Health, Health Insurance, and Inequality*

Chaoran Chen† Zhigang Feng‡ Jiaying Gu§

August 8, 2022

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 introduce this feature into the prototypical macro-health model and estimate the baseline economy by matching the observed joint distribution of health insurance purchase, health status and income over the life cycle. Quantitative analysis reveals that an individual's insurance status has significant and persistent impact on health, which will be reinforced by and subsequently amplify the feedback effect of health on labor earnings and income inequality. Providing *"Universal Health Coverage"* would narrow health and life expectancy gaps, with a mixed effect on income distribution in absence of any additional redistribution of income or wealth.

Keywords: Health Insurance, Health Disparity, Income Distribution.
JEL classification: E21, E60, I14, O15.

*We thank Zhifeng Cai, Luca Dedola, Hanming Fang, Kevin Huang, Greg Kaplan, Sagiri Kitao, Dirk Krueger, James MacGee, Serdar Ozkan, Gajendran Raveendranathan, Jose-Victor Rios-Rull, Manuel Santos, David Wiczer, Kai Zhao, three referees, and seminar participants at Jinan University, the 2019 (10th) Shanghai Macro Workshop (SHUFE), York University, UIBE Macro Workshop, UIBE Conference on Inequality and Open Economy, SAET 2022, University of Delaware, University of Nebraska at Lincoln, University of Utah, VMACS Junior Conference 2020, 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.

†Department of Economics, York University, 1034 Vari Hall, 4700 Keele Street, Toronto, ON M3J 1P3, Canada, chenecon@yorku.ca.

‡Department of Economics, University of Nebraska-Omaha, NE 68106, United States, z.feng2@gmail.com.

§Department of Economics, University of Toronto, 150 St. George Street, Toronto, ON M5S 3G7, Canada, jiaying.gu@utoronto.ca.

1 Introduction

The U.S., as the only OECD county without universal health insurance, has more than 44 million Americans lacked coverage in 2013, the year before the major coverage provisions of the Affordable Care Act (ACA) went into effect. Even after the ACA expanded comprehensive health insurance to millions of Americans, the uninsured rate still stood at 10.2% in 2017. Staying uninsured can have detrimental effect on one's health due to lack of primary preventive care and screening service, and the inaccessibility to medical service once sick. Such effect has persistent impact on individual's health over life cycle as being unhealthy will face less favorable price in the insurance market and hence more likely to stay uninsured and unhealthy henceforth. Furthermore, such effect will be reinforced by the impact of bad health on individual's productivity and earning ability.[1]

This paper studies the persistent effect of health insurance coverage on health disparity and income inequality over the life cycle. Using data from the Medical Expenditure Panel Survey (MEPS) and the Panel Study of Income Dynamics (PSID), we identify a "health premium" of insurance coverage that the insured is more likely to remain healthy or recover from ill health. To ensure the consistency of our estimation, we develop an empirical triangular model with an instrumental variable (IV), and employ a recently developed semi-parametric estimator. We then introduce this feature into the prototypical macro-health model. Quantitative analysis reveals that an individual's insurance status has a significant and persistent impact on health over the life cycle, which is reinforced by and subsequently amplifies the feedback effect of health on labor earnings and income inequality. Providing *"Universal Health Coverage"* would narrow health and life expectancy gaps, with a mixed effect on income distribution in absence of any additional redistribution of income or wealth.

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 MEPS and PSID. In our data, unhealthy individuals report substan-

[1] Lower wage earned that resulting from reduced productivity associated with poor health implies that the agent is subject to tighter budget constraint and a slimmer chance of being offered with employer-sponsored insurance (EHI), which provides favorable tax treatment and better risk sharing through workplace. Subsequently, this individual is more likely to stay uninsured and unhealthy. Conversely, it also amplifies the impact of health on income (e.g. de Nardi et al., 2017), since one's health is an endogenous variable affected by earning via insurance decision.

tially lower income, and face higher medical expenditure compared with healthy individuals across all ages. Concurrently, low-income workers are less likely to be insured, due to tighter affordability constraints or lower chance of receiving employer-sponsored health insurance (EHI) offer from employers. Furthermore, insurance coverage affects individual health as the insured has better access to primary prevention and screening services; has a regular source of health care; and is more likely to have healthy behavior through wellness program under insurance coverage. These facts indicate significant impact of income on individual health through endogenous insurance choice. We address the endogeneity of health insurance choice and the potential reverse causality through an IV approach—we use the group averages of EHI offer rates within the same occupation among workers in same age as IV for one's insurance status—and employ a semi-parametric estimator to ensure the consistency of the estimation. 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 9 percent more likely to remain healthy next year compared to an uninsured counterpart. There is a even larger difference in the probability of transiting from unhealthy to healthy between the insured and the uninsured.

Guided by these empirical regularities, we develop a life cycle model to study the interaction between health disparity and income inequality. 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. This latter aspect 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.

We estimate parameter values using micro data from MEPS and PSID through innovative econometric techniques that are designed to address the endogeneity of health insurance choice and the simultaneity bias arising from the interaction between health and income.

Our baseline economy reproduces the observed health insurance choice as well as the joint distribution of income and health over the life cycle. 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.

We simulate our model to quantify the impact of health insurance coverage on generating persistent difference in income and health over the life cycle. Numerical analysis indicates that a worker who has health insurance coverage initially is more likely to remain healthy by up to 10 years, compared to an otherwise identical worker without insurance. Such effects would reinforce the impact of health on earning ability, and further widen the income gap. Hence, providing health insurance can be an effective policy in reducing health disparity and income inequality.

In light of the above findings, we conduct counterfactual policy experiment to analyze the welfare implication of "Universal Health Coverage", which has been adapted by most OECD countries. Given health-income interaction, this policy would improve individual health, 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 evolution of health. Moreover, this level effect—individuals are healthier and hence are more productive—increases the tax base, and reduces the potential tax distortion required to finance the policy change compared with the canonical models. Such policy would, however, have countervailing effects on the income distribution. Better health improves productivity and hence labor earnings especially for poor workers, thereby reduces income inequality. Concurrently, poor workers live longer as a result of better insurance coverage, which in turn increases the weight of low-income individuals in the income distribution, consequently enlarges income inequality. In the absence of any further redistribution of income or wealth, the overall effect on income inequality is therefore ambiguous.

1.1 Related literature

Much of the economic literature on the determinants of health starts with Grossman (1972), in which health is modeled 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, see for instance 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 can influence the evolution of health. It is generally challenging to estimate the health production function due to a lack of concrete health metrics, imperfect observability of some contributing factors and the difficulty in estimating the relative price of medical service. Our model focuses on evaluating the impact of health insurance choice on the evolution of a binary health status, and is thereby amenable to parameterization and estimation. We address the endogeneity of health insurance choice and the potential reverse causality, which resolves the identification problem in determining health status due to the bi-directional causality between health and economic outcomes. Our work hence contributes to the recent macro-health researches that endogenize health process by exploring specific mechanisms and rich micro-level data, see for instance Ozkan (2014), Pelgrin and St-Amour (2016), Fonseca et al. (2020), Mahler and Yum (2022), and Greenwood et al. (2022). Nevertheless our results should be interpreted as a conservative lower bound of the impact of income on health, as we omit other factors that affect the evolution of health and are correlated with income.

Our paper complements a strand of literature that studies health and inequality. de Nardi et al. (2017) study 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. This allows us to study the policy implications of health insurance. In this regard, our paper contributes to a broad literature that studies macroeconomic aspects of health policies, including French and Jones (2004), Hall and Jones (2007), Jeske and Kitao (2009), Bruegemann and Manovskii (2010), Pashchenko and Porapakkarm (2013), Hansen et al. (2014), Braun et al. (2017), and others.

Our key mechanism that health insurance causally affects health resonates findings in a large empirical health literature, such as Currie and Gruber (1996), Doyle (2005), Card et al. (2009), and Finkelstein et al. (2012), among others. We differ from them by focusing on its macroeconomic implications on health and income inequality.[2] Our paper is also related to the empirical micro literature studying why so many Americans are uninsured, see, for instance, Cutler and Reber (1998), Card and Shore-Sheppard (2004), Herring (2005), Gruber (2008), and Mahoney (2015), among others. Guided by these works, we build our macro model to include most empirically relevant channels that matters to insurance choice. Taking into account the general equilibrium impacts allows us to understand the importance of these channels at the aggregate level.

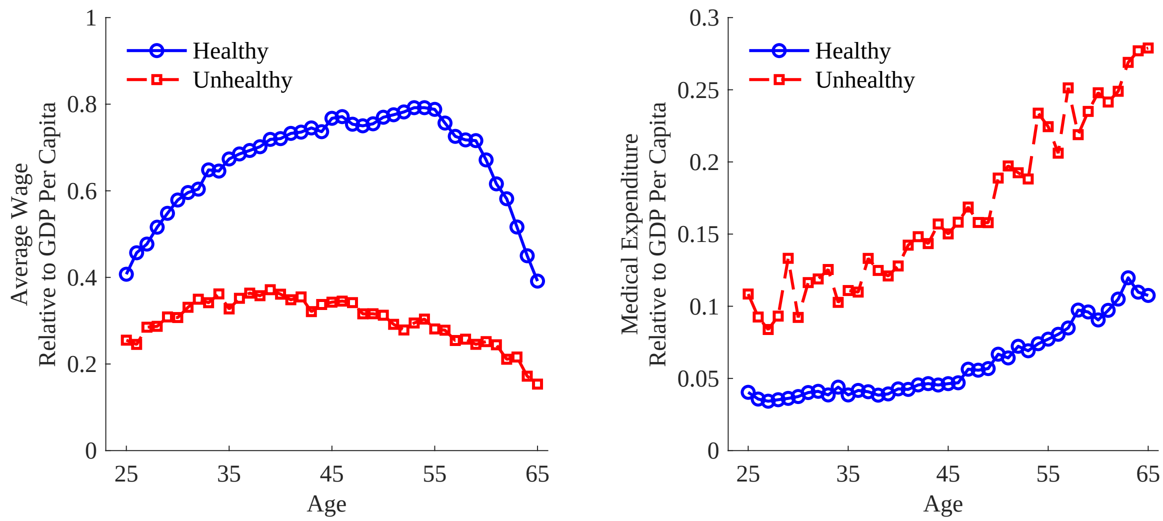
The rest of the paper proceeds as follows. Section 2 describes our identification strategy for estimating the health premium of insurance coverage. Section 3 develops a dynamic equilibrium model. Section 4 details how we estimate the model. Section 5 evaluates the fit and performance of our baseline economy. Section 6 explores the persistence effect of health insurance through the lens of our model, and Section 7 discusses the macroeconomic implications of implementing universal health insurance coverage program. Section 8 concludes the paper.

2 The Effects of Health Insurance on Health

In this section, we identify the crucial role of health insurance in determining the evolution of individual health using data from the Medical Expenditure Panel Survey (MEPS). In the MEPS data, health status is recorded as "excellent", "very good", "good", "fair", and "poor". In line with the literature, we classify an individual as "healthy" if her health status is "excellent", "very good", "good", and as "unhealthy" otherwise. Insurance status can take one of the following five states: no insurance coverage, covered by a private insurance plan, by EHI, by Medicaid, or by Medicare.

