Health, Health Insurance, and Inequality*

Chaoran Chen
York University

Zhigang Feng
University of Nebraska

Jiaying Gu
University of Toronto

March 4, 2020

Abstract

This paper identifies a *"health premium"* of insurance coverage that the insured is more likely to stay healthy or recover from unhealthy status. We develop a quantitative framework to study the interaction between health disparity and income inequality by incorporating such feature into the prototypical macro health model. We estimate the model using MEPS and PSID data. Our baseline economy reproduces the observed joint distribution of income and health status, household health insurance choice, health status, and medical spending by income level and over the life cycle. Quantitative analysis reveals that an individual's early-life income has a significant impact on health in adulthood, which is reinforced by and subsequently amplifies the feedback effect of health on labor earnings and income inequality. We further conduct comparative analysis of health policies intended to alleviate health disparity and income inequality. Providing *"Universal Health Coverage"* would narrow health and life expectancy gaps, with a mixed effect on income distribution. Similarly, implementing *"Universal Basic Income"* reduces income inequality and affects health disparity through its indirect impact on health insurance coverage.

Keywords: Health Insurance, Health Disparity, Income Distribution.
JEL classification: D52, E23, I13, I14.

*We thank Zhifeng Cai, Luca Dedola, Hanming Fang, Kevin Huang, Sagari Kitao, Dirk Krueger, James MacGee, Gajendran Raveendranathan, Jose-Victor Rios-Rull, Manuel Santos, David Wiczer, Kai Zhao, and seminar participants at Jinan University, the 2019 (10th) Shanghai Macro Workshop (SHUFE), UIBE Macro Workshop, University of Nebraska at Lincoln, UIBE Conference on Inequality and Open Economy, and Zhongnan University of Economics and Law for valuable comments and discussion. All errors are our own. Zhigang Feng thanks Natasha Pavlovikj for superb technical assistance, Mammel Family Foundation for financial support. This work was completed utilizing the Holland Computing Center of the University of Nebraska, which receives support from the Nebraska Research Initiative. Contact: Chaoran Chen, Department of Economics, York University, Rm. 1034 Vari Hall, 4700 Keele Street, Toronto, ON M3J 1P3, Canada, chenecon@yorku.ca; Zhigang Feng, Department of Economics, University of Nebraska, Omaha, NE 68106, United States, z.feng2@gmail.com; Jiaying Gu: Department of Economics, University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7, Canada, jiaying.gu@utoronto.ca.

1 Introduction

A large empirical literature documents a strong correlation between individual health and labor earnings, income, and wealth (Chetty et al., 2016). A deterioration in health status has been found to lead to significant reduction in hourly wage, labor force participation, and hours worked.[1] This empirical regularity inspired a recent literature that analyzes the impact of health on income and wealth inequality (de Nardi et al., 2017; Hosseini et al., 2019). It is noteworthy to mention that health is an endogenous variable (Grossman, 1972) while income has been identified as one of the leading factors influencing individual health (Deaton, 2003). Omitting this feedback effect prevents us from fully understanding the interaction between health and income, and hinders a complete assessment of policies designed to reduce health disparity and/or income inequality.

In this paper, we study a novel channel through which income affects the evolution of health: endogenous health insurance choice. We identify a "health premium" of insurance coverage that the insured is more likely to stay healthy or recover from unhealthy status. Our identification addresses the endogeneity of health insurance choice and the potential reverse causality through an instrumental variable approach. Our focus on this particular health insurance aspect is motivated by the following facts. First, income affects health insurance choice, since low-income workers face tighter affordability constraints and are also less likely to receive employer-sponsored insurance (EHI), which provides favorable tax treatment and better risk sharing through workplace. In addition, insurance coverage affects individual health as the uninsured are less likely to receive adequate preventive care and more likely to skip medical follow up (Doyle, 2005; Ward et al., 2008; Levy and Meltzer, 2004).[2] Income hence affects health through the health insurance choices. Second, in practice, government's health care policies are generally designed to target health insurance coverage, as health is imperfectly observable and subject to nondiscrimination rule. Prominent examples include the 2010 Patient Protection and Affordable Care Act and "Medicare for All" proposed by 2020 Democratic presidential hopefuls. Studying this health insurance aspect hence provides rich policy implications.

To explore this channel, we develop a dynamic equilibrium model incorporating the "health premium" and estimate the baseline economy with MEPS and PSID data. Quantitative analysis

[1]See, for instance, Mattke et al. (2007), Attanasio et al. (2011), and de Nardi et al. (2017), among others.

[2]There are other factors contributing to the evolution of health, such as individual's effort to maintain health and access to medical care, among others. Effort and the price for medical care are, however, not perfectly observable, and thereby not amenable to parameterization.

reveals that an individual's early life income has a significant impact on adulthood health, which will be reinforced by the feedback effect of health on labor earnings. Subsequently, the impact of income on health amplifies the effect of health on income inequality. Our model allows us to conduct comparative analysis to investigate the implied health-income interactions on health policies. For example, while direct intervention health policies such as "universal health coverage" improve overall health and reduce health disparity, such policy has countervailing effects on income inequality. This is because while baseline uninsured individuals enjoy higher labor productivity and earnings as their health improves with better insurance coverage, their lower mortality rate also raises the weight of the poor in the aggregate income distribution and offsets the positive effect of improved insurance on income inequality. In contrast, a policy providing "Universal Basic Income" would improve health outcomes indirectly as it renders health insurance more affordable for low-income workers. Its net effect on health disparity would, however, be hindered by the existence of means-tested public health insurance programs such as Medicaid, which may disqualify a subset of low-income workers after receiving universal basic income.

We begin by documenting stylized facts on the interdependence between health and income. We gather data on individual income, health status, medical spending, and health insurance coverage from the Medical Expenditure Panel Survey (MEPS) and the Panel Study of Income Dynamics (PSID). In our dataset, unhealthy individuals report substantially lower income, and face higher medical expenditure compared with healthy individuals across all ages. Better affordability and the presence of a strong correlation between wage and EHI offer rates imply that higher income workers are more likely to obtain insurance coverage. Compared with the uninsured, they are more likely to maintain better health status, after accounting for the endogeity of health insurance choice. Consequently, income may affect individual health status through health insurance choice.

Guided by these facts, we develop a life cycle model to account for the joint distribution of health and income. Our framework departs from standard heterogeneous agent models with incomplete markets, idiosyncratic health expenditure shocks and income risk. We extend these models in the following three dimensions. First, we consider the impact of health on productivity and labor earnings. Second, individual income affects health and mortality risk through endogenous insurance choice. Third, we enrich the modeling of the health insurance market by considering implicit health insurance, including consumption floor and medical bankruptcy. The latter is crucial for understanding individual health insurance choice, especially for low income households. Together with the first two factors, they are essential for reproducing the joint distribution of health and income. As we combine all these ingredients in a dynamic general equilibrium model, we shall impose some simplifying assumptions. We model the impact of income on health as a persistent health transition process depending on insurance choice, which is a function of individual income and other state variables in equilibrium. This dependence approximates the effect of preventive care and access to medical care, which have significant impact on the evolution of individual health (Ozkan, 2014; Hong et al., 2017).

Key parameters are estimated with econometric techniques using micro data from PSID and MEPS. We identify how income affects health through health insurance choice, and how health affects income separately accounting for their simultaneity. To address the endogeneity of health insurance choice, we use an instrumental variable approach and find a significant "health premium"—individuals with health insurance coverage are more likely to stay healthy or recover from an unhealthy status. We also find that healthy individuals have labor income that is on average more than 60 percent higher compared to otherwise identical unhealthy individuals, after addressing the simultaneity bias arising from the interaction between health and income. We estimate other key components, such as medical expenditure shocks, co-insurance rates, and the EHI offer rate, non-parametrically from our data.

We feed these processes, which are estimated independently of the equilibrium, into our dynamic model. The baseline economy reproduces the observed health insurance choice as well as the joint distribution of income and health over the life cycle, despite the fact that they are not explicitly targeted in the calibration. In our model, higher income households are significantly more likely to obtain insurance coverage and stay healthy, with elasticity similar to that of the data. Moreover, we match the insurance coverage rate, fraction of healthy individuals, and aggregate mortality rate over the life cycle.

Our baseline model provides a quantitative laboratory to understand the individual's health insurance choice, especially those that choose to be uninsured. We explicitly model implicit forms of health insurance, including consumption floor and medical bankruptcy, as well as choices among various types of medical insurance, such as EHI, private insurance, Medicaid, and Medicare. Understanding the mechanism that drives these choices has important implications as most health polices focus on insurance coverage. For example, we find that 2.4 percent more individuals would purchase for health insurance in absence of insurance markup, and 14.1 percent would gain coverage if the option of opting in an EHI were made available to all workers. We also find that given the interdependence between health and income, some uninsured rely on social safety nets as an implicit insurance as it not only provides a minimum consumption level to insure against income shocks, but also allows the individual to cover medical expenses through subsidies.

A novel feature of the model is that it allows individuals to endogenize their evolution of health through investment in health insurance. This allows us to simulate our model to identify the relevant factors that determine health and economic outcomes at later stages of the life cycle and influence the joint distribution of income and health. We find that an individual's initial income has a significant and persistent impact on health status over the entire life cycle. Doubling a worker's income temporarily at age 25 would increase the chance of remaining healthy at age 40 from 75 percent to 80 percent. The individual's income influences the evolution of health mainly through the channel of health insurance. A worker who has health insurance coverage initially is more likely to remain healthy by one percent at age 40, compared to an otherwise identical worker without insurance initially. Similarly, a worker who has health insurance initially would be six percent less likely to fall into the bottom quantile of the income distribution at age 40. Such effects would reinforce the impact of health on earning ability, and further widen the income gap. Numerical analysis indicates that an individual entering the economy in good health enjoys a five percent higher income stream relative to an unhealthy counterpart, along with a lower chance of remaining in the bottom quantile of the income distribution at age 40.

In light of the above findings, we conduct comparative analysis of health policies intended to alleviate health disparity and income inequality. We first consider "Universal Health Coverage" proposed by Democratic presidential hopefuls, where government revenues would be used to fund the health coverage of all individuals. Given health-income interaction, this policy would improve health status through the impact of health insurance on health transition, especially for low-income individuals who would otherwise not be able to afford health insurance. This policy would therefore help to narrow health and life expectancy gaps—effects that are absent in the canonical models without endogenous health through insurance choice. Such policy however, would have countervailing effects on the income distribution. While better health improves productivity and hence labor earnings especially for poor workers, thereby reducing income inequality, poor workers living longer as a result of better insurance coverage, in turn increase the weight of low-income individuals in the income distribution, consequently enlarging income inequality. The overall net effect on income inequality is therefore mixed.

Inspired by the significant and persistent impact of income on health, we also consider the policy of providing "Universal Basic Income". In comparison to "Universal Health Coverage", this policy significantly impacts income inequality, while it has complicated effect on health disparity. This is because, while additional income through transfers makes private health insurance more affordable, especially to low-income workers, it is important to keep in mind that a subset may no longer qualify for the means-tested public health benefit, which would counteract the benefit of the basic income transfers on insurance coverage for this subset.

Our study makes several contributions. First, we document a set of new facts indicating a strong correlation between household income and health status through endogenous health insurance choice. We identify a "health premium" of insurance coverage where the insured is more likely to remain healthy or recover from ill health. Second, we develop a quantitative model specifically designed to account for the joint distribution of health and income. To the best of our knowledge, our paper provides a first attempt at reproducing such a distribution endogenously. Third, we simulate the model to identify the defining factors that affect health and economic outcome in adulthood. We find that initial income has a long-lasting and significant impact on health over the life cycle. This effect reinforces the influence of health on labor productivity, earning ability, and income inequality. Lastly, we conduct comparative analysis of reform options to the health care system by comparing direct intervention through improved health insurance coverage and indirect intervention through income redistribution. Our findings shed lights on the design of optimal health policy.

1.1 Related literature

Much of the economic literature on health starts with Grossman (1972). This paper considers health as an investment good whose evolution can be actively managed by investing effort or other resources. A recent quantitative literature embeds this idea into dynamic models with heterogeneous agents and incomplete market to study the macroeconomic and distributional implications of health, health insurance, and health care policies. Within this literature, the most closely related papers are Feng (2010), Jung and Tran (2016), Hong et al. (2017), Prados (2017), and Cole et al. (2019). We depart from these papers by specifying a particular channel in which health insurance and preventive care can influence the evolution of health.[3] It is generally challenging to estimate the health production function due to a lack of concrete health metrics

[3] Ozkan (2014) highlights the impact of preventive care on health status. While he focuses more on the life-cycle behavior of rich and poor individuals investing in preventive care, we study the determination of the joint distribution of health and income, and quantify the reasons that individuals choose to be uninsured.

and difficulty in estimating the price of medical expenditure.[4] Our model focus on evaluting the impact of health insurance choice on the evolution of a binary health status, and is thereby amenable to parameterization and estimation. Nevertheless our results should be interpreted as a conservative lower bound of the impact of income on health, as we omit other factors that affects the evolution of health and are correlated with income, such as lifestyle and nutrition affordability.

Our paper complements a strand of literature that studies health and inequality. The paper by de Nardi et al. (2017), studies the pecuniary and non-pecuniary costs of exogenous bad health shocks, and finds that exogenous health heterogeneity is important in accounting for lifetime inequality. Hosseini et al. (2019) studies the impact of health inequality on lifetime earnings inequality. In these papers, individual's health follows some exogenous path subject to uncertain shock. They focus on the impact of health on health expenditure and labor productivity. Our study adds to this literature by allowing for an endogenous health process: Individuals can invest in their health by purchasing health insurance. In our framework, both health and income distributions are endogenous and jointly determined in the equilibrium. This allows us to study the relative importance of some economic variables in determining various adulthood health and economic outcomes.

Our key mechanism that health insurance causally affects health echoes findings in a large empirical health literature. For instance, Doyle (2005) explores quasi-experiments arising from car accidents and finds that medically uninsured receive 20 percent less health care and have a substantially higher mortality rate when facing unexpected health shocks than those insured. Card et al. (2009) adopts a regression discontinuity design focusing on Medicare eligibility around age 65, and find that Medicare eligible patients receive significantly more medical services and 20 percent reduction in mortality for the severely ill patient group. Similar results are also found in other studies, see for instance Currie and Gruber (1996) and Finkelstein et al. (2012), among others. Our empirical results that health insurance causally affects health confirms findings in this literature. We differ from them by focusing on its aggregate implications on health and income inequality.[5]

Our paper also contributes to a broad literature that studies macroeconomic aspects of health policies, including French and Jones (2004), Hall and Jones (2007), Jeske and Kitao

[4] See, for instance, Suen (2006), among others.

[5] Our instrumental variable approach also allows us to identify the effect for individuals of all ages, as we explained in Section 4.1.2, while their approaches typically apply to a small and special group of individuals, such as those at age of 65 in Card et al. (2009).

(2009), Bruegemann and Manovskii (2010), Pashchenko and Porapakkarm (2013), Hansen et al. (2014), Braun et al. (2017), and others. The main innovation of our paper is that the baseline model allow individuals to endogenize the evolution of individual health status through investment in health insurance. Our paper is also related to the empirical micro literature studying why so many Americans are uninsured.[6] Guided by these works, we build our macro model to include most empirically relevant channels. Taking into account the general equilibrium impacts allows us to understand the importance of these channels at the aggregate level.

The reminder of the paper proceeds as follows. Section 2 shows motivating evidence. Section 3 describes the model. Section 4 explains how we estimate the model, and Section 5 performs quantitative analysis. Section 6 concludes the paper.

2 Empirical Facts

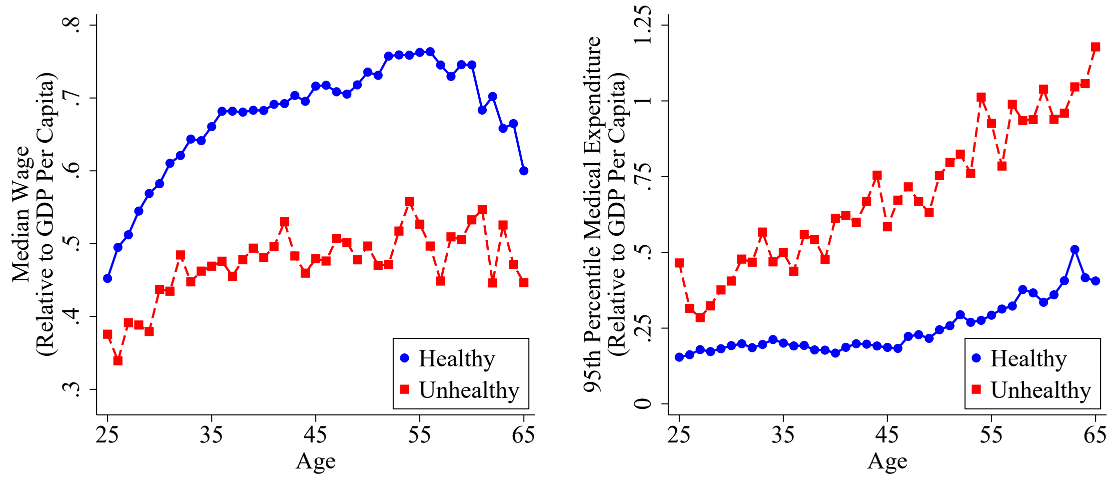
We gather income data from the *Panel Study of Income Dynamics* (PSID) and medical expenditure, health insurance coverage, and health status data from the *Medical Expenditure Panel Survey* (MEPS). The MEPS links people in one household based on coverage eligibility under a typical family insurance plan. This Health Insurance Eligibility Unit (HIEU) defined in the MEPS data set corresponds to our definition of a household. In our sample, we include only the household heads, which we define as the highest income earner in the HIEU.

In the MEPS data, health status is recorded as "excellent", "very good", "good", "fair", and "poor". Following the tradition of the literature, we classify an individual as "healthy" if her health status is "excellent", "very good", or "good", and "unhealthy" otherwise. There are five possible insurance status in the data: no insurance coverage, or covered by a private insurance plan, by EHI, by Medicaid, or by Medicare. For our empirical analysis below, we categorize the population into insured versus uninsured for easier visualization. Note that we explicitly model all the above insurance choices in the quantitative framework. Appendix A includes a detailed description of the data source.

Income and medical expenditure by health Figure 1 reports median income and the 95th percentile medical expenditure by individual's health status over the life cycle. The left panel indicates that unhealthy individuals have substantially lower income compared to healthy

[6] See, for instance, Cutler and Reber (1998), Card and Shore-Sheppard (2004), Herring (2005), Gruber (2008), and Mahoney (2015), among others, for the literature investigating why a substantial fraction of Americans are uninsured.

