nu + i_{ft})^{\zeta}. \tag{41}$$

for $\nu > 0, \zeta \in (0,1)$. Health investment then enters the survival function satisfying the CDF of a type 2 Pareto distribution given by

$$\psi(i_t) = 1 - (1 + i_t)^{-\xi}. \tag{42}$$

for $\xi > 0$. Note that $\psi(\cdot)$ is strictly increasing in ξ, and is strictly increasing in i_t with $\psi(i_{ft} = i_{ht} = 0) = 1 - [1+\nu]^{-\xi} > 0$ and $\lim_{i_t \to \infty} \psi(i_t) = 1$. In addition to giving positive survival for zero health investments, the non-homotheticity parameter $\nu > 0$ prevents the lower Inada condition to hold for final goods investment i_{ft} so that $\lim_{i_{ft} \to 0} \frac{\partial \psi(i_t)}{\partial i_t} \frac{\partial i_t}{\partial i_{ft}} < \infty$ which is a necessary condition for the existence of phase 1 in which $i_{ft} = i_{ft} = 0$. The concavity parameter $\zeta \in (0,1)$ in turn implies that health investment is linear in i_{ht} and concave in i_{ft}. Thus, we have the standard quasi-linear property that agents initially—i.e., once they posses sufficient resources and health spending becomes positive—only buy the good with decreasing marginal benefit i_{ft} and then switch to the linear good i_{ht} forever as soon as marginal benefits over marginal costs of the two are equalized. This property gives rise to the possibility of the model to generate phase 2, where $i_{ft} > 0, i_{ht} = 0$, and subsequently in time phase 3, where $i_{ft}, i_{ht} > 0$.

We 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. See appendix B.1 for the full derivation. 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 $i_{ht} = 0$, $i_{ft} = e_t$: Health investment i_t is given by

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

Interior Solution: In the interior solution the first-order conditions hold with equality, which yields

$$\begin{aligned} i_{ft} &= \tilde{\lambda}_t - \nu \\ i_{ht} &= \frac{e_t - (\tilde{\lambda}_t - \nu)}{p_t} \end{aligned}$$

where $\tilde{\lambda}_t \equiv (\zeta p_t)^{\frac{1}{1-\zeta}}$.

Existence of Phase 2: The corner solution with $i_{ht} = 0$ and $i_{ft} = e_t$ characterized above (phase 2) exists if and only if

$$\begin{aligned} \tilde{\lambda}_t &\geq \nu \\ (\zeta p_t)^{\frac{1}{1-\zeta}} &\geq \nu \end{aligned}$$

Existence of phase 2 requires the non-homotheticity factor ν to be sufficiently small relative to health sector price p_t and ζ.

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

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

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 savings s_t and health expenditures e_t. That is, the household now solves

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

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

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

4.3.1 Balanced Growth Consistent Functional Form Assumptions

To further characterize the optimal level of health investment in the balanced growth path, we need to make a decision on the functional form of the per period utility function $u(\cdot)$ in addition to the assumed functional form assumption of the survival rate in equation (42). 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 life.

Corner Solution: The corner solution with $\vartheta_{t,e} = 0$ corresponds to phase 1 without any health expenditure. We solve for the cash-on-hand level at which the first kickoff happens, that is, at which health expenditure become positive. For $\sigma = 2$ and $\xi = 1$, we obtain an analytical solution which is given by

$$\begin{aligned} x_{\text{kickoff1}} &= \frac{1}{bR}\left(\frac{1}{2}\left(1+\sqrt{1+4bRA_1}\right)\right) \\ &\equiv \frac{1}{bR}\Delta(b,R,A_1) \end{aligned} \tag{43}$$

where

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

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

$$x \in [x_{\text{lowerbound}}, x_{\text{kickoff1}}] \tag{44}$$

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

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

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

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

Interior Solution on the BGP: In the interior solution of the BGP, where both the final goods and the health goods sector are active, we can find the optimal health expenditure share by taking the limit case: $x \to \infty$. Plugging in the functional forms, the first-order condition for the health expenditure share becomes

$$\frac{1}{\xi}(1+i_t)\left((1+i_t)^{\xi} - 1\right) x_t R(1-\vartheta_c^*) = \left(\frac{1}{1-\sigma} + b(c_{t+1}^o)^{\sigma-1}\right) c_{t+1}^o \frac{\partial i_t}{\partial \vartheta_t}. \tag{45}$$

For a BGP with the properties $p_t = p^*$ is constant, and, as x_t converges to infinity, $i_t \to i_{ht} \to \frac{e_t}{p_t} = \frac{(1-\vartheta_c^*)\vartheta_e^* x_t}{p^*}$ so that i_t and c_{t+1}^o are both 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

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

4.4 Transitional Dynamics

In equilibirum, the health price p_t and relative labor allocation $\frac{l_{ht}}{l_{ft}}$ adjust to clear the health goods market and labor market. We analytically derive the demand for labor by final good producers and the supply of labor by households in terms of the wage ratio. We then combine the two which yields the labor market clearing condition and characterizes the relationship between the health price and the quality ratio across the two sectors along the transition.

