

The Lifetime Costs of Bad Health

(Preliminary and Incomplete)

Mariacristina De Nardi, Svetlana Pashchenko, and Ponpoje Porapakkarm*

February 2, 2017

Abstract

How costly is bad health and what makes good health valuable over the life cycle? Answering these questions requires carefully modeling health dynamics, including in the longer run, and a rich model of how health can affect households. We estimate a health shock process that allows for both history-dependence and ex-ante heterogeneity, and we introduce it in a rich life-cycle model that we estimate and that matches three sets of important facts: (i) The dynamics of health; (ii) The quantitative impact of bad health on labor earnings, medical spending, and life expectancy; (iii) The large disparity in accumulated wealth between the healthy and the unhealthy at retirement. We find that the costs of bad health among the working age population are steeply increasing in the number of years spent unhealthy and that the largest component of these costs is the loss in labor earnings. In contrast, the effect of out-of-pocket medical spending is relatively small. To also evaluate the non-pecuniary effects of health, we evaluate the willingness to pay to be healthy and we find that the most valuable aspect of being healthy is a longer life expectancy. *JEL Codes:* D52, D91, E21, H53, I13, I18

Keywords: health, medical spending, wealth-health gradient, life-cycle models

*De Nardi: University College London, Federal Reserve Bank of Chicago, IFS, CEPR, and NBER. Pashchenko: University of Georgia. Porapakkarm: National Graduate Institute for Policy Studies (GRIPS). We thank Liran Einav, Christopher Flinn, Eric French, Selahattin Imrohoroglu, John Kennan, Sagiri Kitao, Rasmus Lentz, Vincenzo Quadrini, Yongseok Shin, and all seminar participants at the Keio University, University of Hong Kong, University of Tokyo for their comments and suggestions. De Nardi gratefully acknowledges support from the ERC, grant 614328 “Savings and Risks.” Porapakkarm gratefully acknowledges support from the JSPS KAKENHI Grant Number 15K03505 and GRIPS’ Research Project Grant. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research, the CEPR, any agency of the federal government, or the Federal Reserve Bank of Chicago.

1 Introduction

Why is it bad to be in bad health, or conversely, why it is good to be in good health? Bad health typically entails a number of negative consequences: reduced labor supply and earnings, higher medical expenses, and shorter lifespans. In addition, one's health status is persistent and these negative effects tend to accumulate over one's life cycle.

While measuring these accumulated effects directly is not easy, some features of the data suggest that they are substantial. Importantly, unhealthy people tend to accumulate substantially less wealth than healthy people (the so-called wealth-health gradient). For example, among the 65 years old males with a high-school degree, the gap in median wealth between the healthy and the unhealthy is more than \$150,000 (in 2015 dollars).¹

Our goal in this paper is to quantify the lifetime consequences of bad health among working-age individuals. More specifically, we aim at investigating the pathways through which bad health affects economic outcomes and to understand their relative importance. We focus on a relatively homogeneous group of men with a high-school degree to abstract from the interactions between health and economic outcomes arising from education and gender.

We proceed in several steps. First, we estimate our health process using a large set of cross-sectional and dynamic moments from the data. Importantly, our parametric model of health allows for both history-dependence and ex-ante heterogeneity. Both forces can generate persistence in health, but distinguishing among them allows us to better understand what generates long-lasting episodes of bad health: bad luck or permanent differences across individuals.

Second, we introduce our estimated health process into a rich structural life-cycle model with endogenous labor supply, health insurance, and saving decisions. In this model, individuals are heterogeneous along several dimensions: a (permanent) type and the realizations of stochastic productivity, health, and medical shocks. An individual's permanent type is characterized by a vector of (i) ex-ante, or fixed, heterogeneity affecting his health dynamics, and (ii) the rate of his time preferences (patience).

The model thus generates a correlation between health and economic outcomes via two mechanisms. First, individuals whose health deteriorates have lower productivity and higher disutility from work, and thus have lower labor supply and lower earnings. In addition, they pay higher out-of-pocket medical costs. These effects decrease their available resources; coupled with a lower life expectancy this produces a decline in savings. Second, the fixed heterogeneity in health dynamics is correlated with the rate of time

¹This number comes from our own calculations, based on the Health and Retirement Study dataset. Wealth represents total net worth. Individuals are classified into healthy and unhealthy based on self-reported health.

preferences: individuals who are more likely to stay unhealthy are also less patient. Thus, the lower savings of the unhealthy are partly due to the fact that this group includes more people with a lower propensity to save. The latter mechanism is important not to attribute the entire correlation between health and economic outcomes to the casual effect of health but, rather, to allow for a third factor that affects health and matters for economic decisions.

We estimate our model using three datasets: the Health and Retirement Study (HRS), the Panel Study of Income Dynamics (PSID), and the Medical Expenditure Panel Survey (MEPS). Our estimated model is consistent with three important sets of facts. First, our model captures the dynamics of health, including the duration dependence of the health transition probability, i.e., the fact that in the data the probability of recovering from bad to good health declines monotonically with the number of years that an individual has been unhealthy. Second, our model reproduces the observed impact of bad health on earnings, labor supply, medical spending, and life expectancy. Finally and importantly, our model also captures the wealth-health gradient both in the cross-section, in that it matches the large difference in accumulated wealth between the healthy and the unhealthy; and over time for an individual, in that it matches the negative relationship between wealth and the number of periods spent unhealthy.

Our results can be summarized as follows. First, both fixed heterogeneity and history-dependence play an important role in driving health dynamics but they play a different role in how individuals get sick versus how they recover. The persistence in bad health is mostly generated by fixed heterogeneity while the persistence in good health is mostly due to history-dependence. Thus, long episodes of bad health are mostly concentrated among individuals with a particular (fixed) health type. Moreover, because of the correlation between the fixed heterogeneity in health dynamics and the rate of time preferences, among the long-term unhealthy a larger fraction have a lower propensity to save. This matters for accounting for the wealth-health gradient: our decomposition exercise shows that around 60% of the wealth gap between the healthy and the unhealthy is due to the latter effect, while the remaining gap is accounted for by the casual effect of bad health.

Second, we find that the costs of bad health are steeply increasing in the number of years that people spend unhealthy over the working stage of their life (ages 20 to 64). Our measure of these costs include both direct (out-of-pocket medical spending) and indirect (loss in labor earnings) costs. We find that the latter component is the largest contributor to the lifetime costs of bad health, and it arises because unhealthy individuals are less productive and work less than healthy ones.

Finally, to capture the non-monetary value of health, we evaluate working-age individuals' willingness to pay to increase the probability to become healthy next period by

one percentage point. We find that this willingness to pay is high on average (more than \$3,000) and it varies significantly with health type, rate of time preferences, and assets of individuals. We provide a decomposition exercise where different channels affecting health are shut down one at a time. We find that the most important aspect of being healthy is that it implies a longer life expectancy. This can make up for 90% of the value that people place on good health.

The rest of the paper is organized as follows. Section 2 reviews the related literature. Section 3 formulates and estimates the health process. Section 4 introduces the full life-cycle model and Section 5 describes its estimation/calibration. Section 6 presents the results. Section 7 concludes.

2 Related literature

A large number of studies have documented the existence of a negative relationship between socio-economic status (SES) and health (see Cutler, Lleras-Muney, and Vogl, 2011 for a review). This relationship is robust to different time periods or geographical regions and different measures of health or socio-economic status. An open question in this literature is what generates the observed gradient: is it the SES affecting health, health affecting the SES, or some third (unobserved) factor affecting both? Our contribution to this literature is in investigation of the latter two channels linking health and socio-economic status in a structural framework.

There are two groups of studies in the literature on the SES-health gradient that are particularly closely related to our work. The first group studies the wealth-health gradient. Smith (1999) uses the HRS data to document the large disparities in wealth between the healthy and the unhealthy. Poterba et al. (2010) uses the same dataset to show that individuals' health status has a significant impact on the subsequent evolution of their assets. Cesarini et al. (2015) use the administrative data on lottery winners in Sweden and conclude that the exogenous change in wealth does not affect subsequent health.

The second branch of the SES-health literature that we are most closely related to studies the economic consequences of health shocks. Dobkin et al. (2016) use the HRS and hospital admissions data and find that hospital admissions result in significant decrease in future earnings and increase in out-of-pocket medical spending. Pohl et al. (2013) use administrative data and hospital records from Chile to show that health shocks measured as accidents significantly reduce subsequent employment. Lundborg et al. (2015) use administrative data from Sweden show that health shocks measured as unexpected hospitalizations have different effects on labor earnings of individuals with

different education. Blundell et al. (2016) adopt the HRS and the English Longitudinal Study of Ageing (ELSA) and estimate a dynamic model of health and employment for individuals between ages 50 and 66. They find that the persistent component of health shocks significantly affects employment.

Methodologically, we relate to structural life-cycle models featuring health shocks. Several studies focus on the implications of medical expense uncertainty on savings after retirement (Ameriks et al., 2015, De Nardi et al., 2010 and 2016, Lockwood, 2015, Nakajima and Telyukova, 2011). French (2005) studies the effects of health on individuals' labor supply over the entire life-cycle while allowing for several channels through which health can affect individuals: productivity, disutility from work and survival uncertainty. Capatina (2015) extends the approach of French (2005) by allowing for uncertain medical expenses. Pashchenko and Parapakkarm (2016) augment the number of channels by allowing health to also affect the access to health insurance. These structural models have been used to answer a wide range of positive and normative questions. A number of studies incorporate life-cycle models with health uncertainty into a general equilibrium framework to study various health reforms in the US (Jeske and Kitao, 2009; Hansen et al., 2014; Pashchenko and Parapakkarm, 2013, 2015, 2016). Several studies focus on explaining historical trends in medical spending and health insurance (Fonseca et al, 2013, Hai, 2015) or on quantifying the distortions implied by the current health insurance system (Pashchenko and Parapakkarm, 2016). In a closely related study, Capatina et al. (2016) measure the effects of health on earnings dynamics over the life-cycle using a model with endogenous human capital accumulation. Our study contributes to this literature along two dimensions. First, we emphasize several dynamic aspects of health evolution that have been overlooked in the existing literature and propose a parsimonious model that can capture these dynamics. Second, to the best of our knowledge, ours is the first structural model that can reproduce the wealth-health gradient, in the cross-section and over time for an individual, which is an essential pre-requisite to appropriately measure the lifetime costs of bad health.

3 Data and sample

An ideal dataset to measure the lifetime effects of bad health would be a large lifelong panel tracking individuals from a young age until the end of their life and containing information on health, medical spending, health insurance, labor earnings, labor supply, and wealth. A data set of this kind does not exist for the US, hence in our study we exploit three datasets to estimate our health shock process and to estimate/calibrate our life-cycle model: the Panel Study of Income Dynamics (PSID), the Health and Retirement

Study (HRS), and the Medical Expenditure Panel Survey (MEPS).

The PSID is a national representative panel survey of individuals and their families. It started in 1968 on an annual basis and from 1997 it is administered bi-annually. The PSID tracks individuals over long period of time but the number of individuals is relatively small. We use the PSID to construct data moments related to health, labor supply, labor earnings, and wealth. For health, labor supply and earnings we use the 1984-1997 waves because in our model a period is one year. The PSID collected wealth information every five years before 1997 and every two years after that. To construct wealth moments, we use waves 1994 and 1999-2011.

The MEPS is a nationally representative survey of households with a focus on medical spending and health insurance variables. It contains individuals of all ages but age is top-coded at age 85. MEPS has a short panel dimension: each individuals is interviewed at most five times over two-year periods. The medical spending reported in MEPS is cross-checked with insurers and providers, which improves their accuracy.² We use waves 1998-2012 of MEPS to estimate medical spending and to construct moments related to insurance coverage.

The HRS is a bi-annual panel survey of a nationally representative sample of individuals over the age of 50. The advantage of the HRS over the PSID is a larger number of older individuals. We use the RAND Version O (waves 1994-2012) of this data set to construct several additional moments related to health and wealth. These moments are not used in our estimation and calibration but are used for external validation. In addition, we use the HRS to estimate the difference in survival probability by health.

In all datasets, we construct a sample of male individuals with a high school degree. We use 1996 as the base year. All level variables are normalized to the base year using the Consumer Price Index (CPI).

4 Health dynamics estimation

To gauge the consequences of bad health over the life-cycle, we have to start by understanding how health evolves. In this section, we first document the cross-sectional and dynamic moments of self-reported health from the PSID (1984-1997). Second, we formulate and estimate a health shock process that is consistent with these moments and discuss its implications.

In our estimation, we use self-reported health for two reasons. First, this variable is available in all three datasets that we use and is consistently measured across them.³

²Pashchenko and Porapakarm (2016) provide more details on the MEPS dataset.

³Below, we show that, for individuals over 50, which is the age group observed in both the PSID and

Second, self-assessed health was shown to be able to capture one’s true underlying health status. In particular, a number of studies find that self-reported health status has significant explanatory power to predict future mortality, even after controlling for other factors that affect mortality (see Idler and Benyamini (1997) for a review and van Doorsaler and Gerdtham (2002) for a more recent examination).

4.1 Data patterns

We construct our measure of health as follows. In both the PSID and HRS, individuals are asked to rank their health on a scale of 1 to 5, where 5 corresponds to excellent, 4 to very good, 3 to good, 2 to fair and 1 to poor. We aggregate these answers into a binary measure of health: individuals who report their health to be excellent, very good and good are classified as healthy or in good health, while individuals who report being in fair and poor health are classified as unhealthy or in bad health.

The top panel of Figure (??) displays the percentage of unhealthy individuals by five-year age brackets. The dots in this figure correspond to the statistics constructed from the PSID, while the crosses refer to the statistics constructed from the HRS. Note that the two datasets are consistent: the percentage of unhealthy individuals over the age of 50 computed from the HRS is similar to that computed from the PSID. The bottom panel of Figure (??) displays the transition between health statuses by five-year age bracket.⁴ These figures show that, as people age, they are more likely to become unhealthy and less likely to recover from bad health.

To better understand the dynamics of health, we next analyze how the transition probabilities to good and bad health depend on the duration of the current health status. Specifically, we compute the transition probability to move to good (bad) health conditional on being in bad (good) health for at least τ consecutive years.⁵ Due to the small sample size we group observations into three relatively large age groups: 30-54, 55-69, and older than 70. Figure (??) plots (shaded bars) the resulting duration-dependent

HRS, the measure of self-reported health is consistent in these two datasets. Attanasio et al. (2011) compare health variable in HRS and MEPS and show that the two datasets are consistent.