Figure 1: Income and medical expenditure by health



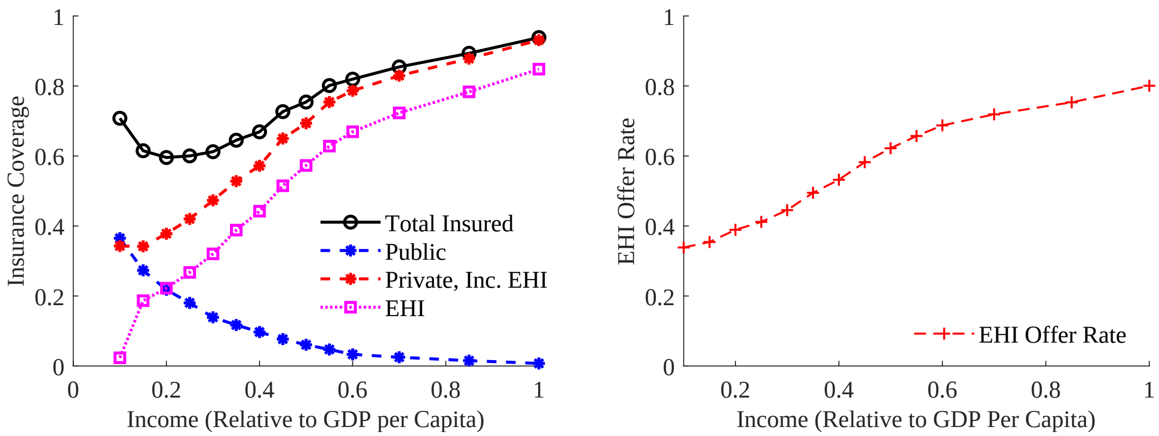
Note: The left panel shows the median annual income for healthy and unhealthy individuals for different ages, where annual income is normalized relative to GDP per capita. The right panel shows the 95th percentile of medical expenditure for healthy and unhealthy individuals for different ages, also normalized relative to GDP per capita. The sample is from the Medical Expenditure Panel Survey (MEPS).

individuals for all ages, consistent with findings in Aizawa and Fang (2019). The right panel shows that unhealthy individuals also have higher medical expenditure, especially among the elderly. The above two facts imply that unhealthy individuals tend to be poorer, which echos findings in de Nardi et al. (2017) that health disparity matters for income inequality.

Insurance coverage by income Data from the *National Health Interview Survey* show that around 15 percent of nonelderly Americans were uninsured during the period from 1970 to 2010.[7] The percentage of uninsured remains largely unchanged over time, until the substantial expansion of public coverage in recent years. The left panel of Figure 2 shows that the proportion of insured individuals between the age of 25 and 65 increases with income level, except for very low income individuals. Among the uninsured, public insurance coverage declines with income while private insurance coverage (including EHI) increases with income. Furthermore, the right panel of Figure 2 documents that the EHI offer rate increases with income. For individuals earning an annual income of 20 percent of the GDP per capita, less than 40 percent receive an EHI offer from employers, compared to over 80 percent for individuals whose annual income is around or above the GDP per capita.

[7]The data (accessed July 12, 2017) can be found here.

Figure 2: Insurance Coverage by Income



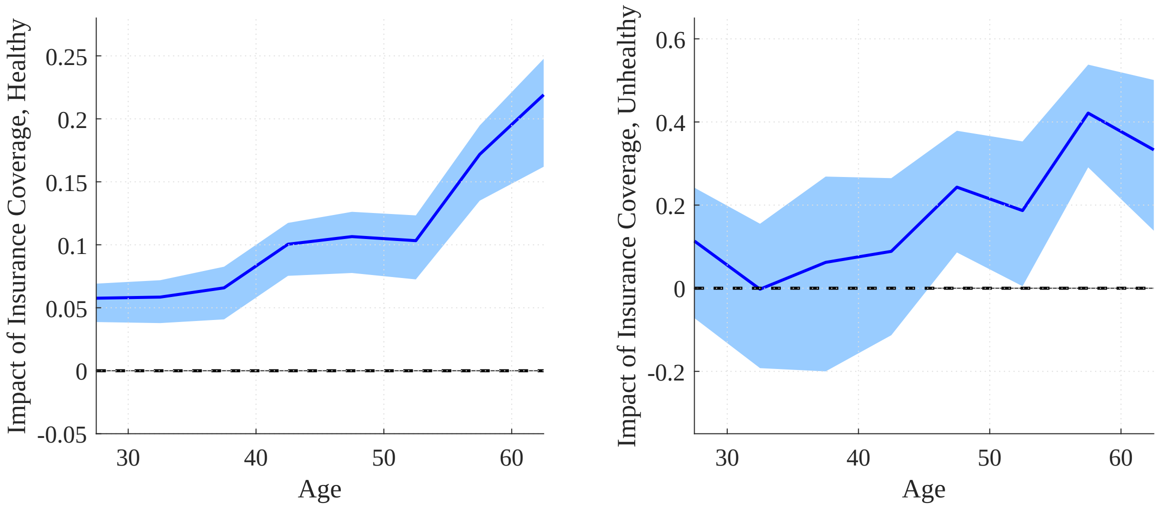
Note: The left panel shows the insurance status of individuals by income levels. Public insurance includes Medicaid and Medicare. The right panel plots the average EHI offer rate by income levels. Annual income is normalized relative to GDP per capita.

Health by insurance coverage. Assume health status follows a first order Markov process with transition probabilities governed by $\pi_{Hh}^{i_{hi}} = \text{Prob}[h' = H|h]$, where $h \in \{Healthy, Unhealthy\}$ and $i_{hi} \in \{Insured, Uninsured\}$. We estimate these probabilities using an instrumental variable approach to address the endogeneity of health insurance status, while estimation details are relegated to Section 4.1.2. We find that health insurance coverage carries a significant "health premium" for the working age population. Conditional on being healthy, an insured 40 year old is eight percent more likely to remain healthy next year compared to an uninsured counterpart. There is a similar difference in the probability of transiting from unhealthy to healthy between insured and uninsured. Our estimation resonates with empirical findings that the uninsured receive less medically necessary treatment and have inadequate access to preventive care, which are essential for the maintenance of good health and for recovery from a bad health shock.[8]

To sum up, we find rich inter-dependence between income and health. Health disparity affects income and wealth inequality, as unhealthy individuals have both lower income and higher medical expenditure. Income inequality also affects health disparity through the insurance market: High-income individuals are more likely to be insured, while insurance coverage positively affects health status evolution. Guided by these facts, we build a model that features health-income interactions where health affects income and income affects health through

[8]See, for instance, Doyle (2005), Ward et al. (2008), Card et al. (2009), and Finkelstein et al. (2012), among others.

Figure 3: Health Premium of Insurance



Note: This figure plots the "health premium", defined as the advantage of the insured over the uninsured in terms of the probability of transiting from healthy to healthy (top panel) or unhealthy to healthy (bottom panel). The solid blue line indicates the point estimates while the shaded light blue area indicates 95 percent confidence intervals.

individual insurance choice.

3 Model

We extend the framework of Aiyagari (1994) with overlapping generations of households who live to a maximum of J periods. Households are endowed with one unit of time in each period that is supplied inelastically to the labor market, work until retirement age J_R, and maximize discounted lifetime utility from consumption. They face idiosyncratic labor productivity shocks z and medical expense shocks m in addition to health shocks in each period over the life cycle. There is no aggregate uncertainty. The financial market is incomplete with a risk-free bond in the economy. Households can purchase health insurance to hedge against health expenditure shocks. In what follows, x' denotes the value of x in the next period.

3.1 Demographics, preference, and endowment

Preference Preferences are represented by

$$\mathbb{E}\sum_{j=1}^{J}\left[\beta^{j-1}\prod_{t=0}^{j-1}\rho_{h,t}u(c_j)\right], \tag{1}$$

where β is the time-invariant discount factor, $\rho_{h,t}$ is the specific age and health survival probability, and c_j is consumption at age j. We assume that $u(\cdot)$ is strictly increasing and concave.

Endowment Labor income e_j^i of household i at age $j \leqslant J_R$ depends on household productivity and wage rate $\tilde{w}$:

$$e_j^i = \tilde{w}\hat{z}_j^i, \tag{2}$$

where $\hat{z}_j^i$ is the product of realized labor productivity z_j^i and health shock impacts $g_j(h)$, which capture individual health impacts $h \in \{H, U\}$ on productivity, where H and U denote healthy and unhealthy, respectively. In particular, we normalize productivity by that of the healthy worker such that the unhealthy worker has a productivity loss of d_j relative to a healthy counterpart. Hence, we have

$$g_j(h) = \begin{cases} 1 & \text{if } h = H; \\ 1 - d_j & \text{if } h = U. \end{cases} \tag{3}$$

Health In each period, agent's health status evolves according to a Markov process, whose transition matrix varies with endogenous health insurance coverage:

$$\pi^{j,i_{hi}} = \begin{bmatrix} \pi_{HH}^{j,i_{hi}} & \pi_{HU}^{j,i_{hi}} \\ \pi_{UH}^{j,i_{hi}} & \pi_{UU}^{j,i_{hi}} \end{bmatrix}. \tag{4}$$

Here, $\pi_{hh'}^{j,i_{hi}}$ denotes the probability that an age-j agent's health status changes from $h \in \{H, U\}$ to $h' \in \{H, U\}$ conditional on health insurance status $i_{hi} \in \{0, 1\}$. This insurance-dependent health transition process reflects the fact that individuals with insurance coverage have access to adequate preventive care and face a lower probability of experiencing severe health shocks (Ozkan, 2014). Uninsured agents may receive insufficient medical service that undermines their health (Doyle, 2005).

Given health status, agents receive an idiosyncratic health expenditure shock m, whose distribution varies with the agent's age, health status, and insurance coverage. To insure against this risk, agents can purchase health insurance $i_{hi} \in \{0, 1\}$, where $i_{hi} = 1$ indicates that the agent is insured.[9] Health status also affects her survival probability in that, at age j, an agent's survival probability is $\rho_{h,j}$ where $\rho_{H,j} \geqslant \rho_{U,j}$.

3.2 Market arrangement

3.2.1 Financial and insurance market

There is a state non-contingent risk-free bond, the quantity of which is denoted as a. Households save by purchasing a' units of this bond at the common market rate of r. Borrowing is limited. Firms rent capital at a competitive rate.

Agents have access to the employer sponsored health insurance (EHI) with probability $p_E(e)$, which depends on the agent's income.[10] Once an offer is made to the employee, denoted as $i_E = 1$, he/she decides whether or not to obtain the coverage ($i_{hi} = \{0, 1\}$). The EHI costs a premium of π_E and covers a fraction $\phi_E(m)$ of the realized medical expenditures m. The premium is not dependent on prior health history or any individual states. This accounts for the practice that group health insurance does not price-discriminate the insured by individual characteristics. The employer subsidizes a fraction $\psi \in [0, 1]$ of the insurance premium.

Without an EHI offer from the employer, the worker has the option to purchase health insurance in the private market at premium $\pi_P(m)$ with co-insurance rate $\phi_P(m)$. Health insurance companies are competitive. The premiums for EHI and private plans are determined as the expected expenditures for each contract plus a proportional markup η.

Medicaid, a means-tested public health insurance program, covers working-age agents who earn income below certain thresholds. We assume that the agent is automatically enrolled in the Medicaid program if $e \leqslant \Theta_\mathrm{e}$ and $a \leqslant \Theta_\mathrm{a}$ and receive coverage for a fraction $\phi_{md}(m)$ of medical expenditure. In addition, the Medicare program covers all retirees for whom there is a fixed premium π_{mr} and coverage of a fraction $\phi_{mr}(m)$ of the medical expenditures.

[9]Note that agents have several options for obtaining health insurance coverage depending on income, assets, jobs and age. See the discussion in Section 3.2.1.

[10]The MEPS data show that about 97% of firms with over 100 employees offer health insurance, about 80% of firms with 25-99 employees offer insurance, and only 40% of firms with less than 25 employees offer coverage (Aizawa and Fang, 2019). Instead of modelling the firm's EHI offer decision and the search and matching process between workers and firms, we assume a shock that determines the EHI offer status for employees. Note that we allow for probability p_E to vary with income as suggested by the empirical patterns in the benchmark estimation of the model.

3.2.2 Medical bankruptcy

We interpret a health expenditure shock m as medical bills from the hospital. This accounts for the fact that federal law prohibits hospitals from denying emergency treatment to uninsured patients. In the course of bad health shock, the households can always go to the emergency room. When this shock m hits, the household is responsible for paying $[1 - i_{hi}\phi(m)]m$, where i_{hi} represents the household's insurance coverage and $i_{hi} = 0$ for the uninsured.

Households have the option to declare bankruptcy over their medical spending. Denote the status of medical bankruptcy as ι. During the bankruptcy period ($\iota = 1$), agents are subject to a linear garnishment of earnings $\lambda = \gamma \max\{y - \bar{y}, 0\}$ following Livshits et al. (2007). Here, λ denotes the total amount garnished and transferred to the hospital (and eventually to the government), $\bar{y}$ is an earnings exemption that cannot be seized and $\gamma \in [0, 1]$ is the marginal rate of garnishment. The garnishment technology is costless. After declaring bankruptcy, the agent's medical bill $[1 - i_{hi}\phi(m)]m$ is wiped out and they are temporary excluded from credit markets ($a' = 0$). The cost of uncompensated care, $[1 - i_{hi}\phi(m)]m - \lambda$ is covered by the government.

3.3 Production

In each time period, a representative firm produces final good y using capital k and labor n using a neoclassical production function: $y = Ak^{\alpha}n^{1-\alpha}$. The firm maximizes one-period profit by setting its marginal products equal to the cost of input factors:

$$r + \delta = AF_k(k, n), \quad w = AF_l(k, n). \tag{5}$$

If offering EHI, the firm adjusts the wage to ensure the zero-profit condition holds by subtracting the cost c_E of providing EHI from the wage rate. Therefore, an agent's wage is given by

$$\tilde{w} = \begin{cases} w, & \text{if } i_E = 0, \\ w - c_E, & \text{if } i_E = 1. \end{cases} \tag{6}$$

The production of the final good can be used for private consumption c and investment i. The law of motion for capital k is given by $k' = (1 - \delta)k + i$, where δ is the depreciation rate of capital.

3.4 Government programs

The government runs four programs: a means-tested welfare program, Medicaid for low income workers, Medicare for retirees, and a pay-as-you-go Social Security program.

The welfare program provides transfers to households whose after-tax disposable assets fall below $\underline{c}$, as in Hubbard et al. (1995). The transfer amount is given by the difference between $\underline{c}$ and the the total after-tax disposable income the agent has available for consumption. The welfare program parsimoniously captures unemployment insurance and food stamps. The retired individuals receive social security benefit $ss(\bar{e})$, which depends on their average lifetime earnings $\bar{e}$. Medicaid provides public health insurance to agents whose income and assets are below certain thresholds. Medicare provides health insurance for all retirees.

Workers pay proportional social security tax τ_{ss} and Medicare tax τ_{mr}. The social security tax rate for earning above y^{ss}_{max} is zero. The government also levies a progressive income tax $T(\cdot)$ and a proportional consumption tax τ_c to finance its expenditures G and the above programs. The government balances its budget every period.

3.5 Decision problems

The timing of the economy is given as follows: (1) Households enter a new period with assets a and health insurance status i_{hi}; (2) Idiosyncratic shocks z, m, h, and i_E are drawn for survivors and newborns; (3) Each household makes a decision on health insurance i'_{hi}, medical bankruptcy ι, labor supply n, consumption c, and borrowing/saving a'; (4) Firm production takes place and all markets clear.

3.5.1 Households

The state of households can be summarized by vector $s_w = \{j, a, z, \bar{e}, m, h, i_{hi}, i_E\}$ for workers and $s_r = \{j, a, \bar{e}, m, h\}$ for retirees.[11] Here, j denotes age, a denotes assets, z and m are idiosyncratic shocks to labor productivity and health expenditure, respectively, $\bar{e}$ is the household's average lifetime earnings, h is the health status, and i_{hi} and i_E are indicator functions for health insurance coverage and availability of an EHI offer, respectively.

Let $\varphi(s)$ be the population density function of individuals at the beginning of each period

[11] Following Jeske and Kitao (2009), we distinguish newly retired agents from the rest of the retired agents as new retiree health bills are covered by insurance and not by Medicare if $i_{hi} = 1$. Hence, this age group's state variable is given by $\{j, a, \bar{e}, m, h, i_{hi}\}$.

and $S = \{\varphi(s), \pi, \phi, r, w, \tau\}$ the aggregate variables. The household's problem is given as follows.

The young agent's problem ($j < J_R$)

$$\mathbf{V}^j(s_w) = \max_{\{c,a',i'_{hi},\iota\}} \left\{ \; u(c,n) + \beta\rho_{h,j}\mathbb{E}\mathbf{V}^{j+1}(s'_w) \; \right\} \tag{7}$$

subject to

$$(1+\tau_c)c + a' + i'_{hi}\tilde{\pi}^j + (1 - i_{hi}\phi(m))m = \tilde{w}\hat{z}n + (1+r)a - \text{Tax} + \text{TR}, \text{ if } \iota = 0; \tag{8}$$

$$(1+\tau_c)c + a' + i'_{hi}\tilde{\pi}^j = \tilde{w}\hat{z}n + (1+r)a - \text{Tax} + \text{TR} - \lambda \text{ and } a' = 0, \text{ if } \iota = 1; \tag{9}$$

$$a' \geqslant -\underline{a}, \tag{10}$$

where

$$\tilde{\pi}^j = \begin{cases} \pi_E(1-\psi), & \text{if } i_{hi} = 1, i_E = 1; \\ \pi_P^j(m), & \text{if } i_{hi} = 1, i_E = 0; \\ 0, & \text{if } i_{hi} = 0 \text{ or } e \leqslant \Theta_e \text{ and } a \leqslant \Theta_a; \end{cases} \tag{11}$$

$$\text{Tax} = T(y) + \tau_{mr}(\tilde{w}\hat{z}n - i_E\tilde{\pi}^j) + \tau_{ss}\min\left\{(\tilde{w}\hat{z}n - i_E\tilde{\pi}^j), y^{ss}_{max}\right\}; \tag{12}$$

$$y = \max\{\tilde{w}\hat{z}n + ra - i_E\tilde{\pi}^j, 0\}; \tag{13}$$

$$\text{TR} = \max\left\{0, (1+\tau_c)\underline{c} + (1-\iota)(1-i_{hi}\phi(m))m + T(\tilde{y}) - \tilde{w}\hat{z}n - (1+r)a\right\}; \tag{14}$$

$$\tilde{y} = \tilde{w}\hat{z}n + ra. \tag{15}$$

The budget constraint (8) states that the household finances consumption c, savings a' (subject to a borrowing constraint), the purchase of health insurance $\tilde{\pi}$, and out-of-pocket health expenditure $(1 - i_{hi}\phi(m))m$ using after-tax capital and labor income $\tilde{w}\hat{z}n + (1+r)a$. Wages adjust with EHI offering as in equation (6). The health insurance premium varies with the policy the household chooses as in equation (11). The budget constraint (9) states a similar condition when the household declares bankruptcy. The government subsidizes EHI purchases based on $i_{hi}i_E\tilde{\pi}$. The household pays consumption tax, income tax, and Medicare and Social Security tax given by equation (12). The social insurance provides a minimum consumption floor $\underline{c}$ through a lump-sum transfer given by equation (14).