Labor demand: Combining the first-order condition for labor from final good producers with the intermediate good producer solution yields for wages in each sector j

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

This delivers the following relationship for the wage ratio

$$\begin{aligned} \frac{w_{ht}}{w_{ft}} &= p_{ht}^{\frac{1}{1-\alpha_j}} \frac{1-\alpha_h}{1-\alpha_f} \alpha_h^{\frac{2\alpha_h}{1-\alpha_h}} \alpha_f^{\frac{2\alpha_f}{\alpha_f - 1}} \frac{q_{ht}}{q_{ft}} \\ &\equiv p_{ht}^{\frac{1}{1-\alpha_j}} \frac{q_{ht}}{q_{ft}} C(\alpha_f, \alpha_h) \end{aligned}$$

where $C(\alpha_f, \alpha_h) \equiv \frac{1-\alpha_h}{1-\alpha_f} \alpha_h^{\frac{2\alpha_h}{1-\alpha_h}} \alpha_f^{\frac{2\alpha_f}{\alpha_f - 1}}$.

Labor supply: The first-order condition for households' labor supply is given by

$$\frac{w_{ht}}{w_{ft}} = \left(\frac{l_{ht}}{l_{ft}} \right)^{\frac{1}{\epsilon}}$$

Equilibrium: In equilibrium labor markets have to clear. Setting the demand and supply condition for the wage ratio equal to each other yields the following equilibrium condition:

$$\left(\frac{l_{ht}}{l_{ft}} \right)^{\frac{1}{\epsilon}} = p_{ht}^{\frac{1}{1-\alpha_j}} \frac{q_{ht}}{q_{ft}} C(\alpha_f, \alpha_h). \tag{46}$$

Note that qualities are determined endogenously in the R&D sector and depend on previous quality q_{jt-1} and the price p_{jt} in each sector. Thus, the only period t endogenous variable on the right hand side is the health price p_{ht}.

Let us provide some intuition for the role of the equilibrium relationship (46). It disciplines the relationship between the health price and the allocation of labor allocation across the two sectors along the transition and thereby pins down the price of health p_t. With perfect labor mobility, $\epsilon = \infty$, the left hand side is constant and equal to 1 as wages have to be equalized across sectors for households to optimally supply labor to both sectors. As a result, the left hand side is not growing along the transition which limits how much the quality ratio $\frac{q_{ht}}{q_{ft}}$ can grow. With imperfect labor mobility, $\epsilon < \infty$, the quality ratio $\frac{q_{ht}}{q_{ft}}$ has to grow faster along the transition in order to generate growth in the wage ratio $\frac{w_{ht}}{w_{ft}}$ which is necessary to incentivize labor reallocation from households.

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 growth of the health sector essentially is the consequence of research and development efforts, which in turn are triggered by economic developments (the treatment of soldiers in the second world war might be an additional trigger; our model has nothing to contribute to this observation), rather than interpreting the introduction of Medicare as the trigger as in Frankovic and Kuhn (2018), who develop their theory on the basis of a conjecture in Weisbrod (1991).

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 final consumption hygiene health good. 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 simply fixing their values (first-stage parameters). Others are calibrated to match selected moments in the data. We then intend to endogenously calibrate the following parameters $b, \zeta, \epsilon, \lambda_f, \lambda_h$.

5.1 External Calibration

We set $\sigma = 2$ to generate an intertemporal elasticity of substitution of 0.5. To ensure the existence of a BGP, we impose $\xi = \sigma - 1 = 1$, see 4.3.1 for the derivation of the BGP. We set the weight on second period utility relative to first period utility, β, to match an annualized discount factor of 0.92. We choose a real interest rate in the small open economy of 1%. The capital intensity in the final good sector is $\alpha_f = 1/3$, consistent with estimates for the U.S. According to Frankovic, Kuhn, and Wrzaczek (2017), based on Acemoglu and Guerrieri (2008), an appropriate estimate for the capital intensity in the health sector is $\alpha_h = 1/5$.

