Earnings Dynamics and Selection in Health Insurance Markets*

Yaming Cao†

November 14, 2022

Latest version can be found here.

Abstract

This paper investigates how incorporating earnings dynamics into an insurance demand model affects estimates of adverse selection in health insurance markets. Using a novel dataset that links Utah All-payer Claims Data to earnings records, this paper empirically estimates a structural demand model that jointly considers earnings dynamics and medical risks. Relative to a model in which health insurance demand only depends on medical risks, I find that incorporating earnings dynamics causes the estimated equilibrium take-up rate to increase by 5.2%, premiums to decrease by 17.2%, and deadweight loss from selection to decrease by 40%. Because earnings dynamics play an important role in the demand model, including them in the model dampens the connection between willingness to pay and insurers' average costs, reducing estimated deadweight loss from selection. Counterfactual analyses show that ignoring earnings dynamics in insurance demand leads to overestimates of the importance of subsidies as a policy tool to combat adverse selection. Targeting subsidies towards median earners reduces more deadweight loss than targeting subsidies towards low earners. These results suggest that the evaluation of public policies to reduce adverse selection can be improved by considering the joint dynamics of health and earnings.

*I am grateful to my advisors Kurt Lavetti, Bruce Weinberg, and Tamar Oostrom for their guidance and support. I would like to extend a special thanks to Alex Hollingsworth, Meta Brown, Audrey Light, and other participants of Applied Micro Seminars and Micro Lunch Workshops at The Ohio State University, seminar participants of the Health Policy Workshop at the O'Neill School of Public and Environmental Affairs, and CSWEP Health Econ/Health Policy Mentoring Workshop. This paper also benefited from discussions with Ashley Swanson, Keith Ericson, Ben Casner, and my colleagues. All errors are my own. I am grateful to the Utah Office of Health Care Statistics for the data to make the research possible.

†The Ohio State University, Department of Economics, Email: cao.851@osu.edu

I Introduction

In 2018, the federal government spent around $49 billion to provide health insurance subsidies under the premium tax credit (PTC) program, which was introduced under the 2010 U.S. healthcare reform bill.[1] One rationale for spending this enormous amount of money is that subsidies are a textbook policy response to adverse selection (Einav and Finkelstein, 2011). The eligibility and generosity of subsidy policies usually depend on people's income. However, how joint earnings and medical spending dynamics affect adverse selection is not fully understood. In contrast to many textbook models of insurance demand (Akerlof, 1970; Einav, Finkelstein, and Cullen, 2010; Einav and Finkelstein, 2011), real-world insurance demand depends on more than just medical risk. Individuals also face considerable uncertainty about their earnings and are often exposed to economic risk after experiencing negative health shocks (Dobkin et al., 2018).

Knowledge about how earnings dynamics affect adverse selection is essential for designing health insurance policies, including subsidy policies. In the "textbook" adverse selection model, subsidies reduce adverse selection by motivating healthier individuals to buy health insurance because demand only depends on medical risks. However, in models incorporating earnings dynamics, individuals' willingness to pay also depends on earnings uncertainty and its correlation with medical risks. Therefore, subsidies attract sicker people because medical risks no longer directly impact willingness to pay, as in the textbook model.

This paper aims to investigate the impact of earnings dynamics on adverse selection in health insurance markets. Using a novel dataset that links Utah All-payer Claims Data to earnings records, I empirically estimate a structural demand model that jointly considers earnings dynamics and medical risks. In this model, individuals face uncertainties about health, job mobility, earnings levels, and earnings volatility. Then, I study adverse selection by aggregating individual-level willingness to pay to the market level. I find that incorporating earnings dynamics in the model causes a weaker relationship between willingness to pay for insurance and expected medical costs. By reducing the influence of private information about medical risks in insurance markets, earnings dynamics (along with their connection to medical risks) tend to attenuate deadweight loss from adverse selection relative to models that abstract from earnings dynamics. Finally, the counterfactual simulations performed reveal evidence that subsidies are less effective in reducing deadweight loss in a model with earnings dynamics. This evidence suggests that incorporating joint earnings and medical spending dynamics could improve policy evaluations.

I begin by presenting a binary insurance choice model. Risk-averse individuals choose

[1] Federal Subsidies for Health Insurance Coverage for People Under Age 65: 2019 to 2029 — Congressional Budget Office. 2019. 2 May 2019. https://www.cbo.gov/publication/55085.

between being uninsured or fully insured to maximize expected utility. Individuals are not only uncertain about medical risks but also uncertain about earnings and the correlation between earnings and medical spending.

Conceptually, even if individuals face the same medical risk, individuals' willingness to pay for health insurance may differ if they face different earnings dynamics. The direction is theoretically ambiguous. The expected utility of being uninsured is lower for individuals whose earnings are more volatile. Thus, they have a higher incentive to purchase health insurance. At the same time, individuals with more volatile earnings are more likely to face low-resource states. A fixed nominal insurance premium reduces consumption utility by a greater amount in lower resource states, therefore causing reductions in individuals' willingness to pay for health insurance.

Furthermore, the correlation between earnings and medical spending influences willingness to pay. This discussion is motivated by the dependence between earnings and medical spending documented in the literature (e.g., Dobkin et al., 2018; Cochrane, 1991; Charles, 2003; Chung, 2013; Meyer and Mok, 2013; Poterba, Venti, and Wise, 2017; Lockwood, 2022). Individuals who face negatively correlated earnings and medical spending tend to be more willing to pay for health insurance than those whose earnings and medical spending are independent. The intuition is that the negative correlation reallocates resources from low-resource states to high-resource states, which is undesired by risk-averse individuals. However, a positive correlation can reduce the willingness to pay because it works as implicit insurance by reallocating consumption from high-resource states to low-resource states.