The retired agent's problem ($j \geqslant j_R$)

$$\mathbf{V}^j(s_r, S) = \max_{\{c,a',\iota\}} \left\{ \ u(c,n) + \beta\rho_{h,j}\mathbb{E}\mathbf{V}^{j+1}(s_r', S') \ \right\} \tag{16}$$

subject to

$$(1+\tau_c)c + a' + \pi_{mr} + (1-\phi_{mr}(m))m \leqslant ss(\bar{e}) + (1+r)a - T(ra) + \mathrm{TR}, \ \text{ if } \iota = 0, \tag{17}$$

$$(1+\tau_c)c + a' + \pi_{mr} \leqslant ss(\bar{e}) + (1+r)a - T(ra) + \mathrm{TR} \text{ and } a' = 0, \ \text{ if } \iota = 1, \tag{18}$$

$$a' \geqslant -\underline{a}, \tag{19}$$

where

$$\mathrm{TR} = \max\left\{0, (1+\tau_c)\underline{c} + (1-\iota)(1-\phi_{mr}(m))m + \pi_{mr} + T(ra) - ss(\bar{e}) - (1+r)a\right\}, \tag{20}$$

and $\bar{e}$ is the average lifetime labor income before an agent retires. Since all retired agents are enrolled in the Medicare program, retirees have no health insurance choice. Their income comes from the social security benefit and their own savings. A social welfare transfer may apply based on Equation (20).

3.5.2 Health insurance company

There is free entry in the health insurance market. To maintain the zero-profit condition, the health insurance company sets the insurance premium π_E such that it is enough to cover the insurance company's share of the total medical expenditure $\phi_E(m) \cdot m$ and a proportional markup η due to administrative cost:

$$\pi_E = (1+\eta)\frac{\int i_{hi}(s) i_E \phi_E(m) m \varphi(s) ds}{\int i_{hi}(s) i_E \varphi(s) ds}. \tag{21}$$

The cost to the representative firm for providing EHI to its workers is given by

$$c_E = \frac{\int \psi \pi_E i_E i_{hi}(s) \varphi(s) ds}{\int i_E z(s) g(h(s)) n(s) \varphi(s) ds}, \tag{22}$$

where $\varphi(s)$ denotes the equilibrium distribution of households.

In the individual health insurance market, the insurer sets the premium to satisfy the zero-profit condition for each contract; that is,

$$\pi_p^j(m) = \frac{(1+\eta)\mathbb{E}\{\phi_p(m')m'|m\}}{1+r}. \tag{23}$$

3.6 Definition of equilibrium

We consider a stationary competitive equilibrium defined as the following.

Definition 1 *Given government policies, including income tax function $T(\cdot)$, consumption tax τ_c, Medicare, social security, and social insurance program, a stationary competitive equilibrium consists of factor prices w, r; aggregate labor and capital N, K; allocation functions for workers $\{c(\cdot), a'(\cdot), i'_{hi}(\cdot), n(\cdot), \iota(\cdot)\}$ and for retirees $\{c(\cdot), a'(\cdot), \iota(\cdot)\}$; value functions $\mathbf{V}^j(\cdot)$ for all age j; health insurance contracts $\left\{\pi_E, \phi_E(\cdot); \pi_P^j(\cdot), \phi_P(\cdot)\right\}$; and distribution of households $\varphi(s)$ over state space $\mathbb{S}$ such that*

1. *Given prices, government policies, and health insurance contracts, the allocations solve the individual's problem;*

2. *N, K solve the firm's optimization problem;*

3. *The health insurance company behaves competitively;*

4. *Labor market clears: $N = \int z(s)g(h(s))n(s)\varphi(s)ds$;*

5. *Capital market clears: $K = \int a(s)\varphi(s)ds$;*

6. *Good market clears: $C + K' - (1-\delta)K + M + G = Y$;*

7. *Government's budget is balanced:*

$$\begin{aligned} G &+ \int_s [ss(\bar{e}(s)) + \phi_{mr}m(s) - \pi_{mr}]\mathbf{1}\left[j \geqslant j_R\right]\varphi(s)ds + \int_s TR(s)\varphi(s)ds \\ &+ \int_s \iota(s)[(1-\phi(m))m(s) - \lambda(s)]\varphi(s)ds + \int_s \phi_{md}(m)m(s)\mathbf{1}_{e\leqslant\Theta_e, a\leqslant\Theta_a}\varphi(s)ds \\ = &\int \left[\tau_c c(s) + T(y(s))\right]\varphi(s)ds + \int_s \left[(\tau_{mr}(y(s)) + \tau_{ss}(y(s))\right](1-R(s))\varphi(s)ds; \end{aligned} \tag{24}$$

8. *The distribution of agents is stationary: $\varphi(s) = \mathbb{L}\left[\varphi(s)\right]$.*

4 The Benchmark Economy

4.1 Estimation strategy

We estimate the model using two data sets: The income process is estimated from the Panel Study of Income Dynamics (PSID) data, and the medical expenditures and insurance parameters are estimated using the Medical Expenditure Panel Survey (MEPS) data. Appendix A provides an in-depth description of the MEPS and PSID samples.

We restrict the key parameters of our model using rich micro data. We start by describing our strategy for estimating the interdependence between income and health, especially the causal effect of health on income, and the reverse causation of income on health through insurance choice, using the instrumental variable approach. We then describe how we estimate other components of the model, including the non-parametric estimations of medical expenditure shocks, co-insurance rates, and EHI offer rate, from the MEPS data. Parameters governing the demographic process and welfare and taxation programs can be determined directly from relevant macro data, such as life tables and tax documents. Finally, we calibrate all other parameters, such as the discount factor, following the standard procedures in the literature by comparing the equilibrium model moments with the data.

4.1.1 Income process

We estimate the income process directly from the data, independent of the general equilibrium. Recall that the labor income of an individual i of age $j \leqslant j_R$ is given by

$$e_j^i = \tilde{w}\hat{z}_j^i,$$

where $\hat{z}_j^i$ is a product of the health component $g_j(h)$ and a component representing the underlying income process z_j^i. We now describe how we estimate the process of z_j^i, and in particular, how to separately identify $g_j(h)$ from the underlying income process z_j^i, allowing for simultaneity between health and income, which sets us apart from existing literature.

Since MEPS only has a panel of two periods, we estimate the income process using PSID, which maintains a longer panel of information on U.S. households. There is a long list of literature on income dynamics using PSID data.[12] Unfortunately, we cannot use the litera-

[12] See, for instance, Meghir and Pistaferri (2004), Guvenen (2009), Hospido (2012), and Gu and Koenker (2017), among others.

ture's estimates of labor income processes directly as they do not account for individual health dynamics. In other words, the estimated income process conflates the income process of the healthy with that of the unhealthy, which may lead to bias. For instance, if an individual whose health status changes from healthy to sick suffers a drop in her earnings as a result of missing work, ignoring the health transition information would downward bias the income persistence parameter. In our PSID sample of 2052 individual households, 563 household heads (about 27% of the sample) report staying healthy throughout the observed periods and 75 report remaining unhealthy throughout. The remainder (about 69% of the sample) experience health status transition, which is a non-negligible proportion.

To construct the underlying income process z_{it} independent of the health component $g_j(h)$, we consider the following econometric model. Denote y_{it} as the (log) observed income of an individual i for year t, denominated by the nominal GDP per capita of year t, and $g_j(h_{it})$ as the age-specific effect of health status on labor income. We assume

$$\begin{aligned} y_{it} &= g_j(h_{it}) + X_{it}\beta_e + \varepsilon_{it}, \\ \text{where } \varepsilon_{it} &= \alpha_i + \rho\varepsilon_{it-1} + u_{it}, \end{aligned} \tag{25}$$

with $u_{it} \sim N(0, \sigma_i^2)$ and the joint distribution of (α_i, σ_i^2) is denoted as F.

The control variables X_{it} include nine age group dummies (20–24, 25–29, 30–34, ..., 60–64) and year fixed effects. We follow Meghir and Pistaferri (2004) and the standard practise in the literature to first eliminate the effects of $g_j(h_{it})$ and X_{it} and then focus on the log earning residuals ε_{it} to estimate the income process. The impact of X_{it} can be easily estimated using an ordinary least squares (OLS) estimation. The impact of health $g_j(h_{it})$, however, is more complicated. In particular, the OLS estimator of $g_j(h_{it})$ suffers from simultaneity bias: Health affects income, and income also affects health through endogenous insurance choice. Appendix A discusses this issue in detail. We then take an instrumental variable (IV) approach to address this issue.

We argue that information on hospital stays of the household head serves as a good instrumental variable for health. Hospital stays directly correlate with health status, yet its correlation with income operates solely through health status. A further complication in our case here is that since the health variable is binary (healthy vs. unhealthy), we cannot apply the two-stage least squares (2SLS) directly due to the forbidden regression (see Angrist and Pischke (2008) p. 142). We follow the suggestions in Angrist and Pischke (2008) and use a probit model for the first stage and then use the fitted value as an instrument to estimate the impact of health on income. Due to the fact that in PSID hospital stays data exist only for three years (every two years during 1983–1987), we cannot accurately estimate $g_j(h)$ for the full interaction of age and the health indicator. We hence restrict $g_j(h) = g(h)$ for all age j, i.e., the age profile of income for the healthy to be a parallel shift of those for the unhealthy. We then obtain a 2SLS estimate $\hat{g}(h = H) - \hat{g}(h = U) = 0.674$ which is significant at the one percent level. We highlight that this is in contrast to an OLS estimate of 0.329, which is evidence of simultaneity bias. Again, since the hospital stay information only spans three years, so as not to sacrifice the rest of the data, we obtain the log earning residuals from the following regression on the whole sample

$$Y_{it} - \hat{g}(h_{it}) = X_{it}\beta_e + \varepsilon_{it}.$$

We then use the estimated residual term, $\hat{\varepsilon}_{it}$, to estimate the underlying income process z_{it} including dynamic parameter ρ and the individual fixed effect for both location and scale (α_i, σ_i^2). Again, note that this z_{it} should be independent of health status. In contrast to the existing literature on income dynamics, not only do we allow for individual fixed effect α_i here, but our variance of income shock u_{it} is also individual specific. This dimension of heterogeneity is important for several reasons. First, it is well known that scale mixtures of normal distributions are very flexible for fitting long tail behavior of income, which we do observe in the log earnings residual. Figures 16 and 17 in Appendix A provide evidence that the marginal distribution of u_{it} deviates from the normal distribution and presents long tail behavior. Indeed, the kurtosis of our estimated u_{it} is 5.2, which is much higher than the normal distribution's kurtosis of 3. We provide additional discussion concerning this point in Appendix A. Second, ignoring the variance heterogeneity in u_{it} also leads to bias in the dynamic parameter. See Gu and Koenker (2017) for a more detailed discussion.

To estimate ρ and distribution F, we follow the methodology in Gu and Koenker (2017). In particular, the joint distribution is estimated via a non-parametric maximum likelihood (NPML) method (see Heckman and Singer (1984) for its use in duration models). The advantage of the NPML method is that it does not impose any parametric or smoothness assumptions on the class of distributions for (α_i, σ_i^2) and does not involve any tuning parameters, which is also more natural to be brought to model simulation. The non-parametric maximum likelihood estimator for F is discrete as shown by Lindsay (1983) and its consistency is established in

Kiefer and Wolfowitz (1956). The parameter ρ is estimated through the profile likelihood estimator. That is, on a grid of values for $\rho \in [0, 1]$, we apply the NPML to obtain $\hat{F}(\rho)$ and maximize the resulting profile log-likelihood. The log-likelihood peaks at $\hat{\rho} = 0.25$, while the joint distribution F of (α_i, σ_i^2) are reported in Table 5 of Appendix A. The lower persistence of the income shock resonates with the findings in Guvenen (2009) that once richer heterogeneity is permitted in the income process, we no longer see close to unit root behavior of ε_{it}.

We also consider the possibility of unemployment. We estimate the probability of unemployment conditional on the event that the individual is working in the last period as a function of log real income and health status. Given the Markov condition, we create a rolling two year panel of individual's employment status. Conditional on working, the probability of unemployment or being laid off as a function of income and health status is estimated through the following probit model:

$$\mathbb{P}(\text{unemp}_{it} = 1) = \Phi(\alpha_{\text{unemp}} + \beta_{\text{unemp}} \log(y_{i,t-1}) + \gamma_{\text{unemp}} h_{it}),$$

where Φ is the normal CDF, $y_{i,t-1}$ is the real labor income of the previous period, $h_{it} \in \{H, U\}$ is the indicator for health, and unemp_{it} is the indicator for unemployment status. We find that $\hat{\beta}_{\text{unemp}} = -0.174$ and $\hat{\gamma}_{\text{unemp}} = -0.176$, both of which are significant at one percent level. Hence, we allow the probability of being unemployment to depend on income level of the previous period and health status. We also estimate the probability of an individual regaining employment after having been out of a job in the last period to be 0.659.

We then use the estimated process of ε_{it} specified in Equation (25) to construct our income process z_{it}. For each type of (α_i, σ_i^2), process z_{it} is then approximated using a Markov chain, augmented with one additional state of unemployment of which the probability is previously estimated. We highlight that our estimated income process deviates from the literature in three important ways: We estimate the income process allowing for heterogeneous variances among individuals; we consider the possibility of unemployment as a worst state of the Markov process; and most importantly, we separately identify the health component $g(h_{ij})$ and underlying income process z_{it} allowing for the possibility of simultaneity bias. Finally, note that one period in our model represents 5 years and hence we adjust the income process accordingly.[13]

[13] Note that assuming the model period to be 5 years affects the size of the simultaneity bias between health and income, which is in turn determined by the autocorrelation of the income process. A 5-year modelling period implies a much lower autocorrelation than a 1-year modelling period. Hence, if we estimate the effect of health on income using OLS, the resulting simultaneity bias in the model simulated data will be smaller than that from the real data. This potentially explains the fact that in the benchmark estimation of our model, the

4.1.2 Health transition matrix

The previous section describes how we estimate the impact of health on income, summarized by $\hat{g}(h)$. In this section, we investigate reverse causality. As we may expect, income could affect health in many different ways. As emphasized in the introduction, we focus on one of such channels: How does income affect health through individuals' insurance choice. In particular, we need to estimate the impact of insurance coverage status on the transition of health.

Recall that the transition matrix of health status is written as

$$\pi^{j,i_{hi}} = \begin{bmatrix} \pi_{HH}^{j,i_{hi}} & \pi_{HU}^{j,i_{hi}} \\ \pi_{UH}^{j,i_{hi}} & \pi_{UU}^{j,i_{hi}} \end{bmatrix},$$

where j is age and i_{hi} is the insurance status. It is estimated from MEPS directly and is also independent of the general equilibrium. In particular, MEPS reports the perceived health status of each individual in five categories, from "excellent" to "poor". We treat an individual as healthy (H) if the perceived health status falls within the first three categories ("excellent", "very good", and "good"), and unhealthy (U) otherwise. This criterion is common in the literature, such as in de Nardi et al. (2017) and Hong et al. (2017), among others. By this criterion, around 83% of individuals are healthy and the rest are unhealthy.

We separately estimate the transition matrix for individuals of each age j, using the following probit setup:

$$\text{Prob}[h' = H|h, j] = \Phi(\beta_\pi i_{hi} + \gamma_\pi X_i), \quad h \in \{H, U\}, \tag{26}$$

where i_{hi} is the health insurance coverage dummy, and the control variables X_i include standard demographic controls—education, marriage, race, gender—and particularly income. $\pi_{Hh}^{j,i_{hi}}$ in the transition matrix is then simply this probit function evaluated at $X_i = \bar{x}$ where $\bar{x}$ is the mean of X_i for individuals with age j independent of h and i_{hi}:

$$\pi_{Hh}^{j,i_{hi}} = \text{Prob}[h' = H|h, j, X_i = \bar{x}] = \Phi(\beta_\pi i_{hi} + \gamma_\pi \bar{x}), \quad h \in \{H, U\}.$$

The key parameter of interest is β_π, which governs how medical insurance coverage affects health transition. If $\beta_\pi > 0$, then individuals with insurance coverage are more likely to have good health in the next period. It is important to highlight that medical insurance coverage is endogenous. An individual may *choose to* opt in an medical insurance plan if she expects poor

income gap between healthy and unhealthy is somewhat larger than that in the data.

health in the following period. In other words, if we were to directly estimate β_π using maximum likelihood, then we would overestimate the probability of π_{HU} for the insured. Indeed, a naive maximum likelihood estimator of β_π yields $\hat{\beta}_\pi < 0$ for many age groups; i.e., individuals with health insurance coverage are more likely to have bad health status in the next period, which reflects the selection bias.

To address this endogeneity issue, we again use the instrumental variable approach. The challenge here is that not only the health insurance dummy is endogenous, but the income variable included in X is also likely to be endogenous. This departs us from the usual probit model with an endogenous variable. Income directly affects health status, and also indirectly affects health through the insurance choice. The parameter of interest β_π is called mediation effect and the health insurance choice is the mediator variable. If we want to identify both coefficients β_π and γ_π, then we need to look for two instruments to separately address the endogeneity from both income and the mediator variable—health insurance status; see for example Flölich and Huber (2017). However, it is extremely challenging to find a valid instrument for income based on observational data. Fortunately, since we are not interested in the direct effect of income but rather just focus on the indirect mediation effect through health insurance status, we can use one instrument for the insurance status to identify β_π. In Appendix B.2, we discuss in detail on the necessary assumptions on this instrument for us to identify the parameter β_π. Briefly speaking, the exclusion restriction condition requires that conditional on income, the instrument does not depend on any other factors that jointly influence health insurance choice and health status. The rank condition requires that conditional on income, there is still correlation between the IV and the health insurance choice.