Parameter	Description	Value	Target
		Small open economy	
R-1	Rate of return	1.5	1 % annual return
		Households	
$1/\sigma$	IES	0.5	Standard
β	Discount factor	0.078	$\beta_{\text{annualized}} = 0.94$
ξ	Curvature survival function	$\sigma - 1 = 1$	Ensures BGP existence
		Firms	
α_f	Capital intensity final	0.33	Standard
α_h	Capital intensity health	0.2	FKW (2017), AG (2008)

Table 2: External Calibration

5.2 Internal Calibration

We aim at expressing in any period t, x_t in terms of the state variables of the problem in order to derive expressions for the conditions for an inoperative health sector. The state variables in any period t are q_{ft-1}, q_{ht-1}, T_t.

Initial Conditions. We choose 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.

Conditional on the modern health sector being inactive in period 1 (which we verify below), we have $l_{f1} = 1$ which allows us to solve for the production side in period 1 analytically. Given R and some q_{f0}, which remains to be determined, we can use (29) and (32) to compute the quality in period 1 as:

$$q_{f1} = (1 + (\lambda_f - 1)\mu_{f0})\, q_{f0} = \Upsilon_t(R) q_{f0-1} \tag{47}$$

$$= \left(1 + (\lambda_f - 1)\varphi_f^{\frac{1}{1-\gamma_f}} \left[\frac{1-\alpha_j}{\alpha_j} \gamma_f \left(\frac{\alpha_f^2}{R^{\alpha_f}} \right)^{\frac{1}{1-\alpha}} \right]^{\frac{\gamma_f}{1-\gamma_f}} \right) q_{f0} = \Upsilon_t(R) q_{f0}. \tag{48}$$

Further note that for $l_{f1} = 1$ we obtain from (15) that the capital stock employed in production is

$$k_{f1} = \left(\frac{\alpha_f^2}{R} \right)^{\frac{1}{1-\alpha_f}} q_{f1}. \tag{49}$$

1. **Initial assets**: We assume that the small open economy has a zero net foreign asset position in period 1 so that $k_{f1} = s_0$.

   Now use (47) in (49) and set $k_{f1} = s_0$ to get

$$s_0(q_{f0}, R) = \left(\frac{\alpha_f^2}{R}\right)^{\frac{1}{1-\alpha_f}} \Upsilon_1(R) q_{f0}$$
$$= \left(\frac{\alpha_f^2}{R}\right)^{\frac{1}{1-\alpha_f}} \left(1 + (\lambda_f - 1)\varphi_f^{\frac{1}{1-\gamma_f}} \left[\frac{1-\alpha_j}{\alpha_j}\gamma_f \left(\frac{\alpha_f^2}{R^{\alpha_f}}\right)^{\frac{1}{1-\alpha}}\right]^{\frac{\gamma_f}{1-\gamma_f}}\right) q_{f0} \tag{50}$$

2. **Initial income and first kickoff**: We calibrate the initial quality in the final good sector $q_{f,0}$ and the value of life b jointly targeting two moments. First, an initial income level of the economy such that households in period 1 are indifferent between suicide and survival. Second, the timing of the first kickoff in 1860, meaning the period in which households start consuming basic health goods, which corresponds to period 2 in the model. Note, this implies that there are no investments in basic and modern health goods, $i_0 = i_{f0} = i_{h0}$ in the initial period. As a consequence, $n_0^o = \psi(0)$.

   Given the value of life b, we can solve analytically for the initial quality in the final good sector that ensures that households in period $t = 1$ are indifferent between survival and suicide. Observe that for $l_{ft} = 1$ wages in the final goods sector and transfers are given by

$$w_{ft} = (1-\alpha_f) k_{ft}^{\alpha_f} q_{ft}^{1-\alpha_f} \tag{51}$$
$$T_t = R s_{t-1}(1 - n_t^o), \tag{52}$$

   Now, use (40) for $l_{ft} = 1$ to obtain

$$w_{ft} = R\frac{1-\alpha_f}{\alpha_f^2} k_{ft}$$

   and thus cash-on-hand $x_t = w_{ft} + T_t$ is given by

$$x_t = R\frac{1-\alpha_f}{\alpha_f^2} k_{ft} + R s_{t-1}(1 - n_t^o)$$

and thus we can express the period 1 cash-on-hand in terms of initial savings using that $k_{f1} = s_0$ as

$$x_1 = R\left(\frac{1-\alpha_f}{\alpha_f^2} + (1-n_1^o)\right) s_0(q_{f0}, R)$$

where $s_0(q_{f0}, R)$ is given by (50). Finally, recall that for the lower bound (to insure that households do not commit suicide) we need