Savings provide an alternative mechanism to insurance for smoothing consumption over time. To incorporate this idea, I estimate optimal life-cycle savings and asset accumulation in the model. To incorporate the impacts of safety nets, I also assume a consumption floor. The literature has provided evidence of the impact of the consumption floor on demand for health insurance, including protection from bankruptcy (Mahoney, 2015) and uncompensated care (Garthwaite et al., 2015). Building on this insight, this model points out that people with different earnings dynamics receive different amounts of protection from the consumption floor. For example, an individual with higher earnings uncertainty expects a higher probability of receiving transfers from the consumption floor, implicitly reducing his incentive to purchase insurance.

To make insurance decisions, individuals first predict joint dynamics of earnings and medical spending, which are necessary for calculating expected utility of being uninsured or fully insured. I empirically estimate a model of how individuals predict joint distribution of earnings and medical spending using Utah All-payer Claims Data and earnings records derived from the UI database. The model allows the correlation between earnings and medical spending because health status affects not only medical spending prediction but also individuals' earnings prediction. I assume that people predict earnings using deterministic information, such as age, gender, and a fixed person earning type. The person earning types represent general skills or human capital levels that are rewarded equally across employers. In this model, individuals are uncertain about their health status in the next period. Conditional on each possible realization of health type, individuals predict their job mobility status and the destination firm types if they change employers. Finally, conditional on each possible type of realization, individuals are uncertain about transitory earning shocks. Individuals predict their medical spending using past medical spending, health-type transition, and whether they have chronic conditions such as diabetes and hypertension. My paper relates to the literature that models earnings dynamics using employer-employee-matched databases (Abowd et al., 1999; Addario et al., 2022; Bonhomme et al., 2019). This paper is also closely related to Blundell et al. (2020), which models the household's health and income as a transitory-permanent processes and allows the health and income shocks to be correlated. Because this dataset provides probabilistic information about the network structure of the labor market, I can model the permanent component of earnings as changes in job mobility status and employer characteristics.

To investigate how heterogenous earnings dynamics affect adverse selection in health insurance markets, I aggregate individual-level willingness to pay to a market-level analysis. Beginning with a textbook model where individuals only differ in medical risks, I sequentially add differences in assets, earning means, and earnings uncertainty. My results show a lower average willingness to pay when estimated in models incorporating earnings dynamics, especially for individuals with higher medical risks. Sicker individuals are more likely to earn less and face higher earnings uncertainty. Therefore, they expect to receive protection from the consumption floor in low-resource states. Lower average willingness to pay leads to a downward-shifted demand curve, which causes increases in adverse selection. However, because earnings dynamics reduce the influence of private information about medical risks in insurance markets, the average cost curve is steeper, which reduces adverse selection. On net, relative to a model in which demand only depends on medical risks, the estimated equilibrium take-up rate is 5.2% higher, the premium is 17.2% lower, and the deadweight loss is 40% lower. The estimated socially efficient coverage rate also drops significantly by 17% because many individuals have a lower willingness to pay than their expected medical costs, making them socially inefficient to cover. This result reveals that individual mandates may not improve welfare.

Because incorporating earnings dynamics reduces the correlation between willingness to pay and expected medical costs, uninsured individuals are not always healthier than the insured. Theoretically, subsidies can attract sicker individuals to buy health insurance, potentially increasing adverse selection. Counterfactual analyses of subsidies reveal that when applying uniform subsidies, the share of reduced deadweight loss is lower in models with earnings dynamics. Moreover, counterfactual simulations show that offering more subsidies to the low earners reduces deadweight loss less than with uniform subsidies or when targeting the median earners. Due to the negative correlation between earnings and medical spending, low earners are also more likely to be sicker. Subsidies targeting low earners can motivate these sicker people to enroll. Therefore, adverse selection is not a sufficient reason for targeting low earners in health insurance markets. Designing subsidy policies that efficiently reduce deadweight loss without hurting low earners is an important question.

Related literature. — My work shows the importance of considering the joint dynamics of earnings and medical spending. However, obtaining data on earnings and medical utilization for a sample with a wide range of heterogeneity in earnings is difficult. The literature has found ways to reduce the potential negative impact of the under-modeled joint distribution of earnings and medical spending. For example, some studies exclude the income effect by assuming the CARA utility function and incorporate a limited degree of income heterogeneity in risk preferences (e.g., Einav et al., 2013; Handel, 2013; Marone and Sabety, 2022). Some other studies follow Einav, Finkelstein, and Cullen (2010) and use price variations to estimate the willingness to pay for insurance. One advantage of this method is that it does not require the researcher to make assumptions about consumer preferences or ex-ante information. Thus the distribution of earnings is not necessary when applying this method. However, these methods limit our ability to investigate how earnings dynamics affect adverse selection.