One possible instrument is EHI offer status, since whether or not a firm offers EHI is the firm's decision and independent of individual characteristics of its workers, given the non-discriminative nature of group insurance. However, this EHI offer status may still contain some endogeneity, as workers who expect bad health shocks may choose to work in firms that offer EHI (Feng and Zhao, 2018). To further address this endogeneity of firm-worker matching, we use the average EHI offer rate of firms within the same industry and the same worker age group as an instrument of the insurance coverage status. The rank condition is also satisfied as we find a strong first stage effect between the instrument and the individual health insurance choices conditional on income, with details in Appendix B.3. Note that the endogenous regressor in this probit model i_{hi} is a binary choice variable, and therefore we cannot use the standard two-stage least square procedure of IV estimators. We instead estimate the model specified in (26) using a bi-probit setup with the group-average EHI offer rate as an instrument for insurance status.

This IV estimator provides us with a consistent estimate of β_π for each age j and health status h. Without relying on the estimate of γ_π, we can consistently estimate the *difference* in the transition between the insured and the uninsured:

$$\begin{aligned} &\Delta\text{Prob}[h' = H|h, j, X = \bar{x}] \\ &= \text{Prob}[h' = H|h, j, X = \bar{x}, i_{hi} = 1] - \text{Prob}[h' = H|h, j, X = \bar{x}, i_{hi} = 0], \end{aligned} \tag{27}$$

which is exactly the "health premium" of insurance—individuals with insurance coverage are more likely to remain healthy or to recover from unhealthy status—evaluated for each age group at the mean. The details of this procedure is in Appendix B.1. The results are illustrated in Figure 3. It is clear that insured individuals have significant and positive health premium, consistent with the evidence that having insurance coverage incurs more preventive care and hence improves health status (Ozkan, 2014). Note again that allowing for insurance status to affect health is also a reduced-form way to capture the health production function estimated in Hong et al. (2017).

With this health premium $\pi^{j,1}_{Hh} - \pi^{j,0}_{Hh}$ together with the observed unconditional transition probability π^{j}_{Hh} and the fraction of insured, we can then separately identify the transition probability conditional on insurance choice: $\pi^{j,1}_{Hh}$ and $\pi^{j,0}_{Hh}$ that are needed for calibration. Again we describe the details in Appendix B.1. We then repeat this process for all age group j and the associated transition probabilities are illustrated in Figure 4. Note that in the calibration, we use a polynomial of age to smooth the transition probability as the red dashed line in the figure.

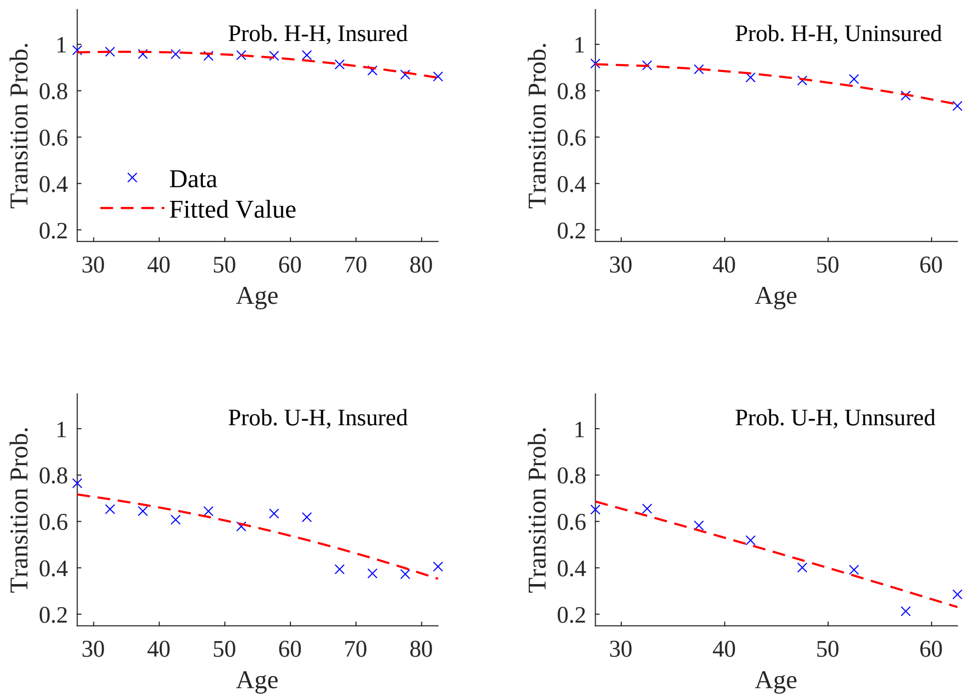
4.1.3 Demographics, preferences, and endowments

Individuals are born at age 25 ($j = 1$) and retire at age 65 ($j = 8$, given each period in our model is 5 years). For simplicity, we assume that the labor supply is inelastic. The utility function takes the form of Constant Relative Risk Aversion (CRRA):

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

We follow Attanasio et al. (2011) and set $\sigma = 2$, which implies an intertemporal elasticity of substitution of 0.5. We follow Hall and Jones (2007) and Ozkan (2014) by assuming that each

Figure 4: Health Transition



Note: This figure plots the probability of transiting from healthy to healthy (H-H) or unhealthy to healthy (U-H), for individuals with and without insurance coverage. For uninsured individuals, we only show the probability up to age 65 since older individuals are covered by the Medicare program and hence are all insured. The dashed lines show the fitted values that we use to simulate the model.

individual gains a flow utility b in every period. This assumption guarantees a positive flow utility in every period so that individuals prefer a *longer* life expectancy. We hence choose b such that, in equilibrium, all individuals have positive value of life $V^j(.)$, with $\min\{V^j(.)\} = 0$.

Recall that the survival rate $(\rho^j_{h,t})$, and therefore the mortality rate, differs between healthy and unhealthy individuals. We estimate the survival rates following Hall and Jones (2007). In particular, mortality may arise from issues unrelated to health, such as accidents and homicides/suicides, whose rate we denote as δ_a, and health-related mortality (non-accident mortality) we denote as δ_d. We assume that healthy individuals are only subject to accident mortality and not health-related mortality, while unhealthy individuals suffer from both. We estimate these two mortality rates using data from the National Center for Health Statistics publication *Health, United States 2016*. See Appendix C.1 for details. In addition, we assume that the longest life span is age 85 ($j = 12$). However, our results are not sensitive to this age of death as all retired individuals are covered by the Medicare and therefore do not need to make health insurance choice in our framework.

We set annual discount factor β to 0.87 to match the private sector's capital output ratio of 2.4 in the year 2000. The capital share in the production function is set to 0.36, and the depreciation rate is set to 0.06. The borrowing limit $\underline{a}$ is set to 0.

4.1.4 Medical expenditure and health insurance

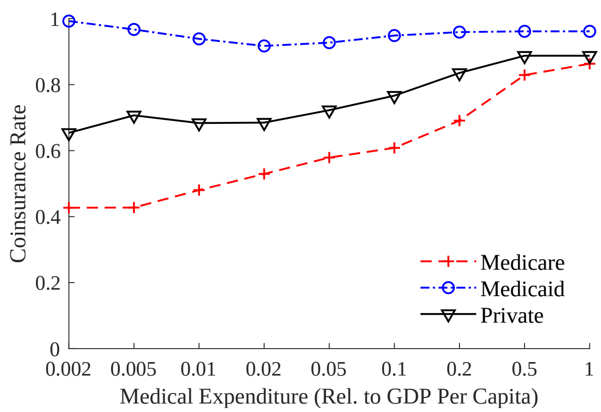
Medical expenditure shocks. We estimate medical expenditure shocks non-parametrically. Medical expenditure shocks, conditional on age and health status, are assumed to be independent and identically distributed among individuals and over time. For each age-health cohort, we discretize health expenditure shocks according to the sample observations in the MEPS data using seven grid points, which represent the 10th, 25th, 50th, 75th, 90th, 95th, and 99th percentile of all medical expenditure observed among individuals within that cohort.[14] Also note that we only use observations from insured individuals to estimate medical expenditures to avoid the endogeneity that uninsured individuals may purchase less medical services than they need due to budget constraints, and hence incur lower medical expenditures.

Co-Insurance rates and EHI offer rates. We estimate co-insurance rates non-parametrically from MEPS: $\phi_E(m)$, $\phi_P(m)$, $\phi_{md}(m)$, and $\phi_{mr}(m)$. In particular, we calculate the fraction of total medical expenditures paid by insurance and fit a piecewise linear function of the co-insurance rate that varies by the amount of medical expenditure m. The estimated co-insurance rates are presented in Figure 5. We also estimate EHI offer rate $p_E(e)$ non-parametrically from MEPS and similarly fit a piecewise linear function of the probability of receiving EHI as observed in the data, varying by income levels. The left panel of Figure 2 shows that jobs with higher wage payments are more likely to offer EHI, consistent with Aizawa and Fang (2019). See Appendix C.2 for details.

Health insurance premium. We set the firm's share of health insurance premium ψ to 80%, which is in the empirical range of employer contribution rates. Gruber (2008) documents

[14]Note that we add the 95th and 99th percentile grid points to better capture the right tail of medical expenditures which generates substantial risk for individuals. For example, for an unhealthy individual of age 60, the 99 percentile of medical expenditure can be over \$100,000 USD (in 2015) or roughly two times of GDP per capita.

Figure 5: Coinsurance Rates



Note: The figure shows the co-insurance rates of Medicare, Medicaid, and private insurance, respectively, as functions of medical expenditures, which are normalized by the median household income.

that the administrative cost of private insurance is about 12% of the premium, so we set the markup on the insurance premium η to $0.12/(1-0.12) = 13.6\%$. The Medicare premium π_M is determined in equilibrium such that the annual Medicare premium is about 2.1% of aggregate GDP based on the MEPS data.

An individual is eligible for Medicaid if $e \leqslant \Theta_e$ and $a \leqslant \Theta_a$. The income threshold varies across states and depends on family size. We follow Pashchenko and Porapakkarm (2017) and set the income threshold Θ_e to 0.311 of GDP per capita, and the asset threshold Θ_a to 0.538 of GDP per capita.[15]

4.1.5 Government programs

Medical bankruptcy We follow Livshits et al. (2007) and assume, in the event that an individual files for bankruptcy, that individuals are subject to a linear garnishment of earnings $\lambda = \gamma \max\{y - \bar{y}, 0\}$ during the bankruptcy period. In particular, we set the marginal garnishment rate γ to 0.35 and the earning exemption $\bar{y}$ to 20% of GDP per capita.

Consumption floor. In line with Jeske and Kitao (2009), we interpret the minimum consumption floor $\underline{c}$ as the poverty line, which we set to 0.103 of GDP per capita; i.e., $\bar{c}$ is \$5,000 USD in 2008. Note that in our model, Medicaid can also be viewed as a part of the consumption

[15]The asset threshold in reality also depends on family size and characteristics, but is generally low (below \$2,000 USD in cash per individual, plus some exclusions, such as one car, for most states in 2005).

floor in addition to $\underline{c}$ as it is used to cover medical expenditure while its cost is close to zero.

Social security. The U.S. social security system is progressive. The benefit payment function is given as follows:

$$ss(\bar{e}) = \begin{cases} s_1\bar{e}, & \text{for } \bar{e} \leqslant \tau_1 \\ s_1\tau_1 + s_2(\bar{e} - \tau_1), & \text{for } \tau_1 < \bar{e} \leqslant \tau_2 \\ s_1\tau_1 + s_2(\tau_2 - \tau_1) + s_3(\bar{e} - \tau_2), & \text{for } \tau_2 < \bar{e} \leqslant \tau_3 \\ s_1\tau_1 + s_2(\tau_2 - \tau_1) + s_3(\tau_3 - \tau_2), & \text{for } \bar{e} > \tau_3 \end{cases} \tag{28}$$

where, respectively, marginal replacement rates s_1, s_2, and s_3 are set to 0.90, 0.33, and 0.15; threshold levels τ_1, τ_2, and τ_3 are set to 20%, 125%, and 246% of the median income; and annual Medicare premium π_{mr} is about 2.6% of GDP based on MEPS data.

Taxes and government expenditure. We set the consumption tax rate τ_c to 5.67%, following Mendoza et al. (1994). The social security tax is paid by both the employer and the employee; with each side paying 6.2 percent, we therefore set τ_{ss} to be 12.4%. Similarly, we set Medicare tax rate τ_{mr} to be 2.9%. While the maximum taxable income for social security tax increases over time, it is roughly two times of GDP per capita.[16] We therefore set y^{ss}_{max} to 2.0. We use non-linear income tax function $T(y) = y - \lambda_p y^{1-\tau_p}$ as specified by Heathcote et al. (2017). Using PSID data, they estimate income tax to be progressive with $\tau_p = 0.151$ with a standard error of 0.003. λ_p is determined in the equilibrium to balance the overall government budget. We assume a fixed government expenditure G equal to 18% of GDP in the benchmark economy.

To summarize, we use micro data to restrict the income process, the health process, and in particular, the inter-dependence between health and income, which are key to our analysis. We also estimate medical expenditure shocks, co-insurance rates, and mortality rates directly from data. Other parameters commonly found in the literature are calibrated in the standard way following the literature, and the values of these parameters are summarized in Table 9 of Appendix C.

[16] For example, in 2005, the maximum taxable income is \$90,000 USD and GDP per capita is \$44,237 USD, the ratio of which is 2.03.

4.2 Joint distribution of health and income

In our framework, both health and income are endogenous: Health directly affects income through function $g_j(h)$, and income affects health through the individual's endogenous choice of insurance. This makes matching the joint distribution of health and income a non-trivial task. For instance, if the estimated model over-predicts the percentage of individuals who take up health insurance, then the model would over-predict the percentage of healthy individuals and thereby affect the income distribution as well. We identify the income process z_{it}, the effect of health on income $g_j(h)$, and the effect of insurance coverage on health transition matrix $\pi_{Hh}^{i_{hi}}$ separately by applying econometric techniques to micro data independent of the general equilibrium. Despite the fact that we do not target any dimensions of the joint distribution in the general equilibrium, our model's rich structure and estimation accuracy endogenously generate a joint distribution that replicates well the key facts regarding to the health-income interaction as summarized in Section 2.

In our model, 78.1 percent of individuals aged 25–64 have health insurance, compared with 77.3% in the data. The model generates a reasonable composition of insurance, in terms of EHI, other private insurance, and public insurance; see Table 1. In addition to the overall level of insured, our estimated model matches the distribution of insurance over the life cycle and by income level; see Figure 6. Similar to the data, our model's percentage of insurance coverage increases with age. The model also replicates the pattern that individuals with higher income are more likely to be insured.

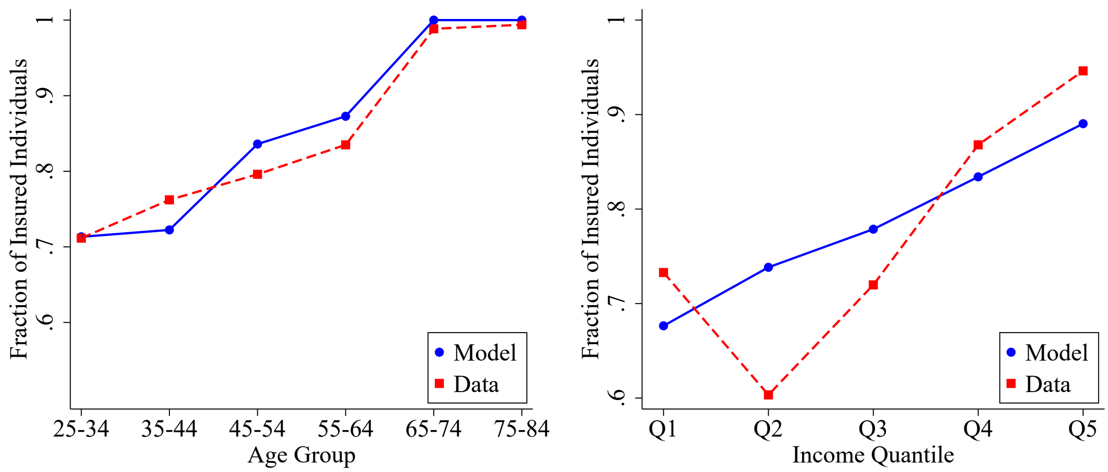
Table 1: Overall Insurance Coverage Rate

Insurance Rates	Model	Data
Percentage of Insured	78.1	77.3
EHI	53.4	46.8
Other Private	12.0	16.4
Public	12.7	14.1

Note: This table compares the model's prediction on insurance coverage with MPES data for the working age population, with a breakdown into different categories of insurance.

Next, we report the distribution of health status in our baseline economy, which is not targeted in the estimation. One novelty of our model is that the health transition depends crucially on insurance status. Hence, having a correct distribution of insurance is necessary

Figure 6: Distribution of Insured

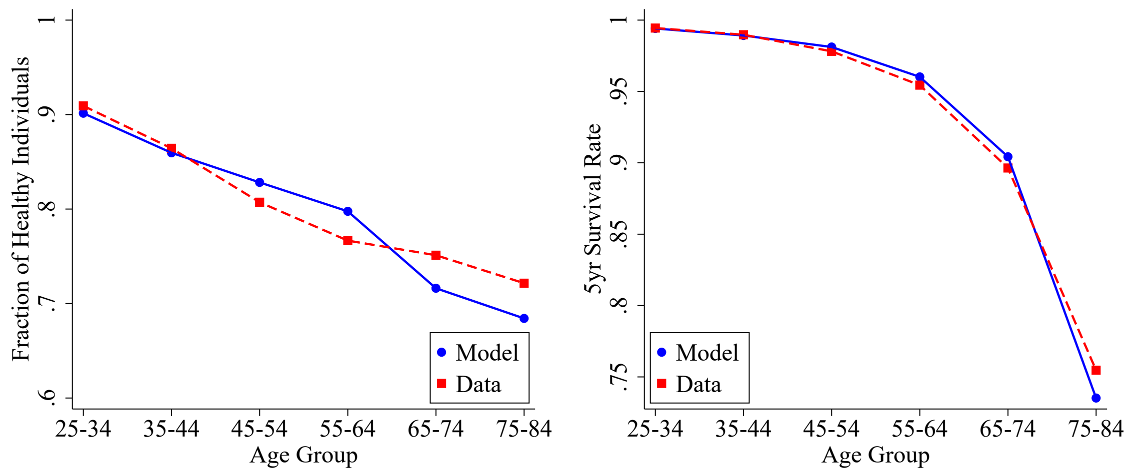


Note: The figure shows the insurance coverage rate by age group (left panel) and over income (right panel). The blue solid line represents simulated outcomes, and the red dashed line represents empirical moments from MEPS data.

to obtain a correct health distribution. On top of that, we would need the health transition matrix, conditional on insurance coverage, to be estimated properly. The left panel of Figure 7 shows the fraction of healthy individuals by age group. As we do not directly target the health distribution in our calibration, our model's goodness of fit in this dimension indicates that our health transition matrix estimation is accurate. We estimate the mortality rate by health for each age group. Given the simulation accuracy of the health distribution over the life cycle, it is not surprising that our model is also able to replicate overall mortality rates over the life cycle, as shown in the right panel of Figure 7.