⁴These transition probabilities were constructed as follows. Denote health status of an individual i at age t as $h_{i,t}$. For a group of individuals aged 20 to 24, the probability to move to good health conditional on currently being unhealthy can be expressed as

$$\frac{\sum_{t=20}^{24} \sum_i \mathbf{1}(h_{i,t} = B \cap h_{i,t+1} = G)}{\sum_{t=20}^{24} \sum_i \mathbf{1}(h_{i,t} = B \cap h_{i,t+1} = \{B, G\})},$$

where $\mathbf{1}(\cdot)$ is the index function equal to one if its argument is true; otherwise it is zero.

⁵Denote the sequence of health status in the past τ years up to age t as $h_{i,t}^\tau$. For age group 30-54, we compute the probability to move to good health conditional on being unhealthy for at least τ years

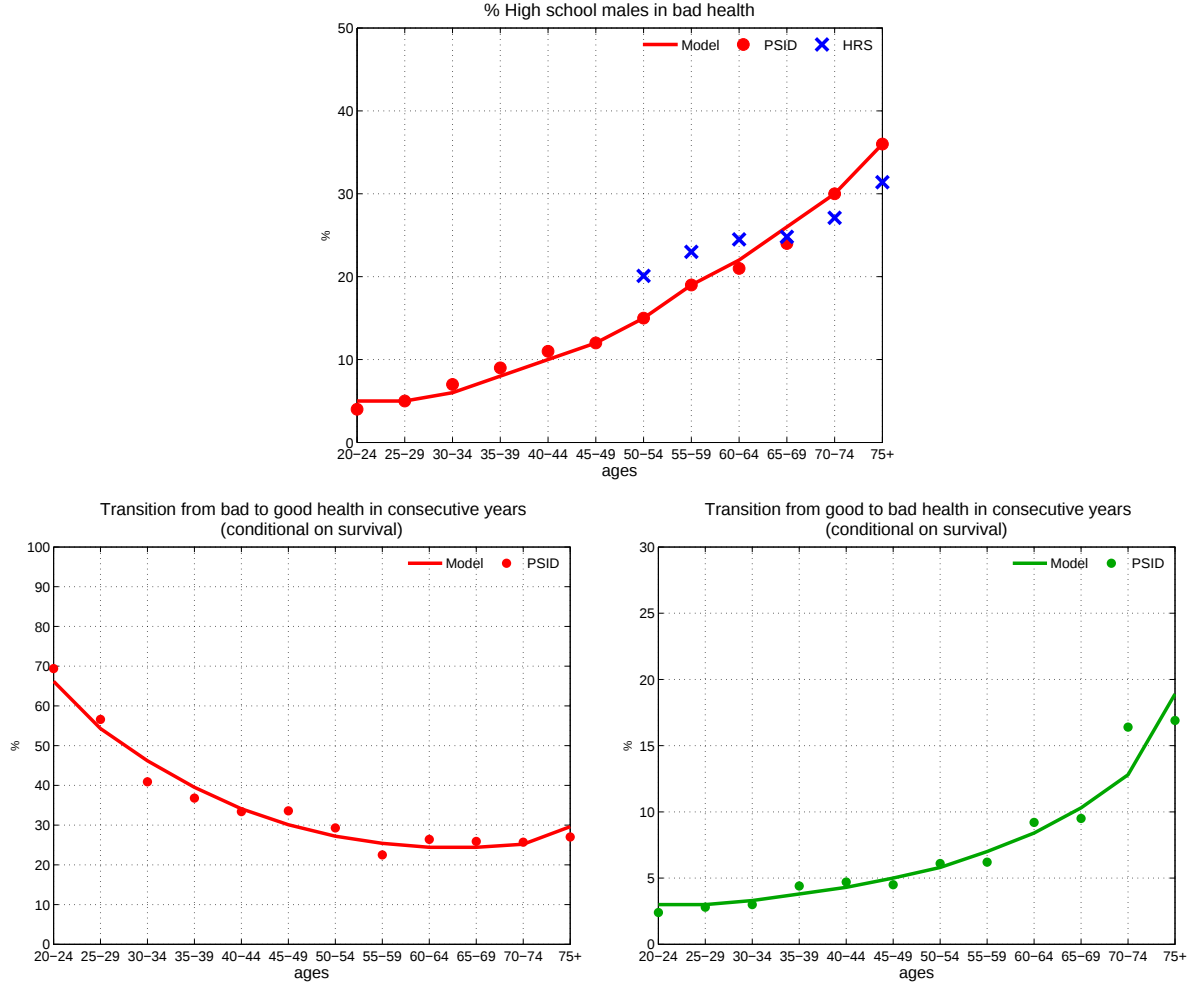


Figure 1: Cross-sectional health moments from the data. Top panel: Percentage of individuals in bad health by age. Bottom left panel: Percentage of individuals moving from bad to good health. Bottom right panel: Percentage of individuals moving from good to bad health. Dots: PSID. Crosses: HRS. Solid lines: model.

transition probabilities from bad to good health (top panel) and from good to bad health (bottom panel).

A key feature of the probability of recovering from bad health is that it declines monotonically with duration: the longer an individual is unhealthy, the less likely he is to become healthy, and this pattern holds for all age groups.⁶

as follows:

$$\frac{\sum_{t=30}^{54} \sum_i \mathbf{1}(h_{i,t}^\tau = B \cap h_{i,t+1} = G)}{\sum_{t=30}^{54} \sum_i \mathbf{1}(h_{i,t}^\tau = B \cap h_{i,t+1} = \{B, G\})}.$$

⁶This negative duration dependence is a robust pattern even when we use smaller age groups, for example, based on a 10-year age bracket. As an additional robustness check, we also compute the

It is important to point out that this decline cannot be captured with the low-order Markov process for health dynamics that is commonly used in the literature (e.g., French, 2005; French and Jones, 2011; Capatina, 2015). For example, a first order Markov process implies that the transition probability does not depend on how long one has been in bad health, while a second order Markov process would imply that this probability is the same for durations longer or equal to two years. In the next section, we discuss how this observation motivates our parametrization of the health process.

In contrast to the transition from bad to good health, the transition from good to bad health does not display noticeable duration dependence, especially at younger ages, as can be seen from the bottom panel of Figure (??). More specifically, there is a noticeable difference in the probability of moving into bad health after having been healthy one- and two-years but, after that, the probability profile is rather flat. In other words, individuals who are healthy for two years have almost the same probability to become sick compared to individuals who are healthy for longer than two years. This lack of duration dependence suggests that the probability to become sick can be well described by a low-order Markov process.

4.2 Health process specification and estimation

The negative duration dependence in the probability of recovering from bad health shown in the top panel of Figure (??) can be explained by two different mechanisms. First, the effects of bad health can be compounding, i.e., individuals who stay sick for a long period of time might have a smaller recovery probability compared with those who are sick for a short period of time. This mechanism is consistent with a high order Markov process. Second, individuals may differ in terms of their ability to recover, i.e., some individuals have lower recovery probability than others. In the latter case, people who are more likely to recover from good health move out of the bad health state faster, therefore the pool of the long-term unhealthy is predominantly composed of individuals who are inherently less likely to recover. This mechanism is consistent with the existence of fixed heterogeneity in health transition probabilities.

The choice of our health shock model is guided by two criteria. First, the model must capture the cross-sectional and dynamic moments of health that we documented in the previous subsection. Second, the model must be parsimonious so that a structural life-cycle model augmented with this health shock process is computationally feasible. Based on this criteria, we formulate our health shock process as a second order Markov

transition probability from bad to good health, where we include in the bad health category only people who report their health being fair, thus excluding individuals with poor health (the worst self-reported health status). The observed pattern still holds in this case.

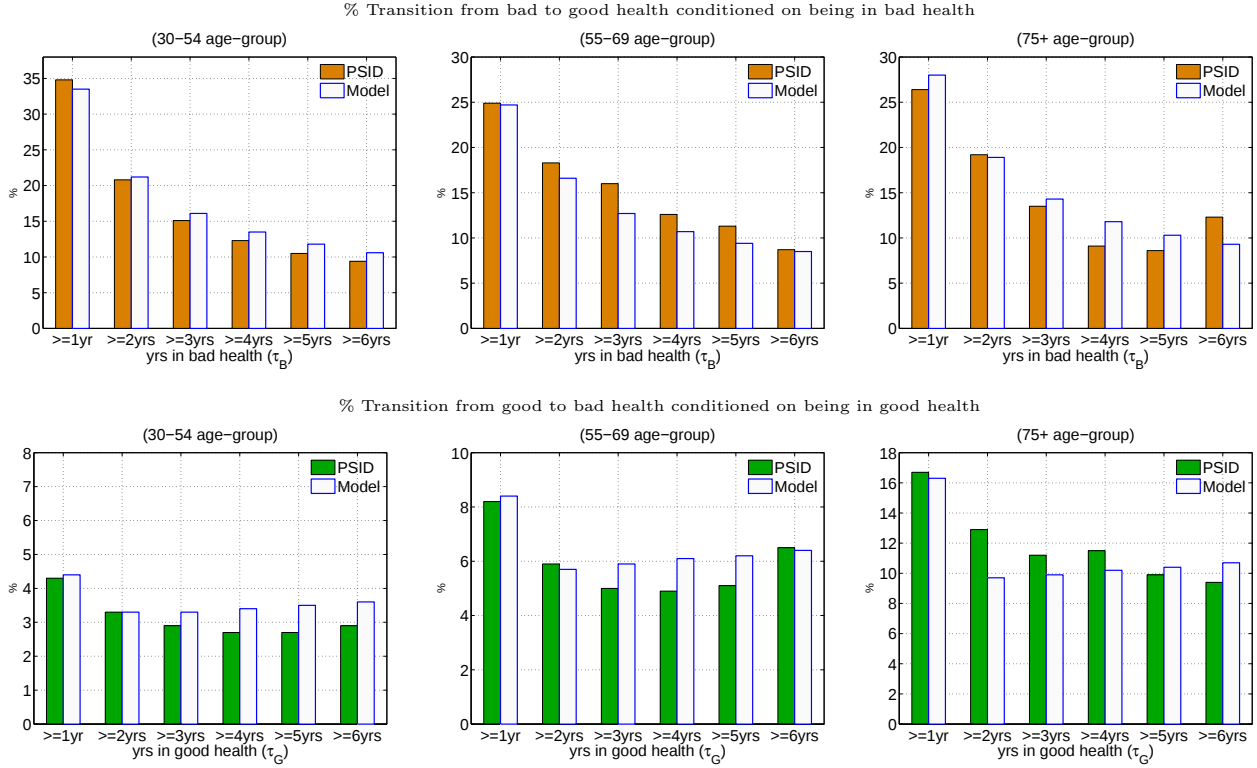


Figure 2: Dynamics of health conditional on duration.

process with fixed heterogeneity. Specifically, the probability of being in good health at age $t + 1$ conditional on surviving to age $t + 1$ and being in bad health for τ_B years is formulated as a following logit function:⁷

$$\text{logit}\left(\pi_{i,t}^G(\tau_B)\right) = \left(a_1 \mathbf{1}(\tau_B = 1) + a_2 \mathbf{1}(\tau_B \geq 2)\right) + \left(b_1 t + b_2 t^2\right) + \eta_i. \quad (1)$$

The first bracket is the second order Markov process, the second bracket is the degree two polynomial in age, and η_i is the fixed heterogeneity or *health type*. We assume that η_i is uniformly distributed over five discrete points that are symmetric around zero, i.e., there are five distinct health types. Note that an individual with low η_i has a lower probability to recover.

In a similar fashion, we model the probability to be in bad health at age $t + 1$ conditional on surviving to age $t + 1$ and being in good health for τ_G years as follows:

$$\text{logit}\left(\pi_{i,t}^B(\tau_G)\right) = \left(a_3 \mathbf{1}(\tau_G = 1) + a_4 \mathbf{1}(\tau_G \geq 2)\right) + \left(b_3 t + b_4 t^2\right) + b_5 \eta_i. \quad (2)$$

⁷The proposed specification is similar to a proportional hazard model commonly used in survival models, where the first bracket is a baseline hazard function.

We allow the health type to have a different effect on the probabilities to get sick and recover by introducing the coefficient b_5 in Equation (??).

We use the Method of Simulated Moments to estimate our health shock process and we target the moments documented in Figures (??) and (??). Since the transition probabilities in Equations (??) and (??) and the targeted moments are conditional on surviving from age t to $t + 1$, we first need to estimate the health-dependent survival probability.

Following Attanasio, et al. (2011), we use the HRS (1994-2012) to estimate a probit model of the two-year survival probability as a function of a cubic polynomial of age separately for the healthy and the unhealthy. The one-year survival probability is computed as the square root of the estimated two-year survival probability. To make our average survival probability consistent with the Social Security Life tables (for the year 1996) we proceed as follows. First, we compute the survival premium which is the difference in survival probabilities between the healthy and the unhealthy estimated from HRS. Second, we compute the fraction of unhealthy individuals from the PSID (1984-1997). Next, using the survival premium, the fraction of unhealthy people, and the survival probability from the Social Security life tables, we compute the implied survival probability for the healthy and for the unhealthy.⁸

Given our estimated survival probabilities and parameters $\theta = \{a_{1-4}, b_{1-5}, \eta_{1-5}\}$, we can simulate the realized health status over the life-cycle for a large number of individuals. The initial distribution of health status is computed based on a sample of individuals aged 19-22 in the PSID, and we assume that the initial health status is orthogonal to the health type η_i .⁹

our algorithm searches for the parameters that minimize the following function:¹⁰

$$\min_{\theta} \left(\mathcal{M}^d - \mathcal{M}^m(\theta) \right) \left(\mathcal{M}^d - \mathcal{M}^m(\theta) \right)',$$

⁸Denote the average probability to survive from age t to $t + 1$ as ζ_t , and the survival probability conditional on health as ζ_t^{bad} and ζ_t^{good} for individuals in bad and good health, respectively. The survival premium Δ_t can be expressed as $\zeta_t^{good} - \zeta_t^{bad}$. Denote the fraction of unhealthy individual of age t as μ_t . The following should hold:

$$\zeta_t = \mu_t \zeta_t^{bad} + (1 - \mu_t) \zeta_t^{good}$$

This can be transformed to find the survival probability for the healthy:

$$\zeta_t^{good} = \zeta_t + \mu_t \Delta_t$$

The survival probability for the unhealthy is then $\zeta_t^{bad} = \zeta_t^{good} - \Delta_t$.

⁹This assumption is innocuous since the absolute majority (96%) of individuals are healthy at this age; thus most people become unhealthy later on, after receiving a health shock.

¹⁰We first do a grid search over the possible values of θ ; and we then use the simplex method to find the minimum using the parameters from the grid search as our initial guess.

where $\mathcal{M}^d$ and $\mathcal{M}^m$ are the targeted moments from the PSID and the simulated data, respectively. The targeted moments in our estimation are listed below.