This paper also contributes to the literature on how actual insurance markets differ from the textbook adverse selection models. For example, the multidimensional private information can cause advantageous selection (Finkelstein and McGarry, 2006; Fang et al., 2008). Other factors considered by researchers include administrative costs of providing insurance and preference heterogeneity (Einav and Finkelstein, 2011), uninsurable background risk (Doherty and Schlesinger, 1983), consumer inertia (Handel, 2013), selection on moral hazard (Einav et al., 2013), and hospital networks (Ho and Lee, 2017). My paper discusses how earnings dynamics cause adverse selection to deviate from the prediction of the textbook model. Indeed, understanding the impact of earning dynamics is uniquely essential. First, it helps to improve the design of many policies that target people based on income levels, such as subsidies and individual mandates in the Affordable Care Act (ACA). Second, modeling earnings dynamics enables us to discuss the spillovers of safety nets or labor market shocks in the health insurance market. For example, individuals' earnings dynamics might be unevenly affected by a financial crisis. How policymakers modify health insurance policies when a financial crisis occurs requires knowledge of the impact of earnings dynamics on adverse selection.

My work also relates to the growing literature on subsidy policies for health insurance — including in the ACA context (Tebaldi, 2022; Jaffe and Shepard, 2020), in the Massachusetts healthcare reform (Finkelstein et al., 2019; Aizawa and Fang, 2020), and in Medicare Part D (Decarolis, 2015; Decarolis et al., 2020). This paper contributes to this literature by studying a nearly population-level sample in Utah with rich labor and health dynamics heterogeneity, and stresses the importance of considering heterogeneous earnings dynamics among consumers when designing subsidy policies.

II Conceptual Framework

This section presents a model of individual insurance choices when earnings dynamics are incorporated. This model allows me to discuss why in models with earnings dynamics, individuals who face the same medical risk can have different willingness to pay for insurance.

II.1 A Model of Insurance Choice

This is a model of individual behavior,thus, I omit i subscripts to simplify notation. To investigate the impact of earnings dynamics on adverse selection, I discuss in Section VI how to aggregate individuals' insurance decisions into market demand. At the beginning of period t, individual i is characterized by two objects: $f(w_t, m_t)$ and A_t. The individual is uncertain about the possible pair of earnings w_t and medical spending m_t that can be realized in period t. The first, $f(w_t, m_t)$, represents the probability density function (PDF) of the joint distribution of earnings and medical spending this individual expects for the period t. I further denote the PDF of the marginal distribution of earnings and medical spending as $f(w_t)$ and $f(m_t)$, respectively. The second object is A_t, which represents the nonstochastic assets individual i holds at the beginning of period t.

Insurance choice. — Before earnings and medical spending for the period t are realized, individuals face a binary insurance choice $I_t \in \{0, 1\}$. Individuals can purchase full-coverage health insurance at p or stay uninsured. Earning reductions are not insurable in this model because health insurance will only cover the medical spending the individual can face. After choosing insurance, the earnings and medical spending for period t are realized.

I assume that individuals are risk-averse expected utility maximizers, with the von Neumann Morgenstern (vNM) utility function as $u(.)$. The individual also faces the consumption floor $\underline{c}$. Whenever his resources fall under the consumption floor, he receives a money transfer to guarantee his resources are above $\underline{c}$. Following the above assumptions, if the individual i

chooses to be uninsured ($I_t = 0$), the expected utility is:

$$EU_{I_t=0} = \int_{w_t} \int_{m_t} u(\max[A_t + w_t - m_t, \underline{c}]) f(w_t, m_t) dm dw \tag{1}$$

However, if he purchases full-coverage health insurance ($I_{it} = 1$) priced at p, his expected utility is:

$$EU_{I_t=1}(p) = \int_{w_t} u(\max[A_t + w_t - p, \underline{c}]) f(w_t) dw \tag{2}$$

The individual's willingness to pay for fully covered health insurance g_t is given by

$$g_t = \max\{p : EU_{I_t=1}(p) \geq EU_{I_t=0}\} \tag{3}$$

The individual will buy the plan if his willingness to pay is larger than or is equal to the price offer he receives from the insurers.

To estimate the individuals' willingness to pay in this binary choice model we need to estimate the joint distribution of earnings and medical spending $f(w_t, m_t)$ and assets A_t. Section IV discusses how I model and empirically estimate the joint distribution of earnings and medical spending, including how individuals predict the possible combinations of (w_t, m_t) and the probability of realizing each combination. Finally, in Section V, I introduce how to estimate the assets individuals hold when they choose insurance.

In Section VI, I aggregate the estimated individual-level willingness to pay into market-level analysis and discuss how earnings dynamics affect adverse selection. Furthermore, in Section VII, the willingness to pay estimates are applied to the counterfactual analysis of subsidies when earnings dynamics are incorporated.

II.2 How Earnings Dynamics Influence Willingness to Pay

This section discusses, why conceptually, if earnings dynamics are incorporated into the insurance choices, individuals' willingness to pay may differ even if they face the same medical risk. First, I consider three parameters of the joint distribution of earnings and medical spending: (1) the mean of earnings $\mu_w = \int w_t f(w_t) dw_t$, (2) the variance of earnings $\sigma_w^2 = \int (w_t - \mu_w)^2 f(w_t) dw_t$, and (3) the correlation between earning and medical spending ρ. I introduce the joint distribution details and how to empirically estimate them in Section IV. Second, I discuss the impact of assets and the consumption floor on individuals' willingness to pay. In the following discussion, individuals are assumed to face the same marginal distribution of medical spending.