A key feature of our framework is that health and income are jointly determined in equilibrium. The left panel of Figure 8 shows the percentage of healthy individuals over income quantiles. The baseline model replicates the pattern that individuals with higher income are more likely to be healthy. The right panel of Figure 8 shows the income differences between healthy and unhealthy individuals. As we document before, unhealthy individuals have lower income and higher medical expenditures. We hence plot the median income gap between healthy and unhealthy individuals over the life cycle. In spite of not being targeted, the model replicates the data pattern that healthy individuals tend to have higher income than unhealthy individuals. While we target how health affects income with function $g_j(h)$, health itself is endogenous and depends on income, and hence the equilibrium relationship between health and income is

Figure 7: Health and Mortality Rate over Life Cycle



Note: The left panel of the figure shows the fraction of healthy individuals by age group. The right panel shows the mortality rate by age group. The blue solid line represents simulated outcomes, and the red dashed line represents empirical moments from MEPS data and life tables.

more complicated than the function of $g_j(h)$.

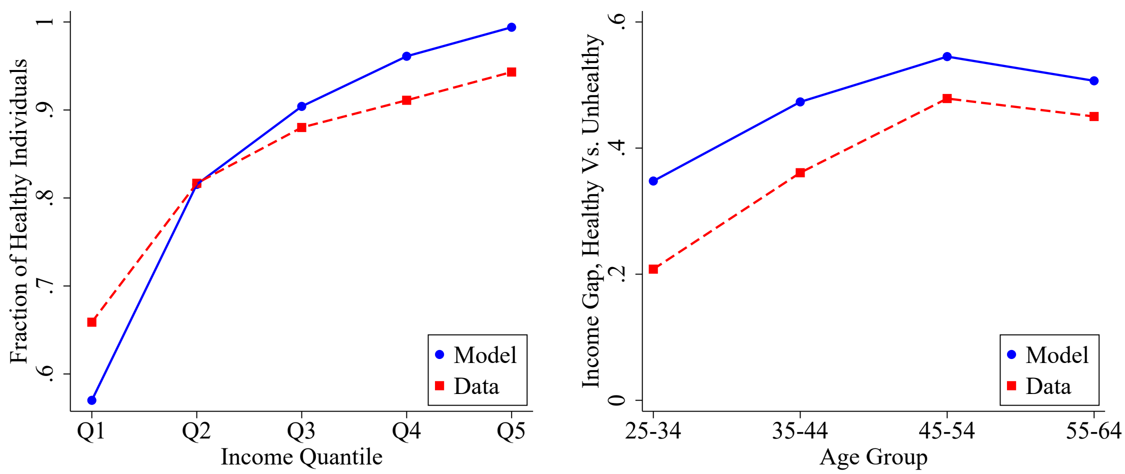
Another way to look at the joint distribution is to redraw the left panel of Figure 7, which is the fraction of healthy individuals by age group, for individuals with top quartile income and for those with below median income, rather than of all individuals. The results are displayed in Figure 9. We can conclude that the model replicates the empirical health distribution over the life cycle, for both the rich and poor.

Our model also replicates the income inequality as we observed in the data. Table 2 reports the level of inequality measured by the income share of the top 10th, 25th, and 50th percent of the income distribution, as well as the Gini index. The moments on inequality level are not explicitly targeted in the calibration. This provides further evidence that our estimated income and health processes are accurate, and the joint distribution in the equilibrium captures the salient features in the data.

4.3 Decomposition of uninsured

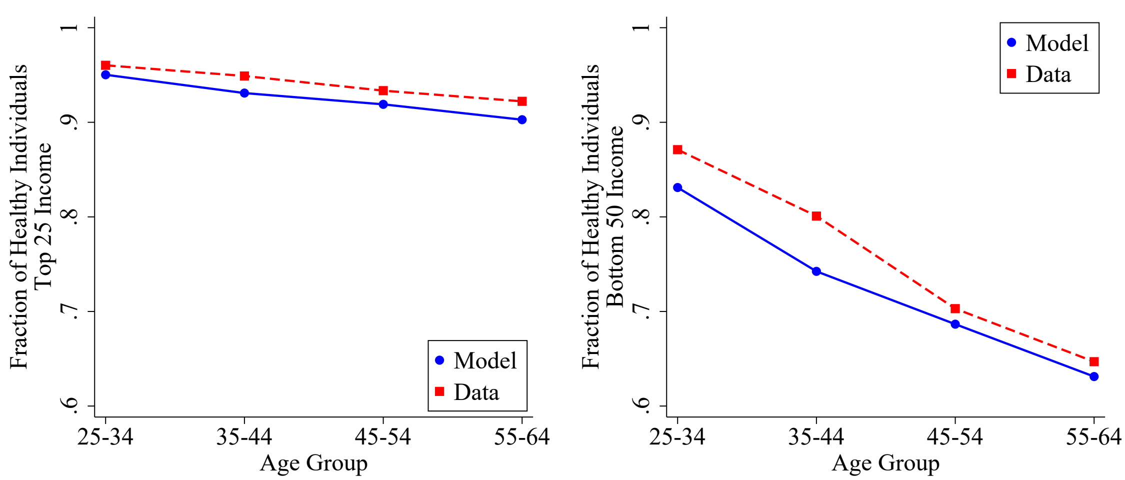
After years of steady decline in the number of uninsured since the passage of the Affordable Care Act, there are still 27.5 million people without health coverage (United States Census Bureau 2019). We use our benchmark model to analyze the factors that contribute to the observed uninsured rate, despite the considerable cost associated with being uninsured.

Figure 8: Joint Distribution of Health and Income



Note: The left panel shows the fraction of healthy individuals by age group. The right panel shows the mortality rate by age group. The blue solid line represents simulated outcomes, and the red dashed line represents empirical moments from MEPS (left panel) and PSID (right panel).

Figure 9: Health over Life Cycle, by income



Note: The left panel of the figure shows the fraction of healthy individuals by age groups for individuals with top quartile income, and the right panel shows that for individuals with below median income. The blue solid line represents the model outcome, and the red dashed line represents the data moments from MEPS.

Table 2: Income Inequality

Moments	Model	Data
Income Share of		
Top 10 Percentile	0.27	0.27
Top 25 Percentile	0.51	0.49
Top 50 Percentile	0.78	0.75
Gini Index	0.39	0.36

Note: The table displays the level of income inequality, measured as the income share of top 10th, 25th, and 50th percentiles of the distribution as well as the Gini index. The empirical moments are calculated using PSID data.

The literature documents various reasons as the main causes of uninsurance, including the minimum consumption floor or medical bankruptcy as implicit insurance, high markups in the private insurance market, and the lack of EHI offer for some employees. Our baseline model explicitly models all of these channels so that we may test the extent to which any one of the aforementioned factors affects the decision to insure. We then experiment by shutting down these channels one at a time to obtain the change in uninsurance rate compared to baseline. We find that a reduction in the minimum consumption floor from $5,000 in the benchmark case to only $100 leads to a substantial decline in the uninsurance rate from 21.9 percent to 16.7 percent. Similarly, prohibition of medical bankruptcy would reduce uninsurance rate to 19.7 percent. These results confirm that the minimum consumption floor and the option to declare medical bankruptcy can be used as implicit insurance against medical expenditure shocks and hence lower the incentive to buy insurance. In addition, high insurance markups are also relevant: Eliminating the markup in the health insurance market would reduce uninsurance rate to 19.7 percent, a reduction of 2.4 percentage points. Finally, making EHI available to all workers substantially reduces uninsurance rate to 7.8 percent. These experiments suggest that insurance affordability is another major reason why so many workers stay uninsured. Furthermore, EHI expansion is associated with additional friction: When an employer offers EHI, all worker salaries are adjusted regardless of whether he decides to take up the EHI offer or not. This creates additional incentive for individuals to take up the EHI offer.

Note that an individual's income level has a large effect on insurance choice, and implicit insurance mechanisms, such as the consumption floor and medical bankruptcy, mainly appeal to low income individuals. In addition, the consumption floor and medical bankruptcy are alternatives of implicit insurance mechanisms. If we reduce the consumption floor to $100, the

Table 3: Decomposition of Uninsurance

Setup	Percentage of Uninsured	Changes in Percentage Points
Full Model	21.9	—
Lower Consumption Floor	16.7	−5.2
No Medical Bankruptcy	19.7	−2.2
No Markup	19.5	−2.4
EHI Expansion	7.8	−14.1

Note: This table compares the model's simulated uninsurance rate under different setups. We start from the full model with all the mechanisms of uninsurance, and then make changes one at a time to study their impact on uninsurance.

medical bankruptcy rate of the lowest income quantile (Q1) would increase substantially from less than two percent in the benchmark economy to more than 20 percent.

5 Quantitative Analysis

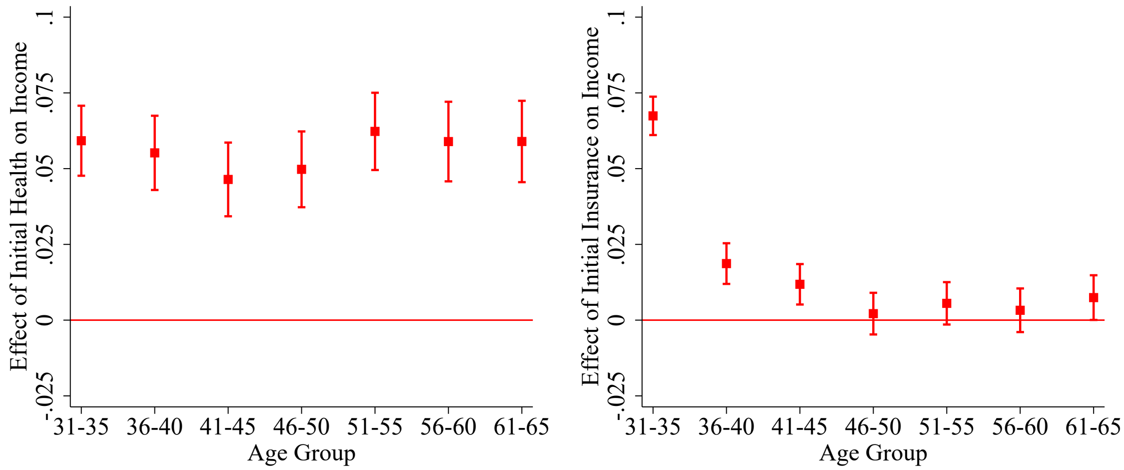
5.1 Determinants of health and income distribution

In this section, we analyze the impact of an individual's initial condition on their income and health status at later stages of the life cycle. We study which of the agent types (in terms of initial condition) are more likely to stay in certain regions of income-health distribution (poor health, low income, low wealth). Our model is particularly suitable for performing this task since we allow for a rich interaction between health and income. For instance, we can study how the initial income status affects one's health status later in life, a question that has been overlooked by the standard macro health literature which treats health as an exogenous process.

To perform these studies, we simulate a long panel data using the equilibrium decision rules from the model for different types of agents, given a carefully chosen initial distribution at the age of 25 when agents enter the economy. In particular, we choose the initial distribution of agent income as estimated directly from the data. We then determine the initial distribution of EHI offer, health insurance coverage status, and health status (all of which are determined in the previous period before an individual enters the economy) such that their correlation with income matches the data.

Given that the model replicates the joint distribution of health and income reasonably well, the results from analyzing the simulated data should also be comparable to those from the

Figure 10: Determinants of Lifetime Income



Note: The figure shows the effect of initial health (left panel) and initial health insurance status (right panel) on future income, estimated using model simulated panel data. The regression is specified in Equation (29). The red squares represent the point estimates, and the bands show the 95 percent confidence interval.

actual data. Moreover, using our simulated data to study the determinants of the health-income distribution has several advantages. We can simulate a panel which is much longer in time dimension compared to the survey data, with no attrition bias. Furthermore, we have better control over unobserved heterogeneity in the simulated data.

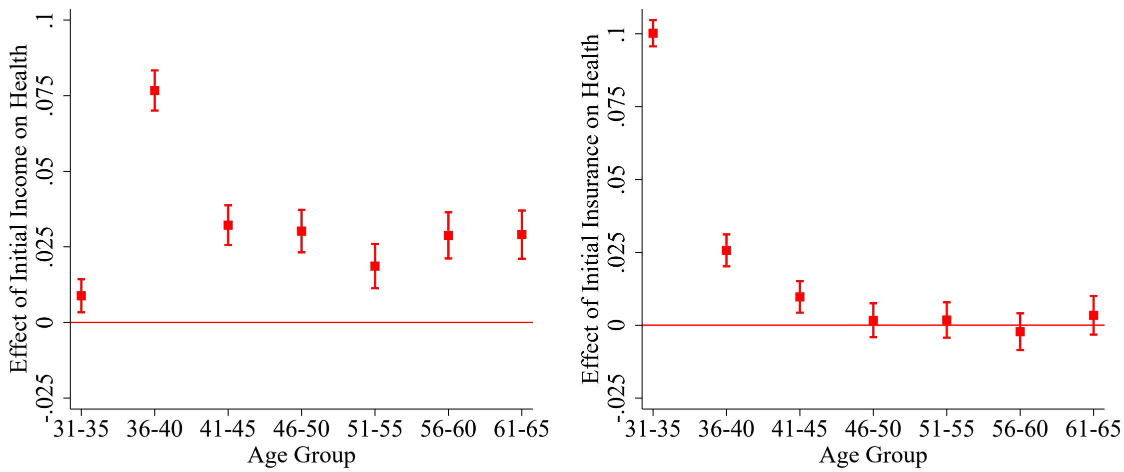
We start by investigating the impact of initial conditions on income at later stages of life by considering the following regression:

$$e_j^i = \gamma_h h_1^i + \gamma_e e_1^i + \gamma_{hi} i_1^{hi} + \sum_k \psi_k \mathbb{1}[(\alpha^i, \sigma^i) = (\alpha_k, \sigma_k)] + u_i^j, \tag{29}$$

where e_j^i is our variable of interest—log labor income of individual i including the health component $g(h)$ at age $j \geqslant 2$; h_1^i, e_1^i, and i_1^{hi} are initial health status, initial income, and initial insurance status, respectively. We also control for fixed effects of individual types (α, σ) through $\sum_k \psi_k \mathbb{1}[(\alpha^i, \sigma^i) = (\alpha_k, \sigma_k)]$.

The left panel of Figure 10 reports how simulated initial health (h_1) affects lifetime income for $j \geqslant 2$. We find that initial health affects income persistently through the life cycle. If an individual entering the economy at age 25 is healthy, her income would be around six percent higher at age 31–35 than if she had entered unhealthy, all else equal. The effect is significant at five percent level, albeit the magnitude tends to be smaller compared with the contemporaneous

Figure 11: Determinants of Lifetime Health



Note: The figure shows the effect of initial income (left panel) and initial health insurance status (right panel) on future health, estimated using model simulated panel data. The linear probability model is specified in Equation (29). The red squares represent point estimates, and the bands show the 95 percent confidence interval.

effect (which is over 30 percent). Furthermore, this effect lingers around: It remains significant even at the age of 61–65. The persistence of such impact can be attributed to the following factors. First, the health process itself is persistent as we documented in Section 2; see also de Nardi et al. (2017). Second, good initial health leads to higher labor productivity and income. This individual is more likely to obtain insurance coverage which improves her chance of staying healthy. This new channel amplifies the persistence of the impact of health on income.

To further elaborate on this point, we examine the effect of initial insurance status, which is represented by the estimated γ in Equation (29), with results shown on the right panel of Figure 10. An individual experiences roughly six percent higher income at age 31–35 if she begins with health insurance coverage at age 25, compared to an otherwise identical individual without insurance initially. This is because the initial insurance coverage makes her more likely to have good health later in the next period, and hence a higher income. Note again that this effect is persistent: We find significant effects even 15 years later at age 41–45.[17]

Next, to study the determinants of lifetime health, we consider a linear probability model, specified similarly to Equation (29) with health as the dependent variable, to study the deter-

[17]Note that we control for individual type fixed effects (α^i, σ^i) which is the most important component of the income process z_t. As a result, we do not study how initial income affects future income: Once we control for type fixed effects, the residuals of the income process is mainly the transitory component which is mean reversal, and hence the estimated effect of initial income on future income will always be negative.

minants of lifetime health. The left panel of Figure 11 shows that initial income has persistent effects on future health, and the effects remain significant even after 30 years. An individual starting out with one percent higher income at age 25 arising completely from the transitory component or "luck", as we have explicitly controlled for health and type fixed effects, is around 0.08 percent more likely to be healthy at age 36–40. Stated differently, if the worker's initial income doubles due to transitory shocks, her chance of remaining healthy increases by eight percent and her mortality rate drops by over one percentage point age 40. A higher likelihood of being healthy then implies that this individual will have higher income at that age, and is more likely to have insurance coverage later. This is why the initial transitory shock to income may have persistent effects. Conversely, in the literature's canonical models, initial income shocks are unable to impact future health since health is treated as exogenous. Note that for the next period, age 31–35, the magnitude is the smallest overall. This is because the health status of age 31–35 is largely determined by initial insurance status, which is not a choice variable but is exogenously given when individuals enter the economy at age 25.

Again we can better illustrate this point by studying the effect of initial insurance status on future health. Having health insurance initially means that an individual is around 10 percent more likely to remain healthy, reflecting precisely the "health premium" documented in Section 2. This effect is again persistent, significantly affecting health even 15 years later at age 41–45.

Finally, we determine the agent types that are most likely to end up in the bottom quantile of the income distribution at age 36–40, more than 10 years after entering the economy. For example, if one enters the economy as a healthy individual, the probability of being in the bottom quantile of the income distribution at age 36–40 is 19.9 percent, compared to 21.2 percent for starting off unhealthy. Similarly, this probability is 18.1 percent for entering with health insurance coverage, which is substantially lower than 24.4 percent for being without insurance initially. This comparison again confirms our previous results that initial insurance coverage has persistent effects on lifetime income even after more than 10 years. Note that this comparison differs from the previous regression analysis, as the regression analysis controls for other characteristics and hence compares two otherwise identical individuals who differ solely in one dimension, say, initial insurance coverage. In this analysis, however, individuals with initial insurance coverage likely differ in other dimensions, such as having higher initial income arising from a favorable individual fixed effect (α_i, σ_i^2), compared to those without coverage, and hence are less likely to fall into the bottom quantile of the income distribution at age 36–40 for reasons other than the initial insurance.

5.2 Improving insurance coverage vs. alleviating income disparity

The analysis in the previous section indicates that both initial health and initial income have significant and persistent impacts on health and economic outcome. In this section, we consider two distinct policy experiments in order to understand how we should design policy to mitigate the long lasting negative impact of health inequality. In the first experiment, the government implements "universal health coverage" (UHC). In the second experiment, we introduce a "universal basic income" (UBI). Both experiments are financed by increasing consumption tax.[18] For comparability between the two experiments, we first determine the additional consumption tax needed to finance "universal health coverage" and we raise the consumption tax by the same amount under "universal basic income" to distribute as lump sum transfers.