[2] Our instrumental variable approach also allows us to identify the effect for individuals of all ages, as we explained in Section 2.2, 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).

Figure 1: Income and medical expenditure by health



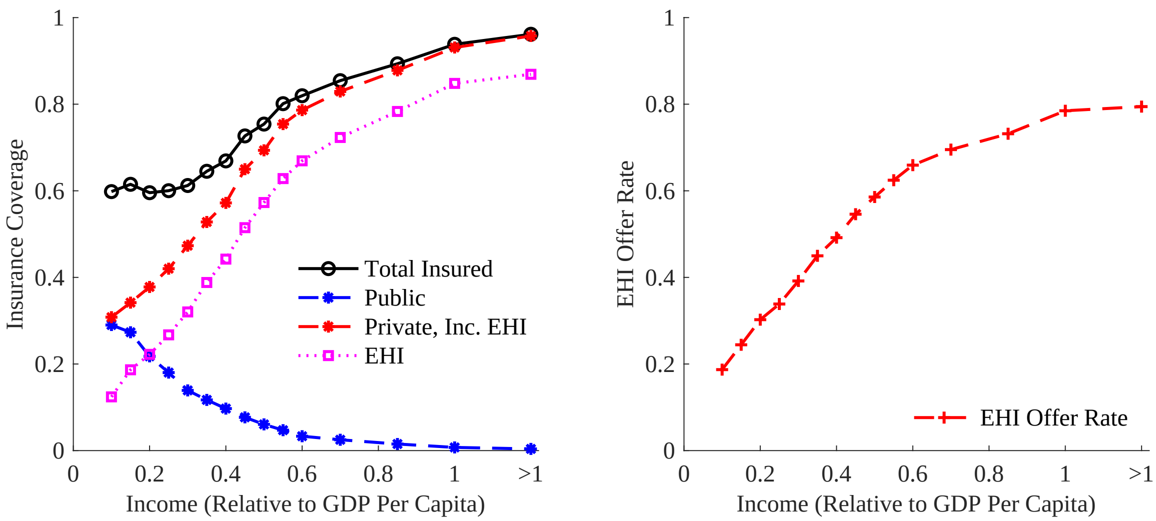
Note: The left panel shows the average annual income for healthy and unhealthy individuals for different ages, where annual income is normalized relative to per capita GDP. The right panel shows the average medical expenditure for healthy and unhealthy individuals for different ages, also normalized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

2.1 Interdependence Between Health and Income

The most noticeable empirical facts based on our data resonate findings by a burgeoning literature that studies the impact of health on income inequality. Figure 1 reports average income and medical expenditure by individual's health status over the life cycle. The left panel indicates that unhealthy individuals have substantially lower income compared with healthy individuals for all ages, consistent with findings in Aizawa and Fang (2020). 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 echoes findings in de Nardi et al. (2017) that health disparity matters for income inequality.

Health is an endogenous outcome variable determined by multiple factors. In a landmark research, Chetty et al. (2016) document strong correlation between individual income and health proxied by mortality risk. However, income itself doesn't improve health and it must work through some intermediaries, such as medical service, healthy behavior, exercise, dietary choice etc. It is generally challenging to estimate the health production function

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 per capita GDP.

due to a lack of concrete health metrics, imperfect observability of some contributing factors and the difficulty in estimating the relative price of medical service (e.g. Suen, 2006). To address these issues, we focus on the impact of health insurance choice on the evolution of a binary health status, which are observable and thereby amenable to parameterization and estimation. The focus on insurance choice is also motivated by its observability and its strong correlation with income. 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. It is clear that private insurance coverage and especially EHI increases with income. 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 per capita GDP, less than 30 percent receive an EHI offer from employers, compared to over 80 percent for individuals whose annual income is around or above the per capita GDP. More importantly, focusing on health insurance choice reflects the well-documented health benefit of insurance coverage, see discussion below. Furthermore, government's health care policies, such as the Affordable Care Act, are generally designed to target health insurance coverage.

2.1.1 Health Insurance and Health Outcome

A large body of literature studies the relationship between health insurance and mortality and other health outcome. A report by the Institute of Medicine (Institute of Medicine Committee on the Consequences of Uninsurance, 2002) finds that uninsurance is associated with 25% increase in mortality risk. In a recent study, Goldin et al. (2019) evaluate a randomized pilot study in which the IRS sent informational letters to 3.9 million taxpayers who paid a tax penalty for lacking health insurance coverage under the Affordable Care Act. They find that the increased coverage induced by the intervention reduces mortality among middle-aged adults by approximately 0.06 percentage points in the two years following the treatment.

In health economics literature, several factors have been recognized as to mediate the relationship between health insurance and health-related outcomes. These include having access to primary prevention and screening services; being able to see a provider when one believes care is needed and having a regular source of health care; and incentives to lead healthy life through wellness program under insurance coverage.

Insurance and preventive care. The insured are more likely to have preventive care and screening services, which improve the likelihood of disease screening and early detection and the management of chronic illness. In our data sample, 70.6% of insured individuals have their last routine health check-up by doctors or health professionals within a year, while only 36.7% of the uninsured individuals have done so. Additional to overall health assessment, this difference also applies to more specific items of health check, such as blood cholesterol, blood stool test, and extended healthcare like dental check, see Table 1 for details.[3]

Insurance and health-care services once sick. The uninsured visit the doctor and are admitted to a hospital less often, and are less likely to have prescription drugs filled due to financial reasons. Doyle (2005) studies a subpopulation with strong need for emergency medical care (victims of auto accidents who are alive when they reach the hospital) and finds significantly higher adult mortality rates for uninsured persons in Wisconsin during

[3] Moreover, this difference remains robust for all preventive care measures after controlling for observables such as income, age, health status, and education (see Appendix A.8 for further details).

Table 1: Preventive Care: Insured Vs. Uninsured

Check-up Items	% among Insured	% among Uninsured
Within last year:		
Overall health assessment	70.6	36.7
Dental	61.9	31.7
Blood cholesterol	68.1	34.3
Flu shot	41.2	15.1
Prostate specific antigen (male only)	42.7	12.7
Pap smear test (female only)	54.8	41.1
Breast exam (female only)	63.9	40.8
Mammogram (female only)	47.2	23.7
Ever had:		
Blood stool test	20.7	6.8
Sigmoidoscopy or colonoscopy	29.0	7.3

Note: This table documents the percentage of individuals with certain preventive care by insurance status, calculated from the MEPS data. The insured are more likely to have preventive care, and this difference remains robust for all items after controlling for observables such as income, age, health status, and education (see Appendix A.8 for details).

the period of 1992–1997. This difference was attributed to disparity in treatment intensity, rather than pre-accident heterogeneity in individual health. Card et al. (2009) explore the fact that the insurance status changes at the age of 65 for most people due to the eligibility to Medicare. With a regression discontinuity design, they find that, among patients admitted to hospital, those with insurance receive more health care services and have 20% lower 7-day mortality rate.

Insurance and wellness program. The 2014 Kaiser Family Foundation (KFF) Employer Health Benefits Survey finds that most large employers provide wellness benefits through group health insurance plan. These wellness programs usually offer employees with help on quitting smoking, diabetes management, weight loss programs, and preventative health screenings, all of which are designed to promote better health among employees and to help reduce health care costs. A recent report by Financial Times found that smarter and widely available technology allow the insurance industry to intervene earlier and in a greater variety of ways. Many life and health insurers have developed a range of innovations and prevention scheme that encourage fitter, longer-living customers to reduce risk and cost.[4] Government provided public insurance has similar programs (e.g. Medicaid Incentives Initiative) to promote healthy behavior among beneficiaries and reduce costs. Based on our MEPS data, we also find a statistically significant difference in weekly physical activity between the insured and the uninsured population.

2.2 An Econometric Framework for Estimation

Our goal is to estimate the effects of health insurance on health. More specifically, we want to identify a "health premium", which measures the difference in the probability of maintaining favorable health outcome between insured and uninsured given identical health status in the last period. The difficulty is that the choice of health insurance is endogenous: individuals may *choose to* opt in medical insurance plans if they expect poor health in the following period. This selection bias hence prevents us from directly estimating health transition separately for the insured versus the uninsured.

We employ an instrumental variable (IV) approach to address this endogeneity issue. Denote an individual's health status as y_i (the outcome variable), the insurance status as d_i (the treatment variable), and our instrument variable as z_i. Given the binary nature of the endogenous treatment variable, we consider the following triangular model:

$$\begin{aligned} Y &= \mathbb{1}\left\{X'\beta + \delta D \geqslant U\right\}, \\ D &= \mathbb{1}\left\{X'\alpha + \gamma Z \geqslant V\right\}. \end{aligned} \tag{1}$$

Here δ is the key parameter of interest, and X consists of demographic controls, including gender, race, region, log wage, marriage status, and education.[5] Given that the effects of insurance on health transition may differ by age, we estimate δ separately for four age groups: 25–34, 35–44, 45–54, and 55–64. We also estimate δ separately for those who were healthy in the previous period versus those who were unhealthy to reflect the facts that the effects of

[4] "Rethinking insurance: how prevention is better than a claim", Financial Times, July 23 2022. https://on.ft.com/3cCv4VD.

[5] Note that we include wage as a control variable in X which may suffer from endogeneity and it is difficult to find a valid instrument for income. We discuss in Appendix A.2.1 on the identification of δ and the health premium in such situation.

insurance may differ by existing health condition and unhealthy individuals may potentially benefit more from having health insurance.

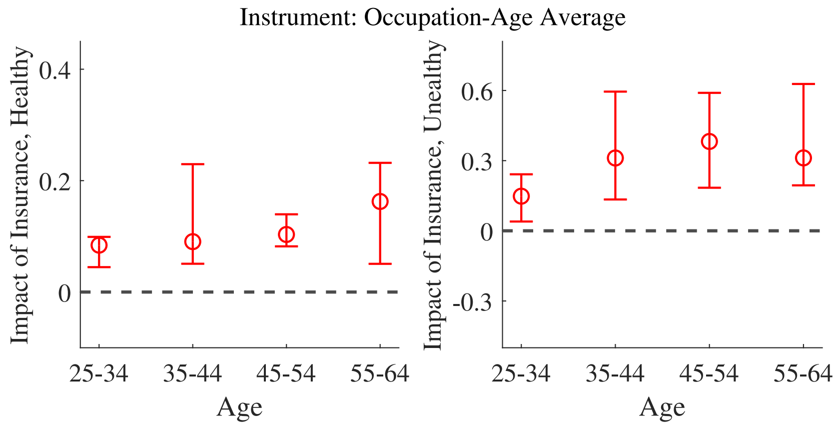
One possible instrument is EHI offer status, since whether or not a firm offers EHI is the firm's decision and independent of individual characteristics of its workers, given the non-discriminative nature of group insurance. However, this EHI offer status may still contain some endogeneity, as workers who expect bad health shocks may choose to work in firms that are more likely to 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 occupation and the same worker age group as an instrument of the insurance coverage in our baseline analysis. Intuitively, workers could change jobs but they are very unlikely to change occupations merely to search for an EHI offer. We discuss the validity of this instrument in details in Section 2.3.