- The percentage of unhealthy individuals for each five-year age group, as shown in the top panel of Figure (??) (12 moments).
- The health transition probabilities between two consecutive years for each five-year age group, as shown in the bottom panel of Figure (??) (24 moments).
- The duration-dependent profiles of the transition probabilities, as shown in Figure (??) (36 moments).

The identification of θ is straightforward, given the relatively simple specification of our health shock process. The percentage of unhealthy individuals and the age-dependent transition probabilities in Figure (??) help to pin down the coefficients $\{b_1, b_2, b_3, b_4\}$. As discussed in the previous subsection, our Markov process of order two implies that the transition probabilities do not change with duration longer than two years. Thus, the coefficients $\{\eta_1, \eta_2, \eta_3, \eta_4, \eta_5, b_5\}$ are identified from the declining transition probabilities over the duration exceeding two years, as plotted in Figure (??). Finally, the coefficients $\{a_1, a_2, a_3, a_4\}$ are used to capture the difference in the transition probabilities between those in bad (or good) health for at least one year vs. two years. The solid lines in Figure (??) and white bars in Figure (??) show that our parsimonious model of health captures well both the cross-sectional and dynamic moments of health over the life-cycle.

4.3 Estimation results

The implications of our estimated health process are illustrated in Figure (??). The left (right) panel of the figure plots the probability of moving from bad to good (good to bad) health conditional on one's fixed health type and duration of current health status (respectively, bad and good). Comparing the two panels reveals a striking difference in what generates persistence in good and bad health. The left panel shows that fixed heterogeneity has a large impact on the recovery probability. For example, at age 60, an individual of health type η_1 (the "worst" type), who is in bad health, has a probability of recovering of about 5%, while an individual of type η_5 (the "best" type) and the same age, has a probability of recovering of almost 80%. At the same time, duration dependence plays little role, once fixed heterogeneity is controlled for: Individuals who spend one year being unhealthy have almost the same probability to recover as individuals who spend more than two years being unhealthy *conditional on being of the same health type*

(see the comparison of the dashed and solid lines for each health type).¹¹

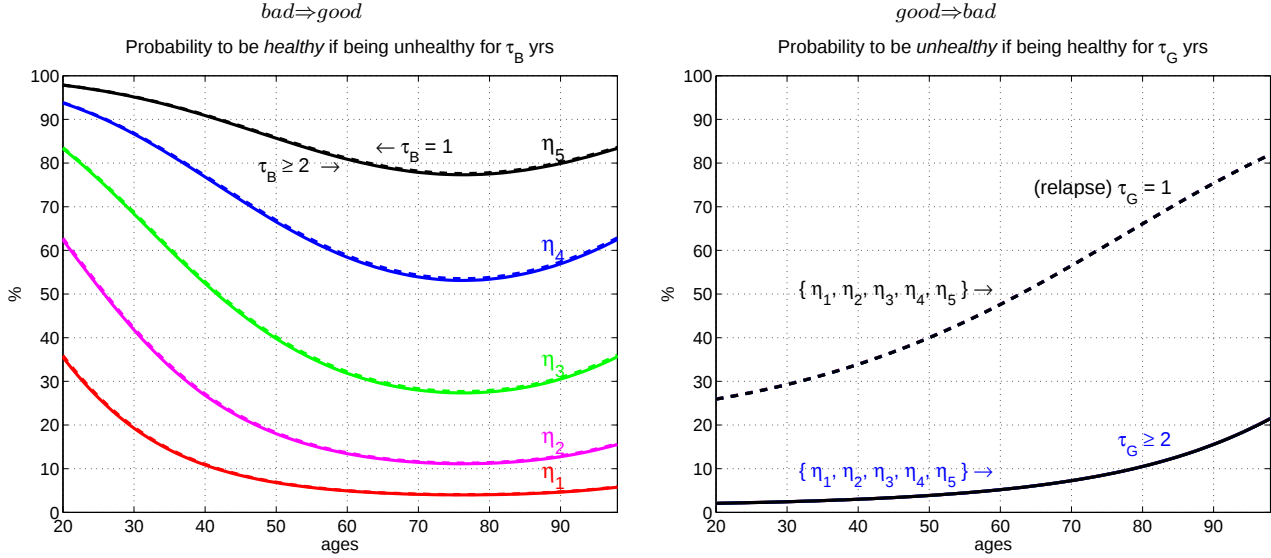


Figure 3: Estimated health process. Dotted line: Conditional on the duration of the current health status being one year. Solid line: Conditional on the duration of the current health status being at least two years.

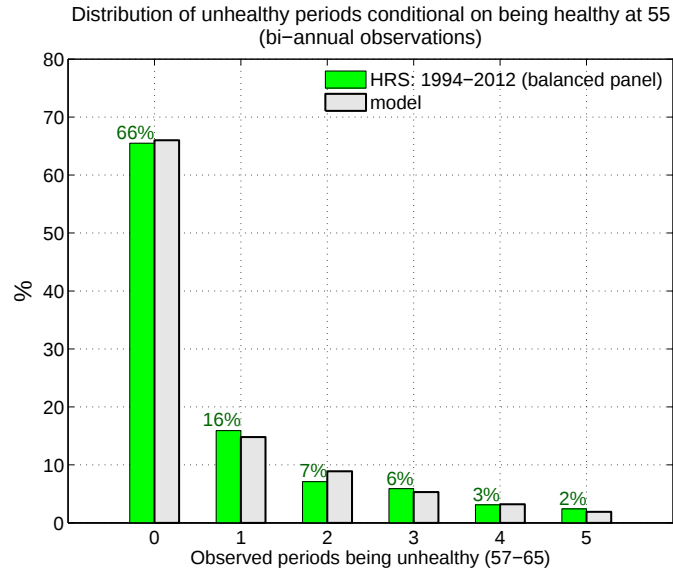


Figure 4: Distribution by unhealthy periods: HRS vs model

¹¹Figure (??) in Appendix ?? compares the duration-dependent transition probabilities for the PSID data and the version of the model where we remove the fixed health type (i.e., assume that $\eta_i = 0$ for all i) but allow for Markov process of order three. As can be seen from the top panel of the figure, in this case we cannot capture the continuous decline in the probability to move to good health after the duration of bad health exceeds three years.

The right panel of Figure (??) shows that, in contrast to the recovery probability, the probability to get sick is influenced very little by the health types: What plays important role in this case is duration dependence. For example, a 60-year old individual who has been healthy for two or more periods has less than a 10% probability to become unhealthy, while an individual of the same age who just recovered (is healthy for only one period) has close to a 50% probability to relapse back to bad health.

For an external validation of our estimated health process, we turn to the HRS and select a sample of healthy males with a high-school degree, aged 55-56, and whom we observe until they are 65-66. This leaves us with 958 individuals or 5,748 individual-year observations. We then compute the distribution of the number of unhealthy periods that these individuals report over this ten-year interval. Since HRS is a bi-annual survey, an individual can report being unhealthy for at most five periods. We then construct a comparable distribution using simulated data from our model. Figure (??) shows that our simulated data and the data from the HRS look very similar.

4.4 What accounts for the long spells of bad health?

Using our estimated model, we can construct the lifetime distribution of unhealthy years. The left panel of Figure (??) plots the distribution of individuals by the total number of years that they have spent being unhealthy between ages 20 and 64, conditional on being alive at age 64. Most people are relatively healthy during their working life: 72% of individuals experience less than 5 years of bad health. However, a non-trivial number of individuals spend more than a third of their working period being unhealthy. For instance, 5% of individuals experience 16 or more years in bad health. The right panel of Figure (??) illustrates how this distribution differs by health type by comparing two extreme health types: individuals born with the best health type (η_5) and those born with the worst health type (η_1). The difference in these distributions is large. Among individuals with η_5 , 92% spend less than 5 years being unhealthy and almost none of them experience more than 11 unhealthy years. In contrast, among η_1 -type individuals, 24% experience between 11 and 20 unhealthy years, and 13% are unhealthy for 20 years or longer. Thus, long spells of bad health are primary concentrated among individuals with the the worst health types, or in other words, long spells of bad health are mostly due to fixed heterogeneity, rather than repeated draws of bad realizations from a persistent stochastic process.

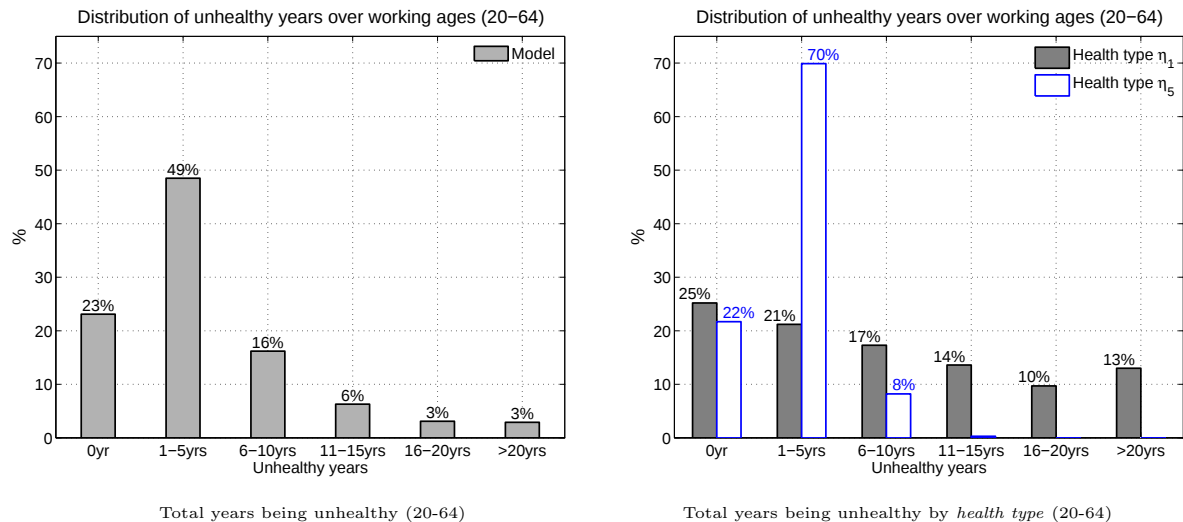


Figure 5: Distribution by lifetime unhealthy years. Left panel: all individuals. Right panel: individuals with η_1 and η_5 health types

4.5 How should the health type be interpreted?

As our previous discussion shows, the health type (η) plays an important role in the persistence of bad health: our specification allows for the possibility that people have different ability to cope with a disease and our estimation shows that this heterogeneity is large.¹²

The origins of this heterogeneity can be due to genetic differences and/or childhood circumstances.¹³ In other words, individuals can recover differently from sickness due to genetic predisposition and/or life-style. To understand whether we can find support for these mechanisms in the data, we use the same sample of individuals from the HRS that we used to construct Figure (??). Specifically, we consider healthy individuals aged 55-56 and whom we observe until they are 65-66. Table ?? sorts these individuals based on the total number of unhealthy periods that they report over this ten-year interval. An interesting observation is that there is a correlation between their *future* number of unhealthy periods and factors that can be linked to genetics (whether their parents are still alive) or lifestyle (smoking and body mass index) *recorded at age 55*. In particular, individuals who report four and five unhealthy periods between ages 57 and 65, at age 55 are significantly more likely to smoke, have higher body mass index (BMI) and less likely to have living parents.

¹²Halliday (2008) estimates the dynamics of health allowing for history-dependence and fixed heterogeneity. He also finds that for a large part of his sample persistence is mostly generated by fixed heterogeneity.

¹³Case et al. (2002) show that early childhood circumstances have a long-lasting effects on health later in life.

# unhealthy periods (57-65)	At age 55			
	% father alive	% mother alive	% smoking	BMI (median)
0-1	21.2	49.5	23.2	27
2-3	20.2	46.7	25.9	28
4-5	15.2	36.9	43.5	30

All individual are healthy at age 55.
 BMI stands for Body Mass Index.

Table 1: Selected characteristics of individuals at age 55 by the number of unhealthy periods between ages 57 and 65, HRS.

That is, even among a relatively homogeneous sample of healthy males in the same educational group, we observe some heterogeneity in unhealthy spells, likely due to fixed or long-lasting factors that are correlated with the future evolution of their health. This features of the data are consistent with our stylized model of health dynamics.

5 The life-cycle model

In this section, we construct a life-cycle model with health uncertainty. Health affects individuals through multiple channels and evolves according to the process that we describe in the previous section.

5.1 Demographics, preferences, and labor income

A model period is one year and each individual lives at most N periods. During the first $R - 1$ periods of life, an individual chooses whether to work or not, while at age R , all individuals retire.

At age t , an agent's health, h_t , can be either good ($h_t = 1$) or bad ($h_t = 0$). Health evolves according to the process defined in Equations (??) and (??), i.e., current health status depends on health status in the previous two periods and on health type $\eta_i \in (\eta_1, \dots, \eta_5)$. One's current health status affects one's medical spending, productivity, disutility from work, access to health insurance, and survival probability. We denote the probability of surviving from period t to $t + 1$ as ζ_t^h .

The individual's discount factor is β_i . We assume the discount factor can take two values: $\beta_i \in (\beta_{low}, \beta_{high})$, where $\beta_{low} < \beta_{high}$. We allow for correlation between one's discount factor and health type, specifically, at age 20, $0 \leq Pr(\beta_j | \eta_m) \leq 1$, where $j \in (low, high)$ and $m \in (1, \dots, 5)$.

An individual is endowed with one unit of time that can be used for either leisure or work. Labor supply (l_t) is thus indivisible; $l_t \in \{0, \bar{l}\}$. Work implies a fixed utility cost ϕ_W for healthy individuals and $\phi_W + \phi_B$ for unhealthy ones. We assume that the preferences of individuals over consumption and leisure take the following form:

$$u(c_t, l_t, h_t) = \frac{(c/\bar{n}_t)^{1-\rho}}{1-\rho} - \phi_W \mathbf{1}_{\{l_t > 0\}} - \phi_B \mathbf{1}_{\{h_t=0, l_t > 0\}} + \bar{\mathbf{b}}, \quad (3)$$

where $\mathbf{1}_{\{\cdot\}}$ is an indicator function mapping to one if its argument is true, ρ is risk-aversion, and $\bar{n}_t$ is age-specific household size. Because the first three terms in Equation (??) are negative, we add a positive term $\bar{\mathbf{b}}$ to ensure that individuals always prefer being alive (and receiving positive utility flow) to being deceased (and receiving zero utility flow). In this approach we follow Hall and Jones (2007).