The variance of earnings σ_w^2 has an ambiguous effect on willingness to pay. — I begin by discussing the impact of earning variance when earnings and medical spending are independent. I compare individuals who face the same earning mean but have a different earning variance. It is tempting to think that people with a higher earning variance are willing to pay more for health insurance because they face more volatile consumption. However, the impact of earning variance on the willingness to pay for health insurance is ambiguous. I explore two opposing forces: changes in (1) the expected utility of being uninsured; (2) the expected utility cost of a premium.

First, a higher earning variance leads to higher consumption volatility. [2] Risk-averse individuals derive lower expected utility from volatile consumption. Figure 1 helps to illustrate this channel. For individuals who face a higher earning variance, the uninsured option can lead to two consumption realizations with equal probability: c_1^H and c_2^H. For those whose earning variances are lower, the possible consumption realizations change to c_1^L and c_2^L. In both cases, the average consumption is $\bar{c}$. When consumption is more volatile, the expected utility is lower: $E(u(c^H)) < E(u(c^L))$. Therefore, individuals with a higher earning variance are worse off if they choose to be uninsured, making them more willing to purchase health insurance.

Second, a higher earning variance increases the expected utility cost of premiums. Individuals with a higher earning variance are more likely to face a low-resource state. In a low-resource state, a fixed nominal insurance premium reduces the consumption utility by a greater amount. Therefore, these individuals are expected to give up more utility for the same premium. The "expensive" insurance that is more expensive in terms of utility leads to a lower willingness to pay.[3]

The mean of earnings μ_w has an ambiguous effect on willingness to pay. — How the mean of the earnings affects the willingness to pay can also be explained by two opposing forces. First, individuals with a lower mean of earnings derive a lower expected utility from the uninsured choice. Therefore, they have a higher incentive to purchase health insurance.[4] Second, individuals with a lower mean of earnings consider insurance more expensive in terms of utility. [5]

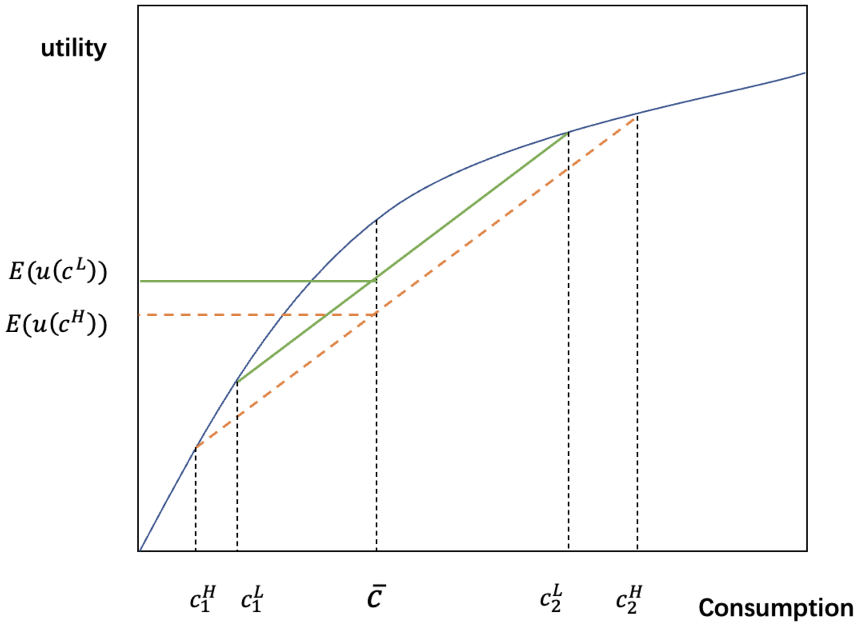
[2] To see this mathematically, the variance of consumption when choosing to be uninsured is $\sigma^2_{w-m} = \sigma^2_w + \sigma^2_m$, which increases with the variance of earning σ^2_w. This assumes that both earnings and medical spending are independent and normally distributed.

[3] When the utility function is differentiable, the utility given up to pay for insurance when the resource is w can be represented by $u'(w)$. Because individuals are risk averse, utility function $u(.)$ is concave and the marginal utility $u'(.)$ is convex. Thus, $E((u'(w^H)) > E((u'(w^L))$ holds, where w^H is the case that earning variance is higher.

[4] As illustrated in Figure A.1 in Appendix A, when facing the same level of earning volatility, if the average consumption equals $\bar{c}_L$, the expected utility is $E(u(c^L)$, which is lower than $E(u(c^H))$, which is the expected utility when the average consumption is $\bar{c}_H$.

[5] The marginal utility cost of insurance premium at earning w can be represented by $u'(w)$. Because

Figure 1: Higher earning variances lead to higher consumption volatility



Note: This figure illustrates the impact of earnings variance on the expected utility of being uninsured. For Individuals who face higher earning variance, the uninsured option can lead to two consumption realizations with equal probability: c_1^H and c_2^H. For those whose earning variances are lower, the possible consumption realizations change to c_1^L and c_2^L. In both cases, the average consumption is $\bar{c}$. When consumption is more volatile, the expected utility is lower: $E(u(c^H)) < E(u(c^L))$.

Correlation between earning and medical spending ρ — If earnings and medical spending are not independent, what would happen to the willingness to pay for health insurance? The answer is important because the literature has provided evidence that these two factors are interdependent of (Dobkin et al., 2018; Cochrane, 1991; Charles, 2003; Chung, 2013; Meyer and Mok, 2013; Poterba, Venti, and Wise, 2017).