5.2.1 Level effects

With endogenous health transition and health-dependent labor productivity, these experiments generate level effects on aggregate demographic variables and human capital supply. As displayed in Table 4, the proportion of healthy individuals increases to 81.7 under universal health coverage, compared with 78.6 in the baseline economy. As a result, mortality declines, and the total population increases by 2.2 percent. Human capital, measured by efficiency units of labor $\tilde{z}$, which is the product of underlying productivity and health component $g(h)$ for the working age population, increases by 2.5 percent on average. Aggregate human capital increases by 4.7 percent, due to the increase in average human capital together with the increase in population. We highlight that all these changes arises from the channel that improved insurance coverage promotes individual health, a mechanism that is novel in our framework. With the implementation of universal basic income, we find qualitatively similar results, albeit of a smaller magnitude. UBI affects individual health only indirectly through the affordability channel that allows some individuals to buy health insurance who would have chosen to be uninsured in the baseline.

5.2.2 Joint distribution of health and income

Next, we analyze how these policies affect the joint distribution of health and income. We start with their impact on individual health insurance choice. Clearly, all individuals are insured un-

[18]We find that our results below are robust to how policy changes are financed. To avoid the discussion on distortion caused by income tax, we focus on financing with consumption tax.

Table 4: Level Effects

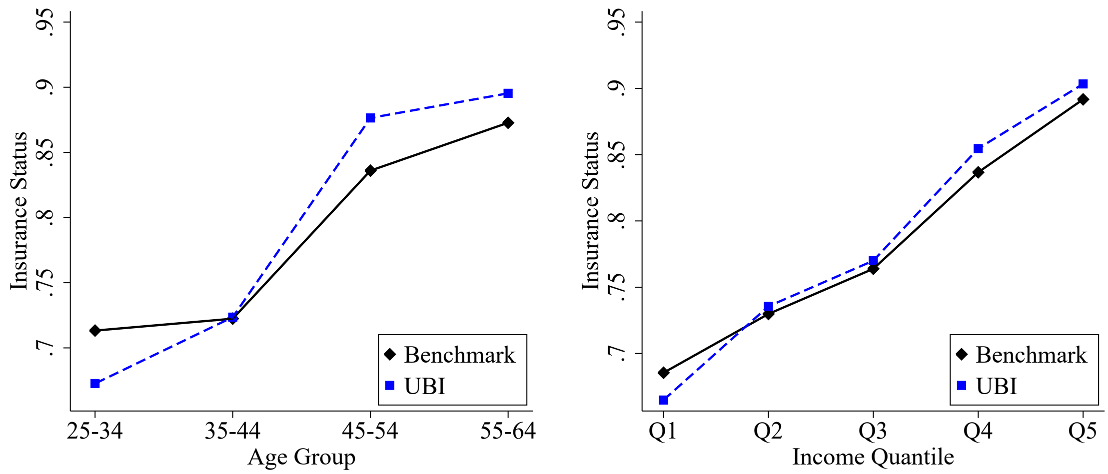
	Universal Health Coverage	Universal Basic Income
	(% Change Compared to Baseline)	
Healthy Individuals	+3.12	+0.22
Population	+2.24	+0.26
Human Capital (Efficiency Units of Labor)		
Aggregate	+4.73	+1.61
Average	+2.43	+1.35

Note: This table shows the outcome of two experiments, universal health coverage and universal basic income, as percentage changes compared to the baseline outcome. Human capital is measured by efficiency units of labor $\tilde{z}$, which is the product of underlying productivity and health component $g(h)$ for the working age population.

der UHC. Figure 12 presents the insurance coverage under UBI. The left panel shows insurance coverage by age, while the right panel shows it by income quantiles. Health insurance coverage apparently drops among young individuals, especially for the poorest income quantile. This is mainly because UBI increases income for individuals, and hence many of them no longer qualify for the means-tested insurance program, Medicaid. If we focus on private insurance and EHI, then under UBI more individuals choose to pick up EHI or buy private insurance, especially among the poor. For instance, for the poorest quantile (Q1), the private and EHI insurance rate increases substantially from 43.3% in the baseline to 48.9% with UBI. Intuitively, UBI helps the poor to afford private insurance and hence improve the coverage rate. Although UBI does not directly target the health insurance market, it affects health insurance coverage indirectly through affordability.

We now discuss how the joint distribution of income and health changes with these two experiments. Figure 13 shows the distribution of health by age and by income. Under UHC, individuals are healthier across all age and income groups, since everyone is covered by health insurance which is beneficial to maintaining health, but more so for poor individuals than rich ones, as rich individuals are more likely to have insurance coverage in the baseline economy. The effect of UBI on insurance coverage is less straightforward. Since many young individuals no longer qualify for Medicaid due to higher income from UBI, young individuals are now less likely to be insured overall. As a result, the health transition of young individuals is negatively affected and hence the health of those aged 35–44 deteriorates. For older individuals, the insured rate is higher than that of the baseline economy and hence the health status is generally better. We can also investigate the effect of UBI on health by income groups. For the poorest quantile,

Figure 12: Insurance Coverage under UBI



Note: The figure shows our UBI experiment's insurance coverage rates by age group (left panel) and income quantile (right panel).

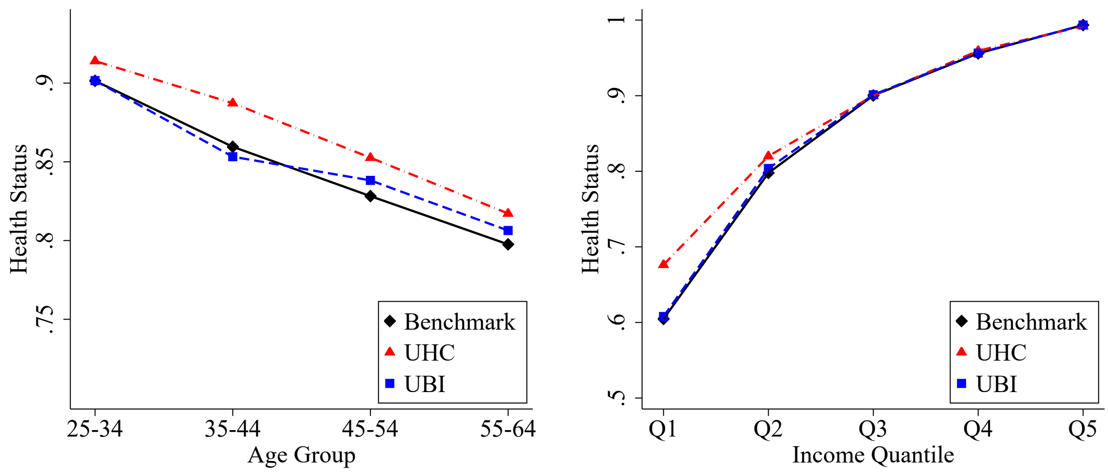
since some no longer qualify for Medicaid while at the same time more individuals choose to pick up private insurance, the overall healthy percentage is largely unchanged. For the middle income quantiles, the channel of Medicaid qualification does not apply while private insurance take up rate increases, and hence healthy rates increase. For the richest quantile, the rate barely changes as their healthy rate is high already in the baseline. Again, although UBI is not directly affecting health, it does so indirectly through improving the affordability of the health insurance market, especially for the poor.

5.2.3 Income inequality

It is well documented in the literature that health disparity and income inequality are closely related. Given that these two policy experiments affect the joint distribution of health and income, we now investigate how income inequality responds to these two policy experiments.

Figure 14 presents income inequality under these two policy experiments as well as in the baseline economy. Note that income inequality is measured by the income share of individuals within the top 25 percent of the income distribution (the rich) and that of the bottom 50 percent of distribution (the poor). Based on Figure 14, UBI leads to a substantial drop in income share for the rich and a rising share for the poor, while UHC has little impact on income inequality, as the income shares of the rich and poor remain largely unchanged. In the previous section, however, we find that the health of the poor improves substantially under UHC. This

Figure 13: The Distribution of Health under Two Experiments



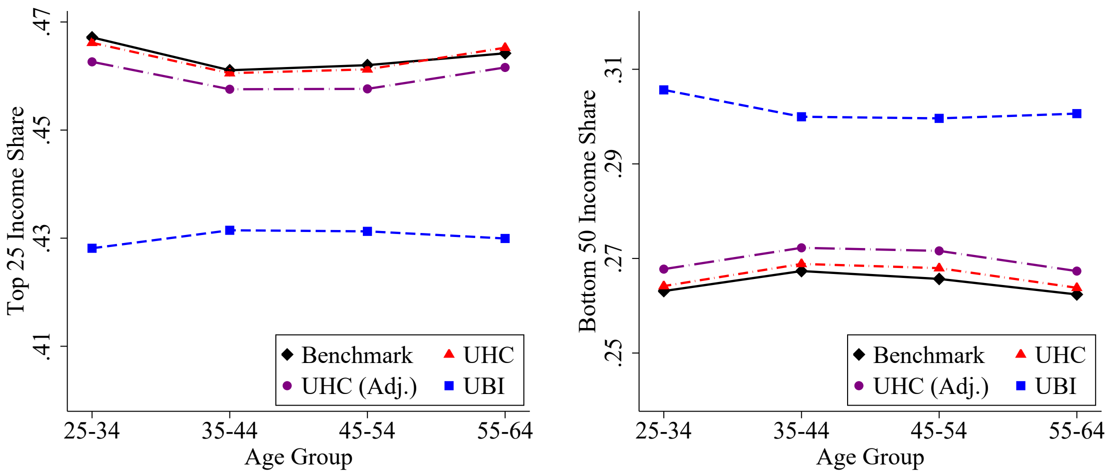
Note: The figure shows health status in our two experiments (UHC and UBI) by age group (left panel) and by income quantile (right panel).

health-income channel implies that the level of inequality should decline, as the poor's income increases with their health. The key to this seemly puzzling outcome is the change in mortality rate. Since UHC improves the health status of the poor, their mortality rate declines. More poor people survive in equilibrium, which increases the level of inequality. This mortality channel almost completely offsets the previously mentioned health-income channel, and hence the equilibrium inequality level is largely unchanged. Figure 15 suggests that the mortality rate declines substantially, especially for the individuals from the poorest 50 percent of the income distribution.

To further elaborate this aforementioned mortality channel, we do a back-of-the-envelope calculation for a hypothetical level of inequality assuming the mortality rate remains unchanged as in the baseline economy. In particular, we start from the invariant distributions after policy change. Then we adjust the distributions to mimic the mortality rate seen in the baseline economy. For example, for the poorest quantile of individuals, their mortality rate is around one percent higher in the baseline economy than in UHC. We hence reduce the sample weights of the poorest quantile by one percent, so that the composition is the same as in the baseline economy. Using this adjusted distribution, we then calculate the inequality level and report it in Figure 14 with the purple longdashed line. With this adjustment, UHC reduces the level of inequality compared to the baseline economy.

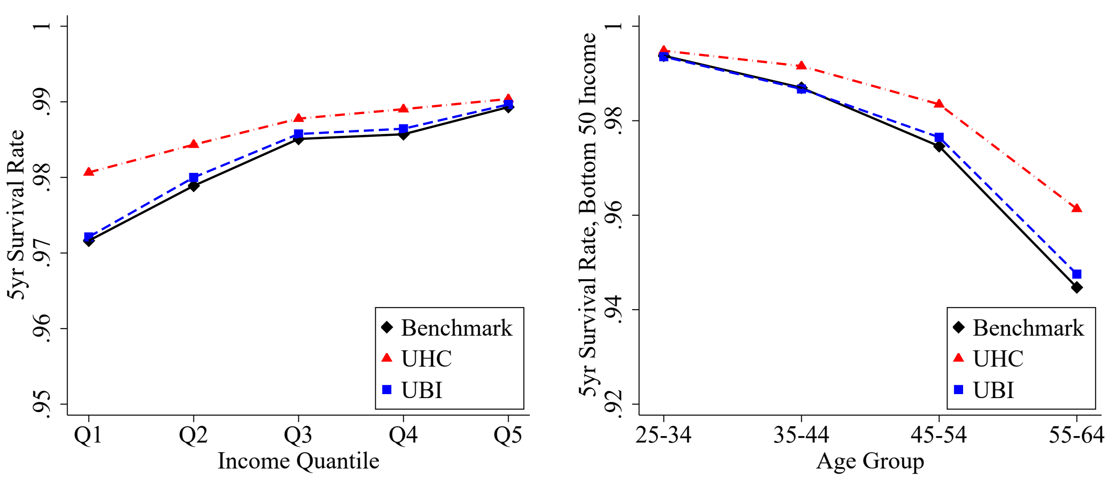
Lastly, we remark on the tension of intervening in the health insurance market. While pro-

Figure 14: Mortality under Two Experiments



Note: The figure shows the income share of top 25 percent earners (left panel) and bottom 50 percent earners (right panel) for our two experiments (UHC and UBI). Note that mortality rate changes substantially under UHC. For the curve "UHC (Adj.)", we calculate a hypothetical level of inequality by assuming that the mortality rate remains unchanged from the benchmark economy.

Figure 15: Mortality under Two Experiments



Note: The figure shows total 5-year survival rates by income quantile for all individuals (left panel) and by age group for bottom 50 percent earners (right panel) for our two experiments (UHC and UBI).

viding UHC clearly eliminates some frictions in the health insurance market (adverse selection), it does distorts individual inter-temporal choice: As UHC increases the life expectancy of the poor, it also effectively changes their discount factor and raises the new concern of whether the poor have enough savings for their longer lifespan. If given the choice, the poor would be better off by liquidating the provided insurance coverage in exchange for more consumption. In this sense, universal health coverage could be viewed as a forced saving mechanism, requiring the poor to save (in a special form of human capital, i.e., health). The welfare impact is hence complicated to evaluate and we leave it for future work.

6 Concluding Remarks

In this paper, we study the determination of the joint distribution of health and income. We develop a quantitative framework that encompasses endogenous evolution of health status and health-dependent labor productivity. Our baseline economy reproduces the observed joint distribution of health, health insurance, medical expenditure, and income over the life cycle. These results rest on a rich dynamic equilibrium model featuring a novel channel through which income affects health, and on econometric methods that provide consistent estimation of the income and health processes. We use our model to evaluate the factors affecting household choice of health insurance. Our study confirms that affordability and implicit insurance mechanisms such as medical bankruptcy are the main reasons why a large fraction of Americans stay uninsured. We also simulate the model to identify the defining factors that affect health and economic outcome in adulthood. Quantitative analysis reveals the significant impact of an individual's early life income on their health in adulthood, which is reinforced by and subsequently amplifies the effect of health on labor earnings and income inequality. We conduct comparative analysis of health policies intended to alleviate health disparity and income inequality. Providing "Universal Health Coverage" would narrow health and life expectancy gaps, benefiting low-skilled workers the most. However, such policy has an ambiguous overall effect on income inequality, as its reduction in mortality of the poor and the resultant increase of their weight in the income distribution almost completely offsets the positive effect of improved health on income inequality. In contrast, while "Universal Basic Income" would reduce income inequality, means-tested public insurance would attenuate its effect on health disparity. Our comparative analysis also highlights a trade-off between i) distorting individual inter-temporal decisions when we directly intervene in health insurance coverage, and ii) leaving aside the frictions in the health insurance market when we indirectly intervene through income redistribution.

There are a number of potentially interesting extensions. Given the availability of data for health, health investment, and the tractability of the model, we focus on the impact of income on health through endogenous health insurance choice. Explicitly modeling other resources or effort invested by households to impact the evolution of health would likely make the interaction between income and health more dynamic. For instance, in addition to the fact that higher income individuals are more likely to obtain insurance coverage as in our model, they are also more likely to lead a healthier lifestyle, such as afford more nutritious food and gym memberships. Our results quantifying the interaction between health and income should therefore be interpreted as a lower bound as modeling additional channels would likely further strengthen this interaction. Secondly, we find that both the worker's initial health and income of workers have significant and long-lasting impact on the evolution of their health over the life cycle. The current study is silent on estimating the cost of providing better health. An answer to this question would allow us to find the optimal policy intervention that improves the distribution of health.

References

Aiyagari, S. R. (1994). Uninsured idiosyncratic risk and aggregate saving. *Quarterly Journal of Economics*, 109(3):659–684.

Aizawa, N. and Fang, H. (2019). Equilibrium labor market search and health insurance reform. *Working Paper*.

Angrist, J. D. and Pischke, J.-S. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Attanasio, O., Kitao, S., and Violante, G. L. (2011). *Financing Medicare: A General Equilibrium Analysis*. University of Chicago Press.

Braun, R. A., Kopecky, K. A., and Koreshkova, T. (2017). Old, sick, alone and poor: A welfare analysis of old-age social insurance programs. *Review of Economic Studies*, 84(2):580–612.

Bruegemann, B. and Manovskii, I. (2010). Fragility: A quantitative analysis of the U.S. health insurance system. *Working Paper*.

Card, D., Dobkin, C., and Maestas, N. (2009). Does medicare save lives? *Quarterly Journal of Economics*, 124(2):597–636.

Card, D. and Shore-Sheppard, L. D. (2004). Using discontinuous eligibility rules to identify the effects of the federal medicaid expansions on low-income children. *Review of Economics and Statistics*, 86(3):752–766.

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., and Cutler, D. (2016). The Association Between Income and Life Expectancy in the United States, 2001–2014. *JAMA*, 315(16):1750–1766.

Cole, H. L., Kim, S., and Krueger, D. (2019). Analyzing the effects of insuring health risks. *Review of Economic Studies*, 86(3):1123–1169.

Currie, J. and Gruber, J. (1996). Health Insurance Eligibility, Utilization of Medical Care, and Child Health*. *The Quarterly Journal of Economics*, 111(2):431–466.

Cutler, D. M. and Reber, S. J. (1998). Paying for health insurance: The trade-off between competition and adverse selection. *Quarterly Journal of Economics*, 113(2):433–466.

de Nardi, M., Pashchenko, S., and Porapakkarm, P. (2017). The lifetime costs of bad health. *NBER Working Paper 23963*.

Deaton, A. (2003). Health, inequality, and economic development. *Journal of Economic Literature*, 41(1):113–158.

Doyle, J. J. (2005). Health insurance, treatment and outcomes: Using auto accidents as health shocks. *Review of Economics and Statistics*, 87(2):256–270.

Feng, Z. (2010). Macroeconomic analysis of universal coverage in the U.S. In Chai, S.-K., Salerno, J. J., and Mabry, P. L., editors, *Advances in Social Computing*, pages 87–96. Springer, New York.