Individuals also have bequest motives; the utility individuals derive from leaving a bequest in the amount k is specified as follows:

$$\theta_{Beq} \frac{(k + k_{Beq})^{1-\rho}}{1-\rho},$$

where θ_{Beq} determines the strength of the bequest motive and k_{Beq} is a shift parameter that determines to what extent bequests are a luxury good. In this approach we follow De Nardi (2004).

The earnings of individuals are equal to $z_t^h l_t$, where z_t^h is an idiosyncratic productivity component, which takes the following form:

$$z_t^h = \lambda_t^h \Upsilon_t. \quad (4)$$

Here, λ_t^h is a deterministic function of age and health, and Υ_t is the stochastic shock that will be described in Section ??.

5.2 Medical expenditures and health insurance

During each period, every agent receives a medical expenditure shock, (x_t^h) , which depends on age and health. We denote the distribution of medical shocks as $\mathcal{G}_t(x_t^h|h_t)$.

Every working age individual can buy health insurance against medical shocks in the individual health insurance market. The price of health insurance in the individual market depends on one's age and health. We denote the individual market price as $p_I(h_t, t)$.

During every period, a working-age individual receives an offer to buy employer-sponsored health insurance (ESHI) with probability $Prob_t$, which depends on age (t),

productivity (z_t^h), and health (h_t). The variable $g_t^{h,z}$ characterizes the status of the offer: $g_t^{h,z} = 1$ if an individual gets an offer, and $g_t^{h,z} = 0$ if he does not. We assume that an employer who offers ESHI fully covers the premium, i.e., the employer contribution is 100%.¹⁴

All retired individuals are covered by public health insurance, Medicare. We denote the Medicare premium as P_{MCR} .

We index the insurance status of an individual using the variable i_H ; $i_H = 0$ corresponds to the absence of any insurance, $i_H = 1$ corresponds to individual insurance, $i_H = 2$ corresponds to group (or ESHI) insurance, and $i_H = 3$ corresponds to Medicare. All types of insurance only provide partial medical expenses coverage. We denote by $cvg(x_t^h, i_H)$ the fraction of medical expenditures covered by insurance, and we allow it to be a function of the medical shock and insurance type. Note that $cvg(x_t^h, 0) = 0$.

5.3 Taxation and social transfers

All individuals pay an income tax $\mathcal{T}(y_t)$ that consists of two parts: a progressive tax and a proportional tax. Taxable income y_t is based on both labor and capital income. Working households also pay payroll taxes, which include the Medicare tax (τ_{MCR}) and the Social Security tax (τ_{ss}). The Social Security tax rate for earnings above $\bar{y}_{ss}$ is zero. Consumption is taxed at a proportional rate of τ_c .

We also assume a public safety-net program, $T_t^{SI}(\bar{c})$. This program guarantees every household a minimum consumption level $\bar{c}$, which is a reduced form representation of several means-tested programs existing in the US, including Medicaid, food stamps, and Supplement Security Income. In addition, the consumption floor captures the existence of uncompensated care.¹⁵

Retired households receive Social Security benefits ss . In practice, these payments depend on the previous history of earnings. To reduce computational costs, we allow ss to depend only on the fixed productivity type, which is part of the stochastic component of productivity Υ_t (see Section ??), and the fixed health type η_i . Since the labor supply decisions of individuals over the life-cycle are affected by fixed productivity and health types, this approach allows us to capture the resulting heterogeneity in pension benefits without introducing an additional state variable.

¹⁴On average, employers who offer ESHI contribute around 80% of the premium for single coverage and around 70% for family coverage (Kaiser Family Foundation, 2009); we abstract from workers's contribution for simplicity, this assumption does not affect our results but helps to lower the computational costs since individuals with an ESHI offer always buy insurance.

¹⁵In 2004 85 percent of the uncompensated care was paid by the government. The largest portion was coming from the disproportionate share hospital (DSH) payment (Kaiser Family Foundation, 2004).

5.4 Timing of the model

The timing of the model is as follows. At the beginning of the period, individuals learn their productivity, health and ESHI offer status. Based on this information, an individual decides his labor supply (l_t) and insurance choice (i_H). At the end of the period, the medical expenses shock (x_t^h) is realized. After paying the out-of-pocket medical expenses, an individual chooses his consumption (c_t) and savings for the next period (k_{t+1}). The problem of retired individuals is simpler; they only choose consumption and savings for the next period.

5.5 The optimization problem

Working age individuals ($t < R$). At the beginning of each period, the state variables for an individual are capital ($k_t \in \mathbb{K} = R^+ \cup \{0\}$), health status in the current and previous periods ($h_t, h_{t-1} \in \mathbb{H} = \{0, 1\}$), idiosyncratic labor productivity ($z_t^h \in \mathbb{Z} = R^+$), ESHI offer status ($g_t^{h,z} \in \mathbb{G} = \{0, 1\}$), age ($t \in \mathbb{T} = \{1, 2, \dots, R-1\}$), health type ($\eta_i \in \{\eta_1, \dots, \eta_5\}$) and discount factor ($\beta \in \{\beta_{low}, \beta_{high}\}$). We denote the vector of state variables as $\mathbb{S}_t$, $\mathbb{S}_t = (k_t, h_{t-1}, h_t, z_t^h, g_t^{h,z}, \eta_i, \beta_i)$.

The value function of a working age individual at the beginning of period t is:

$$V_t^i(\mathbb{S}_t) = \max_{l_t, i_H} \sum_{x_t^h} \mathcal{G}_t(x_t^h | h_t) W_t^i(\mathbb{S}_t; l_t, i_H, x_t^h)$$

where

$$W_t^i(\mathbb{S}_t; l_t, i_H, x_t^h) = \max_{c_t, k_{t+1}} u(c_t, l_t, h_t) + \beta_i \left[\zeta_t^h E_t(V_{t+1}^i(\mathbb{S}_{t+1})) + (1 - \zeta_t^h) \theta_{Beq} \frac{(k_{t+1} + k_{Beq})^{1-\rho}}{1-\rho} \right] \quad (5)$$

subject to

$$k_t(1+r) + z_t^h l_t - x_t^h (1 - cvg(x_t^h, i_H)) - P_t^h - Tax + T_t^{SI}(\bar{c}) = c_t + k_{t+1} \quad (6)$$

$$P_t^h = \begin{cases} 0 & ; \text{ if } i_t \in \{U, G\} \\ p_I(h_t, t) & ; \text{ if } i_t \in \{I\} \end{cases} \quad (7)$$

$$T_t^{SI}(\bar{c}) = \max(0, \bar{c} + Tax + P_t^h + x_t^h(1 - cvg(x_t^h, i_H)) - k_t(1+r) - z_t^h l_t)$$

$$Tax = \mathcal{T}(y_t) + \tau_{MCR} z_t^h l_t + \tau_{ss} \min(z_t^h l_t, \bar{y}_{ss}) \quad (8)$$

$$y_t = k_t r + z_t^h l_t \quad (9)$$

$W_t^i(\mathbb{S}_t; l_t, i_H, x_t^h)$ is the interim value function conditional on the labor supply and insurance choices after the medical shock is realized. The conditional expectation on the right-hand side of Equation (??) is over $\{h_{t+1}, z_{t+1}^h, g_{t+1}^{h,z}\}$. Equation (??) is the budget constraint; in this constraint P_t^h is insurance premium, which is described in Equation (??). In Equation (??), the first term is the income tax and the last two terms are payroll taxes. Equation (??) describes taxable income.

Retired individuals ($t \geq R$). The state variables for retired people are assets (k_t), health in the current and previous periods (h_t, h_{t-1}), medical shock (x_t^h), age (t), health type (η_i), and discount factor (β_i).¹⁶ We denote the vector of state variables as $\mathbb{S}_t^R$, $\mathbb{S}_t^R = (k_t, h_{t-1}, h_t, \eta_i, \beta_i)$.

The value function of a retired household is:

$$V_t^i(\mathbb{S}_t^R) = \sum_{x_t^h} \mathcal{G}_t(x_t^h | h_t) W_t^i(\mathbb{S}_t^R; x_t^h)$$

where

$$W_t^i(\mathbb{S}_t^R; x_t^h) = \max_{c_t, k_{t+1}} u(c_t, 0, h_t) + \beta_i \left[\zeta_t^h E_t V_{t+1}^i(\mathbb{S}_{t+1}^R) + (1 - \zeta_t^h) \theta_{Beq} \frac{(k_{t+1} + k_{Beq})^{1-\rho}}{1 - \rho} \right] \quad (10)$$

subject to

$$k_t(1 + r) + ss - x_t^h(1 - cvg(x_t^h, 3)) - P_{MCR} - \mathcal{T}(y_t) + T_t^{SI}(\bar{c}) = c_t + k_{t+1} \quad (11)$$

$$T_t^{SI}(\bar{c}) = \max(0, \bar{c} + \mathcal{T}(y_t) + P_{MCR} + x_t^h(1 - cvg(x_t^h, 3)) - k_t(1 + r) - ss)$$

$$y_t = k_t r + ss \quad (12)$$

$W_t^i(\mathbb{S}_t^R; x_t^h)$ is the interim value function conditional on medical shock realization. The conditional expectation on the right-hand side of Equation (??) is over h_{t+1} . Equation (??) is the budget constraint.

6 Model parameters estimation

In this section, we explain our strategy to estimate our model parameters and exogenous shocks from the PSID and MEPS. The information related to earnings, labor

¹⁶As explained in Section ??, Social Security payments ss depend on the fixed productivity type; thus, fixed productivity is also part of the state variables for retired households. We omit it from the description of the optimization problem to simplify the notation.

supply, and wealth is taken from the PSID, while the information about medical expenses and health insurance is taken from MEPS.¹⁷

We adopt a two-step strategy. In the first stage, we estimate the shock processes and some parameters that can be directly identified from the data without using our model; for example, the health shock (discussed in Section ??), the medical expense shock and labor productivity shock. In the second step, we estimate the preference parameters of the model $\{\phi_W, \phi_B, \beta_{low/high}, Pr(\eta_i|\beta_{low}), \theta_{Beq}, k_{Beq}\}$, the consumption floor $\bar{c}$, the age-dependent labor productivity λ_t^h , and loads on individual insurance premium $\{\xi^h, \varphi^h\}$ using our structural model to match the targeted moments from the data. We summarize the model's parameters in Appendix ??.

6.1 The medical expenses shock

The medical costs in our model correspond to total paid medical expenditures in the MEPS dataset. The medical expense shock is approximated by a 3-state discrete health- and age-dependent stochastic process. For each age and health status, these three states correspond to the average medical expenses of three groups: those with medical expenses below the 50th, 50th to 97th, and above the 97th percentiles, respectively.¹⁸

More details about the estimation of our medical shock process are in Appendix ??.

6.2 Health insurance status and coverage

We define a person as having employer-based insurance if he reports having ESHI for at least eight months of the year. The same criterion is used when defining a person as having individual insurance. If a person reports having both ESHI and individual insurance in one year and each coverage lasts for eight months or less, but the total duration of coverage lasts for more than eight months, we classify this person as individually insured.

We use MEPS to estimate the fraction of medical expenses covered by insurance

¹⁷MEPS has five rounds of interviews over two years and in each round, during which individuals are asked to rank their health on a scale of 1 to 5 as in the PSID and HRS. We compute the average score of self-assessed health for each year and categorize individuals with the average score above 3.0 as healthy and below 3.0 as unhealthy. We choose the cutoff point of 3.0 to make the fraction of unhealthy individuals in MEPS comparable to this fraction in PSID.

¹⁸MEPS tends to underestimate aggregate medical expenditures (Pashchenko and Porapakkarm, 2016). To bring aggregate medical expenses computed from the MEPS in line with the corresponding statistics in the National Health Expenditure Account (NHEA), the estimated medical expenses were multiplied by 1.60 for people younger than 65 years old and by 1.90 for people 65 or older. These numbers correspond to the ratio of aggregate medical spending in NHEA divided by aggregate medical spending in MEPS for people younger and older than 65 years old, respectively, averaged over the years 2002, 2004, 2006, 2008, and 2010 (the years when NHEA provides the aggregate statistics by age).

policies $cvg(x_t^h, i_H)$. We report these estimates in Appendix ??.¹⁹ For retired households, we set $cvg(x_t^h, 3)$ to 0.5, following Attanasio et al. (2011).

To compute the premium for individual insurance, we assume that it is based on the expected medical costs of an individual plus administrative loads:

$$p_I(h_t, t) = \xi^h EM_t(h_t, t) + \varphi^h \quad (13)$$

The term ξ^h is a proportional load, while φ^h is a fixed load. We allow the loads to depend on health to capture the fact that unhealthy individuals may face more frictions when purchasing insurance through the individual market, for example, through search costs or a larger probability of being denied coverage due to pre-existing conditions.

The expected medical costs covered by insurance are determined as follows:

$$EM_t(h_t, t) = \sum_{x_t^h} x_t^h cvg(x_t^h, 1) \mathcal{G}_t(x_t^h | h_t)$$

The estimated fixed load (φ^h) is \$541 for both health status; the proportional load (ξ^h) is 1.06 for the healthy and 1.12 for the unhealthy. Given other parameters of the model, we estimate the proportional load and fixed costs by matching the life-cycle profile of individual insurance coverage among the healthy and the unhealthy.

6.3 ESHI offer rate

We assume that individuals receive ESHI offer with probability $Prob_i$, which is estimated from the following logit regression:²⁰

$$logit(ESHI_i) = \sum_{h=0}^1 \left(a_0^h + a_1^h \log(inc_i) + a_2^h (\log(inc_i))^2 + a_3^h (\log(inc_i))^3 + a_4^h ESHI_{i,-1} \right) \quad (14)$$

$ESHI_i$ is one if an individual i is insured through ESHI, otherwise it is zero; inc_i is labor income normalized by the average income; $ESHI_{i,-1}$ is ESHI status in the previous period. This specification allows for the positive relationship between labor income and ESHI coverage that is observed in the data. In MEPS, unhealthy people have a noticeably lower ESHI coverage. To capture this, we allow for the ESHI offer probability to be

¹⁹Due to the small sample size of those with individual insurance, we assume that ESHI and individual insurance provide the same coverage, i.e., $cvg(x_t^h, 1) = cvg(x_t^h, 2)$.

²⁰We use only individuals who earn more than \$2,470 per year in base year dollars. Note that we use the ESHI status as opposed to ESHI offer status in our logit regression. Since everyone in our model always buys employer-based insurance if offered, there is no difference between being insured through ESHI and receiving an ESHI offer in our model. In MEPS, 95% of individuals who receive an ESHI offer take it.

different for the healthy and the unhealthy. For the the initial distribution at age 20, we run a separate logit regression among individuals aged 19-22 without including $ESHI_{i,-1}$ in the regression specification.