$$\begin{aligned} & x_1 \geq \frac{[b(\sigma-1)]^{\frac{1}{1-\sigma}}}{R} \\ \Leftrightarrow \qquad & R\left(\frac{1-\alpha_f}{\alpha_f^2} + (1-n_1^o)\right) s_0(q_{f0}, R) \geq \frac{[b(\sigma-1)]^{\frac{1}{1-\sigma}}}{R} \end{aligned}$$

which simplifies to $x_1 \geq \frac{1}{bR}$ for $\sigma = 2$. Using (50) then gives

$$\begin{aligned} & R\left(\frac{1-\alpha_f}{\alpha_f^2} + (1-n_1^o)\right) \\ & \cdot \left(\frac{\alpha_f^2}{R}\right)^{\frac{1}{1-\alpha_f}} \left(1 + (\lambda_f - 1)\varphi_f^{\frac{1}{1-\gamma_f}} \left[\frac{1-\alpha_j}{\alpha_j}\gamma_f \left(\frac{\alpha_f^2}{R^{\alpha_f}}\right)^{\frac{1}{1-\alpha}}\right]^{\frac{\gamma_f}{1-\gamma_f}}\right) q_{f0} \geq \frac{[b(\sigma-1)]^{\frac{1}{1-\sigma}}}{R} \end{aligned}$$

and thus we parametrise (breaking indifference as survival) the initial quality in the final goods sector as

$$\begin{aligned} q_{f0} = & \frac{[b(\sigma-1)]^{\frac{1}{1-\sigma}}}{R} \cdot \\ & \left[R\left(\frac{1-\alpha_f}{\alpha_f^2} + (1-n_1^o)\right)\left(\frac{\alpha_f^2}{R}\right)^{\frac{1}{1-\alpha_f}} \left(1 + (\lambda_f - 1)\varphi_f^{\frac{1}{1-\gamma_f}} \left[\frac{1-\alpha_j}{\alpha_j}\gamma_f \left(\frac{\alpha_f^2}{R^{\alpha_f}}\right)^{\frac{1}{1-\alpha}}\right]^{\frac{\gamma_f}{1-\gamma_f}}\right)\right]^{-1} \end{aligned} \tag{53}$$

3. **Kickoff of modern health sector**: We calibrate the initial quality in the health sector $q_{h,0}$ to match the kickoff timing of the modern health sector in 1940. We assume that during phases 1 and 2 while the modern health sector is inactive, health quality q_{ht} grows exogenously at the same rate as quality in the final goods sector, q_{ft}.

4. **Survival function and initial life expectancy**: We calibrate ν in the survival function to match initial life expectancy during phase 1. Observe that ν^ζ pins down the initial life

expectancy (before the kick-off). Recall that

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

and

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

so that with zero health investments in both goods and with the calibration of $\xi = 1$ we get

$$\psi(i_t = \nu^{\zeta}) = 1 - \left(1 + \nu^{\zeta}\right)^{-1} \tag{54}$$

Remaining life-expectancy at economic birth is given by

$$y = (1 + \psi) \cdot 40. \tag{55}$$

Using (54) in (55) we get as value for ν under the maintained parametric restriction $\xi = 1$ and for given ζ that

$$\nu = \left[\left(\frac{40}{80 - y} - 1\right)\right]^{\frac{1}{\zeta}}. \tag{56}$$

To calibrate y in the above we take our estimate of cohort life expectancy in year 1790 (the beginning of our sample), of $y = 40.43$ years.

Calibration of Remaining Parameters. We calibrate the curvature of basic health spending in the health investment function, ζ, to minimize the distance between life expectancy in the model and in the data from 1860 onwards. We further calibrate the step size for innovations in the final good sector, λ_f, targeting average GDP per capita growth between 1820 and 2020, and the step size for modern health, λ_h, targeting the average growth rate of the output share of the modern health sector between 1940 and 2020. Lastly, we fix the scaling and curvature parameters of the innovation probability, γ and ϕ, to 0.5 in the first stage and calibrate the remaining parameters conditional on that.

Parameter	Description	Value	Target
	Households		
b	Value of life	129	LE20 in 1860
ζ	Investment curvature	0.6	LE20 after 1860
ϵ	Labor mobility	2	Growth: Q-adj price
	Firms and R&D		
λ_f, λ_h	Growth factor	120, 100	GDP per capita
γ_f, γ_h	Innovation probability: curvature	0.5, 0.5	First stage
ϕ_f, ϕ_h	Innovation probability: scaling	0.5, 0.5	First stage
	Initial Conditions		
$q_{f,0}$	Initial quality: final	0.71	Initial income
$q_{h,0}$	Initial quality: health	0.01	Kickoff 1940

Table 3: Internal Calibration

6 Results

In this section we contrast the positive predictions of the model concerning the time paths of income, 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}.