In addition, parametric assumptions on the joint distribution of (U, V) are needed to point identify this triangular model. We employ a semi-parametric estimator proposed by Han and Lee (2019). Particularly, we assume non-parametric marginal distributions for both U and V and a parametric copula for the dependence structure. This is a substantial generalization from the bi-probit model that is often used: If we restrict the marginal distributions to be Gaussian and we assume a Gaussian copula, then we are back to the bi-probit model. When joint normality of (U, V) is misspecified, Han and Lee (2019) show that the bi-probit model estimates can exhibit substantial bias.[6]

To summarize, our baseline triangular model identifies the impact of health insurance coverage on the evolution of health status. The left panel of Figure 3 presents the estimated "health premium" for individuals who are healthy in the previous period, while the right panel plots the estimation for unhealthy counterparts. Based on our estimation, a healthy individual at age 40 is about 9 percent more likely to stay healthy in the next period if she has insurance coverage. The effect is even larger for the unhealthy, since the unhealthy

[6]Regarding the correlation structure, we do not have direct evidence on how to choose the parametric form for copulas. Fortunately, Han and Lee (2019) show that the estimation results are insensitive to the choice of copulas once we allow for non-parametric marginal distributions. We hence choose frank copula which allows us to interpret the dependence structure more easily. In Appendix A.3, we report estimates based on Gaussian copula with non-parametric marginal, as well as the estimates assuming the Gaussian marginals on (U, V), which includes the bi-probit model estimates as a special case (Figure 4). Results are similar to those in Figure 3.

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 (left panel) or unhealthy to healthy (right panel). The circle indicates the point estimates while the bar indicates 95% confidence intervals.

individuals usually require more medical services and hence insurance coverage plays a bigger role. The 95 percent confidence intervals are obtained through bootstrap repetitions. Our estimations suggest that there exists a positive and significant health premium for all age-health groups.

2.3 Validity

Our strong first-stage results for all age-health groups, reported in Appendix A.6, guarantee the relevance of the instrument (the average EHI offer rate within the occupation-age group). To ensure the validity of the instrument, we will assess the exclusion restriction assumption, in both the first stage ($Z \perp V$) and the second stage ($Z \perp U$).

2.3.1 Second Stage ($Z \perp U$)

This assumption may be violated if the EHI offer rate is correlated with health through mechanisms other than insurance coverage. For instance, if unhealthy individuals self-select into less demanding occupations which are lower in pay and less likely to offer EHI, then this instrument would be invalid. We start by arguing that this is unlikely to be true, since we find similar results using an alternative instrument: the average EHI offer rates within industry-age groups. The intuition is that the pattern of sorting of individuals into occupations may differ from that between individuals and industries, and hence, if the sorting plays a key role then we expect to find different outcome with these two instruments. We note that occupations and industries are not necessarily highly correlated, and the rank correlation between these two instruments is only 0.37. In addition, we also construct our third and fourth instruments by further refining the categories by education, where the intuition is that, for most adults, the level of education stays unchanged. Again the results are very similar, and we relegate the details to Appendix A.3.

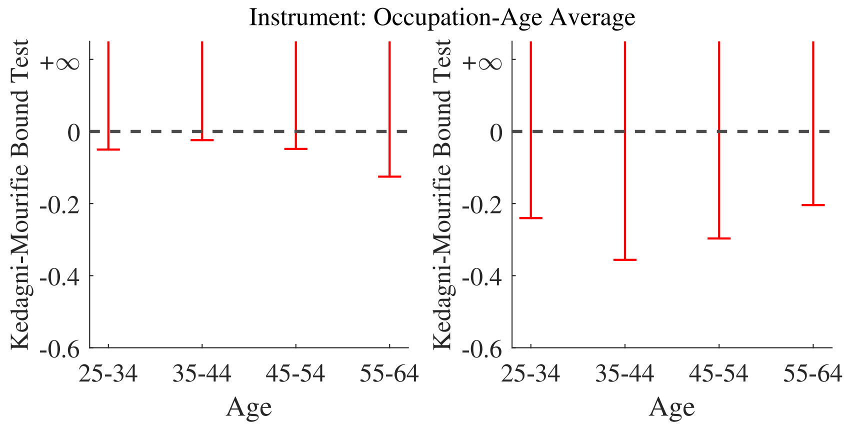
We also employ a formal statistical test recently proposed by Kédagni and Mourifié (2020). They propose a sharp bound test to detect all observable violations of the instrument variable's independence assumption for models where treatment variable is binary but instruments are unrestricted, exactly like our scenario here.[7] As shown in Kédagni and Mourifié (2020), exclusion restriction assumptions of the instruments can be formed as moment inequalities. They then construct the one-sided confidence interval using the intersection bound approach proposed in Chernozhukov et al. (2013). The exclusion restriction assumption would be rejected by data if the lower bound of the one-sided confidence interval excludes zero. As we see from Figure 4, for all health-age groups, the lower bound of the 95 percent one-sided confidence interval is negative, and hence the data do not reject the hypothesis that our instrument is independent to U. Note that all our alternative instruments survive this test, see Appendix A.4 for details.

2.3.2 First Stage ($Z \perp V$)

This assumption may be violated if our instrument is correlated with unobservables that affects one's insurance status. For instance, workers who are more eager to opt in insurance may choose to work for jobs that are more likely to offer EHI. For this first stage exclusion restriction assumption, unfortunately we do not have a formal statistical test to assess its validity. We instead employ a potential outcome framework that does not require a first stage

[7] To provide some examples of the power of this test, Kédagni and Mourifié (2020) assess the validity of various instruments used in the returns to college literature, and find that parental education—which is often used as an instrument for one's college education status—is not a valid one even conditional on experience and a measure of ability. On the contrary, the college tuition fees—another commonly used instrument—is valid when controlling on measures of ability.

Figure 4: Kedagni-Mourifie Test for Second-Stage Exclusion Restriction



Note: This figure plots the results of the sharp bound test proposed by Kédagni and Mourifié (2020) for healthy (left panel) or unhealthy individuals (right panel) by age group. The bounds are evaluated at the 95% one-sided confidence interval. The exclusion restriction would be rejected by data if the lower bound excludes zero.

model to assess its robustness. Briefly speaking, we do not restrict the relationship between the insurance status and the instrument (i.e. the model on D in equation (1)), and hence do not rely on the assumption of $Z \perp V$ for identification. As a result, our identification will not be contaminated by any potential violation to the $Z \perp V$ assumption. The cost is that we lose point estimation of the health premium. Yet, we are able to find informative bounds of the premium and for most of the age-health group, the lower bound is larger than zero. This provides further evidence on the positive effect of insurance on health outcome even with less restrictive model assumptions. This approach has recently been used in the labor literature, see for instance De Haan and Leuven (2020).

We now outline the potential outcome framework as follows. To ease the exposition, we do not use any additional control variables with the understanding that more control variables result in tighter bounds. The treatment equation can be written as

$$Y = DY(1) + (1 - D)Y(0),$$

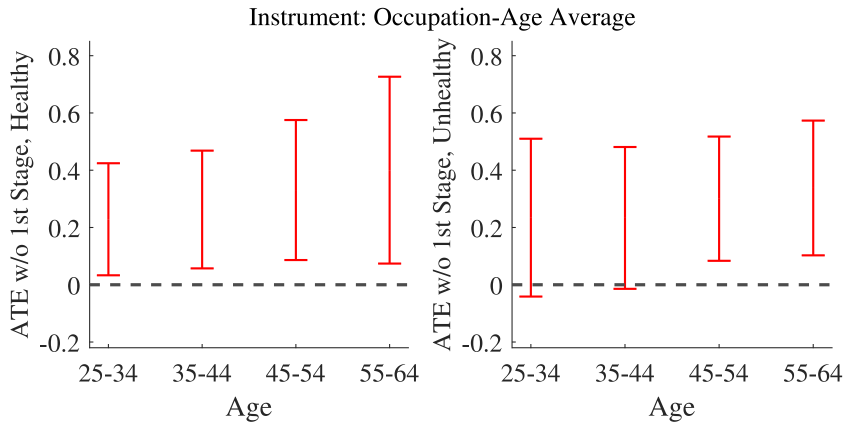
where $Y(1)$ and $Y(0)$ are potential outcomes. $Y(1)$ indicates the health status with health insurance coverage, and $Y(0)$ indicates health status without insurance. The average treatment effect (ATE), or our health premium, can then be written as

$$\text{ATE} = \text{Prob.}[Y(1) = 1] - \text{Prob.}[Y(0) = 1].$$

The difficulty of identification is that for each individual, only one of the potential outcome $Y(1)$ and $Y(0)$ is observed. Under this setting, we ask whether the data provide any information on the ATE. As expected, without further assumption, the bound for ATE is not informative. However, if we have an instrument variable Z that is correlated with D and is assumed to be independent from $(Y(1), Y(0))$, or equivalently $Z \perp U$ in equation (1) (but not necessarily $Z \perp V$), then the joint distribution (Y, D, Z) may impose enough restrictions on the joint distribution of $(Y(1), Y(0))$ and thus yields an informative bound on ATE. This potential outcome model with instrument variables has been well studied in the literature (e.g. Balke and Pearl, 1994). To further tighten the bounds on ATE, we follow Manski (1997) and Manski and Pepper (1998) and assume monotone treatment response, i.e., either $Y(1) \geqslant Y(0)$ or $Y(1) \leqslant Y(0)$. Our data reject $Y(1) \leqslant Y(0)$, hence we maintain the assumption $Y(1) \geqslant Y(0)$, which can be interpreted as health insurance may have no effects on one's health but it should not negatively affects one's health. We then again use the intersection bound approach to construct a confidence set for ATE following Acerenza et al. (2020). More details are provided in Appendix A.5.

Based on results shown in Figure 5, we find that, without the first stage model, there exists a positive and significant ATE for most of the age-health groups, i.e., we find $\text{Prob.}[Y(1) = 1] > \text{Prob.}[Y(0) = 1]$ with strict inequality for most of our population. The advantage of this approach is that, without any first stage specification, we are no longer concerned about the exclusion restriction assumption of $Z \perp V$ and thus provide robustness evidence of a positive effect of health insurance on health outcome. Note that we assume $Y(1) \geqslant Y(0)$ but not $Y(1) > Y(0)$, hence the bounds of ATE excluding zero is not by construction. Since we get non-empty bound for ATE, it implies the data supports the assumption of $Y(1) \geqslant Y(0)$ as well as $Z \perp U$. Further results using three alternative IV's are reported in Appendix A.5.

Figure 5: Bounds of Health Premium without First Stage Assumptions



Note: This figure plots the results of our potential outcome framework. Specifically, we estimate the health premium without a first stage model. As a result, we do not have any concern regarding the exclusion restriction assumption ($Z \perp V$). Without a first stage specification, we cannot obtain a point estimate. We report the 95 percent confident set for the identified set of the health premium.