6.4 Taxes and government transfers

In specifying the tax function $\mathcal{T}(y)$ we use a combination of the nonlinear functional form formulated by Gouveia and Strauss (1994), and a linear income tax τ_y :

$$\mathcal{T}(y) = a_0 [y - (y^{-a_1} + a_2)^{-1/a_1}] + \tau_y y$$

The first term captures the progressive income tax; in this functional form, a_0 controls the marginal tax rate faced by the highest income group, a_1 determines the curvature of marginal taxes, and a_2 is a scaling parameter. Following Gouveia and Strauss (1994), we set a_0 and a_1 to 0.258 and 0.768, respectively. The parameters a_2 and τ_y are set to 0.6160 and 0.066 percent following Pashchenko and Porapakkarm (2016).

The Medicare, Social Security and consumption tax rates were set to 2.9 percent, 12.4 percent and 5.67 percent, respectively. The maximum taxable income for Social Security ($\bar{y}_{ss}$) is set to \$62,700.

For retired individuals, Social Security pension payments ss are calculated as follows. We first group individuals based on their fixed productivity and health types. For each group, we compute the average labor income over the 35 highest-earning years and then apply the Social Security benefit formula for 1996.

6.5 The labor productivity process

Labor productivity takes the following form:

$$z_t^h = \lambda_t^h \Upsilon_t = \lambda_t^h \exp(y_t) \exp(\gamma) \quad (15)$$

where λ_t^h is a deterministic component that depends on age and health, while the stochastic component of productivity Υ_t consists of a persistent shock y_t and a fixed productivity type γ :

$$y_t = \rho_y y_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \quad (16)$$

$$\gamma \sim N(0, \sigma_\gamma^2)$$

We use the PSID (1984-1997) to estimate the parameters of the stochastic component Υ_t . For that, we use a sample of male workers with a high-school degree, where a worker is defined as a person who earns at least \$2,470 per year in our base year dollars (this

corresponds to working at least 10 hours per week and earning a minimum wage of \$4.75 per hour). This leaves us with 1,036 individuals or 10,778 individual-year observations. Our estimation strategy closely follows French (2005). We first run the following fixed effect regression:

$$\log(inc_{it}) = \sum_{h_{it}=0,1} (b_1^h age_{it} + b_2^h age_{it}^2) + \sum_{t=1984}^{1997} b_t D_t + f_i + u_{it},$$

where inc_{it} is income defined as the household's head labor earnings and labor income from business converted into dollars of 1996 using CPI; and D_t is a set of year dummy variables. Then we construct the empirical autocovariance matrix using the residuals from this regression ($f_i + u_{it}$). We estimate the parameters of the productivity shock in Equation (??) by minimizing the distance between the autocovariance matrix computed from our model and the empirical autocovariance matrix.²¹ Our resulting estimates are as follows: ρ_y is 0.9275, σ_ε^2 is 0.0209, and σ_ξ^2 is 0.042. In our computation, we discretize the shock processes using 9 gridpoints for y_t and 3 gridpoints for γ . The grid of y_t is expanding over ages to capture the increasing cross-sectional variance. Because our AR(1) process is highly persistent, we use the method in Floden (2008) for our discretization.

To estimate the deterministic part of productivity λ_t^h , we need to take into account the fact that in the data, we only observe labor the income of workers and we do not know the potential labor income of non-workers. To correct for the possible selection into employment, we follow French (2005) and Pashchenko and Porapakarm (2015 and 2016) and estimate λ_t^h together with the disutility from work parameters ϕ_W and ϕ_B inside the model as follows. We start by estimating the labor income profiles of working individuals from the PSID separately for the healthy and for the unhealthy. Then given other parameters discussed in Subsection ??, we solve our model using an initial guess of λ_t^h , ϕ_W and ϕ_B , and estimate labor income profile on our simulated data separately for healthy and unhealthy workers. We compare these model-generated labor income profiles with the data and update our guess of λ_t^h . Simultaneously, we compare the fraction of workers in our simulated data with the PSID, separately for the healthy and the unhealthy and update our disutility from work parameters ϕ_W and ϕ_B . For example, if too many individuals work in our model we increase the disutility from work parameters. Then we solve the model again and compare labor income profiles and employment profiles constructed from a new simulated data with PSID. We continue until the average labor income of workers and the fraction of workers in our simulated data are the same as in PSID both for the healthy and the unhealthy. Figure (??) compares the resulting

²¹This is a standard procedure commonly used in macroeconomic literature (see for example, Storesletten et al., 2004), French (2005).

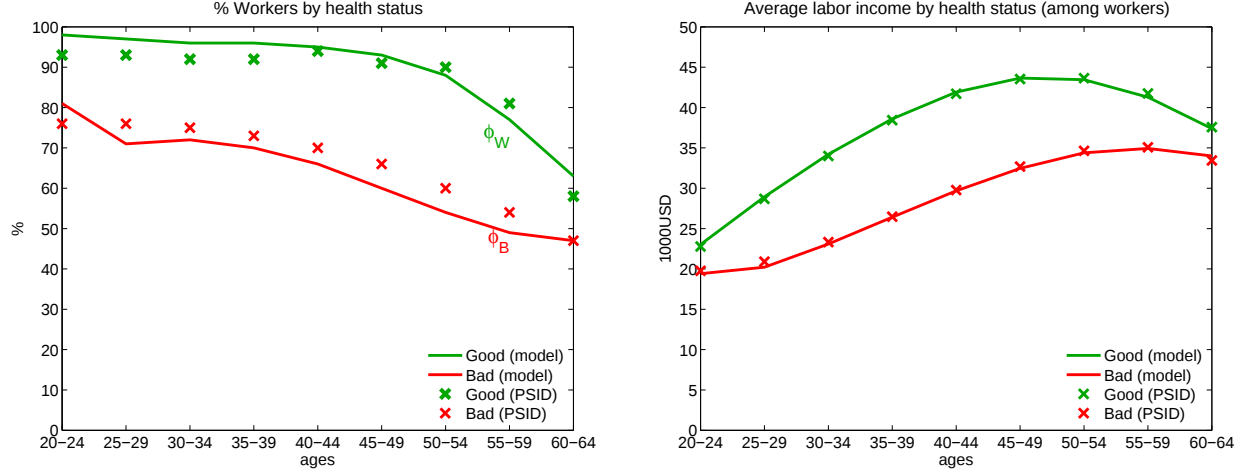


Figure 6: Employment by health (left panel) and average labor income among workers by health (right panel): data vs model

employment (left panel) and the average labor income for workers (right panel) generated by our model with the targeted profiles from the data.

6.6 Demographics, preferences, and the consumption floor

An individual enters the model at the age of 20 and can live at most until the age of 99. We set the risk aversion ρ to 3, which is the common value in macroeconomic and structural life-cycle models. The age-dependent family size $\bar{n}_t$ is the average family size from PSID.

To identify the remaining preference parameters $\{\beta_{low/high}, Pr(\beta_{low}|\eta_i), \theta_{Beq}, k_{Beq}\}$ and consumption floor $\bar{c}$ we use the information from the wealth profiles for the healthy and the unhealthy constructed from PSID (1994, 1999-2013).²² Note that the wealth variable is measured at the household level but our model abstracts from the heterogeneity in family size. To make the data and our model comparable, we first construct the net wealth after controlling for the family size. Specifically, we estimate the following regression equation:

$$wealth_{it} = \sum_{h=0}^1 \left(d_{age}^h D_{it}^{age} + d_1^h n_{it} + d_2^h n_{it}^2 \right) + \sum_{t=1994}^{2013} d_t D_t + res_{it}, \quad (17)$$

where $wealth_{it}$ is the net wealth, D_{it}^{age} and D_t are age and year dummy variables, and n_{it}

²²More specifically, we use the net wealth variable which is derived from the value of business/farm, checking/saving accounts, real estate, stock, vehicles, other assets, annuity/IRA accounts, and home equity net of the value of mortgages/debts. We convert it to 1996 dollars using the CPI.

is the number of individuals in a family unit.²³ Given the estimated coefficients and the residual res_{it} we replace n_{it} in the above equation with the average family size of each age $\bar{n}_{age}$, and compute the net wealth after controlling for family size.²⁴ Then we use our measure of net wealth to construct the targeted 25th, 50th, 75th percentile of wealth distribution by health status as reported in Figure (??) (dashed lines). As a comparison, we also apply the same method to the net wealth variable in HRS (1994-2012), and the results are plotted as crossed marks in Figure (??). Note that the constructed net wealth variables from the two datasets are remarkably similar.

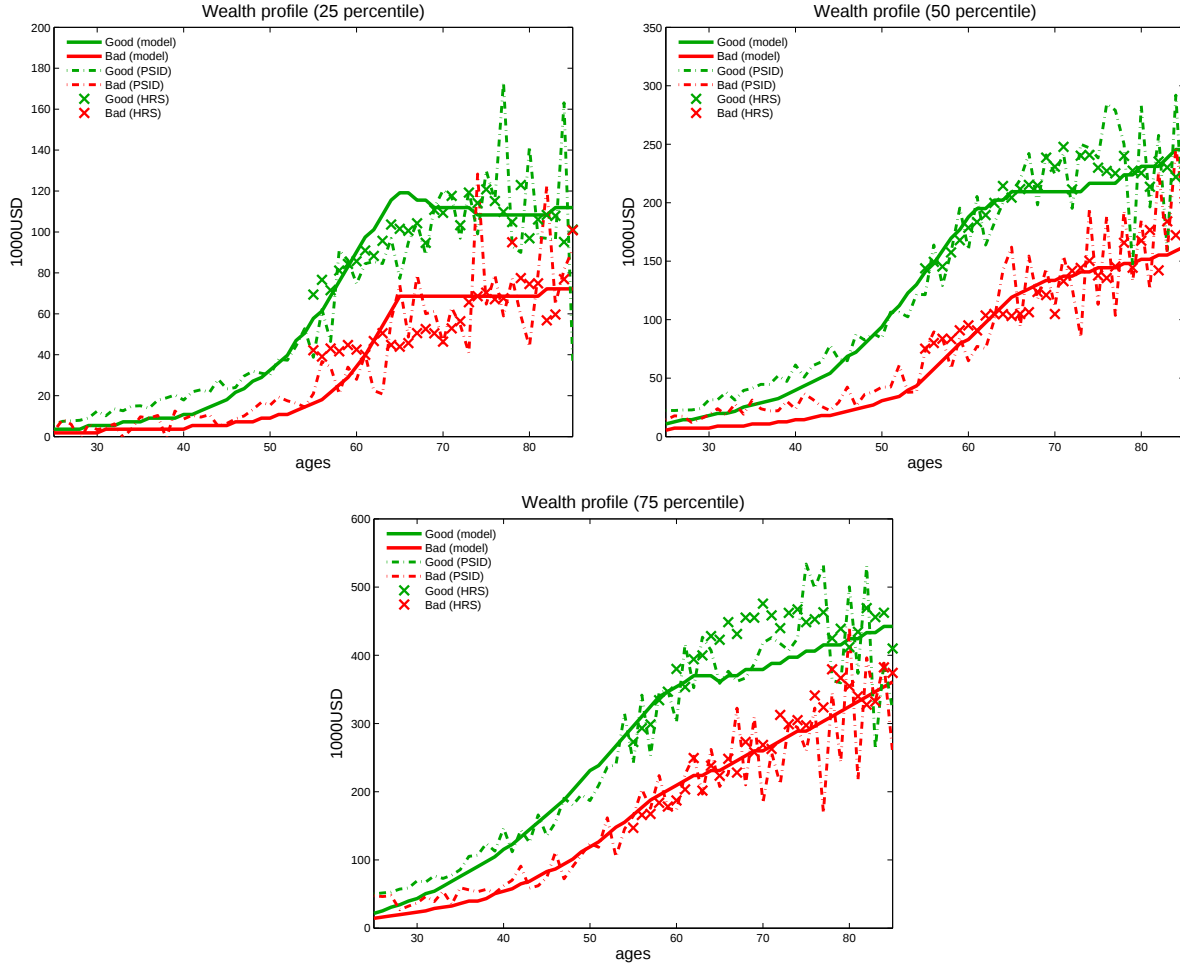


Figure 7: Wealth profiles: data vs model

We use the simplex method to search for the set of parameters $\{\beta_{low/high}, Pr(\beta_{low}|\eta_i), \theta_{Beq}, k_{Beq}, \bar{c}\}$ that minimizes the difference between the wealth profiles by health status from age 30

²³In the regression, we truncate the maximum number of individuals to nine.

²⁴Specifically, we compute $\sum_{h=0}^1 (d_{age}^h D_{it}^{age} + d_1^h \bar{n}_{age} + d_2^h \bar{n}_{age}^2) + d_{1999} + res_{it}$. By construction, we remove the variation in net wealth due to the variation in family size that is orthogonal to health status and age. And we choose d_{1999} which is the closest year to our base year.

to 85 constructed from our simulated data and PSID. We discard the wealth moments below age 30 because we assume that individuals enter the model with zero assets. The solid lines in Figure (??) display the wealth profiles from our model and show that they are close to the ones from the data. In addition, the model can match the wealth gap between healthy and unhealthy individuals for the 25th, 50th, and 75th percentiles as in the data.

Parameters		(1) baseline	(2) no correlation	(3) no health type
consumption floor	$\bar{c}$	\$3,593	to be filled	\$3,335
bequest parameter	θ_{Beq}	4,464	...	4,305
"	κ_{Beq}	246,371	...	233,642
discount factor	$\beta_{low}, \beta_{high}$	0.904, 0.995	...	0.901, 0.988
% β_{low} by η_i at age 20	$Pr(\beta_{low} \eta_1)$	89%	50%	-
"	$Pr(\beta_{low} \eta_2)$	81%	50%	-
"	$Pr(\beta_{low} \eta_3)$	66%	50%	-
"	$Pr(\beta_{low} \eta_4)$	36%	50%	-
"	$Pr(\beta_{low} \eta_5)$	12%	50%	-

Table 2: Preference parameters and the consumption floor

The estimated preference parameters and the consumption floor are displayed in Table ??. The consumption floor, which mostly affects the savings of the low-income group, is \$3,593. This is comparable with estimates in other structural models featuring full life-cycle, medical spending uncertainty, and endogenous labor supply.²⁵ The bequest parameters θ_{Beq} and κ_{Beq} , which mostly affect wealth decumulation after retirement, are set to 4,464 and 246,371, respectively. In the one-period consumption-saving model with the risk aversion of 3, these values imply that bequest becomes operational at the asset level of \$15,000 and the marginal propensity to bequeath (MPB) is 0.94. Put differently, individuals with assets below \$15,000 do not leave bequests, while individuals with assets above \$15,000 leave 94 cents out of every dollar for bequests. These numbers are in the range estimated in other studies.²⁶

The discount factors play important role in wealth accumulation before retirement; the estimated β_{low} and β_{high} are 0.904 and 0.995, respectively. The correlation between the discount factor and the health type was estimated to match the wealth-health gradient. The resulting conditional probability $Pr(\beta_{low}|\eta_i)$ is displayed in Table ??. Importantly, we find a strong correlation between the discount factor and health type: almost 90% of individuals in the worst health type (η_1) have low patience, while among the best

²⁵Capatina's (2015) estimate of the consumption floor is \$4,114 (\$ of 2006); Pashchenko and Porapakarm's (2016) estimate is \$1,540 (\$ of 2003).