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 the 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.

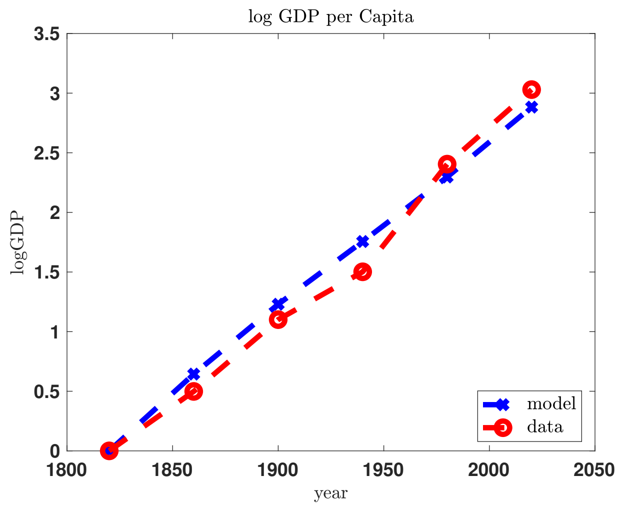
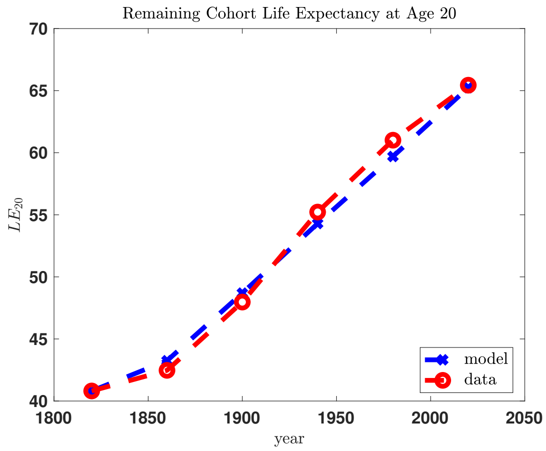


Figure 5: GDP per Capita [Logarithmic Scale]

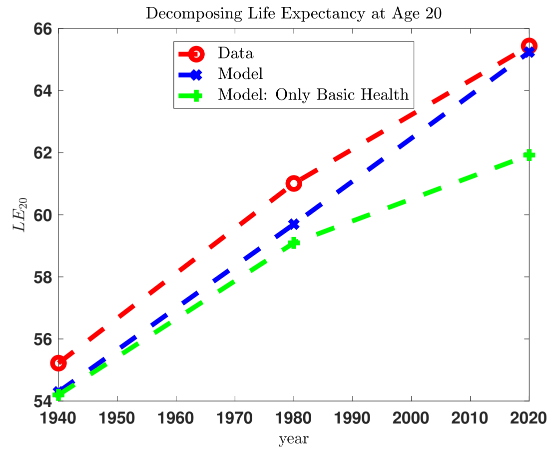
Notes: Natural logarithm of real GDP per capita in model and data. *Source:* Data sources as in Figure 2, own calculations.

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. The left panel shows the entire time series starting from 1820, and the right panel zooms into the period after the modern health sector emerged (from 1940 on) and additionally displays the component of life expectancy that is driven by investments into only basic health goods (green dots). The gap between the blue and the green line can be interpreted as the contribution of the modern health, by comparing the evolution of life expectancy when household optimally allocate health investments between the basic and the modern health goods (blue line) and a counterfactual economy where the modern health sector is simply absent.

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 1940's, implies that life expectancy grows continually throughout the last 2 centuries and matches the data on cohort life expectancy well.



(a) Remaining Cohort Life Expectancy at Age 20



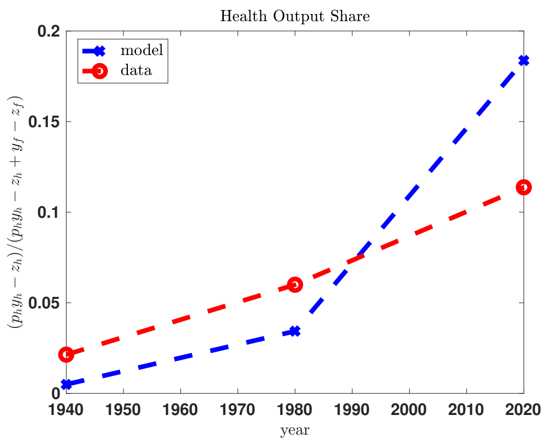
(b) Decomposition of Life Expectancy

Notes: Remaining cohort life expectancy at age 20 in model (blue x mark) and data (red circles). The figure also shows a decomposition setting spending on modern health goods to zero (black dots). *Sources:* Data sources as in Figure 1, own calculations.