Consumption volatility increases when earnings and medical spending are negatively correlated and decreases when the correlation is positive.[6] Moreover, this correlation does not affect the expected utility cost of insurance premiums. Therefore, a negative correlation between earnings and medical spending unambiguously increases the willingness to pay.

Intuition. — The negative correlation reallocates resources from low-resource states to high-resource states, which risk-averse individuals do not favor. However, the positive correlation increases individuals' expected utility of being uninsured because it is a form of

people are assumed to be risk averse, the utility function is concave. Therefore, $u'(w)$ decreases in w.

[6]The consumption volatility $\sigma_{w-m}^2 = \sigma_w^2 + \sigma_m^2 - \rho\sigma_w\sigma_m$ increases if earnings and medical spending are negatively correlated ($\rho < 0$) and it decreases if the correlation is positive ($\rho > 0$).

implicit insurance that reallocates resources from high-resource states to low-resource states.

Assets A_t. — Assets affect the willingness to pay because individuals buy insurance to protect assets by reducing out-of-pocket medical spending and medical debt (Finkelstein et al., 2018). Moreover, individuals who have different earnings dynamics can accumulate different levels of assets for two reasons. First, they have different saving motivations. Second, negative earning shocks can reduce wealth levels. Individuals also find it harder to accumulate wealth after persistent shocks, such as unemployment.

Consumption floor $\underline{c}$. — So far, my discussion of how earnings dynamics affect willingness to pay assumes that the consumption floor is never hit. The consumption floor transfers wealth to individuals when they face extremely low-resource states. The ability to claim bankruptcy policy is one example. The consumption floor further influences the willingness to pay for health insurance because individuals with different earnings dynamics differ in how much protection they obtain from the consumption floor. For example, individuals with a lower mean of earnings, more volatile earnings, or a negative correlation between earnings and medical spending are more likely to face states with lower resources than the consumption floor.

III Data

As discussed in Section II, earnings dynamics can theoretically affect the willingness to pay for health insurance. Thus, empirical analysis of how earnings dynamics affect adverse selection would require panel data on individual-level earnings and medical utilization. I use data from the 2013-2015 All-Payer Claims Database (APCD) that are linked to earnings records derived from the Utah unemployment insurance (UI) database. Data from the APCD provide information about the medical spending and service utilization of Utah residents from 2013 to 2015, including insurance coverage, diagnosis of patients, and medical utilization records for inpatient, outpatient, physician office visits and prescription drug consumption. I use Johns Hopkins ACG software to calculate annual health risk scores in the APCD. Researchers and commercial insurers widely use health risk scores to describe or predict patients' healthcare costs and set insurance premiums. For each worker-quarter-year, the earnings data contain the thousand-tile in which each worker's total quarterly earnings from all jobs fell. It also reports the average earnings level of all the workers in that thousand-tile of quarter-year. Moreover, the earnings file also reports a measure of compensation at the employer level, which is calculated by dividing the firm-specific total payroll in the quarter by the number of employees in the firm in the same quarter. To protect confidentiality, the average payroll at the firm-quarter-year level is reported after a white noise term is added. The white noise component is constructed by randomly drawing from a normal distribution with a mean of zero and a standard deviation of $50. All workers at the same firm in the same quarter have the same firm-level average payroll, which is rounded to the nearest cent. Blocks of coworkers have the same average payroll in each quarter over time. Therefore, this dataset provides probabilistic information about the network structure of the labor market.

Sample Selection. — To construct my sample, I begin with individuals aged 26 to 64 during the years 2013 to 2015. I restrict the sample to individuals under 65 because most people over 65 are retired and are eligible for Medicare. Furthermore, I focus on the individuals who are always enrolled in plans with a large enough number of enrollees because we can estimate the plan actuarial value, which is the percentage of total costs that insurers cover on average, only for these plans.[7] One reason to focus on always-insured workers is that I can only observe medical utilization for insured people in the data. In this model, I use their medical utilization information to study their demand for health insurance in a hypothetical market in which individuals purchase insurance for themselves. In the real world, most individuals receive insurance from their employers or they receive large subsidies from ACA markets. I further focus on the workers for whom I can construct person earning types and their employers' types. I introduce the details of these types in Section IV. Moreover, I focus on people who earned a positive amount for at least one quarter from 2013 to 2015 because people who earned nothing may have exited the labor force. Table 1 reports how the sample size changes when I step-by-step select the sample based on individuals' (1) age, (2) whether they are covered by insurance, (3) whether they are enrolled in plans with estimated plan actuarial value, (4) whether I can construct types for the workers and (5) whether the worker is employed for at least one quarter.[8]

[7]The plan actuarial value is defined for every person who is in a plan that pays out a positive number of claims to someone else (the focal person is left out).

[8]How the detailed summary statistics change with different sample selection criteria is given in Table B1 in Appendix B.

Table 1: Number of observations change with sample selection criteria

	Person observation	Person-quarter observation
Starting Sample	1,283,539	14,118,929
Age (26-65)	788,655	8,675,205
Insured	783,104	7,008,764
With plan characteristics	382,122	4,203,342
No missing types	348,146	3,829,606
Potential stayers in labor force	314,685	3,461,535

Notes: This table shows how the number of observations changes by sample selection criteria. This table reports the number of unique individuals observed and the number of person-quarter observations. I begin with the whole sample in Row 1. In Row 2, I restrict the sample to people aged 26 to 64. In Row 3, I restrict the sample to insured people in all quarters during the period 2013 to 2015. In Row 4, the sample is restricted to people who are enrolled in plans with a large enough number of enrollees. In Row 5, the sample is restricted to people who are linked to employers, which allows me to construct firm earning level types and uncertainty types (details introduced in Section IV). In Row 6, I consider the sample size change if we focus on those employed for at least one quarter from 2013-2015, which allowed me to limit the impact of including individuals who have exited the labor force.