2.3.3 The Sign of Endogeneity Bias

In our semi-parametric estimation strategy, we employ the frank copula for dependence between U and V. An advantage of assuming a parametric (frank) copula in the correlation structure is that we can interpret the correlation between (U, V) and hence recover the sign of endogeneity bias. In all our settings and health-age groups, we find negative and significant correlation between (U, V). Without much restriction to generality, let us consider an observation with negative v_i and positive u_i to illustrate this correlation. From equation (1), a more negative v_i means that the inequality of $x_i'\alpha + \gamma z_i \geqslant v_i$ is more likely to hold ceteris peribus, and hence this individual is more likely to have health insurance coverage. A more positive u_i means that the inequality of $x_i'\beta + \delta d_i \geqslant u_i$ is less likely to hold and hence this individual is more likely to be unhealthy. This correlation then suggests that individuals who will be unhealthy are more likely to obtain health insurance, and hence the endogeneity bias is negative.

This estimated negative correlation between U and V suggests that a plausible selection mechanism is that individuals expecting negative health shocks may choose to opt in health insurance. It follows immediately that the actual health premium will be even larger if the instrument is inadequate in addressing the endogeneity issue. Moreover, mechanisms that require a positive correlation between U and V are unlikely to be important here—for instance, certain jobs are mentally and physically demanding, while they provide generous compensations including EHI—which implies that individuals who will be healthy are more likely to obtain health insurance.

2.3.4 Discussions

Health premium by income. Income affects health through health insurance choice; in addition, it also directly affects health (Pijoan-Mas and Rios-Rull, 2014). We note that in our estimation we explicitly control for income in both stages, but there is no interaction effect of income and insurance status. To assess the interaction, we also separately estimate the health premium for poor and rich individuals. We find positive and significant health premium in both groups, with details relegated to Appendix A.7. With smaller sample sizes the magnitudes are less tightly estimated and hence we choose not to use this specification as our baseline.

Public insurance and non-workers. Our identification above uses EHI offer rate as an instrument for health insurance coverage and hence the results mainly apply to workers, while the health transition for non-workers may be different. In our sample, slightly more than one third of non-workers have private health insurance, most of which obtain it through EHI coverage of spouse or other family members. For the remaining non-workers, more than half of them are covered by public insurance (Medicaid or Medicare) only. Public insurance typically has less generous coverage compared to private insurance, and the pool of people on Medicaid could differ from the pool of workers with private insurance. Hence the health premium estimated above may not apply to non-workers with public insurance. To address this issue, we distinguish these two types of insurances and their health premium when we estimate our baseline model.

3 Model

We now present our baseline model that extends the prototypical macro-health model by incorporating the "health premium" of insurance identified in previous section. Households are endowed with one unit of time in each period that can be supplied to the labor market, work until retirement age J_R, and maximize discounted lifetime utility. They live to a maximum of J periods, face idiosyncratic labor productivity shocks z and medical expense shocks m in addition to health shocks in each period over the life cycle. The financial market is incomplete with a risk-free bond traded in the economy. Households can purchase health insurance to hedge against health expenditure shocks, where their endogenous insurance choice affects the evolution of health as detailed below.

A representative firm produces final good Y using capital K and efficiency units of labor N through a neoclassical production function: $Y = AK^{\alpha}N^{1-\alpha}$. At an interior solution rental prices equal their respective marginal productivity: $r + \delta = AF_K(K, N), w = AF_N(K, N)$, where δ is the depreciation rate of capital. If EHI is offered, the firm adjusts the wage to ensure the zero-profit condition by shifting the cost of providing EHI c_E to the employees. The production of the final good can be used for private consumption, investment, medical service and public spending. The law of motion for capital K is given by $K_{t+1} = (1-\delta)K_t + I_t$. To ease exposition, we may denote x' the value of variable x in the next period.

3.1 Demographics, preference, and endowment

Preference Preferences are represented by

$$\mathbb{E}\sum_{j=1}^{J}\left[\beta^{j-1}\prod_{t=0}^{j-1}\rho_{h,t}u_h(c_j, n_j)\right], \tag{2}$$

where β is the time-invariant discount factor, $\rho_{h,t}$ is the age and health specific survival probability, c_j and $n_j \in \{H, U\}$ are consumption and labor supply, respectively, and H and U represent healthy and unhealthy. We assume that $u_h(\cdot,\cdot)$ depends on health status $h \in \{H, U\}$, is strictly increasing (decreasing) in consumption (labor supply) and concave in consumption.

Endowment Labor income e_j^i of household i at age $j \leqslant J_R$ depends on household productivity, labor supply n_j^i, wage rate w, and EHI offer status i_E:

$$e_j^i = w\hat{z}_j^i n_j^i - i_E c_E, \tag{3}$$

where $\hat{z}_j^i$ is the product of realized labor productivity z_j^i and $g_j(h_j^i)$ that captures the impact of individual health h on labor productivity.

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

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

Here, $\pi_{hh'}^{j,i_{hi}}$ denotes the probability that an age-j agent's health status changes from h to h' conditional on their health insurance status $i_{hi} \in \{$*uninsured, private, EHI, Medicaid, Medicare*$\}$.

At each period, agents receive an idiosyncratic health expenditure shock m, whose distribution varies with the agent's age, health status, and health insurance coverage. Health status also affects her survival probability at age j, with $\rho_{H,j} \geqslant \rho_{U,j}$.

3.2 Market arrangement

Financial and insurance market Households can save by purchasing a' units of risk-free bond at the common market rate of r. Borrowing is limited, i.e., $a \geqslant -\underline{a}$. Depending on individual's health insurance choice i_{hi}, they will pay a premium $\tilde{\pi}(i_{hi})$ and the insurance will cover a fraction $\tilde{\phi}(m, i_{hi})$ of the realized medical expenditure m. For agents with current income (e) and EHI status in the past (η_E), they will have access to EHI with probability $p_E(e, \eta_E)$.[8] Once an offer is made to the employee, he/she may decide whether or not to obtain the coverage. The EHI has a premium π_E and a co-insurance rate of $\phi_E(m)$. The

[8]EHI offer rate increases in firm size and wage (Aizawa and Fang, 2020). 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 and to be history-dependent.

premium is independent of prior health history or any other individual states, since U.S. law requires employers that offer health plans to use a price common to all employees. The employer subsidizes a fraction $\psi \in (0, 1)$ of the insurance premium and the remaining is paid by the employee who obtains the coverage.

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 population with low income and saving: $e \leqslant \Theta_\mathrm{e}$ and $a \leqslant \Theta_\mathrm{a}$, where Θ_e and Θ_a represent the income and asset thresholds for Medicaid eligibility respectively. Medicaid beneficiaries will receive coverage for a fraction $\phi_{md}(m)$ of medical expenditure with zero premium. In addition, the Medicare program covers all retirees for whom there is a fixed premium π_{mr} and coverage ratio of $\phi_{mr}(m)$ for the medical expenditures.

Medical bankruptcy Households have the option to declare bankruptcy over the out-of-pocket medical expenditure. Upon declaring bankruptcy ($\iota = 1$), agents are subject to a linear garnishment of earnings $\lambda = \gamma \max\{y - \bar{y}, 0\}$, and an one-time non-pecuniary utility penalty ν, 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 cost-less. After declaring bankruptcy, the agent's out-of-pocket medical bill $[1 - \phi(m, i_{hi})]m$ is forgiven and they will be temporary excluded from the credit market ($a' = 0$). The cost of uncompensated care, $[1 - \phi(m, i_{hi})]m - \lambda$, will be covered by the government.

3.3 Government programs

Besides Medicaid and Medicare, the government also runs a means-tested welfare program and a pay-as-you-go Social Security program. The welfare program provides transfers to households whose after-tax disposable incomes falls below $\underline{c}$, as in Hubbard et al. (1995).

The welfare program parsimoniously captures unemployment insurance and food stamps. The retired individuals receive social security benefit $ss(\bar{e})$, which depends on the economy-wide average earnings $\bar{e}$.[9]

Workers pay proportional social security tax τ_{ss} and Medicare tax τ_{mr}. The worker's earning above y^{ss}_{max} is exempted from social security tax. The government also levies a progressive income tax $T(\cdot)$ and a proportional consumption tax τ_c to finance its expenditures G and the programs described above. The government runs a balanced budget every period.

3.4 Optimization and equilibrium

The timing of the economy is given as follows: (1) Households of age j enter a new period with asset position a, health insurance status i_{hi} and past EHI offer status η_E; (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 savings a'; (4) Firm production takes place and all markets clear.

Households The state of households can be summarized by vector $s_w = \{j, a, z, m, h, i_{hi}, i_E, \eta_E\}$ for workers and $s_r = \{j, a, m, h\}$ for retirees.[10] Let $\varphi(s)$ be the population density function of individuals at the beginning of each period and $S = \left\{\tilde{\pi}, \tilde{\phi}, r, w, T(\cdot)\right\}$ the aggregate variables. The young worker's ($j < J_R$) solves for the following optimization problem.

$$\mathbf{V}(s_w) = \max_{\left\{c, a', i'_{hi}, n, \iota\right\}} \left\{ \; u_h(c, n) - \iota\nu + \beta\rho_{h,j}\mathbb{E}\mathbf{V}(s'_w) \; \right\} \tag{5}$$

[9] We assume the social security payment to be homogeneous among individuals to reduce computational cost. As our model is mainly about health insurance among the working-age population, this simplification will has little effect on quantitative results.

[10] 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 $\{j, a, m, h, i_{hi}\}$.

subject to

$$(1+\tau_c)c + a' + \tilde{\pi}^j(i'_{hi}) + (1-\phi(m, i_{hi}))m = w\hat{z}n - c_E i_E + (1+r)a - \text{Tax} + \text{TR}, \text{ if } \iota = 0; \quad (6)$$

$$(1+\tau_c)c + \tilde{\pi}^j(i'_{hi}) = w\hat{z}n - c_E i_E + (1+r)a - \text{Tax} + \text{TR} - \lambda, \text{ if } \iota = 1; \quad (7)$$

$$a' \geqslant -\underline{a} \quad (8)$$

where

$$\tilde{\pi}^j(i'_{hi}) = \begin{cases} \pi_E(1-\psi), & \text{if } i'_{hi} = \textit{EHI}, \ i_E = 1; \\ \pi_P^j(m), & \text{if } i'_{hi} = \textit{private}; \\ 0, & \text{if } i'_{hi} = \textit{uninsured}, \ \text{ or } e \leqslant \Theta_e \text{ and } a \leqslant \Theta_a; \end{cases} \quad (9)$$

$$y = \max\{w\hat{z}n - c_E i_E + ra - i'_{hi} i_E \tilde{\pi}^j, 0\}; \quad (10)$$

$$\text{Tax} = T(y) + \tau_{mr}(y - ra) + \tau_{ss} \min\left\{y - ra, y^{ss}_{max}\right\}; \quad (11)$$

$$\text{TR} = \max\left\{0, (1+\tau_c)\underline{c} + (1-\iota)(1-\phi(m, i_{hi}))m + T(\tilde{y}) - (w\hat{z}n - c_E i_E) - (1+r)a\right\}; \quad (12)$$

$$\tilde{y} = w\hat{z}n - c_E i_E + ra. \quad (13)$$

The budget constraint (6) states that the household finances consumption c, savings a' (subject to a borrowing constraint (8)), the purchase of health insurance $\tilde{\pi}$, and out-of-pocket health expenditure $(1-\phi(m, i_{hi}))m$ using after-tax capital and labor income $w\hat{z}n - c_E i_E + (1+r)a$. Wages adjust with EHI offering. The health insurance premium varies with the policy the household chooses as in equation (9). The budget constraint (7) 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 (11). The social insurance provides a minimum consumption floor $\underline{c}$ through a lump-sum transfer which is governed by equation (12).