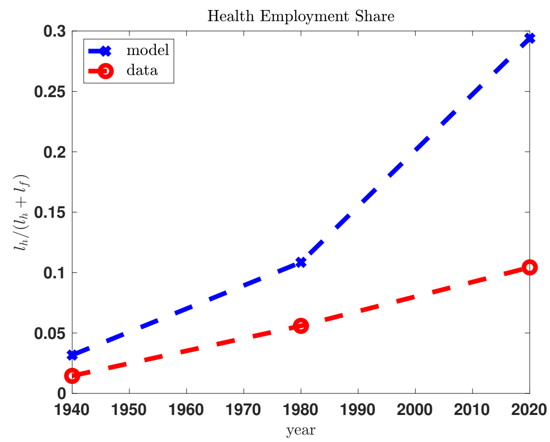
Figure 6: Health Shares (Spending, Employment, R&D) in the U.S.

Turning to the decomposition of life expectancy improvements driven by the increase in traditional health goods and the modern high-tech health sector, we observe that according to the model, modern health goods become an important driver of life expectancy after 1980 and account for 30% of the increase in life expectancy at age 20 since 1940, translating into 3.3 additional expected years of life. The complementary implication is noteworthy, too: even in the absence of the emergence of modern medicine, income-induced growth in basic health goods (a richer, more balanced diet, better hygiene) would still have led to an increase in life expectancy between 1940 and 2020 by close to 8 years, and basic health goods are the main driver of increased longevity, according to the model, until 1980.

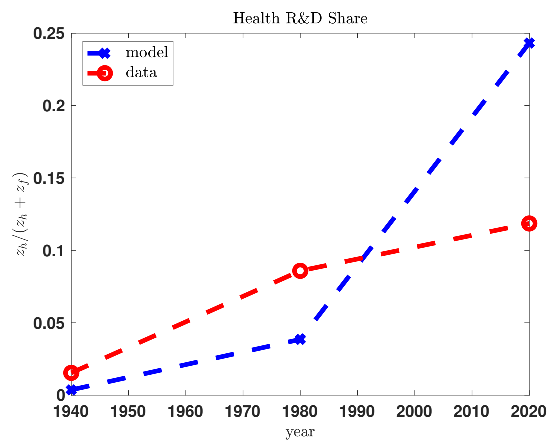
The health-related output-, labor- and R&D shares underlying the emergence of the modern health sector are displayed in Figure 7. Note that we plot these shares only from 1940 onward as they are not available empirically, and are equal to zero in the model. The figure shows that qualitatively, the model reproduces the increasing shares of labor, R&D and consequently output accruing to the modern health sector. Quantitatively, in the model the takeoff is initially somewhat too slow between 1940 and 1980 then accelerates too much between 1980 and 2020, relative to what the data suggests. Note that these shares were not targeted in the calibration, and therefore some divergence between model and data is to be expected and the ability of the model to qualitatively match the relevant time series from the data should be a considered a qualified success.



(a) Health Output Share



(b) Employment Share



(c) R&D Spending Share

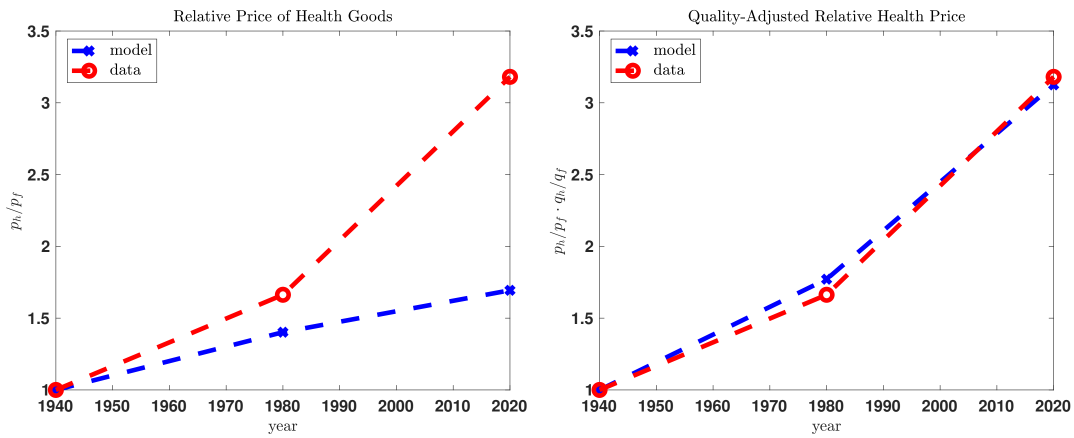
Figure 7: Health Shares (Spending, Employment, R&D) in the U.S.