IV Model of Earnings and Medical Spending

I assume that individuals have constant relative risk aversion (CRRA) preferences with a coefficient of 2 and face the same consumption floor when they make insurance choices. Therefore, as discussed in Section II.1, two important objects are needed to estimate willingness to pay. The first object is the joint distribution of earnings and medical spending predicted by individuals for period t. The second object is the assets individuals hold prior to making insurance choices.

In this section, I discuss the estimation of the first object. I specify a model of how individuals predict earnings and medical spending in period t and empirically estimate the model using the data introduced in Section III. The estimation of assets also requires the estimation of the joint distribution of earnings and medical spending, which is introduced in Section V.

I assume that individual i knows the PDF of the joint distribution of earning and medical spending $f(w_t, m_t)$ for the period t.[9] From Section IV.1 to Section IV.3, I discuss the estimation of each possible combination of earning and medical spending (w_t, m_t). In Section IV.4, I show how to estimate the probability of realizing each combination.

IV.1 Earning Prediction

I begin with a model of individual i's prediction of log earnings in each possible state of the world. I assume that individual i's log earning in the state of world ξ is given by

$$ln(w_t^{\xi}) = R_w(\theta_{wt}^{\xi}) + \epsilon_t^{\xi} \tag{4}$$

where $R_w(\theta_{wt}^{\xi})$ is the component of the log earning that is predicted by the type $\theta_{wt}^{\xi} \in \Theta$ that individual i faces in the world ξ.[10] The types θ_{wt}^{ξ} include demographic characteristics like age and gender, which are directly observable in the data. The types also capture individual i's health status, skills that are rewarded equally across employers, job mobility status, and employer characteristics. ϵ_t^{ξ} is the random draw of log-earning residuals. First, I introduce how I construct the health and earning types. Second, I discuss how individuals use these types to predict earnings in each possible state of the world.

[9]Because this is an individual prediction model, I omit i subscripts to simplify notation.

[10]Given the massive number of heterogenous workers and employers, predicting the probability of each world (introduced in Section IV.4) can be challenging. Therefore, I follow Abowd et al. (2019) and Bonhomme et al. (2018) using a latent-type framework.

IV.1.1 Health and Earning types

Health types. — I group the employees into four health-type categories on the basis of their annual health risk scores from 2013 to 2015. Health type takes values of 1, 2, 3, and 4, which represent people with risk scores below 1, between 1 to 1.5, between 1.5 and 2.5, and above 2.5, respectively. Because the distribution of risk scores is highly right-skewed, grouping observations into health types rather than into equal quartiles fits the data better. Moreover, because only annual risk scores are observed, I assume that individuals face the same health type in each quarter of the year.

Person earning types and Firm earning level types. — In the labor market, the agents are workers, indexed by $i \in \{1...I\}$, and employers, indexed by $j \in \{1...J\}$. I assume that on entry into the labor market, individual i samples his person earning type from six latent ability types $a_i \in \mathcal{A}$. The person earning type is interpreted as a combination of skills and other factors that are rewarded equally across employers. Likewise, employer j samples its firm earning level type from four latent types $k_j^{\mu} \in \mathcal{K}^{\mu}$. I interpret this earning as the pay premium that is paid by the employer j to all employees. One example of this premium is an efficiency wage premium.

To construct these earning types, I follow Abowd et al. (1999, also known as AKM) by estimating a linear model with the additive person and firm fixed effects. I run the following regression on the sample of workers from 2011 to 2017.

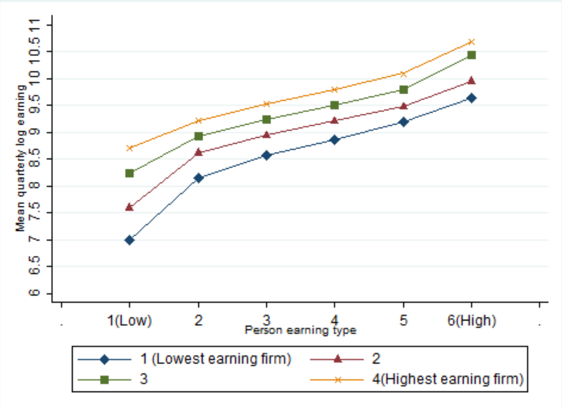
$$ln(w_{ijt}) = \alpha_i + \psi_{j(it)} + \tau_t + \eta_{ijt} \tag{5}$$

w_{ijt} stands for the observed earning of individual i who works for employer j in period t. α_i is the person fixed effects. $\psi_{j(it)}$ is the firm fixed effects. τ_t is the year-quarter fixed effects. η_{ijt} is an error component.