The retired agents solve for a similar optimization problem, except that they are enrolled in the Medicare program with a fixed premium π_{mr}. Their incomes include the social security benefit $ss(\bar{e})$ and own savings $(1+r)a$.

Health insurance company There is free entry in the health insurance market. Perfect competition implies that the insurance premium π_E equals to the expected total medical expenditure $\phi_E(m) \cdot m$ among insured and a proportional markup η associated with the administrative cost:

$$\pi_E = (1+\eta)\frac{\int i_{hi}(s) i_E \phi_E(m) m \varphi(s) ds}{\int i_{hi}(s) i_E \varphi(s) ds}. \tag{14}$$

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

$$c_E = \frac{\int \psi \pi_E i_E i_{hi}(s) \varphi(s) ds}{\int i_E z(s) g(h(s)) n(s) \varphi(s) ds}. \tag{15}$$

In the individual health insurance market, the insurer sets the premium to satisfy the zero-profit condition for each contract indexed by worker's age and health expenditure shock; that is,

$$\pi_p^j(m) = \frac{(1+\eta)\mathbb{E}\{\rho^j \phi_p(m')m' | m\}}{1+r}. \tag{16}$$

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 $\mathbf{V}(\cdot)$; 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. *All markets clear: $N = \int z(s)g(h(s))n(s)\varphi(s)ds$, $K = \int a(s)\varphi(s)ds$; $C + K' - (1-\delta)K + M + G = Y$, and (14, 16) are always satisfied;*

3. *Government's budget is balanced:*

$$\begin{aligned} & G + \int_s [ss(\bar{e}(s)) + \phi_{mr} m(s) - \pi_{mr}]\mathbf{1}_{j \geqslant j_R} \varphi(s) ds + \int_s TR(s)\varphi(s)ds \\ & + \int_s \iota(s)[(1-\phi(m))m(s) - \lambda(s)]\varphi(s)ds + \int_s \phi_{md}(m)m(s)\mathbf{1}_{e \leqslant \Theta_e, a \leqslant \Theta_a}\varphi(s)ds \qquad (17) \\ = & \int [\tau_c c(s) + T(y(s)) + \tau_{mr}(y(s) - ra) + \tau_{ss}\min\{(y(s)-ra, y^{ss}_{max}\}]\varphi(s)ds; \end{aligned}$$

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

4 Calibration and Estimation

We have described our identification strategy for health premium, as well as results estimated using the MEPS data in Section 2.2. In this section, we explain our strategy for estimating other parameters of the model. We use the MEPS data to estimate the medical expenditure shocks and parameters on health insurance. The income process, especially the causality of health on income, will be estimated by using the PSID data. Parameters governing the demographic process, welfare and taxation programs are determined directly from relevant macro data, such as life tables and tax documents. All other parameters, such as the discount factor, depreciation rate, and parameters for preference are calibrated by matching moment conditions from the data. The values of these parameters are summarized in Table 6 of Appendix C.3.

4.1 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

$$e_j^i = \tilde{w} g_j(h) z_j^i n_j^i,$$

where $g_j(h)$ captures the effect of health on income and z_j^i represents the underlying income process net of effect from health. We now describe how we 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 provides a short panel (two periods for most individuals), 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.[11] Unfortunately, we cannot use the literature's estimates of labor income processes directly as

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

they do not account for individual health dynamics. In other words, their 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 hence downward bias the income persistence parameter. In our PSID sample, a majority of households experienced transition between healthy and unhealthy throughout the observed period.

Consider the following econometric model that we use to separately identify the underlying income process z_{it} from the health component $g_j(h_{it})$. Denote y_{it} as the (log) observed income of an individual i for year t, denominated by the nominal per capita GDP of year t, and $g_j(h_{it})$ as the effect of health status on labor income. We assume

$$\begin{aligned} y_{it} &= y_{it}^* n_{it}, \\ y_{it}^* &= g_j(h_{it}) + \beta_{j1} j + \beta_{j2} j^2 + u_{it}, \\ u_{it} &= \alpha_i + \rho u_{it-1} + \varepsilon_{it}, \end{aligned} \tag{18}$$

with $\varepsilon_{it} \sim N(0, \sigma_u^2)$ and $\alpha_i \sim N(\mu_\alpha, \sigma_\alpha^2)$. y_{it}^* is the latent variable of y_{it} and hence $y_{it} = y_{it}^*$ if an individual supplies labor ($n_{it} = 1$) and $y_{it} = 0$ otherwise. The term $\beta_{j1} j + \beta_{j2} j^2$ controls for life-cycle profiles of earning.

The estimation process is, however, complicated due to the following two challenges. The first challenge is to deal with sample selection—unhealthy individuals may choose not to work and hence we do not observe their income process. We follow Wooldridge (1995) and Semykina and Wooldridge (2010) who propose a panel-data version of the Heckman correction approach (Heckman, 1974). Specifically, we estimate a probit specification of labor supply, as a function of health (h_{it}) and the age profile and then calculate the inverse Mills ratio, denoted as $\mathbb{M}_{it}$. We then explicitly control for this calculated inverse Mills ratio in the following steps of estimating income process such that our estimates are not biased by sample selection.

The second challenge is that the ordinary least square (OLS) estimator of $g_j(h_{it})$ suffers from simultaneity bias: Health affects income, and income also affects health through endogenous insurance choice. We take an instrumental variable approach to address this issue, and Appendix B.3 provides a detailed discussion on this endogeneity bias.[12]

We find that information on hospital stays of the household head can be a good instrument for health. Hospital stays directly correlate with health status, yet its correlation with income operates solely through health status. A potential concern is that economically well-off individuals may afford to having a longer hospital stay. As a robustness check, we estimate the empirical model with a binary indicator of hospital stay (taking value 1 if stayed in hospital irregardless of how long and taking value 0 otherwise). The estimated coefficient of health in the second stage regression remains similar. This result rules out the possibility that our results are driven by some outliers with extended hospital stays.

In addition, we note that hospital admissions are usually recommended and administered by doctors rather than choices of patients. It is hence unlikely that individuals with insurance choose to stay in hospital. Indeed, we find that among the 6.4 percent of individuals who ever stayed in hospital in a given year, around 70% of them are covered by some sorts of insurance while the remaining pay for their own medical bills. This fraction is comparable to health insurance coverage among the general population; if anything, the uninsured are slightly more likely to stay in hospital probably because they are likely to be unhealthy.

Based on our data, individuals who stay in hospital are not necessary the most sick ones among the unhealthy population. Compared to those who do not stay in hospital, those who do stay in hospital are indeed three times as likely to be unhealthy, but the majority of them are still healthy. Hospital stays can occur among healthy people for a variety of reasons such as childbirth.

In PSID data, we only observe both health and hospital stays for four years (1984–1987), which doesn't allow us to accurately estimate $g_j(h)$ for the full interaction of age and the health indicator. We instead restrict $g_j(h) = g(h) = \beta_H h$ 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 also use the information on hospital stays as an instrument for health when we estimate the selection

[12]The standard macro literature on health and income inequality treats health as exogenous process and does not consider the endogeneity issue. Our instrumental variable approach, admittedly with some limitations, contributes the existing literature by providing first evidence and estimate on the causality from health to income, and on the simultaneity bias.

equation described above as suggested by Semykina and Wooldridge (2010).

With the constructed inverse Mills ratio $\hat{\mathbb{M}}_{it}$ and predicted health $\hat{h}_{it}$ from the first stage and selection regressions, we rewrite our econometric model as follows:

$$\begin{aligned} y_{it} &= \beta_{\mathbb{M}}\hat{\mathbb{M}}_{it} + \beta_H \hat{h}_{it} + \beta_{a1}\text{age}_{it} + \beta_{a2}\text{age}_{it}^2 + u_{it}, \\ u_{it} &= \alpha_i + \rho u_{it-1} + \varepsilon_{it}. \end{aligned} \tag{19}$$

Note that our model in Equations (19) is now a standard income process without complications from selection and simultaneity bias. Importantly, we find that the key parameter, β_H, to be 0.468 which is significant at the one percent level. In other words, health positively affects income with a large magnitude: Being healthy on average increases income by 46.8 percent. We obtain the residual $\hat{u}_{it}$ to estimate the autocorrelation parameter ρ through a standard dynamic panel data approach, cf. Arellano and Bond (1991). We discretize u_{it} for the simulation of our baseline model. Furthermore, the effects of age, individual fixed effects α_i, and the idiosyncratic shocks ε_{it} are attributed to z_j^i, and the effect of health corresponds to $g(h)$ in the model. We relegate the details of this process to Appendix B.

4.2 Demographics, preferences, and endowments

Individuals enter the model economy at age 25 ($j = 1$) and retire at age 65 ($j = 41$ since one period in the model corresponds to one year). To contain the computational cost, we model the labor supply decision as a binary choice with $n \in \{0, 1\}$. The utility function takes the form of Constant Relative Risk Aversion (CRRA):

$$u_h(c, n) = b + \frac{c^{1-\sigma} - 1}{1 - \sigma} - [\gamma_l + \mathbf{1}_{(h=U)}\gamma_h]\frac{n^{(1+\chi_l)}}{1 + \chi_l}.$$

In line with Attanasio et al. (2011), we choose $\sigma = 3$, which implies an inter-temporal elasticity of substitution of 0.3. This σ also roughly reconciles for the fraction of individuals with health insurance. We set $\chi_l = 1.0$ to get a Frisch labor supply elasticity equals to 2.0, and let $\gamma_l = 2.7$ and $\gamma_h = 4.0$ to match the labor force participation rates of the healthy and unhealthy, respectively, for the working age population.[13] We add a constant b in utility function to guarantee the positive value of flow utility to ensure that all individuals prefer a *longer* life expectancy (e.g. Hall and Jones, 2007). The value of b is chosen such that the model-implied value of statistical life (VSL) among working age population equals to \$6.5 million, which is within the range of its empirical estimations.[14]

The survival rate ($\rho^j_{h,t}$), and therefore the mortality rate, differs between healthy and unhealthy individuals. In line with Hall and Jones (2007), we differentiate between mortality that may arise from issues unrelated to health (such as accidents and homicides/suicides) and health-related mortality (non-accident mortality). We denote them as δ_a and δ_d respectively. 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*, with details documented in Appendix C. In addition, we assume a maximum life expectancy of age 85 ($j = 61$). However, our results are not sensitive to this age of death as all retired individuals are covered by the Medicare.

We set the annual discount factor β to 0.92 to match the equilibrium interest rate of 0.04. The capital share in the production function is set to 0.33, and the depreciation rate is set to 0.06. Total factor productivity A is normalized to one in the baseline model. The borrowing limit $\underline{a}$ is set to 0.

4.3 Medical expenditure and health insurance

Medical expenditure shocks. To better estimate the medical expenditure shocks, we allow them to vary with age, health, and insurance coverage status. We also assume that they are independent and identically distributed among individuals and over time. For each age-health-insurance 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,

[13] We focus on household head in the PSID data when we calculate the labor force participation rate to be consistent with our sample in the income process estimation.