²⁶For example, De Nardi's et al. (2010) estimates imply bequest threshold of around \$36,000 and the MPB of 0.88. Pashchenko (2013) provides a comparison of the MPBs and bequest thresholds across several structural life-cycle studies.

health type (η_5) only 12% of individuals have low discount factor. The unconditional average of the discount factor in our model $E(\beta)$ is 0.944.

To illustrate the importance of the correlation between the rate of time preference and health type in accounting for wealth-health gradient, we re-estimate two alternative models. In the first one, we restrict $Pr(\beta_j|\eta_m) = 0.50$ for all m and j . In the second one, we substitute our health shock process with a second order Markov process without health type. Note, that the two models still feature all the channels through which bad health can affect individuals' savings. We report the estimated parameters from these two models in the last two columns of Table ?? . Table ?? compares the wealth differences between the healthy and the unhealthy at age 60-69 for the baseline model and for the two alternative models. The latter two fall short of replicating the large wealth-health gap as in the data. This suggests that the casual effect of health on wealth explains only part of the wealth-health gradient, and the third factor, parsimoniously formulated in our model as correlation between health type and patience, is important.

$k_{good} - k_{bad}$	PSID	(1) baseline	(2) no correlation	(3) no health type
25 th pct	To be filled	...	...	...
50 th pct				
75 th pct				

Table 3: The wealth ratio between unhealthy and healthy people at age 60-69

The scaling constant $\bar{\mathbf{b}}$ in Equation (??) is set so that even individuals with the worst possible situation have non-negative utility. The worst outcome an individual can have in our model is to end up receiving the consumption minimum floor, thus $\bar{\mathbf{b}}$ is determined from the following equation:

$$\frac{(\bar{c}/\bar{n}_t)^{1-\rho}}{1-\rho} + \bar{\mathbf{b}} = 0$$

The resulting number is 16.9. This implies the statistical value of life equal to \$6 million, which is in the range estimated in micro literature.²⁷

We also construct several measures of SES-health gradients from our model and compare them with the data. Tables ?? and ?? display the percentage of unhealthy individuals in each earnings and wealth tercile by age group. Note that among those in low terciles of both earnings and wealth, significantly more people are unhealthy at each age. Our model can reproduce these aspects of income-health and wealth-health gradients as

²⁷The measure of the statistical value of life emerged from the literature on compensating differentials for risky occupations. Viscusi (1993) provides an extensive review of these studies and concludes that majority of estimated wage differentials imply the statistical value of life in the range between \$3 and \$7 million.

in the data.

% of unhealthy by earning tercile (Income-health gradient)						
	Data: PSID (HRS)			Model		
	bottom 1/3	middle 1/3	top 1/3	bottom 1/3	middle 1/3	top 1/3
25-34	12%	5%	2%	16%	2%	0%
35-44	21%	8%	4%	22%	4%	2%
45-54	22%	12%	8%	28%	9%	5%
55-64	30% (36%)	15% (20%)	14% (13%)	33%	24%	11%

Table 4: Percentage of unhealthy individuals in each earnings tercile: data vs model

% of unhealthy by wealth tercile (Health-wealth gradient)						
	Data: PSID (HRS)			Model		
	bottom 1/3	middle 1/3	top 1/3	bottom 1/3	middle 1/3	top 1/3
25-34	10%	10%	5%	8%	5%	3%
35-44	17%	10%	5%	14%	7%	5%
45-54	23%	13%	9%	24%	10%	8%
55-64	33% (36%)	17% (21%)	12% (14%)	34%	17%	13%
65-74	36% (38%)	26% (24%)	17% (16%)	41%	27%	19%
75+	46% (41%)	37% (29%)	24% (25%)	47%	38%	29%

Table 5: Percentage of unhealthy individuals in each wealth tercile: data vs model

7 Results

In this section, we use our calibrated life-cycle model to answer three sets of questions. First, we provide several decomposition exercises to better understand what generates the wealth-health gradient. Second, we construct a comprehensive measure of costs of bad health over the life-cycle. Third, we evaluate how much individuals value good health and what aspects of being healthy are most valuable.

7.1 Decomposing the wealth-health gradient

As was shown in the previous section (Figure ??), our model can accurately capture the gap in accumulated wealth between the healthy and the unhealthy. The relationship between wealth and health in our model is generated by two mechanisms. First, unhealthy individuals have lower labor earnings and higher medical spending, thus have less resources to save. In addition, they have lower life expectancy that also lowers their marginal propensity to save. The second mechanism arises because fixed health type is

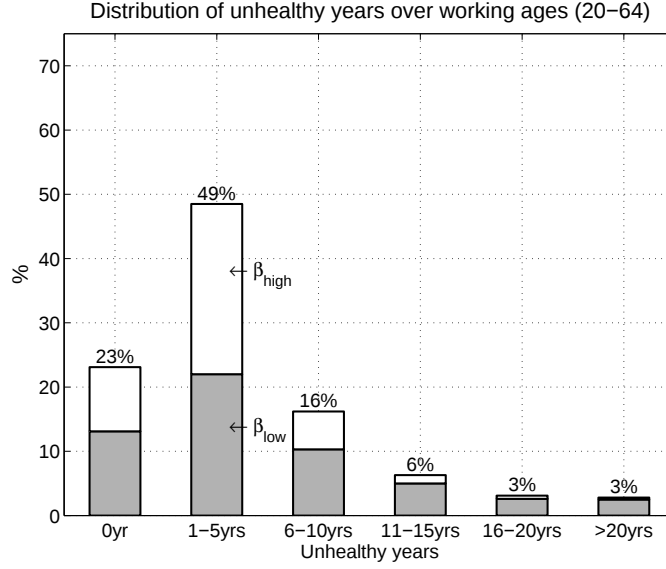


Figure 8: Distribution of individuals by the number of unhealthy years over ages 20 to 64

correlated with the discount factor: individuals in the worst health type who spend longer time being unhealthy also have lower patience. As shown in Table ?? in the previous section, there is a strong correlation between η and β . To provide better illustration of the implication of this correlation, Figure (??) displays the composition of individuals in terms of their discount factor conditional on how many years they spend being unhealthy between the ages of 20 and 64. As can be seen from this figure, among those who spend more than 6 years being unhealthy, most individuals have low discount factor. Moreover, more than 90% of individuals with the history of more than 11 unhealthy years discount the future at the rate β_{low} .

To get a better understanding of how this correlation between the discount factor and the number of experienced unhealthy years affects the wealth-health gradient, we use the dynamic relationship between health and wealth constructed from HRS. Figure (??) below displays (line with crosses) how median wealth at age 55 (left panel) and wealth at age 65 (right panel) depend on the number of unhealthy periods reported by individuals between ages 57 and 65. As in Section ??, we restrict our sample to individuals who are healthy at age 55 and whom we observe until they are 65. Table ?? summarizes the same information using a regression analysis: every additional period an individual reports being unhealthy between ages 57 and 65 corresponds to decline of \$34,473 in his median wealth at age 65 and \$11,749 in the median wealth at age 55. Note that for the first regression equation (with median wealth at age 65 as dependent variable) the value of the estimated coefficient on the number of unhealthy periods (\$34,473) can be to some

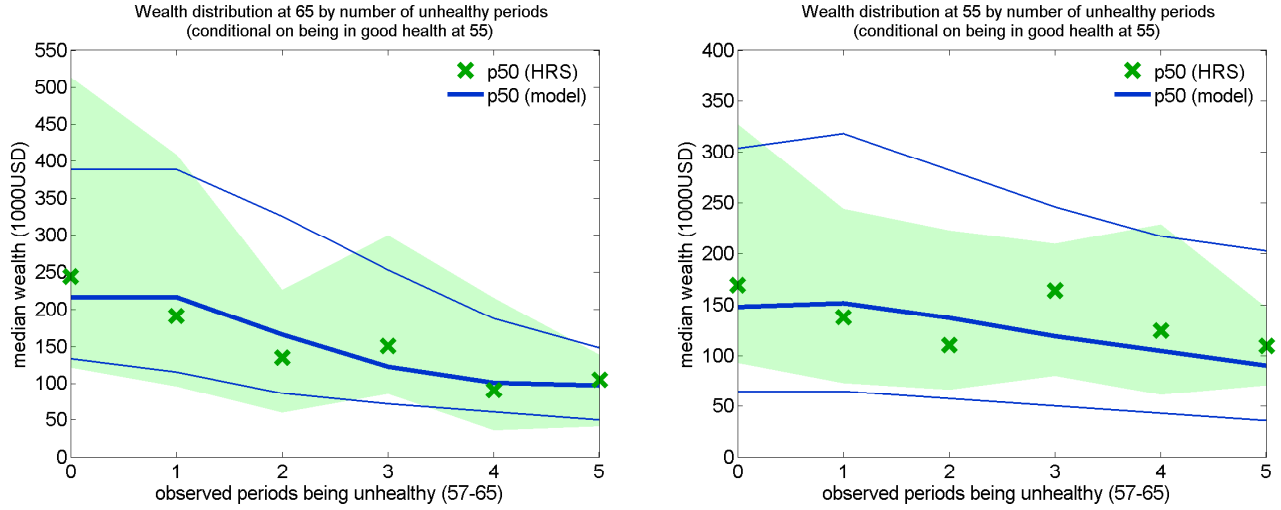


Figure 9: Median wealth at age 65 (left panel) and at age 55 (right panel) depending on the number of unhealthy periods reported between ages 57 and 65 on a bi-annual basis. Line with crosses: HRS. Solid line: model.

extent attributed to the casual accumulated effect of bad health on wealth. The same coefficient in the second regression (with median wealth at age 55 as dependent variable) shows that the wealth *at age 55* is correlated with the number of *future* unhealthy periods suggesting that part of the wealth-health gradient arises from factors not related to the direct casual impact of health on wealth.

We next turn to our model and simulate a similar sample of individuals (healthy at age 55 and alive for the next ten years); the solid lines in Figure (??) show that our model can capture the relationship between the number of unhealthy periods and wealth at age 55 and 65 even though these moments were not targeted. The second column of Table ?? shows that estimating the regression equations of wealth on the number of unhealthy periods on our simulated sample produces very similar coefficients as on the sample from the data.

Median regression: $wealth_{i,age} = const + \alpha_1 \times \text{no of unhealthy-years}_i$			
	HRS	baseline model	baseline model (+ $\alpha_2 D_{\beta_i}$)
α_1 (age = 65)	-34473	-28162	-16849
α_1 (age = 55)	-11749	-10832	-1444

Table 6: The results from regressing median wealth on the number of unhealthy periods: model and data

Given that our model can accurately capture the decline in wealth for each additional unhealthy period, we now turn to our decomposition exercise. On our simulated data, we now run a regression of the median wealth at age 65 and 55 on the number on unhealthy

periods between ages 57 and 65 *conditional on the discount factor of an individual*. The third column of Table ?? reports the resulting coefficients on the number of unhealthy periods. Comparing the second and the third columns of the table shows that once the difference in patience is taken into account, the impact of the length of unhealthy spells drops considerably: from \$28,162 to \$16,849 in case of median wealth at age 65 and from \$10,832 to \$1,444 in case of median wealth at age 55. Thus, around 60% of the correlation between wealth at age 65 and accumulated effect of bad health can be attributed to the casual impact of the latter on the former; the rest comes from the fact that among the unhealthy, a larger fraction has low patience. In case of the correlation between median wealth at age 55 and the number of unhealthy periods between ages 57 and 65, it is almost entirely due to the composition effect: among those healthy at age 55, individuals with low savings are predominantly composed of low-patience and bad-health types who will experience longer unhealthy spells in the future.

To get another angle on the wealth-health gradient decomposition, we next consider a cross-sectional regression of the median wealth on a set of age and age interacted with health dummy variables:

$$k_{age} = \alpha_0 + \sum \alpha_{age} D_{age} + \sum \alpha_{age}^h D_{age} D_{ht} \quad (18)$$

Figure (??) displays coefficients α_{age}^h in this regression: solid line corresponds to the PSID sample, dashed line - to the HRS sample, and shaded bars - to the sample simulated using our model. The model predictions are in line with the data. Next, we use our simulated sample to estimate the modified regression equation that controls for the difference in the discount factor:

$$k_{age} = \alpha_0 + \sum \alpha_{age} D_{age} + \sum \alpha_{age}^h D_{age} D_{ht} + \sum \alpha_{\beta} D_{\beta} \quad (19)$$

The coefficients α_{age}^h from this regression are shown as white bars in Figure (??). Controlling for the fact that relatively large fraction of the unhealthy has low patience and thus lower propensity to save reduces the effect of health on wealth by around 60%. Put differently, around 40% of the wealth-health gradient can be attributed to the casual effect of health on wealth through different channels, while the rest of the wealth-health gap results from the composition effect: the fact that less patient people are inherently prone to be less healthy.

...SUBSTITUTE THIS WITH WEALTH AT 65 ON THE NUMBER OF UNHEALTHY PERIODS BETWEEN 20-64 REGRESSION ...

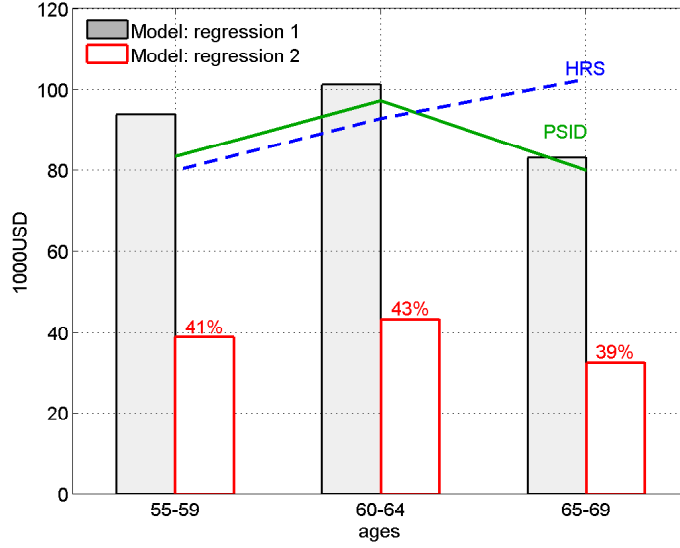


Figure 10: Estimated coefficients from regressing median wealth on dummies of health interacted with age. Solid line: PSID. Dashed line: HRS. Shaded bars: sample simulated from the model. White bars: sample simulated from the model and using regression in Eq ??