Feng, Z. and Zhao, K. (2018). Employer-based health insurance and aggregate labor supply. *Journal of Economic Behavior and Organization*, 154:156–174.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K., and Group, O. H. S. (2012). The Oregon Health Insurance Experiment: Evidence from the First Year*. *The Quarterly Journal of Economics*, 127(3):1057–1106.

Flölich, M. and Huber, M. (2017). Direct and indirect treatment effects: causal chains and mediation analysis with instrumental variables. *Journal of the Royal Statistical Society: Series B*, 79(5):1645–1666.

French, E. and Jones, J. B. (2004). On the distribution and dynamics of health care costs. *Journal of Applied Econometrics*, 19(6):705–721.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2):223–255.

Gruber, J. (2008). Covering the uninsured in the united states. *Journal of Economic Literature*, 46(3):571–606.

Gu, J. and Koenker, R. (2017). Unobserved heterogeneity in income dynamics: An empirical bayes perspective. *Journal of Business and Economic Statistics*, 35:1–16.

Guvenen, F. (2009). An empirical investigation of labor income processes. *Review of Economic Dynamics*, 12(1):58–79.

Hall, R. E. and Jones, C. I. (2007). The value of life and the rise in health spending. *Quarterly Journal of Economics*, 122(1):39–72.

Hansen, G. D., Hsu, M., and Lee, J. (2014). Health insurance reform: The impact of a Medicare buy-in. *Journal of Economic Dynamics and Control*, 45:315 – 329.

Heathcote, J., Storesletten, K., and Violante, G. (2017). Optimal tax progressivity: An analytical framework. *Quarterly Journal of Economics*, 132(4):1693–1754.

Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica*, 52:63–132.

Herring, B. (2005). The effect of the availability of charity care to the uninsured on the demand for private health insurance. *Journal of Health Economics*, 24(2):225–252.

Hong, J. H., Pijoan-Mas, J., and Rios-Rull, J.-V. (2017). Health heterogeneity and the preferences for consumption growth. *mimeo*.

Hospido, L. (2012). Modelling heterogeneity and dynamics in the volatility of individual wages. *Journal of Applied Econometrics*, 27:386–411.

Hosseini, R., Kopecky, K., and Zhao, K. (2019). How important is health inequality for lifetime earnings inequality? *Working Paper*.

Hubbard, R. G., Skinner, J., and Zeldes, S. P. (1995). Precautionary saving and social insurance. *Journal of Political Economy*, 103(2):360–399.

Jeske, K. and Kitao, S. (2009). U.S. tax policy and health insurance demand: Can a regressive policy improve welfare? *Journal of Monetary Economics*, 56(2):210 – 221.

Jung, J. and Tran, C. (2016). Market inefficiency, insurance mandate and welfare: U.S. health care reform 2010. *Review of Economic Dynamics*, 20:132–159.

Kiefer, J. and Wolfowitz, J. (1956). Consistency of the maximum likelihood estimator in the presence of infinitely many incidental parameters. *The Annals of Mathematical Statistics*, 27:887–906.

Levy, H. and Meltzer, D. (2004). *What Do We Really Know about Whether Health Insurance Affects Health*. The Urban Insitute.

Lindsay, B. (1983). The geometry of mixture likelihoods: A general theory. *Annals of Statistics*, 11:86–94.

Livshits, I., MacGee, J., and Tertilt, M. (2007). Consumer bankruptcy: A fresh start. *American Economic Review*, 97(1):402–418.

Mahoney, N. (2015). Bankruptcy as implicit health insurance. *American Economic Review*, 105(2):710–746.

Mattke, S., Balakrishnan, A., Bergamo, G., and Newberry, S. (2007). A review of methods to measure health-related productivity loss. *American Journal of Managed Care*, 13(4):211–217.

Meghir, C. and Pistaferri, L. (2004). Income variance dynamics and heterogeneity. *Econometrica*, 72:1–32.

Mendoza, E. G., Razin, A., and Tesar, L. L. (1994). Effective tax rates in macroeconomics. *Journal of Monetary Economics*, 34(3):297 – 323.

Ozkan, S. (2014). Preventive vs. curative medicine: A macroeconomic analysis of health care over the life cycle. *Working Paper*.

Pashchenko, S. and Porapakkarm, P. (2013). Quantitative analysis of health insurance reform: Separating regulation from redistribution. *Review of Economic Dynamics*, 16(3):383–404.

Pashchenko, S. and Porapakkarm, P. (2017). Work incentives of Medicaid beneficiaries and the role of asset testing. *International Economic Review*, 58(4):1117–1154.

Prados, M. J. (2017). Health and earnings inequality over the life cycle: The redistributive potential of health policies. *Working Paper*.

Suen, R. M. H. (2006). Technological advance and the growth in health care spending. *Working Paper*.

Ward, E., Halpern, M., Schrag, N., Cokkinides, V., DeSantis, C., Bandi, P., Siegel, R., Stewart, A., and Jemal, A. (2008). Association of insurance with cancer care utilization and outcomes. *CA: A Cancer Journal for Clinicians*, 58(1):9–31.

Online Appendix

A Income Process

A.1 MEPS data

The MEPS collects detailed records on demographics, income, medical costs and insurance for a nationally representative sample of households. It consists of two-year overlapping panels and covers the period from 1996 to 2008. For each wave, each person is interviewed five rounds over the two years. We use 13 waves of the MEPS (2002-2014).

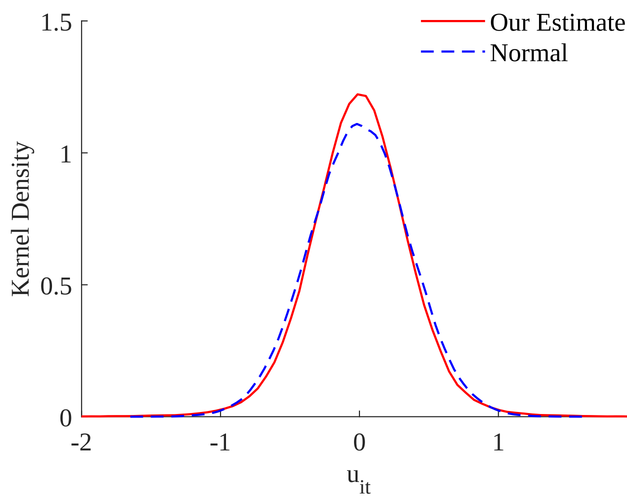
We use the cross-sectional weights and longitudinal weights provided by the MEPS for the cross-sectional and longitudinal pools correspondingly. Since each wave is a representation of the population in that year, when pooling several years (or waves) together the weight of each individual was divided by the number of years (or waves). Note that nominal variables (monetary variables) from different sample years are not directly comparable due to inflation and economic growth. We therefore deflate all nominal variables by the nominal GDP per capita of each year. This normalization not only controls for both inflation and economic growth, but also makes those variables unit free and therefore comparable to model simulations.

The MEPS links people into one household based on eligibility for coverage under a typical family insurance plan. This Health Insurance Eligibility Unit (HIEU) defined in the MEPS data set corresponds to our definition of a household. In our sample we include only the heads of the HIEU. We define the head as the person with the highest income in the HIEU.

A.2 PSID data

We use the PSID data covering 1983–1997, during which information on health status and employment status of the household heads are collected on an annual basis. The main dependent variable is log real earnings, which includes labor income of the household head from all sources and is deflated by the nominal per capita GDP. We distinguish individual households by the health status of household head, which is a dummy variable based on the five levels of self-reported health status (with 1 being excellent and 5 being poor, we define the health dummy to take value one if the reported health status is 1 or 2). The employment status is a categorical variable with 9 possible values (with 1 being working, 2, temporally laid off with a job to return to; 3 as unemployed; 4 as retired; 5 as disabled; 6 as Housewife; 7 as student, 8 as other

Figure 16: Kernel Density of u_{it}



Note: This figure plots the kernel density estimates of u_{it}, defined as $\hat{\varepsilon}_{it} - \hat{\rho}\hat{\varepsilon}_{i,t-1}$ (see Equation 25) in the red solid line, and contrasts it with the kernel density of a normal distribution with the same variance plotted in the blue dashed line.

workfare, prison or jail; and 9 as missing).[19] We include only individuals aged 20 to 64 that have more than 9 consecutive labor income records and who are reported as working in these periods.[20] These sample selection criteria leads to a total of 2052 individual households and 25622 individual-year observations.

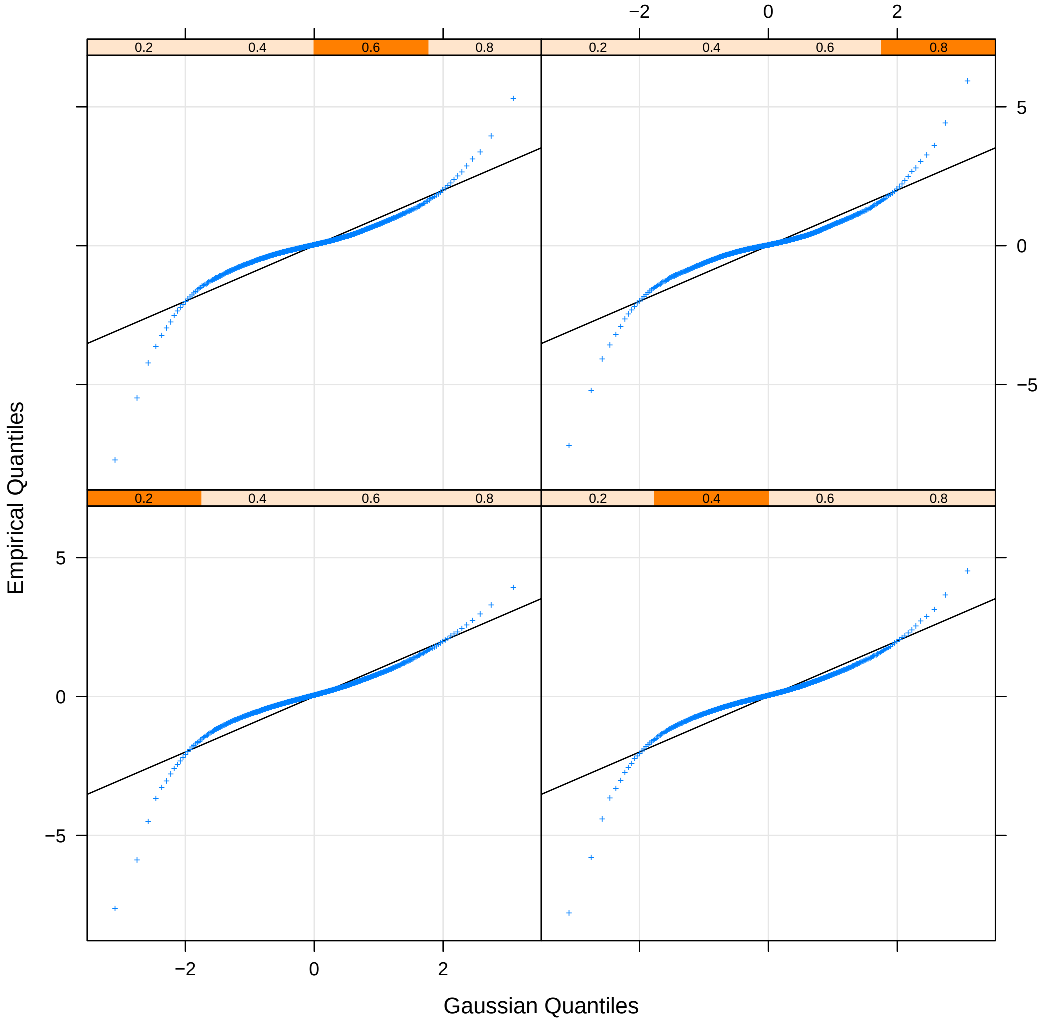
A.3 Estimating the income process z_{it}

Note that, as we described in Section 4, the residual term u_{it} in Equation (25) cannot be well-approximated by a Normal distribution. Instead of imposing normality on the residual process, we use a flexible scale mixture of Normal distributions which allows for heterogeneity in the variance of u_{it} among individuals. This can be seen in the comparison in Figure 16 of the kernel densities with a single Normal distribution. Figure 17 provides further evidence that the marginal distribution of demeaned u_{it} (for different values of ρ) deviate significantly from a normal distribution and present long tail behavior.

[19] In year 1997, the survey question for employment status for household head changes from one question to three sub-questions, which allows individuals to pick three possible answers. We use the primary answer to determine the individual's employment status.

[20] There are a few labor income record that are abnormally low. Hence we discard individuals who report to be employed throughout the year but have an annual labor income below 200 dollar. Only one individual is discarded by this criteria.

Figure 17: QQ Plots



Note: For different values of $\rho \in \{0.2, 0.4, 0.6, 0.8\}$, we plot the empirical quantile of $u_{it} = \hat{\varepsilon}_{it} - \hat{\rho}\hat{\varepsilon}_{it-1}$ against the Gaussian quantile.

We estimate the income process specified in Equation (25) using non-parametric maximum likelihood (NPML). The details are provided as the following. Denote $z_{it} = \epsilon_{it} - \rho\epsilon_{it-1}$, It is easy to see that the sample mean of z_{it}, denoted as $\bar{z}_i = \frac{1}{T_i}\sum_{t=1}^{T_i} z_{it}$, is sufficient statistics for α_i and the sample variance, denoted as $s_i = \frac{1}{T_i-1}\sum_{t=1}^{T_i}(z_{it} - \bar{z}_i)^2$, is sufficient statistics for σ_i^2. Then conditional on α_i and σ_i^2, we have

$$\bar{z}_i \sim \mathcal{N}(\alpha_i, \sigma_i^2/T_i)$$
$$(T_i - 1)s_i/\sigma_i^2 \sim \chi^2_{T_i-1}$$

Combined with the assumption that $(\alpha_i, \sigma_i^2) \sim_{iid} F$ allows us to write down the likelihood function for $(\bar{z}_i, s_i)$

$$L(\bar{z}_i, s_i) = \int\int f(\bar{z}_i|\alpha_i, \sigma_i^2)g(s_i|\sigma_i^2)dF(\alpha_i, \sigma_i^2)$$

The NPMLE for F_ρ is then defined as

$$\hat{F}_\rho = \underset{F\in\mathcal{F}}{\operatorname{argmax}} \prod_{i=1}^{n} \int\int f(\bar{z}_i|\alpha_i, \sigma_i^2)g(s_i|\sigma_i^2)dF(\alpha_i, \sigma_i^2) \tag{30}$$

We index $\hat{F}$ by the parameter ρ because z_{it} depends on ρ, hence for each fixed value of ρ, we can solve for the optimization problem (30). Optimization (30) is a convex programming problem since the objective function is strictly concave and the set $\mathcal{F}$ which contains all probability distribution function is a convex set. Since F is an infinite dimensional object, to make the optimization problem feasible we fix a grid as a discretization for F. The profile likelihood function for ρ is then defined as

$$\ell(\rho, \hat{F}_\rho) = \sum_{i=1}^{n} \log \int\int f(\bar{z}_i|\alpha_i, \sigma_i^2)g(s_i|\sigma_i^2)d\hat{F}_\rho(\alpha_i, \sigma_i^2)$$

Maximizing that gives us the profile likelihood estimator for ρ, which peaks at $\rho = 0.25$ with our data.

Table 5 reports the joint distribution F of (α_i, σ_i^2) with $\rho = 0.25$. We only report the locations of (α_i, σ_i^2) that have their estimated probability mass being more than 1%[21]. We've normalized the reported probability in Table 5 to have them aggregate to 1.

[21] The estimation for $\hat{F}$ is done on a 40 by 40 grid on the support of (α_i, σ_i^2). Our of the 1600 grid points, only 19 grids have their estimated probability mass $\geqslant 1\%$ and their total probability adds up to be about 0.89 and this result does not change when we use a finer grid.

Table 5: Estimates for the Distribution F

		α_i												
		−0.7292	−0.5335	−0.4356	−0.3378	−0.2399	−0.1420	−0.0442	0.0537	0.1515	0.2494	0.3473	0.4451	0.5430
	0.0827	0	0.0273	0.0119	0.0848	0.0371	0.1198	0.0757	0.1166	0.0892	0.0481	0.0855	0.0143	0.0439
σ_i^2	0.1647	0.0224	0	0.0239	0	0	0	0	0	0	0	0	0.0163	0
	0.2467	0	0	0	0	0	0.0540	0	0	0	0	0	0	0
	0.6568	0	0.0263	0	0	0	0	0	0	0	0	0	0	0

Table 6: Distribution F: Reduced Dimension

		α_i							
		−0.6246	−0.5335	−0.3747	−0.1420	−0.1041	0.0961	0.1795	0.3911
	0.0827	0	0	0.1611	0	0.1955	0.2058	0	0.1918
σ_i^2	0.1647	0.0463	0	0	0	0	0	0.1193	
	0.2467	0	0	0	0.0540	0	0.0540	0	0
	0.6568	0	0.0263	0	0	0	0	0	0

There are in total 17 types of combinations of (α_i, σ_i^2), where the number of types is chosen by the estimation algorithm automatically. To reduce the dimension and facilitate computation, we further combine individuals with the same σ_i^2 into fewer types. For example, for individuals with $\sigma_i^2 = 0.0827$, we reduce the number of different α_i's from 13 to 4, denoted as $\tilde{\alpha}_i$. This can be viewed as an ex-post smoothing on the NPMLE of the distribution of (α_i, σ_i). To do so, we choose the second and third $\tilde{\alpha}_i$ to be -0.1041 and 0.0961, which are the weighted mean of the original α_i^5 and α_i^6 and the weighted mean of the original α_i^7 and α_i^8. We further choose the first and last $\tilde{\alpha}_i$ such that the mean and variance of u_{it} simulated from the the reduced 4 types of $(\tilde{\alpha}_i, \sigma_i^2 = 0.0827)$ are the same as those simulated from the original 13 types of $(\alpha_i, \sigma_i^2 = 0.0827)$. The weights are adjusted accordingly as the sum of original corresponding weights. Using a similar method, we smooth the distribution of α_i conditional on $\sigma_i^2 = 0.1647$ to have two mass points. Table 6 displays the smoothed distribution of (α_i, σ_i^2).