[14] In our model, the VSL is measured by the monetary value associated with a marginal reduction in mortality risk that is equivalent to prevent one death on average (or in statistical term). More specifically, we calculate $VSL = \frac{dV/d\rho}{dV/da}$. Most estimates for the VSL typically range from \$1 million–\$10 million. For example, the United States FEMA estimated the value of a statistical life at \$7.5 million in 2020.

75th, 90th, 95th, and 99th percentile of all medical expenditure observed among individuals within that cohort. We add the 95th and 99th percentile grid points to better capture the right tail of medical expenditures distribution.

Health insurance. 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. We also estimate EHI offer rate $p_E(e, \eta_E)$ non-parametrically from MEPS. We similarly fit a piecewise linear function of the probability of receiving EHI as observed in the data, varying by income levels, separately for those with or without EHI offer in the previous period.

We set the firm's share of health insurance premium ψ to 80%, which is in the empirical range of employer contribution rates. We set $\eta = 0.1$ to be consistent with Gruber (2008), who documents that the administrative cost of private insurance is about 10% of the premium. The Medicare premium π_{mr} equals to 2.1% of per capita 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 per capita GDP, and the asset threshold Θ_a to 0.538 of per capita GDP.

4.4 Government programs

Medical bankruptcy In line with Livshits et al. (2007), we set $\gamma = 0.35$ to capture the fact that agents must pay at least some fraction of their medical bills after bankruptcy. The earning exemption $\bar{y}$ equals to 20% of per capita GDP. We set the non-pecuniary penalty ν to match the fraction of population who declares medical bankruptcy in stationary equilibrium.

Consumption floor. We follow Jeske and Kitao (2009) and choose the minimum consumption floor $\underline{c}$ of \$4200 (or 0.09 of GDP per capita) such that the wealth inequality matches the data—especially, the assets of the bottom half individuals is around 10% of the assets of the individuals of the 50–90 percentiles.

Social security. The benefit payment is obtained from evaluating $\bar{e}$ from the following function representing the U.S. social security payment:

$$ss(\bar{e}) = \begin{cases} s_1\bar{e}, & \text{for } \bar{e} \leqslant \tau_1 \\ s_1\tau_1 + s_2(\bar{e} - \tau_1), & \text{for } \tau_1 < \bar{e} \leqslant \tau_2 \\ s_1\tau_1 + s_2(\tau_2 - \tau_1) + s_3(\bar{e} - \tau_2), & \text{for } \tau_2 < \bar{e} \leqslant \tau_3 \\ s_1\tau_1 + s_2(\tau_2 - \tau_1) + s_3(\tau_3 - \tau_2), & \text{for } \bar{e} > \tau_3 \end{cases} \tag{20}$$

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% as in the OECD Tax Database (OECD, 2020). 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 per capita GDP .[15] 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.

We summarize the calibration results in Table 6 of Appendix C.3. Out of all parameters, β, b, γ_l, γ_h, λ_p, $\underline{c}$, and ν are determined in the equilibrium while other parameters and functions, including health process, income process, medical expenditure shocks, mortality rates, co-insurance rates, and EHI offer rates are determined independent of the general equilibrium.

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

5 Model Fit and Validation

This section discusses the model fit for most salient features of the data. We also report statistics that are not directly targeted as a validation of our benchmark model.

Aggregate moments. We start by comparing the benchmark model to the data on the aggregate variables. Our model succeeds in matching those moments targeted directly in our calibration. For instance, the model reproduces the observed gap in the labor force participation rate between healthy and unhealthy individuals, as shown on the top two rows of Table 2. The wealth share of the bottom 50 percent individuals relative to the wealth share of the 50–90 percent individuals is 14%, compared with 10% in the data. There are slightly more people declare medical bankruptcy in the model economy, most of them use it as an implicit health insurance. The model also replicates 0.68 autocorrelation for EHI offer rate as observed in the data. In addition, we match the value of statistical life and the equilibrium interest rate, see the last two rows of Table 2.

Table 2: Model Fit: Aggregate Moments—Directly Targeted

	Model	Data
Labor force participation rate:		
Healthy	0.96	0.97
Unhealthy	0.86	0.86
Statistics on wealth inequality		
Wealth share, bottom 50 vs 50-90	0.14	0.10
Fraction of individuals with medical bankruptcy	0.005	0.004
Autocorrelation of EHI offer	0.68	0.68
Statistial value of life	\$6.6M	\$6.5M
Interest rate	0.04	0.04

Note: This table compares the model's prediction on targeted aggregate moments with data for the working age population.

As a validation of our model, we reports some moments that are not directly targeted by the benchmark economy. A comparison with the data counterparts as listed in Table 3 indicates that our model performs well in several fronts. In our model, there are 77% of working age population choose to obtain some forms of insurance coverage, which is very close to 78% based on the MEPS data. In addition, we match the decomposition into different types of insurances, see the first four rows of Table 3. In the model, 83% of individuals remain in healthy status, compared with 84% in the data. Note that health is endogenous in our model via individual's health insurance choice. As a result, matching the fraction of healthy individuals is reassuring that our health transition matrix conditional on insurance is precisely estimated. We also match the EHI take-up rate and the average out-of-pocket medical expenditure as a fraction of GDP per capita. The next four rows of Table 3 shows that our model matches the level of income inequality reasonably well. For instance, the Gini index of income is 0.45 in our model, compared with 0.43 in the data. The income share of the top 10th, 25th, and the bottom 50th percent of the income distribution is also similar to that of the data. Again, income is endogenous in our model and hence matching income inequality indicates that we correctly estimate the income process, and we get the health distribution right since health affects income.

Table 3: Model Fit: Moments Not Directly Targeted

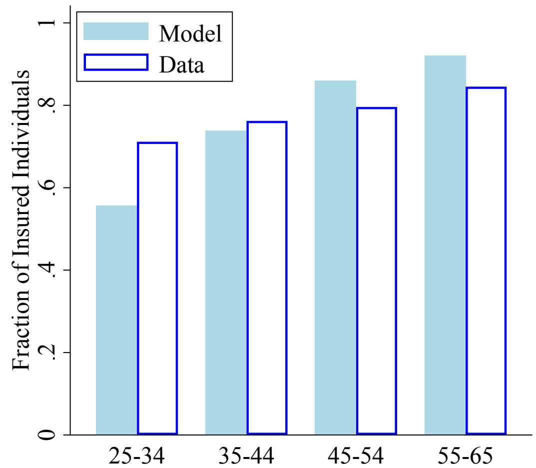
	Model	Data
Fraction of insured individuals	0.77	0.78
EHI	0.47	0.46
Other private	0.18	0.17
Public	0.12	0.14
EHI take-up rate	0.82	0.84
Fraction of healthy individuals	0.83	0.84
Out-of-pocket medical expenditure	0.016	0.014
Statistics on income inequality		
Gini index	0.45	0.43
Income share of top 10 percentile	0.34	0.38
Income share of top 25 percentile	0.58	0.67
Income share of bottom 50 percentile	0.18	0.08

Note: This table compares the model's prediction on aggregate moments with data for the working age population for moments that are not directly targeted in our calibration.

Characteristics of the uninsured. Given our focus on insurance and insurance policy, it is important that the model matches the data along the dimension of who remain uninsured. We assess the model's predictions on the insured versus the uninsured and find that they are broadly in line with the data. Specifically, in our model, the uninsured tend to be younger than the insured, with average age of 37 and 47 respectively, and this pattern also holds in the data (with average age of 41 and 45). The model also predicts that the uninsured are overall poorer, with average income at 65 percent of the economy-wide average, while the insured have average income at 111 percent of the economy-wide average. These statistics are also similar to the data (55 percent and 113 percent). The model also predicts that the uninsured are slightly less healthy—82 percent of the uninsured are healthy while 83 percent of the insured are healthy. These are again similar to the data (83 percent and 84 percent). Despite large health premium, in equilibrium we do not observe the insured to be substantially healthier than the uninsured due to the following two reasons: The unhealthy tends to be poorer and more likely to qualify for public insurance, and they are more likely to opt in EHI.

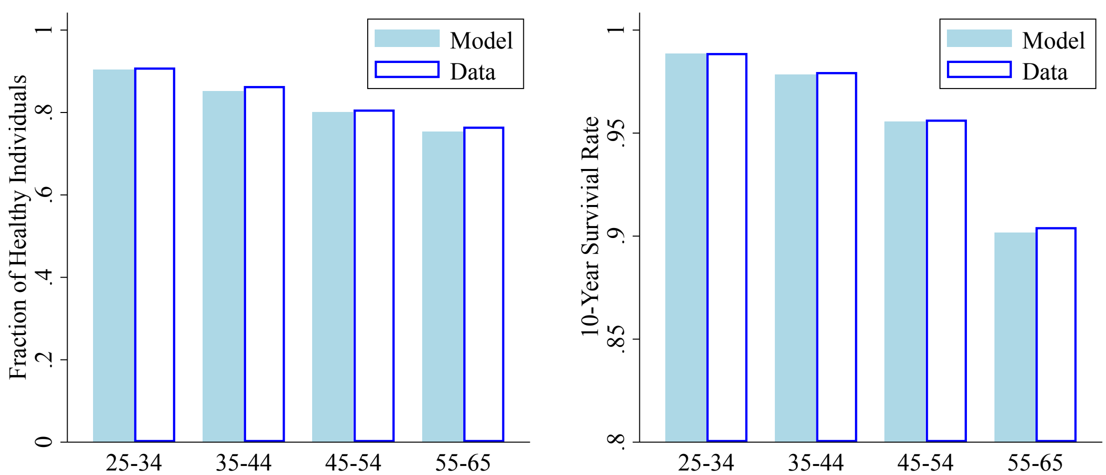
Life-cycle profiles of medical expenditure, insurance, and health. In addition to aggregate moments, our baseline model is able to account for the life cycle profiles in the data despite that we do not directly target them in our estimation. Figure 6 reports the fraction of insured individuals over the life cycle, where the light blue bars represent simulated outcomes, and the hollow blue bars represent empirical moments calculated from the data. Our baseline model replicates the rising health insurance coverage over the life cycle, as both earnings and the value of health insurance rise with age. Furthermore, the model-implied fraction of healthy individuals over the life cycle match the pattern observed in the data, see the left panel of Figure 7. Consequently, the model reproduces the declining aggregate survival rate by age, which corresponds to deteriorating health over the life cycle, see the right panel of Figure 7. Note that although we directly estimate the survival rate by health status, we allow for endogenous health in our model and hence the survival rates over the life cycle are endogenous and not directly targeted in the estimation. Figure 8 shows the model fit for average out-of-pocket medical expenditure over the life cycle. Again, medical expenditure depends on health, insurance status, and co-insurance rates, and hence is also endogenous.