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 quality-adjusted price index implied by the model, $p_t \frac{q_{ht}}{q_{ft}}$. According to the model, the price of one raw unit of health goods is increasing slowly over time, accounting for the improvement in the quality of health care the quality-adjusted price of health goods increases at a rate similar to the data.[12] To decompose the observed relative price increase in the measured health price in the data into a component driven by rising output prices versus falling input prices, we decompose the increase in the quality-adjusted health price $p_t \frac{q_{ht}}{q_{ft}}$ into its two components: 1) growth in the price of health goods p_t, whose main driver is the (income-growth-induced) rising household demand for health goods relative to final goods; and 2) productivity growth in the modern health sector relative to the final goods sector, driven by endogenous technological progress.

[12]This figure implicitly assumes that the empirically observed relative price for health goods has been fully and appropriately quality adjusted.

Between 1940 and 1980, both components contribute roughly half of the overall increase in the quality-adjusted health price $p_t \frac{q_{ht}}{q_{ft}}$; precisely speaking, in rising demand for health goods accounts for 52%. After 1980, technological progress in the modern health sector accelerates and becomes the dominant force, and as a result, 67% of the overall growth in the quality-adjusted health price between 1940 and 2020 is accounted for by greater quality growth in intermediates in the modern health sector (and thus faster productivity growth in that sector) relative to the final goods sector. The relative contribution of rising demand accounts for the remaining 33%.



(a) Health Price

(b) Quality-Adjusted Price

Figure 8: Health Price Index [Quality Adjusted]

Sources: Data sources as in Figure 4, 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.

Since technological progress in the health sector is endogenous, government health policies (such as the funding of basic research in the health sector or the public provision of health goods though government-run health insurance or the direct production of these goods in public hospitals) will impact the timing and speed of the development of a modern health sector. The next steps of our analysis will be to evaluate positively, and to study normatively the importance of government interventions in the health sector in the 20-th century in the U.S.

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. 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.

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

Cardi, O. and R. Restout (2015). Imperfect Mobility of Labor Across Sectors: A Reappraisal of the Balassa-Samuelson Effect. *Journal of International Economics 97*(2), 249–265.

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.

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.

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 (2018). Health Insurance, Endogenous Medical Progress, and Health Expenditure Growth. Working Paper.

Frankovic, I., M. Kuhn, and S. Wrzaczek (2017). Medical Progress, Demand for Health Care, and Economic Performance. Working Paper.

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

Giagheddu, M. and A. Papetti (2019). Demograpics and the Real Exchange Rate. Working Paper.

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.

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.

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

Our data to construct cohort life-expectancy comes from three different sources. First, we use 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 transformations to this data: (i) we hold mortality rates constant within age group and compute age specific mortality rates in each age group i, m_j^i such that m^i is the respective geometric average within this age group, i.e., $m_j^i = 1 - (1 - m^i)^{1/5}$; (ii) to obtain estimates of mortality rats 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 (needed for the computation of cohort life expectancy) we estimate future trends in mortality by a Lee-Carter method—assuming a deterministic trend of the single index—which we apply to the postwar data (years 1950-2021). Finally, over the entire period (including the predicted mortality rates), we filter the data by applying a Hodrick-Prescott filter with a bandwidth parameter of 100.

The bootstrapped confidence intervals shown in Panels (c) to (e) in Figure 1 are computed by bootstrapping along the time dimension of the cross-sectional mortality rates, whereby we ignore the uncertainty of the predicted mortality estimates from the Lee-Carter method. We implement the bootstrap procedure as a block bootstrap and set the width of the blocks according to the standard rule of thumb $bw = T^{1/3}$. We bootstrap on the entire data sequence, starting in year 1790 and ending in 2152[13], thus $T = 363$ and $bw = 7.13$, of which we take the ceil.

[13] Data from the Human Mortality Database range to year 2021 and to age 110. We add 20 additional years. The horizon is thus $2021 + (110 + 1) + 20 = 2152$.