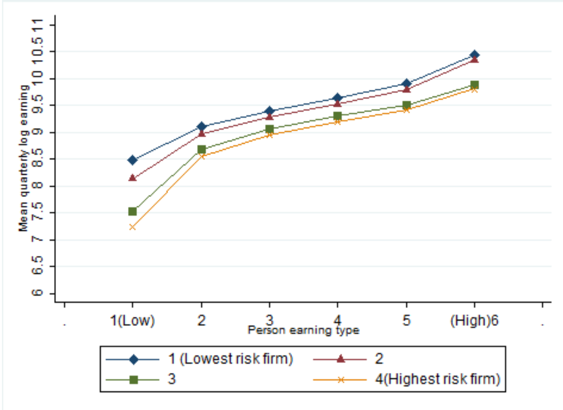
Because the two-way fixed effects model is estimated on a short-term panel from 2011 to 2017, the estimated person fixed effects, $\hat{\alpha}_i$ also reflect the impact of age on the pay premium. To adjust the age effect, I first regress $\hat{\alpha}_i$ on eight age groups in 5-year bins.[11] Then, I estimate the residuals $\tilde{\alpha}_i$ and divide the workers into six groups on the basis of the adjusted person fixed effects, with 1 standing for the lowest earning type and 6 for the highest earning type. I also divide the firms into four groups on the basis of the estimated firm fixed effects $\hat{\psi}_{j(it)}$, with type 1 standing for the lowest firm earning level type and type 4 standing for the highest firm earning level type.

[11] The 8 age groups are as follows: group 1: 26-30; group 2: 31-35; group 3: 36-40; group 4: 41-45; group 5: 46-50; group 6: 51-55; group 7: 56-60; and group 8: 61-64.

Figure 2: Mean log earning of each person type by firm earning level type and risk type



(a) By earning level types



(b) By earning risk types

Note: This figure presents the mean log earning by person type and firm type. In the left graph, I plot the mean log earnings, by person type and firm earning level type. It indicates that firms of higher earning levels type tend to pay more for workers with the same person types. Moreover, people with higher earning types also tend to earn more conditional on the firm type. In the right graph, I plot the mean log earnings, by person type and firm risk type. Conditional on person type, high-risk types tend to pay less.

Firm earning risk types. — Employer j also samples its firm earning risk type from four types $k_j^\sigma \in \mathcal{K}^\sigma$, which I interpret as a proxy for the degree of earning uncertainty that employees face inside firm j. To construct this risk type, I first calculate the log earning difference between subsequent quarters t and $t-1$ for each employee who consistently works in each firm from $t-1$ to t.

$$\Delta ln(w_{i(t-1,t)}) = ln(w_{it}) - ln(w_{i,t-1}) \tag{6}$$

Second, I calculate the standard deviation of the log earning difference for each firm j as SD_j and group the firms into four categories, each of which has an equal number of firms based on SD_j. To reduce the imprecision of standard deviation, I kept only firms that had at least 20 log earning differences during the period 2013 to 2015. The firm earning risk types can take values from 1 to 4, with 1 standing for the lowest risk type and 4 for the highest risk type.

Descriptive statistics: firm types and person earning types. — In Figure 2, I plot the mean log earnings of each person type by firm earning level types and firm earning risk types. Figures 2 (a) and (b) reveal that lower-type workers earn less on average than higher-type workers. Moreover, individuals who work in firms with higher earning levels type tend to earn more. However, workers in firms with a higher earning risk tend to earn less. Figure 3 shows that lower-earning-type firms contain a higher share of higher-risk firms. This negative correlation

Figure 3: Fraction of firms by earning risk type conditional on firm earning type

Note: This figure shows the fraction of firms by earning risk type and level type. It reveals a negative correlation between earning level type and risk type: more firms are classified as higher risk among the firms classified as low earning level types.

between earning level and risk type reveals that workers who are working in lower-earning firms also tend to face higher uncertainty about earnings.

Job mobility types. — I consider four types of job transitions between $t-1$ and t. Individual i is classified as a stayer if he stays in the same firm. If an individual i changes employer, he is considered a mover. I allow the possibility that workers move to a firm that is of same type as their original firm. If the individual was employed in t rather than $t-1$, he is classified as newly employed. Finally, an individual can not employed in t does not receive an earning more than zero.

Descriptive statistics: health types. — Table 2 reports descriptive statistics for person-quarter observations with missing earnings filled as zero.[12]The summary statistics show a potential negative correlation between earnings and medical spending. Individuals who are predicted to spend more on medical spending are also those who earn less per quarter and experience a higher probability of not being employed or changing employers.

[12]The reported number of person-quarter observations differ from Row 6 in Table 1 because Table 1 reports only the observed person-quarter pair when positive earning is observed. In Table 2, I filled the quarters with missing earnings with zero earnings.