Figure 6: Insurance Coverage over Life Cycle



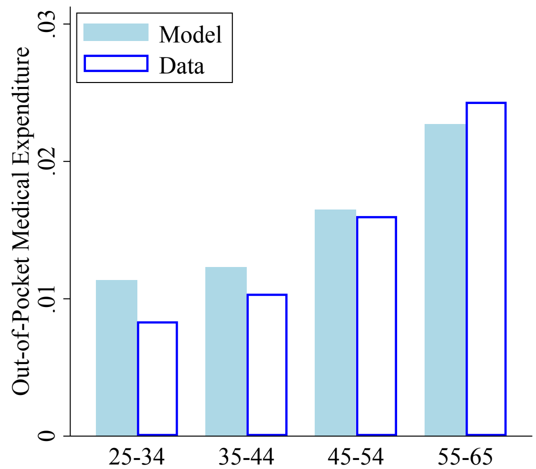
Notes: This figure illustrates the portion of insured individuals over the life cycle. The light blue bars represent simulated outcomes, and the hollow blue bars represent empirical moments calculated from the data.

Figure 7: Health Status and Survival Rates over Life Cycle



Notes: This figure illustrates the portion of healthy individuals (left panel) and the 10-year survival rates (right panel) over the life cycle. The light blue bars represent simulated outcomes, and the hollow blue bars represent empirical moments calculated from the data.

Figure 8: Out-of-Pocket Medical Expenditure over Life Cycle

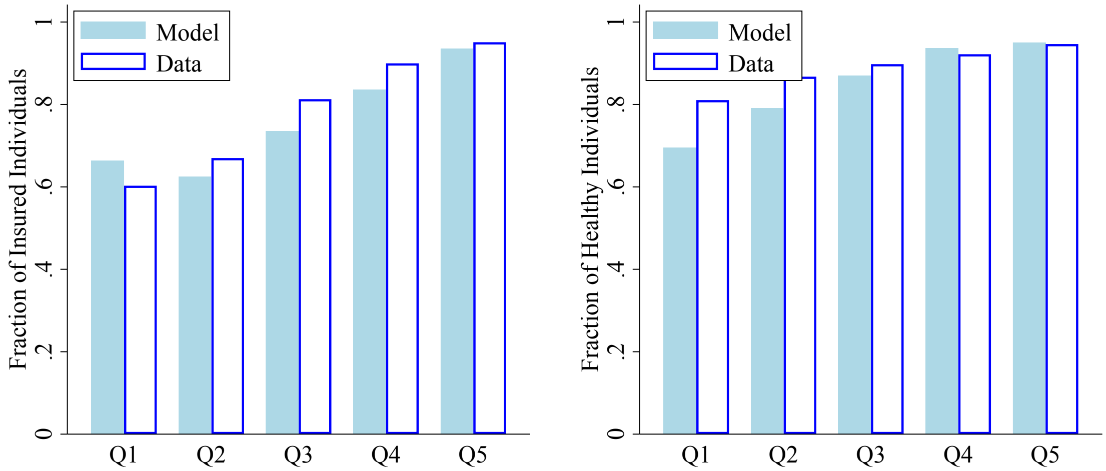


Notes: This figure illustrates the average out-of-pocket medical expenditure over the life cycle normalized by GDP per capita. The light blue bars represent simulated outcomes, and the hollow blue bars represent empirical moments calculated from the data.

Health insurance and health by income. Next, we present the health insurance coverage by income. Note that insurance choice matters to the evolution of health which affects earning and has a persistent impact on future insurance choices. The left panel of Figure 9 shows that health insurance coverage rate by income generated by the model broadly replicates what found in the data: individuals with higher income are more likely to be insured.

A key feature of our framework is that health and income are jointly determined in equilibrium. The right panel of Figure 9 shows the percentage of healthy individuals by income quantiles. The baseline model replicates the pattern that individuals with higher income are more likely to stay healthy. However, the model slightly underestimates the fraction of healthy individuals among the poor. Due to the limitation of data, our identification strategy of health premium only applies to private insurance. As a conservative estimation, we assume that the public insurance has no health premium. A relaxation of this restrictive assumption will improve the performance of the benchmark model along this health-income gradient as public insurance are mainly targeted at lower income households.

Figure 9: Insurance Coverage and Health by Income



Note: The figure shows the insurance coverage rate (left panel) and the portion of healthy individuals (right panel) over income quintiles. The light blue bars represent simulated outcomes, and the hollow blue bars represent empirical moments calculated from the data.

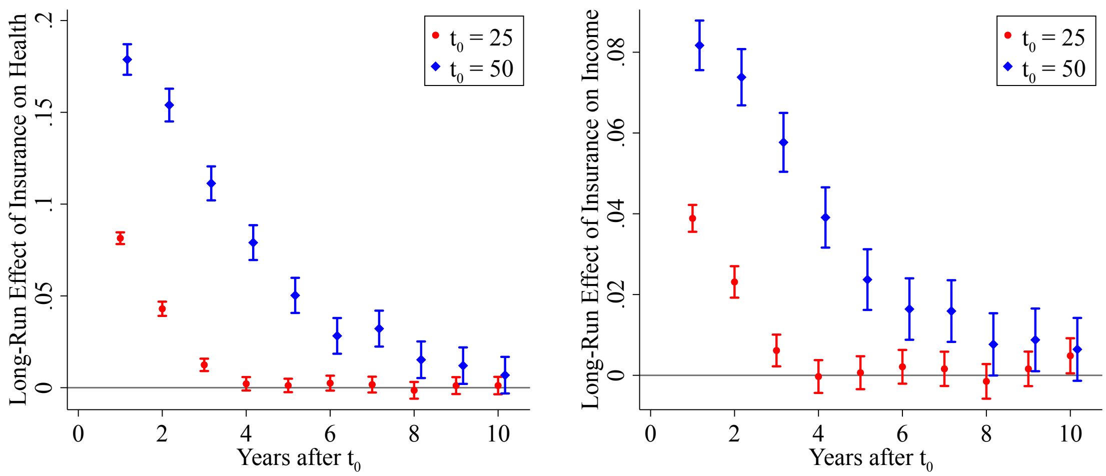
6 The Long-Run Effects of Health Insurance

In this section, we use our baseline model to analyze the long-run impacts of health insurance on individual's health and earning ability over the life cycle. Consider an individual with insurance coverage at age of 25, versus an otherwise identical individual without health insurance, how would their health differ by the age of 30? Since our baseline model reproduces most salient features of the joint distribution of health and income, it allows us to use the simulated data to study such questions. Compared with the survey data, we can simulate a panel with sufficient length but without attrition bias. Furthermore, we have better control over unobserved heterogeneity in the simulated data.

We consider the following regression based on the simulated data. We regress the variable of interest, for instance, an individual's health at the age of t, on this individual's initial health status ($h_{i,t0}$), initial health insurance coverage status ($i^{hi}_{i,t0}$), and initial income shock ($\varepsilon_{i,t0}$), controlling for this individual's type (α_i). Note that controlling for initial income shock and type is important since income is correlated with health insurance coverage, and such controls are only available with model simulated data.

The long-run impacts of health insurance at the age of 25 are shown by red bars in

Figure 10: Long-Run Effects of Health Insurance



Note: The figure shows the long-run effect of insurance on future health (left panel) and income (right panel), estimated using model simulated panel data. The red bars show the effects of insurance at the age of 25, and the blue ones show that at the age of 50. The middle dots represent point estimates, and the bands show the 95 percent confidence interval.

Figure 10. We find that a typical worker who enters the labor market at age 25 with health insurance coverage will enjoy an additional 8 percent chance to stay healthy one year later, compared with an otherwise identical but uninsured individual. This difference is due to the health premium as estimated and modeled in previous sections. Furthermore, such impact of health insurance on health will remain statistically significant within a three year window. Since healthy individuals earn more than unhealthy individuals, this health premium translates into an income difference: Individuals with health insurance at age 25 tend to have around 4 percent higher income at the age of 26, compared with an otherwise identical individual. This effect is shown by the red bars in the right panel of Figure 10.

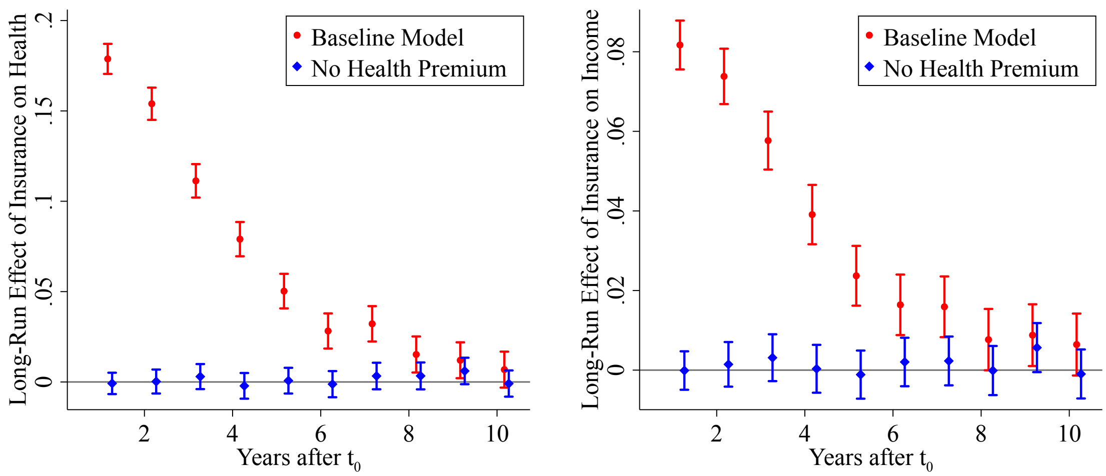
Importantly, the effects of insurance differs by age. The long-run impacts of health insurance are larger in magnitude and more persistent as workers age, see the blue bars in Figure 10. If an 50 years old individual happens to have health insurance coverage, she is around 18 percent more likely to be healthy at age 51. This effect is very persistent: One-year change in health insurance coverage yields difference in health for the subsequent 10 years. Her income is expected to be around 8 percent higher at age 51 since she is more likely to be healthy, and this effect remains significant for the following 10 years.

This heterogeneity in the effect of health insurance by age should be attributed to the rising health premium over the life cycle, see Figure 3. Intuitively, young individuals at the age of 25 are relatively healthy even without health insurance. Providing insurance to these individuals barely change their health evolution and hence their income over the life cycle. On the contrary, health premium is large for the old, and hence providing insurance to the old will substantially benefit their health and income over the life cycle. This is relevant for the policy discussion in Section 7.

We can also use this experiment to illustrate the relative importance between initial insurance and initial health on future health. Specifically, consider a 50-year old individual again. When we assess the impacts on health at the age of 51, the regression coefficient on initial insurance is 0.18 (as in the left panel of Figure 10), while the coefficient on initial health is 0.39. Initial insurance is hence roughly half as important as initial health in explaining future health. As a consequence, initial insurance is also roughly half as important in explaining future income since it affects future income only through future health.

In order to understand the value of modelling health premium and evaluate its impact on equilibrium dynamics, we consider a variation of the baseline model without such premium. To ensure the consistency of our comparison, we re-estimate the health process without conditional on insurance status and re-calibrate the model without health premium to match the health insurance take-up. Once we re-calibrate the model without health premium, these two models are able to generate roughly the same health and income inequality in equilibrium. However, they do differ along two important dimensions. First, our model predicts that initial health insurance has a persistent effect on future health and income, while initial insurance has no impact on health or income in the exogenous health model. Figure 11 illustrates this comparison for individuals at the age of 50. The second difference is that, although the baseline inequality is similar between the two models, any policy change to health insurance coverage will have quite different implications on inequality between these two models. In fact, providing universal health coverage has no impact on inequality in the model without health premium, while it substantially reduces health and income inequality in our baseline economy, which will be the focus of the next section.