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

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

which nests the benchmark specification without consumption in young age for $\beta = 1$. The maximization of the utility function is subject to the constraints:

$$c_t^y + i_{ft} + p_t i_{ht} + s_t = x_t \equiv w_{ft}l_{ft} + w_{ht}l_{ht} + T_t \tag{57a}$$

$$i_t = f(i_{ft}, i_{ht}) \tag{57b}$$

$$1 = g(l_{ft}, l_{ht}) \tag{57c}$$

$$c_{t+1}^o = Rs_t. \tag{57d}$$

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

$$c_t^y + i_{ft} + p_t i_{ht} + \frac{c_{t+1}^o}{R} = w_{ft}l_{ft} + w_{ht}l_{ht} + T_t \equiv x_t. \tag{58}$$

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

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

Define $\tilde{\lambda}_t \equiv (\zeta p_t)^{\frac{1}{1-\zeta}}$. 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 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 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 $i_{ft} = 0$:

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

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

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

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

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

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

$$\frac{1}{\xi}\left(1+\nu^{\zeta}\right)\left(\left(1+\nu^{\zeta}\right)^{\xi}-1\right) \geq \left(\frac{1}{1-\sigma}+b\left(R\tilde{x}_t\right)^{\sigma-1}\right)\tilde{x}_t\zeta\nu^{\zeta-1} \tag{60}$$

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

$$\begin{aligned} \tilde{x}_{\text{kickoff1}} &= \frac{1}{2bR}\left(1+\sqrt{1+4bRA_1}\right) \\ &= \frac{1}{bR}\left(\frac{1}{2}\left(1+\sqrt{1+4bRA_1}\right)\right) \\ &\equiv \frac{1}{bR}\Delta(b, R, A_1) \end{aligned} \tag{61}$$

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

$$\tilde{x} \in [\tilde{x}_{\text{lowerbound}}, \tilde{x}_{\text{kickoff1}}] \tag{62}$$

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.

$$x \in [x_{\text{lowerbound}}, x_{\text{kickoff1}}] = \left[\frac{\tilde{x}_{\text{lowerbound}}}{(1 - \vartheta_{\text{phase }1,c})}, \frac{\tilde{x}_{\text{kickoff1}}}{(1 - \vartheta_{\text{phase }1,c})}\right] \tag{63}$$

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

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

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

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 $\hat{c}^y_t$, $\hat{c}^o_{t+1}$, $\hat{i}$, $\frac{\partial \hat{i}}{\partial \hat{\vartheta}_{t,e}}$ 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

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

$p_t = p$ 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

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

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

$$\frac{\psi}{\psi'} x_t R_{t+1} = \frac{u}{u'} \frac{\partial i_t}{\partial \vartheta_t} \tag{64}$$

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

$$\frac{1}{\xi}\left(1+\nu^{\zeta}\right)\left(\left(1+\nu^{\zeta}\right)^{\xi}-1\right) \geq \left(\frac{1}{1-\sigma} + b\left(R_{t+1} x_t\right)^{\sigma-1}\right) x_t \zeta \nu^{\zeta-1} \tag{65}$$

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

$$\begin{aligned} x_{\text{kickoff1}} &= \frac{1}{2bR_{t+1}}\left(1 + \sqrt{1 + 4bR_{t+1}A_1}\right) \\ &= \frac{1}{bR_{t+1}}\left(\frac{1}{2}\left(1 + \sqrt{1 + 4bR_{t+1}A_1}\right)\right) \\ &\equiv \frac{1}{bR_{t+1}}\Delta(b, R_{t+1}, A_1) \end{aligned} \tag{66}$$

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

$$x \in [x_{\text{lowerbound}}, x_{\text{kickoff1}}] \tag{67}$$

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

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

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

$$i_t = \lambda_t^{\zeta}$$
$$\frac{\partial i_t}{\partial \vartheta_t} = \frac{\lambda_t - \nu}{p_t \vartheta_t}$$

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

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

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

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

The roots of this equation are given by

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

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

$$
\begin{aligned}
x_{\text{kickoff2}} &= (\lambda_t-\nu)+\frac{1}{2bR_{t+1}}\left(1+\sqrt{1+4bR_{t+1}A_2}\right)\\
&=(\lambda_t-\nu)+\frac{1}{bR_{t+1}}\left(\frac{1}{2}\left(1+\sqrt{1+4bR_{t+1}A_2}\right)\right)\\
&\equiv(\lambda_t-\nu)+\frac{1}{bR_{t+1}}\Delta(b,R_{t+1},A_2) \qquad (68)\\
&=i_{ft}+\frac{1}{bR_{t+1}}\Delta(b,R_{t+1},A_2)
\end{aligned}
$$

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

$$\begin{aligned} x_{\text{LB}} &= \frac{1}{bR_{t+1}} = \frac{1}{bR_2} \\ x_{\text{KO1}} &= \frac{1}{bR_{T_1+1}} \Delta(b, R_{T_1+1}, A_1) \\ x_{\text{KO2}} &= (\lambda_t - \nu) + \frac{1}{bR_{T_2+1}} \Delta(b, R_{T_2+1}, A_2) \end{aligned}$$

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

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