A.4 Simultaneity bias

As we explained in Section 4, an OLS estimation of Equation (25) can purge away the impact of health on income and obtain a residual term ε_{it} independent of health to estimate z_{it} as we described before, but it does not correctly estimate the impact of health on income. In particular, this OLS estimator of $g_j(h)$ suffers from simultaneity bias since the causality between income and health runs both ways and the residual in the equation of income follows an AR(1) process. To see the simultaneity bias, consider the following simplified system of equations:

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \varepsilon_{it}$$
$$X_{it} = \alpha_0 + \alpha_1 Y_{i,t-1} + v_{it}$$

where X is health status and Y is income.[22] The specification reflects that health status has a immediate influence on the income in the same period while the effect of income on health is through the health insurance and hence with a lag. The bias of the OLS estimates for β_1 is due to the fact that the residual ε is autocorrelated and the direction of the bias is determined by the autoregressive coefficient of ε_{it} as well as the magnitude of α_1 and β_1.[23]

A remedy, through the instrumental variable method, is to look for an exogenous variable Z that influences X but not directly on Y except through the link of X. We argue that the length of hospital stay in the current period serves as a reasonable instrument for X_t. Hospital stay directly correlates with health status in the current period, yet its correlation with income in current or lag periods is only through the health status.

A further complication in our case here is that since the health variable is binary, we can not directly apply the 2SLS due to the forbidden regression. We hence use a probit model to fit the first stage and then use the fitted value as instrument for health status. The results are included in Table 7 where we consider two specification of the hospital stay variable. The left panel includes a logarithm transformation of the days of hospital stay[24] while the right panel corresponds to result with a linear effect of the length of hospital stay. We prefer the first specification based on the goodness-of-fit. We include age dummies and year fixed effect as additional control variables for which the regression coefficients are omitted from the table. The results of the second stage regression using the fitted value from the first stage probit regression are displayed in Table 8. The magnitude of the effects of health status on income are similar in both specification and are significant for 1% level for both first stage specification. The last

[22] We have excluded other exogenous variables in the system of equations to make the algebra easier to be derived. When other exogenous variables are present in both equations, we can always purge their effect away before we reach to this simplified system of equations.

[23] The OLS bias takes the form $cov(X_{it}, \varepsilon_{it})/V(X_{it})$ and $cov(X_{it}, \varepsilon_{it}) = \rho E(X_{it}\varepsilon_{i,t-1}) = \rho(\alpha_1 V(y_{i,t-1}) - \beta_1 E(X_{it}X_{i,t-1}))$.

[24] Since many observations for the length of hospital stay take value zero, we add 1 to facilitate the log transformation.

column reports the OLS estimates. We again omit regression coefficient for additional control variables including age dummies and year fixed effects.

As discussed in the main text, since we only have data on the length of hospital stay for a shorter period of time, the coefficient of the effect of health status on income has to be estimated based on this subsample. In order to gain efficiency for the other regression coefficients in Equation (25), we run OLS regression of $y_{it} - 0.674h_{it}$ on the age dummies and year fixed effects on the full PSID sample, which then generates the earning residual $\hat{\varepsilon}_{it}$.

Table 7: First stage regression

	Dependent variable:	
	health	
	(1)	(2)
$log(1 + hospital)$	-0.342^{***}	
	(0.041)	
$hospital$		-0.047^{***}
		(0.009)
Observations	6,887	6,887
Log Likelihood	−3,950.976	−3,964.729
AIC	7,991.953	8,019.459

Table 8: Second stage regression

	Dependent variable:		
	$log(y_{it})$		
	instrumental variable		*OLS*
	(1)	(2)	(3)
h_{it}	0.674^{***}	0.552^{***}	0.329^{***}
	(0.154)	(0.178)	(0.017)
Observations	6,887	6,887	6,887
R^2	0.123	0.153	0.175

B The Effect of Insurance on Health

B.1 Estimating the Transition Probability

Recall the transition probability of health is specified as

$$\text{Prob}[h' = H|h, j] = \Phi(\beta_\pi i_{hi} + \gamma_\pi X_i), \quad h \in \{H, U\}.$$

The IV of group average EHI offer rate provides us with a consistent estimate of β_π for each age j and health status h. We cannot, however, recover the *level* of $\text{Prob}[h' = H|h, j, X_i = \bar{x}]$ directly since the estimation of γ_π is biased due to the endogeneity of income. Fortunately, what we focus on is the *difference* in the transition between the insured and the uninsured:

$$\Delta\text{Prob}[h' = H|h, j, X = \bar{x}] = \text{Prob}[h' = H|h, j, X = \bar{x}, i_{hi} = 1] - \text{Prob}[h' = H|h, j, X = \bar{x}, i_{hi} = 0], \tag{31}$$

which is exactly the "health premium" of insurance—individuals with insurance coverage are more likely to remain healthy or to recover from unhealthy status—evaluated for each age group at the mean. With the consistently estimated β_π, we are able to obtain this "health premium" without having a consistent estimate of γ_π.

To better illustrate the problem, let us consider a simple example of a linear model, i.e., $\text{Prob}[h' = H|h, j, x, i_{hi}] = \beta_\pi i_{hi} + \gamma_\pi X_i$, then we can easily obtain the health premium as

$$\begin{aligned}\Delta\text{Prob}[h' = H|h, j, X = \bar{x}] &= \text{Prob}[h' = H|h, j, X = \bar{x}, i_{hi} = 1] - \text{Prob}[h' = H|h, j, X = \bar{x}, i_{hi} = 0] \\ &= \beta_\pi + \gamma_\pi \bar{x} - \gamma_\pi \bar{x} = \beta.\end{aligned}$$

Since our model is non-linear due to the functional form of $\Phi(\cdot)$, we need to apply the following procedures to obtain the health premium. First, note that the health premium for individuals with health status h and age j is written as

$$\Delta\text{Prob}[h' = H|h, j, X = \bar{x}] = \Phi(\beta_\pi + \gamma_\pi \bar{x}) - \Phi(\gamma_\pi \bar{x}). \tag{32}$$

To address the issue that γ_π is not consistently estimated, we assume $\gamma_\pi \bar{x} = \tilde{x}_{h,j,X=\bar{x}}$ for age j

and health status h. We then have

$$\widehat{\Delta \mathrm{Prob}}[h' = H|h, j, X = \bar{x}] = \widehat{\mathrm{Prob}}[h' = H|h, j, X = \bar{x}, i_{hi} = 1] - \widehat{\mathrm{Prob}}[h' = H|h, j, X = \bar{x}, i_{hi} = 0]$$
$$= \Phi(\hat{\beta}_\pi + \tilde{x}_{h,j}) - \Phi(\tilde{x}_{h,j}). \tag{33}$$

The probability for an individual of age $j-1$ and health h whose health status becomes $h' = H$, unconditional on insurance status, is given by

$$\begin{aligned}
&\widehat{\mathrm{Prob}}[h' = H|h, j, X = \bar{x}] \\
&= \mathbb{P}[i_{hi} = 1|X = \bar{x}]\widehat{\mathrm{Prob}}[h' = H|h, j, X = \bar{x}, i_{hi} = 1] + \\
&\quad (1 - \mathbb{P}[i_{hi} = 1|X = \bar{x}])\widehat{\mathrm{Prob}}[h' = H|h, j, X = \bar{x}, i_{hi} = 0] \\
&= \mathbb{P}[i_{hi} = 1|X = \bar{x}]\Phi(\hat{\beta}_\pi + \tilde{x}_{h,j}) + (1 - \mathbb{P}[i_{hi} = 1|X = \bar{x}])\Phi(\tilde{x}_{h,j}).
\end{aligned} \tag{34}$$

Note that, among individuals of health h and age j, we can consistently estimate the unconditional probability ($\mathrm{Prob}[h' = H|h, j, X = \bar{x}]$) together with fraction of insured ($\mathbb{P}[x_{hi} = 1|X = \bar{x}]$) directly from the data. We can then uniquely determine $\tilde{x}_{h,j}$ using Equation (34). We then substitute the $\tilde{x}_{h,j}$ back into Equation (33) to obtain a consistent estimate of our health premium, with results illustrated in Figure 3. It is clear that insured individuals have significant and positive health premium, consistent with the evidence that having insurance coverage incurs more preventive care and hence improves health status (Ozkan, 2014). Note again that allowing for insurance status to affect health is also a reduced-form way to capture the health production function estimated in Hong et al. (2017).

We also use $\tilde{x}_{h,j}$ to evaluate $\Phi(\hat{\beta}_\pi + \tilde{x}_{h,j})$ and $\Phi(\tilde{x}_{h,j})$ which give us the level of $\pi^{j,1}_{hH}$ and $\pi^{j,0}_{hH}$ that are needed for calibration. We then repeat this process for all age group j and the associated transition probabilities are illustrated in Figure 4.

B.2 Mediation Effect

This appendix discusses the requirement on instrument to identify the mediation effect of income on health status through the mediator health insurance. We omit the inclusion of other exogeneous control variable for notation brevity and focus on the causal chain between income, health insurance and health status. Suppose we have the following causal model

$$
\begin{aligned}
T &= f_T(\varepsilon_T) \\
M &= f_M(T, \varepsilon_M) \\
Y &= f_Y(T, M, \varepsilon_Y)
\end{aligned}
$$

where T is income and M is health insurance choice and Y is health status. We are interested in identify the mediator effect of M on Y in a probit regression model

$$\mathbb{P}(Y = 1) = \Phi(\beta_1 M + \beta_2 T)$$

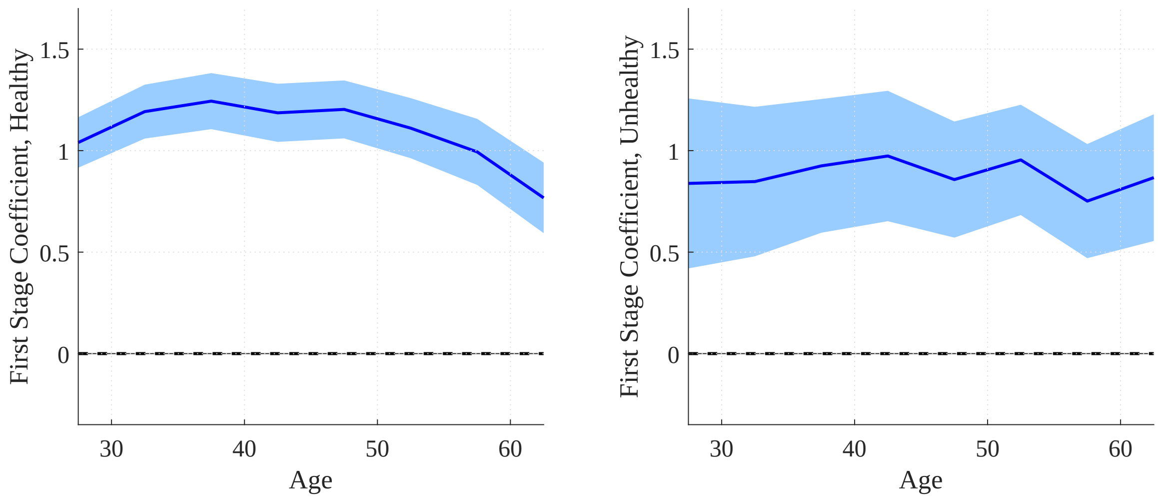
Since individuals' choices for health insurance are not randomly assigned, M is an endogeneous variable. We argue that if there exists an instrument Z which satisfies the following conditions: (i) $Z \perp \varepsilon_M | T$ and $Z \perp \varepsilon_Y | T$ and $\text{cov}(Z, M|T) \neq 0$, then we can identify β_1 and consistently estimate it using the bi-probit method with Z as the instrument for M. We should emphasize that this does not allow us to consistently estimate β_2 because the variable T is also endogeneous and we don't have a separate instrument for that. The intuitive way to understand why Z is a valid instrument for M under the specified assumptions is the following. There are only three types of pathways that induce association between two random variable: confounding, causal and colliding. Conditioning on T blocks all the confounding and causal paths that go through variable T, but may open pathways if T is a collider on the paths.[25] Since T is not on the causal pathway of interest from M to Y (because M is a mediator), we do not need to worry about blocking the causal pathway going through T. Conditioning on T will block all confounding pathways through T, which is what we want. We do not need to worry about opening up new pathway through which T is a collider because this case is ruled out by the assumption $Z \perp \varepsilon_Y | T$. [26]

Note that without a consistent estimator for β_2, we can not obtain a consistent estimator for $\mathbb{P}(Y = 1)$, however, it is possible to consistently estimate $\Phi^{-1}(\mathbb{P}(Y = 1|M = 1, T = t)) - \Phi^{-1}(\mathbb{P}(Y = 1|M = 0, T = t))$, that is the difference of the transformed probability of being healthy conditional on income level being t and the insurance status taking different

[25]The following graph representation illustrates a random variable being a collider in a causal chain. $A \rightarrow T \leftarrow B$. In this graph, T is a collider in the sense that A and B are independent, but they are no longer independence once conditioned on T.

[26]Since ϵ_Y includes all the factors that influences Y except for T and M, hence if T is a collider, it must have a causal parent, which should be included in ε_Y, which will then contradict with the assumption.

Figure 18: First Stage of Insurance Affecting Health



Note: This figure plots the first stage coefficient $\hat{\beta}_{\pi}^{\text{FS}}$, defined as in Equation (35), for healthy and unhealthy individuals and for all age, respectively. The solid blue line indicates the point estimates while the shaded light blue area indicates bootstrapped five percent confidence intervals.

values.

B.3 First Stage

This appendix section discuss on the first stage of the bi-probit model that we specified in Section 4.1.2. Particularly, we use the group average of EHI offer rate as an instrument for individual health insurance status. We therefore need to show that, conditional on income, the group average of EHI offer rate is still correlated with individual insurance status. We hence consider the following probit regression:

$$i_{hi} = \Phi(\beta_{\pi}^{\text{FS}} i_{hi}^{g} + \gamma_{\pi} \log(\text{income}_i)), \tag{35}$$

for all age j and health h, where i_{hi}^{g} is the group average of the EHI offer rate, and log(income) is the income of an individual. The key to the first stage success is that the estimated $\hat{\beta}_{\pi}^{\text{FS}}$ should be significant. Figure 18 illustrates the estimated $\hat{\beta}_{\pi}^{\text{FS}}$ together with its 95 percentile confidence interval. We can see that the estimated $\hat{\beta}_{\pi}^{\text{FS}}$ is highly significant for all age and health status.

C Calibration

C.1 Mortality rate

We estimate these two mortality rates using data from the National Center for Health Statistics publication *Health, United States 2016*. In particular, Table 21 of this publication provides the mortality rate for all causes, while Table 27–31 provides that for drug poisoning, motor-vehicle related injuries, homicide, suicide, and fire-arm related injuries. We take the sum of the mortality rates from Table 27–31 as the (unconditional) mortality rate for accidents, while the difference between that of Table 21 and the sum of Table 27–31 as the (unconditional) mortality rate for health-related issues.[27] We then adjust for the portion of healthy and unhealthy individuals relative to the whole population and obtain our health-related (conditional) mortality rate.

C.2 Co-Insurance rates and EHI offer rates

Here we describe the estimation of $\phi_{mr}(m)$ in detail. Other functions, $\phi_E(m)$, $\phi_P(m)$, and $\phi_{md}(m)$, are estimated in a similar way. Note that since the MEPS data do not distinguish the co-insurance rates between the EHI and the private insurance, we assume $\phi_E(m) = \phi_P(m)$, an assumption that is also made in Jeske and Kitao (2009).

All old individuals are covered by Medicare. In the MEPS data, we are able to observe the total health expenditure as well as the amount that is paid by Medicare. We estimate the coinsurance rates using data from old individuals who *only* have coverage through Medicare and no other coverage to control for the phenomenon that individuals may coordinate multiple health insurance plans in payments. We then estimate the co-insurance rates in two steps. First, we define the co-insurance rates as the ratio of the amount paid by Medicare versus total medical expenditure. We sort all eligible individuals into 10 bins according to their medical expenditure, and calculate the median co-insurance rate of each bin. By doing this, we get the co-insurance rate as a piecewise linear function of medical expenditure with 10 steps. Then, for each of the possible medical expenditure amount that we specified in the previous section (i.e., for each element of the seven possible realizations of medical expenditure for each age-health cohort), we estimate the co-insurance rate from this piecewise linear function.

[27]Note that this source reports mortality rates every ten years, with age breakdowns generally in ten-year intervals. We interpolated by time and age groups to produce estimates for five-year time intervals and age categories.

The other two functions, $\phi_E(m)$ and $\phi_{md}(m)$, are estimated in the same procedure. Note that for Medicaid ($\phi_{md}(m)$), we restrict our sample to individuals who are young and hence ineligible for Medicare and do not have any other insurance coverage. Similarly, for EHI and private insurance ($\phi_E(m)$), we restrict our sample to young individuals ineligible for Medicare or Medicaid and do not have any other insurance coverage. All these functions are plotted in Figure 5. In general, the coinsurance rates are increasing in medical expenditure, which may reflect the lump-sum co-payment of individuals.

The EHI offer rate $p_E(e)$ is also estimated from the MEPS data non-parametrically. We discretize the annual labor income into 15 grid points, which represent for income ranging from 0.1 to 3 times of the nominal GDP per capita, and categorize each individual in our sample into one of these 15 bins. We then separately calculate the EHI offer rate for each bin. The left panel of Figure 2 shows that jobs with higher wage payments are more likely to offer EHI, consistent with Aizawa and Fang (2019). We then fit a piecewise linear function over these grid points. For each realized income level, we calculate the corresponding EHI offer rate from this function.

Table 9: Calibrated Parameters

Parameters		Values	Description and Targets
Technologies			
	α	0.36	Capital share of 0.36
	δ	0.06	Depreciation rate
	A	1	TFP (Normalized)
Preferences			
	β	0.87	K/Y ratio 2.4
	b	7.53	Flow utility (Hall and Jones, 2007)
	σ	2	Relative risk aversion
Insurances			
Medicare	π_{mr}	0.026	Premium as 2.6% of GDP
Medicaid	Θ_e	0.311	Income threshold (Pashchenko and Porapakkarm, 2017)
	Θ_a	0.537	Asset threshold (Pashchenko and Porapakkarm, 2017)
EHI	ψ	0.8	Employer's contribution (Gruber, 2008)
Other Private	η	0.136	Mark-up (Gruber, 2008)
Government			
	G	0.18	Government expenditure of 18% of GDP
	λ_p	0.883	Proportional income tax: government budget balanced
	τ_p	0.151	Progressive income tax (Heathcote et al., 2017)
	τ_c	0.057	Consumption tax (Mendoza et al., 1994)
	τ_{ss}	0.124	Social security tax
	y^{ss}_{max}	2.0	Social security tax limit
	τ_{mr}	0.029	Medicare tax
Social Security	s_1	0.90	Payment rate
	s_2	0.33	Payment rate
	s_3	0.15	Payment rate
	τ_1	0.20	Threshold
	τ_2	1.25	Threshold
	τ_3	2.46	Threshold
	$\underline{c}$	0.103	Consumption floor as the poverty line
Capital Market	$\underline{a}$	0	Borrowing constraint
Med. Bankruptcy	γ	0.35	Garnishment rate (Livshits et al., 2007)
	$\underline{y}$	0.20	Earnings exemption: 20% of GDP per capita