Figure 11: Model Comparison



Note: The figure shows the long-run effect of insurance on future health (left panel) and income (right panel) at the age of 50, estimated using model simulated panel data. The red bars show the results for our baseline model and the blue bars show the results for the model without health premium. The middle dots represent point estimates, and the bands show the 95 percent confidence interval.

7 Quantitative Analysis: Universal Health Coverage

The analysis in the previous section indicates that health insurance has significant and persistent impacts on health and other economic outcomes. In this section, we use our baseline economy to study the macroeconomic effect of implementing "universal health coverage" (UHC). We choose this policy as it has been widely adopted by most OECD countries in various forms, the Obama-era passage of the Affordable Care Act (ACA) sought to move the U.S. closer to universal healthcare, and it was at the center of the discussion on reforming the health care system in the US, during the recent general elections. The analysis will offer valuable insights on how to design policy to mitigate the long lasting negative impact of health inequality. We assume that policy changes will be financed by consumption tax to abstract away from labor market distortion caused by income tax.

Level effects With endogenous health transition and health-dependent labor productivity, any reform to the health care system will generate level effects on aggregate demographic variables and human capital supply. As shown in Table 4, the proportion of healthy indi-

Table 4: Level Effects of Universal Health Care

	Universal Health Coverage (% Change Compared to Baseline)
Working-Age Population	+0.8
Healthy Individuals	+3.8
Human Capital (Eff. Units of Labor)	
Aggregate	+7.1
Average	+6.2

Note: This table shows the change in aggregate moments when we implement universal health care policy. 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.

viduals increases from 82.3 percent to 86.1 percent under UHC, which leads to a decline in mortality rate, and the total working-age population increases by 0.8 percent. The efficiency units of labor $\tilde{z}$ as the average human capital, which is the product of underlying productivity and health component $g(h)$ for the working age population, increases by 6.2 percent. Aggregate human capital increases by 7.1 percent, due to the increase in average human capital together with the increase in population size. Note that all these changes arises from the channel that improved insurance coverage promotes individual health, which has been largely overlooked by standard models without health premium.

Joint distribution of health and income Next, we assess the distributional impacts of UHC. The reform will provide everyone with health insurance coverage which promotes overall health. While the magnitude of benefits may differ among individuals of different age groups and income quintiles. The top panel of Table 5 shows that the poor benefit more from the UHC than the rich ones, as rich individuals are more likely to have insurance coverage in the baseline economy. The bottom panel of Table 5 shows that the UHC benefits the young more than the old. This result may seem surprising since we discuss in the previous section that insurance coverage has larger and more persistent effects among the old individuals. To reconcile these discrepancy, note that the health insurance coverage rate increases with age in the baseline equilibrium. The insurance rate is merely 56 percent among individuals aged 25–34, compared to 92 percent for the age group 55–65. UHC increases insurance coverage

Table 5: Health Inequality

	Percentage of Healthy Individuals (% Change Compared to Baseline)
By Income Quintiles:	
Q1	+8.6
Q2	+7.9
Q3	+2.7
Q4	+0.8
Q5	+0.5
By Age Groups:	
25–34	+4.7
35–44	+4.9
45–54	+3.6
55–65	+2.8

Notes: This table shows the change in the percentage of healthy individuals when we implement universal health care policy, by age groups and by income quintiles where Q1 represents the poorest individuals while Q5 represents the richest individuals.

rate by 44 percent among the youngest group and only by 8 percent among the oldest group, which explain why the policy benefit the young more than the old.

Next, we analyze the effect of UHC on income inequality, which is measured by the income share of the top 10 percent, top 25 percent, and the bottom 50 percent individuals, see Table 6. Under UHC, the income shares of the top 10 percent and top 25 percent individuals reduce slightly, while the income share of the bottom 50 percent individuals increases slightly. These numbers seem to suggest that the UHC barely changes income inequality. One might expect that the expanded insurance coverage will improve individual health and the earning ability for the poor, which reduces health disparity and income inequality. Then why do we observe little change in income inequality?

The key to understand this seemly puzzling outcome is the change in mortality rate. The UHC improves overall health of the poor more than the rich, which implies a larger drop in mortality rate among the lower income group. Hence their share in the total population will rise, and produces a countervailing effect on income inequality. In order to provide a consistent comparison in income inequality, we regress the mortality rate on log income in the baseline economy, and use this relationship to generate a hypothetical mortality rate

Table 6: Income Inequality

	Income Share (%)		
	Baseline	UHC	UHC (mortality adj.)
Top 10	33.6	33.2	31.7
Top 25	58.5	58.2	56.0
Bottom 50	18.0	18.3	20.7

Notes: This table shows the income share of the top 10 percent, top 25 percent, and bottom 50 percent individuals under the baseline economy and the UHC. Note that mortality rate changes substantially under UHC. For the column "UHC (mortality adj.)", we calculate a hypothetical level of inequality by assuming that the mortality rate remains unchanged from the benchmark economy.

for the UHC economy. We then adjust the sample weight in the UHC economy to reflect the difference between the actual and hypothetical mortality rates, which will be used to re-calculate the income shares of the top 10, top 25, and bottom 50 percent individuals. Results are listed in the last column of Table 6.

With this adjustment, we obtain the levels of inequality in the UHC economy holding the sample weight of individuals unchanged from the baseline economy. We find that the UHC has significant impact on income inequality with the income share of the top 10 percent reduces by 1.9 percent, and the income share of the bottom 50 percent increases by nearly 3 percent. To appreciate the magnitude of these changes, not that the difference in income share of the top 10 percent between the U.S. and Canada, where a UHC is in place, is around 5.5 percent. In this regard, the UHC implementation may account for one-third of difference in inequality between US and Canada. It is worth noting that all these effects on health and income inequality will disappear in a model without health premium.

To finance the UHC, the government needs to raise the consumption tax rate to 37 percent, which seems an exceptionally high number. However, one must note that all individuals and firms will be free from all other taxes for health care under this policy. More than 18 percent of the U.S. GDP is devoted to health care, while the consumption share of GDP is around two thirds. That means consumption excluding health care is around half of GDP. If we finance this health care expenditure on consumption tax alone, the tax rate would be increased by around $0.15/0.50 = 30$ percent. This explains why the implementation of UHC requires a 31 percent increase in consumption tax.

In addition, our quantitative exercise also reveals a tension of intervening the health insurance market. While providing UHC eliminates some frictions in the health insurance market (adverse selection), it may also introduce additional distortion into individual intertemporal choice. As UHC increases the life expectancy of the poor, it effectively changes their discount factor. In the absence of any other redistribution, the poor may not have enough savings for the prolonged lifespan. If it is possible, the poor would be better off by liquidating the provided insurance coverage in exchange for more consumption. In this regard, UHC could be viewed as a forced saving mechanism, requiring the poor to save in a special form of human capital, i.e., health. Our analysis suggests that good policy should balance between providing insurance against health risk and income shocks.

8 Concluding Remarks

In this paper, we study the determination of the joint distribution of health and income. We identify a ”health premium” of insurance coverage and 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 innovative 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. We also simulate the model to study the persistent effect of health insurance on individual health and their earnings over the life cycle. Quantitative analysis reveals the significant and persistent impact of an individual's initial insurance status on their health, which is reinforced by and subsequently amplifies the effect of health on labor earnings and income inequality. Finally, we use our benchmark model as a laboratory to analyze the macroeconomic implications of health care reform that is designed 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 a countervailing effect on income inequality, as the subsequent reduction in mortality for the poor and the increase of their weight in the income distribution offsets the positive effect of improved health on income inequality. Our quantitative exercise also sheds lights on 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. Subject to 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 important. 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 these 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

Acerenza, S., Bartalotti, O., and Kedagni, D. (2020). Testing identifying assumptions in bivariate probit models. *Available at SSRN*.

Aizawa, N. and Fang, H. (2020). Equilibrium labor market search and health insurance reform. *Journal of Political Economy*, forthcoming.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *The Review of Economic Studies*, 58(2):277–297.

Attanasio, O., Kitao, S., and Violante, G. L. (2011). *Financing Medicare: a general equilibrium analysis*. University of Chicago Press.

Balke, A. and Pearl, J. (1994). Counterfactual probabilities: Computational methods, bounds and applications. In *Proceedings of the Eleventh conference on Uncertainty in artificial intelligence*, pages 46–54. Elsevier.

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.

Chernozhukov, V., Lee, S., and Rosen, A. M. (2013). Intersection bounds: estimation and inference. *Econometrica*, 81(2):667–737.

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. *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 Haan, M. and Leuven, E. (2020). Head start and the distribution of long-term education and labor market outcomes. *Journal of Labor Economics*, 38(3):727–765.

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

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. *Quarterly Journal of Economics*, 127(3):1057–1106.

Fonseca, R., Michaud, P.-C., Galama, T., and Kapteyn, A. (2020). Accounting for the Rise of Health Spending and Longevity. *Journal of the European Economic Association*, 19(1):536–579.

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

Goldin, J., Lurie, I. Z., and McCubbin, J. (2019). Health insurance and mortality: experimental evidence from taxpayer outreach. *Working Paper*.

Greenwood, J., Guner, N., and Kopecky, K. A. (2022). The downward spiral. *NBER Working Paper*, 29764.

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.

Han, S. and Lee, S. (2019). Estimation in a generalization of bivariate probit models with dummy endogenous regressors. *Journal of Applied Econometrics*, 34:994–1015.

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. (1974). Shadow prices, market wages, and labor supply. *Econometrica*, 42(4):679–694.

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.

Institute of Medicine Committee on the Consequences of Uninsurance (2002). *Care Without Coverage: Too Little, Too Late*. Washington (DC): National Academies Press (US).

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.

Kédagni, D. and Mourifié, I. (2020). Generalized instrumental inequalities: testing the instrumental variable independence assumption. *Biometrika*, 107(3):661–675.

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

Mahler, L. and Yum, M. (2022). Lifestyle behaviors and wealth-health gaps in Germany. *Working Paper*.

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

Manski, C. F. (1997). Monotone treatment response. *Econometrica: Journal of the Econometric Society*, pages 1311–1334.

Manski, C. F. and Pepper, J. V. (1998). Monotone instrumental variables with an application to the returns to schooling. Technical report, National Bureau of Economic Research.

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

OECD (2020). *Consumption Tax Trends 2020: VAT/GST and Excise Rates, Trends and Policy Issues*.

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.

Pelgrin, F. and St-Amour, P. (2016). Life cycle responses to health insurance status. *Journal of Health Economics*, 49:76–96.

Pijoan-Mas, J. and Rios-Rull, J.-V. (2014). Heterogeneity in Expected Longevities. *Demography*, 51(6):2075–2102.

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

Semykina, A. and Wooldridge, J. M. (2010). Estimating panel data models in the presence of endogeneity and selection. *Journal of Econometrics*, 157(2):375–380.

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

Wooldridge, J. M. (1995). Selection corrections for panel data models under conditional mean independence assumptions. *Journal of econometrics*, 68(1):115–132.