7.2 Measuring the costs of bad health

To measure the accumulated costs of bad health, we proceed as follows. We simulate a sample of individuals from our model and divide them into groups based on how many years they have spent being unhealthy between ages 20 and 64. Then we simulate a counterfactual sample of individuals *whose health is always good*. Specifically, we use the decision rules computed from our model but in simulations we always reset realized health to be good. Put differently, we artificially create a sample of exceptionally lucky individuals who expect their health to evolve as in the baseline case but de facto they never become unhealthy. Then we compare their economic outcomes with the sample simulated from our baseline model.

Figure (??) displays the results of this experiment for individuals with low discount factor (β_{low}).²⁸ The dark area in the graph corresponds to the average loss in the annual productivity; the medium-dark area corresponds to the average annual loss in labor earnings, i.e., when labor supply responses to bad health are taken into account; and the light area corresponds to the average annual loss in earnings net of out-of-pocket medical expenses. Several important observations can be made from the figure. First, the costs of bad health are steeply increasing with the number of years individuals spend unhealthy over their working life. For example, an individual who spends between 1 and 5 years being unhealthy loses, on average, less than \$1,000 per year in earnings net of medical

²⁸The graph looks almost identical for individuals with high beta (β_{high}).

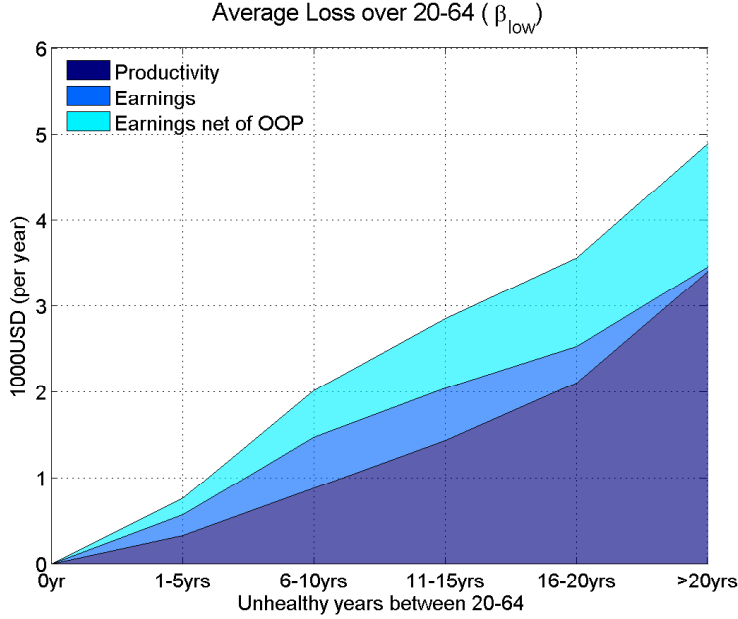


Figure 11: Average annual loss in labor earnings net of medical expenses by the number of unhealthy years between ages 20 and 64.

expenses compared with an individual who is always healthy. In contrast, an individual who spends between 16 and 20 years being unhealthy loses close to \$4,000. Second, the largest component of the costs of bad health is the loss in productivity. In contrast, the increase in out-of-pocket medical spending constitutes less than a third of the total costs of bad health.

Overall, these results suggest that confining the consequences of bad health to high medical spending can significantly underestimate the total losses that unhealthy individuals experience over their lifetime. This also implies that insurance against health deterioration should not be limited to health insurance but also should include more comprehensive measures such as income support.

7.3 Measuring the value of good health

In the previous subsection, we construct a measurement of the monetary losses of bad health. Next, to account for the non-monetary aspect of health, we evaluate the willingness to pay for good health. For each individual in our model, we compute how much he is willing to pay to increase his probability to be healthy next period by one percentage point. Specifically, we construct a counterfactual experiment where the probability to be healthy next period is higher and compare how much an individual in the baseline economy is willing to pay to be moved to this counterfactual economy. Note, we assume

	1 st	2 nd	3 rd	all
β_{low}	\$2,249	\$4,213	\$6,435	\$3,305
β_{high}	\$1,004	\$2,217	\$5,638	\$4,565

Table 7: Dollar value to pay for an increase in probability to be healthy by asset terciles

	η_1	η_2	η_3	η_4	η_5
β_{low}	\$4,473	\$3,573	\$2,552	\$1,891	\$1,589
β_{high}	\$1,0601	\$7,983	\$5,449	\$3,922	\$3,253

Table 8: Dollar value to pay for an increase in probability to be healthy by health type

that this change is temporary: it applies only to probability to move to good health in the next period and after that the situation returns to the status quo. Another way to describe this experiment is that it measures how much people are willing to pay to access medical technology that improves their health dynamics.

Table ?? and ?? display the willingness to pay for individuals with different asset terciles and health type separately by discount factors. On average, individuals with low discount factor are willing to pay around \$3,300, and those with high discount factor - around \$4,500. This difference among people with different patience arises because given the persistence of health process, higher odds of being healthy involve long-term consequences and these are valued differently depending on how the future is discounted. There is also a considerable variation in willingness to pay by health type and assets level. The worst health types (η_1) are willing to pay around three times more than the best health types (η_5) and this is true both for people with low and high discount factor. This is not surprising because individuals with bad health type are less likely to recover from bad health and they are willing to pay more to avoid getting sick. Another important observation from Table ?? is that the rich are willing to pay more than the poor: among those with β_{low} , individuals in the top asset tercile are willing to pay almost three times as much as individuals in the bottom tercile. Among individuals with β_{high} , the ratio of willingness to pay between the top and bottom assets terciles is more than five. This suggests that health is to some extent a luxury good, especially for more patient individuals.

Next, we provide a decomposition exercise to understand for which aspects of good health individuals are willing to pay more. To do this, we consider three counterfactual experiments. In the first experiment, we assume that health does not affect survival probability, i.e., individuals who turn sick do not experience decline in life expectancy. In the second experiment, we assume that health does not affect productivity and disutility from work, i.e., individuals who become sick have the same λ_t^h as the healthy and $\phi_B = 0$.

β_{low} type				
	asset			
	1 st terc	2 nd terc	3 rd terc	all
Baseline	2249	4213	6435	3305
<i>Removed components</i>				
→ survival probability	46%	34%	26%	37%
→ productivity and disutility	73%	76%	86%	76%
→ OOP medical expenses	91%	92%	95%	92%

β_{high} type				
	asset			
	1 st terc	2 nd terc	3 rd terc	all
Baseline	1004	2217	5638	4565
<i>Removed components</i>				
→ survival probability	32%	23%	11%	13%
→ productivity and disutility	84%	93%	99%	98%
→ OOP medical expenses	98%	97%	101%	100%

Table 9: Decomposition the dollar value to pay for an increase in probability to be healthy

In the third experiment, we assume that sick individuals have the same out-of-pocket medical spending as healthy ones. In each of these three experiments, we reevaluate how much individuals are willing to pay to increase the probability to be healthy by one percentage point. Table ?? displays the results of these experiments expressed as a percentage of the willingness to pay in the baseline economy.²⁹

The results show that by far the most valuable aspect of being healthy is longer life expectancy. When bad health does not lower survival probability, individuals with low discount factor decrease their willingness to pay for health improvement by more than 60%, while for individuals with high discount factor the decline constitutes almost 90%. The last result is especially remarkable because it means that for relatively patient individuals, almost 90% of the value of good health is in extending their lifespan. Labor earnings channel contributes from 2% to 24% of the value of good health depending on the discount factor. Finally, medical spending plays relatively small role contributing 8% for individuals with low discount factor and almost nothing for individuals with high discount factor. Put differently, forward-looking individuals mostly care about longevity aspect of good health, while relatively less patient individuals put significant weight on

²⁹Note, that in some instances the resulting number can exceed 100%. This can happen because in the counterfactual experiments, the available resources of individuals can change. For example, if medical spending do not change with health, the average medical spending goes down. Thus, an individual may be willing to pay more to be healthy just because he now has more resources.

such aspects as earnings and medical spending, but even for the latter group longevity plays the dominant role in their evaluation of health. The comparison of results across asset terciles reveals a similar pattern: rich individuals place more weight on longevity and almost none on monetary gains through higher earnings and lower medical spending, while for poorer groups the last two aspects still matter even though less than longevity.

Note, the results of this section provide an important insight into why people want to spend money on health. In our model, we abstract from the possibility to improve health through investments. Recently, several structural papers have attempted to formulate and estimate models of endogenous evolution of health (e.g., Fonseca et al., 2013; Ozkan, 2014). Our results show that in specifying the motives for which individuals are willing to undertake health investments, the improvement in survival has first-order effects. Note also, that through the lenses of our model, one can rationalize the continuous increase in medical spending in the US over the last several decades by people's willingness to pay for increased longevity, which is consistent with the finding of Hall and Jones (2007).

8 Conclusion

In this paper, we investigate the pathways through which health affects individuals over the working stage of their life-cycle. We estimate a model of health dynamics using a rich set of data moments and allowing for both history-dependence and fixed heterogeneity. We find that fixed heterogeneity is important to account for the persistence of bad health. In addition, its correlation with patience plays a key role in accounting for the wealth-health gradient when our health process is embedded in a rich structural life-cycle model.

We find that the costs of bad health quickly accumulate as individuals spend more years being unhealthy. Despite that, accumulated costs of bad health contribute only around 40% to the large gap in wealth between the healthy and the unhealthy. The remaining 60% come from sorting of impatient types into unhealthy group arising from the correlation between fixed heterogeneity in health dynamics and the rate of time preferences.

We find that the largest component of the monetary costs of bad health is the loss in productivity, while the increase in medical spending contributes less than one third. When it comes to non-monetary aspects of health, the most detrimental consequences of being unhealthy is lower longevity. We find that individuals are willing to pay a substantial amount to be able to access technology that makes them healthier, but mostly because they want to live longer. This is especially true for rich individuals who are well-sheltered from monetary losses of bad health, and for individuals with high discount

factor who place higher value on longer lifespan.

We see it as an important future extension to understand to what extent individuals can affect their health evolution through health investments and what role this plays in the evaluation of the consequences of bad health. Our results point out that several directions are particularly interesting. Individuals (especially rich ones) are willing to pay a considerable amount for even relatively small improvement in their health. This means that endogenizing the possibility of health improvements in the model should produce a very strong health investment motive. This raises the question to what extent these investments can be effective and whether this efficiency differs by individuals. Our results show that fixed heterogeneity in health dynamics plays an important role, suggesting that effectiveness to improve one's health can also differ by individuals.

Another important extension of our work is to understand what institutions can provide best insurance against the negative consequences of bad health. Our finding that low labor earnings constitutes the largest part in the monetary costs of bad health suggests that these institutions should not be limited to health insurance but should include some income support programs. One such program, disability insurance, provides support for individuals with permanent deterioration in health. However, an important question is whether other income support programs targeted at people with less serious health problems can substantially improve welfare.

References

- [1] Ameriks, J., Briggs, J., Caplin, A., Shapiro, M., Tonetti, C., 2015. Long-Term Care Utility and Late in Life Saving.
- [2] Attanasio, O., Kitao, S., Violante G., Financing Medicare: A General Equilibrium Analysis. In Shoven, J.B. (Ed.), *Demography and the Economy* (University of Chicago Press, 2011)
- [3] Blundell, R., Britton, J., Dias, M., French, E., 2016. The Dynamic Effects of Health on the Employment of Older Workers. Michigan Retirement Research Center Working Paper 2016-348
- [4] Capatina, E., 2015. Life-cycle Effects of Health Risk. *Journal of Monetary Economics* 74, 67-88
- [5] Capatina, E., Kene, M., Maruyama, S., 2016. Health and Earning Dynamics over the Life-Cycle. Mimeo, University of New South Wales

- [6] Case, A., Lubotsky, D., Paxson, C., 2002. Economic Status and Health in Childhood: The Origins of the Gradient. *American Economic Review*, 92, pp 1308-1664
- [7] Cesarini, D., Lindqvist, E., Ostling, R., Wallace, B., 2016. "Wealth, Health, and Child Development: Evidence from Administrative Data on Swedish Lottery Players". *The Quarterly Journal of Economics*, 687-738
- [8] Cutler, D., Lleras-Muney, A., and Vogl, T., Socioeconomic status and health: dimensions and mechanisms. *Oxford Handbook of Health Economics* (2011)
- [9] De Nardi, M., Wealth Inequality and Intergenerational Links *Review of Economic Studies*, 71(3) (2004), 743-768.
- [10] De Nardi, M., French, E., Jones, J., Why Do the Elderly Save? *Journal of Political Economy*, 118(1) (2010), 39-75.
- [11] Dobkin, C., Finkelstein, A., Kluender, R., Notowidigdo, M., 2016. The Economic Consequences of Hospital Admissions.
- [12] Erosa, A., Fuster, L., Kambourov, G., Towards a micro-founded theory of aggregate labor supply. IMDEA Working Papers N, 2011
- [13] Floden, M., A Note on the Accuracy of Markov-chain Approximations to Highly Persistent AR(1) Processes. *Economic Letters* 99(3) (2008), 516-520.
- [14] French, E., The Effects of Health, Wealth, and Wages on Labor Supply and Retirement Behavior. *Review of Economic Studies*, 72(2) (2005), 395-427.
- [15] French, E., Jones, J., The Effects of Health Insurance and Self-Insurance on Retirement Behavior. *Econometrica*, 79(3) (2011), 693-732.
- [16] Fonseca, R., Michaud, P-C., Kapteyn, A., Galama, T., 2013 Accounting for the Rise in Health Spending and Longevity. IZA Discussion paper 7622
- [17] Hai, R., 2015. The Determinants of Rising Inequality of Health Insurance and Wages. Mimeo, University of Miami
- [18] Hall, R., Jones, C., 2007. The Value of Life and the Rise in Health Spending. *Quarterly Journal of Economics*
- [19] Halliday, T., 2008. Heterogeneity, State Dependence and Health. *Econometrics Journal*, 11(3), pp 499-516

- [20] Hansen, G., Hsu, M., Lee, J., Health Insurance Reform: The impact of a Medicare Buy-In. *Journal of Economic Dynamics and Control*, 45 (2014) 315-329.
- [21] Heathcote, J., Storesletten, K., Violante, G., The Macroeconomic Implications of Rising Wage Inequality in the United States. *Journal of Political Economy*, vol. 118(4) (2010), 681-722.
- [22] Heckman, J., What Has Been Learned About Labor Supply in the Past Twenty Years? *American Economic Review Papers and Proceedings* 83(2) (1993), 116-121.
- [23] Hsu, M., Health Insurance and Precautionary Saving: a Structural Analysis. *Review of Economic Dynamics*, 16(3) (2013), 511-526
- [24] Hubbard, G., Skinner, J., Zeldes, S., The Importance of Precautionary Motives in Explaining Individual and Aggregate Saving. *Carnegie-Rochester Conference Series on Public Policy* 40(1) (1994), 59-125.
- [25] Idler, E., and Benyamini, Y., 1997. Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies. *Journal of Health and Social Behavior*, Vol 38, pp 21-37
- [26] Jeske, K., Kitao, S., U.S. Tax Policy and Health Insurance Demand: Can a Regressive Policy Improve Welfare? *Journal of Monetary Economics*, 56(2) (2009), 210-221.
- [27] Kaiser Commission on Medicaid and the Uninsured, Expanding Medicaid to Low-Income Childless Adults under Health Reform: Key Lessons from State Experiences, 2010. Available at <http://kaiserfamilyfoundation.files.wordpress.com/2013/01/8087.pdf>
- [28] Kaiser Family Foundation, The Cost of Care for the Uninsured. What Do We Spend, Who Pays, and What Would Full Coverage Add to Medical Spending?, 2004. Available at <http://www.kff.org/uninsured/upload/The-Cost-of-Care-for-the-Uninsured-What-Do-We-Spend-Who-Pays-and-What-Would-Full-Coverage-Add-to-Medical-Spending.pdf>
- [29] Kaiser Family Foundation, Employer Health Benefits Survey, 2009. Available at <http://ehbs.kff.org/2009.html>
- [30] Lundborg, P., Nilsson, M., Vikstrom, J., 2015. Heterogeneity in the impact of health shocks on labour outcomes: evidence from Swedish workers. *Oxford Economic Papers*

- [31] Nakajima, M., Telyukova, I., Home Equity in Retirement. *Working paper*, Federal Reserve Bank of Philadelphia, 2012.
- [32] Ozkan, S., 2014. Preventive vs. Curative Medicine: A Macroeconomic Analysis of Health Care over the Life Cycle. *Mimeo*, University of Toronto
- [33] Pashchenko, S., Porapakkarm, P., Quantitative Analysis of Health Insurance Reform: Separating Regulation from Redistribution. *Review of Economic Dynamics* 16 (2013), 383-404.
- [34] Pashchenko, S., Porapakkarm, P., 2015. Welfare costs of reclassification risk in the health insurance market. *Journal of Macroeconomics*, vol. 45(C), pp 21-44
- [35] Pashchenko, S., Porapakkarm, P., 2016. Cross-subsidization in employer-based health insurance and the effects of tax subsidy reform. *National Tax Journal*.
- [36] Pashchenko, S., Porapakkarm, P., 2016. Medical Spending in the U.S.: Facts from the Medical Expenditure Panel Survey Database. *Fiscal Studies*.
- [37] Pashchenko, S., Porapakkarm, P., 2016. Work Incentives of Medicaid Beneficiaries and the Role of Asset Testing. *International Economic Review*, forthcoming
- [38] Pohl, V., Neilson, C., Parro, F., 2016. Health Shocks, Education, and Labor Market Outcomes. *Mimeo*, University of Georgia
- [39] Poterba, J., Venti, S., Wise, D., 2010. "The Asset Cost of Poor Health". NBER Working Paper 16389
- [40] Smith, J., 1999. Healthy Bodies and Thick Wallets: The Dual Relation Between Health and Economic Status. *Journal of Economic Perspectives*, 13(2), pp 144-166
- [41] Storesletten, K., Telmer, C., Yaron, Y., Consumption and Risk Sharing Over the Life Cycle. *Journal of Monetary Economics* 51(3) (2004), 609-633.
- [42] Van Doorslaer, E., and Gerdtham, U., 2002. Does Inequality in Self-Assessed Health Predict Inequality in Survival by Income? - Evidence from Swedish Data. *Social Science and Medicine*, 57, pp 1621-1629
- [43] Viscusi, K., 1993. The Value of Risks to Life and Health. *Journal of Economic Literature*, Vol. XXXI, pp 1912-1946

Appendix

A Summary of the parametrization of the baseline model

Parameter name	Notation	Value	Source
<u>Parameters set outside the model</u>			
Risk aversion	ρ	3.0	
Positive preferences shifter	$\bar{\mathbf{b}}$	16.89	positive utility flow
Consumption scaling	$\bar{n}_t$	-	PSID
Labor supply	$\bar{l}$	0.4	
Tax function parameters	a_0	0.258	Gouveia and Strauss (1994)
	a_1	0.768	"
	a_2	0.616	Pashchenko and Porapakarm (2016)
	τ_y	0.066	"
Medicare premium	P_{MCR}	\$1,055	"
Labor productivity			
- Persistence parameter	ρ_y	0.9275	PSID
- Variance of innovations	σ_ε^2	0.0209	"
- Fixed effects	σ_ξ^2	0.042	"
<u>Parameters used to match some targets</u>			
Discount factors	$\beta_{low}, \beta_{high}$	0.904, 0.995	wealth profiles
Consumption floor	$\bar{c}$	\$3,593	"
Bequest parameters			
- Strength	θ_{Beq}	4,464	"
- Shift	k_{Beq}	246,371	"
Disutility from work			
- Healthy	ϕ_W	XXX	employment profiles (healthy)
- Unhealthy	ϕ_B	XXX	employment profiles (unhealthy)
Proportional load in ESHI/ind ins	γ	1.071	percent individually insured profile
Fixed loads in ind insurance	φ^h		
- healthy		\$101	percent individually insured profile (healthy)
- unhealthy		\$2,100	percent individually insured profile (unhealthy)

Table 10: Parameters of the baseline model

B Alternative specification of the health process

In this section, we consider an alternative specification of the health process: we abstract from fixed heterogeneity in health dynamics and instead assume Markov process of order three. The new health process is formulated as follows:

...Put Alternative SPECIFICATION of Markov(3) Health Shock HERE...

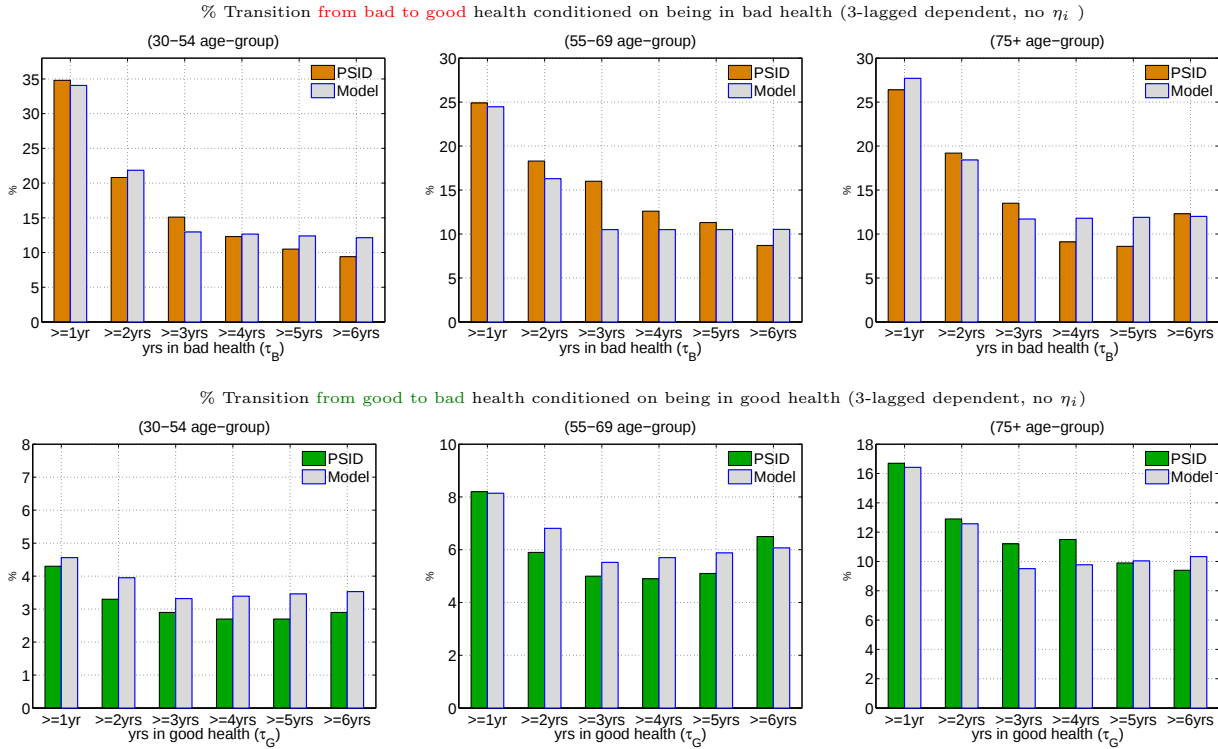


Figure 12: Dynamics of health conditional on duration

Figure (??) compares the duration-dependence transition probabilities simulated by this model with the data. Note that compared with Figure (??) in the main body of the paper, the model performs worse. In particular, it cannot capture continuous decline in the probability to move from bad to good health with duration of bad health status.

C Medical shocks and insurance coverage

To estimate medical expenses, we follow Pashchenko and Porapkkarm (2016). First the medical expenses in MEPS are converted into 1996 price using CPI. Then, we separate our sample into 12 age groups (20-24, 25-29, 30-34, ..., 75+). We assign the age of each

group to the mid-point of the corresponding age interval. For example, 22 for 20-24, 27 for 25-29, 32 for 30-34, etc. For each year, we divide medical expenditures into 3 bins corresponding to the bottom 50th, 50-95th, and top 5th percentiles for each health status and age group. To obtain a value of medical expenses in each bin, we run a regression of medical expenses on a set of age group and year dummies. The coefficients on age dummies in this regression are the average medical expenses for the corresponding age in a particular bin. Then, we fit our estimated coefficients with a quadratic function of age. The resulting numbers are multiplied by 1.60 for people younger than 65 years old and by 1.90 for people who are 65 or older to make medical spending in our model consistent with the aggregate medical spending in NHEA as explained in Section ?? . Figure (??) shows the medical cost for each grid separately for healthy and unhealthy individuals.

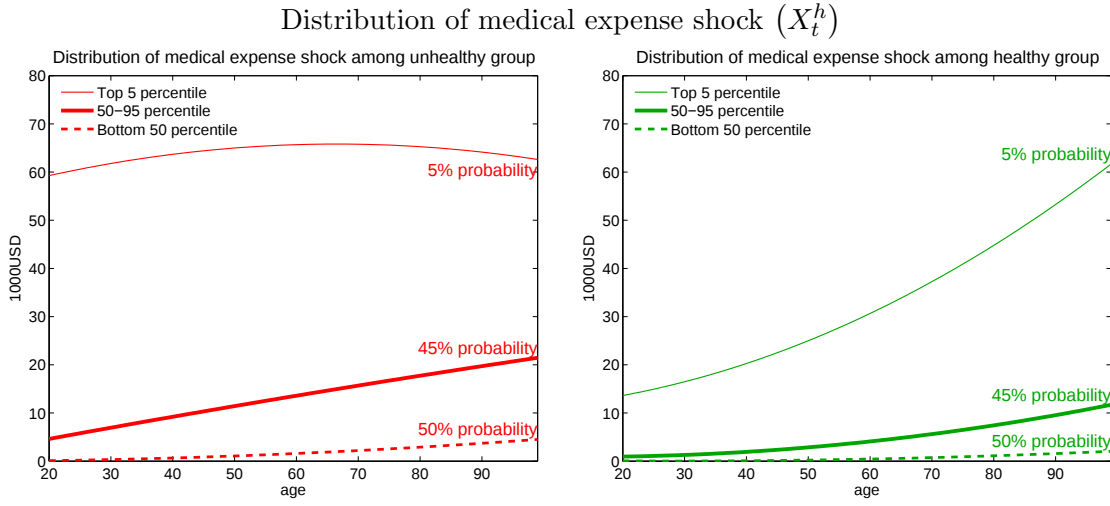


Figure 13: Medical expense grids by health status, x_t^h

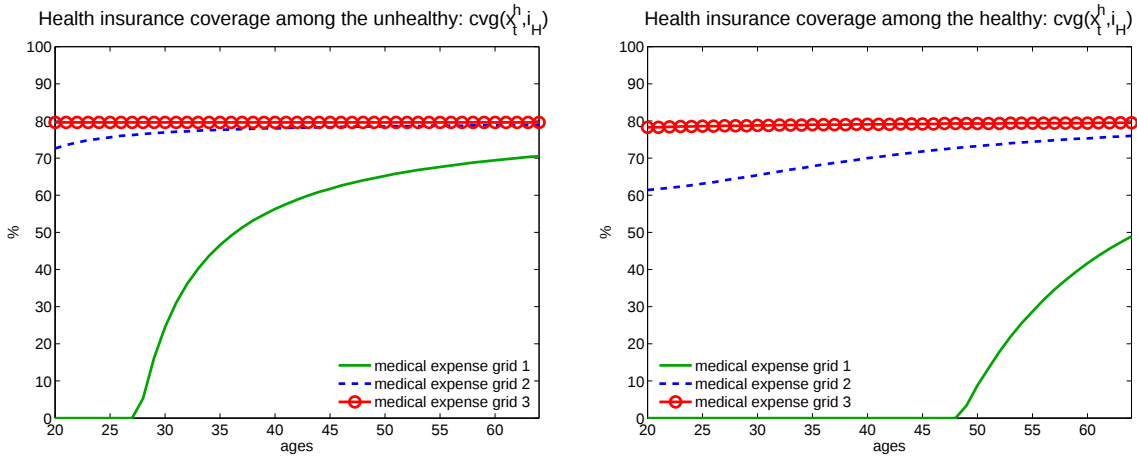


Figure 14: Private health insurance coverages: $cvg(x_t^h, i_H)$

To determine the fraction of medical expenses covered by private insurance $cvg(x_t^h, i_H)$; $i_H \in \{1, 2\}$, we do the following. We estimate medical expenditures paid by private insurers as a function of total medical expenditures and year dummy variables using only individuals who are categorized as individually insured or group-insured. Then, we convert our estimates into the fraction of expenditures covered by insurers. Figure ?? shows the estimated coverages by medical expense grids.