Table 2: Summary Statistics by current health type from 2013 to 2015

	(1) All mean	(2) Health type = 1 (Healthiest) mean	(3) Health type = 2 mean	(4) Health type = 3 mean	(5) Health type = 4(Sickiest) mean
Age	43.5	41.2	47.2	47.4	47.7
Male	0.5	0.6	0.4	0.4	0.4
Quarterly Earning	15875.4	16188.4	15940.9	15465.3	14282.9
Quarterly Earning(Imputed)	14762.2	15123.8	14839.7	14304.4	12959.9
Not Employed	7.0	6.6	6.9	7.5	9.3
Stay in the same firm	96.5	96.6	96.6	96.4	95.8
Change employer	1.9	1.9	1.7	1.8	1.8
Change to earn zero amount	1.6	1.4	1.7	1.8	2.4
Change to earn positive amount	23.5	24.1	23.5	22.6	21.5
Continue to earn zero amount	76.5	75.9	76.5	77.4	78.5
Inpatient Spending(Quarterly)	178.3	3.8	54.2	123.8	1483.4
Outpatient Spending(Quarterly)	294.3	32.0	192.2	441.4	1876.5
Office Visits Spending(Quarterly)	372.5	98.6	340.3	545.4	1896.1
Pharmacy Spending(Quarterly)	217.8	27.5	164.0	377.1	1277.5
Total Medical Spending(Quarterly)	1062.8	161.9	750.7	1487.7	6533.5
Total Medical Spending(Annually)	4251.4	647.7	3002.8	5950.8	26134.1
Acg risk score	1.2	0.5	1.2	1.9	4.8
Has Diabetes	5.7	1.7	7.4	13.8	19.0
Has Hypertension	14.4	4.9	24.2	31.6	38.5
Health Plan Actuarial Value	79.5	79.0	80.0	80.6	81.2
N	3776220	2382740	562828	440348	390304

Notes: This table reports descriptive statistics for person-quarter observations in the whole sample. Column 1 is the whole sample of interest. Health types are constructed on the basis of the annual health risk score from 2013 to 2015. Health type takes values of 1, 2, 3, and 4, which represent people with risk scores below 1, between 1 to 1.5, between 1.5 and 2.5, and above 2.5, respectively. Columns 2 to 5 report the summary statistics of individuals with health types equal to 1,2,3,4, respectively. Quarterly Earning (Imputed) means that earnings in quarters with missing numbers are imputed as zero.

IV.1.2 Earning determination

As stated in Section IV.1, individual i predicts the log earning in one state of the world by summing two components: the component of the log earning predicted by the types θ_{wt} (mean of log-earnings); and a random draw of the log earnings. The log of earnings for employed workers in period t is given by

$$ln(w_{it}) = a_i\beta_a + d_{it} \times k^{\mu}_{it}\beta_d + k^{\sigma}_{it}\beta_k + H_{it}\beta_h + X_{it}\beta_x + \epsilon_{it} \tag{7}$$

a_i is the 1×6 vector that describes the category of the six person earning types the individual i samples when entering the labor market. $d_{it} \times k^{\mu}_{it}$ has a dimension of 1×12, which is an interaction term between job mobility types and the destination firm earning level type. It describes whether the individual is a job stayer, mover, or newly employed from $t-1$ to t. Moreover, it further describes, conditional on the job mobility type, the category of firm earning level type to which his destination employer belongs. k^{σ}_{it} is of dimension 1×4 and describes the firm earning risk type the individual works for in period t. H_{it} has a dimension of 1×16, which documents the health type transitions from $t-4$ to t. The vector X_{it} includes demographic characteristics, such as age groups, gender, and year-quarter dummies. β_a, β_d, β_k, and β_h are parameters with dimensions 6×1, 12×1, 4×1, and 16×1, respectively, that describe the effect on the level of log earnings associated with membership in the various heterogeneity types. The ϵ_{it} is heteroscedastic and normally distributed with mean 0. I assume that ϵ_{it} is uncorrelated with person earning type a_i, the interaction term between job mobility types and destination firm earning level types $d_{it} \times k^{\mu}_{it}$, firm earning risk types k^{σ}_{it}, health type transitions H_{it}, and time-varying covariates X_{it}.

I further assume that individuals believe that the log earning residuals are random draws from a normal distribution $N[0, var(\hat{\epsilon}(k^{\sigma}_{it}))]$, where $var(\hat{\epsilon}(k^{\sigma}_{it}))$ is the sample variance of the estimated log-earning residuals by firm earning risk type.

IV.2 Medical Spending Prediction

In this section, I discuss how individuals predict the log medical spending in each possible state of the world. The log medical spending is predicted by two components: the component of the log medical spending predicted by health types and a random draw of the log medical spending.[13] The log annual medical spending is given by[14]

$$ln(m_{iy(t)}) = \gamma_m ln(m_{i,y(t)-1}) + X_{iy(t)}\gamma_x + H_{it}\gamma_h + \phi_i\gamma_\phi + r_{iy(t)}\gamma_r + \nu_{iy(t)} \tag{8}$$

$y(t)$ denotes the year that contains the quarter of interest. $m_{i(y(t)}$ and $m_{i,y(t)-1}$ are continuous variables that represent the annual log total medical spending in year y and the past year $y-1$. $X_{iy(t)}$ includes time-varying observables, including gender, age groups, and year dummies. H_{it} has a dimension of 1×16, which documents the health type transitions from $t-4$ to t. ϕ_i is a vector that documents whether the worker has diabetes or hypertension.[15] $r_{iy(t)}$ represent the average actuarial value of their health insurance in the year y.[16] γ_m, γ_x, γ_h, γ_ϕ, and γ_r are parameters that describe the effect on the level of log medical spending associated with the lagged log medical spending, time-varying covariates, health type transitions, chronic conditions indicators, and the average actuarial values of insurance plans, respectively. $\nu_{iy(t)}$ is heteroscedastic and normally distributed with mean 0. I assume that $\nu_{iy(t)}$ is uncorrelated with lagged log medical spending $ln(m_{i,y(t)-1})$, time-varying observables $X_{iy(t)}$, health type transitions h_{it}, chronic conditions indicators ϕ_i, and the average actuarial value of their health insurance $r_{iy(t)}$.

I further assume that individuals believe that the log medical residuals are random draws from a normal distribution $N[0, var(\hat{\nu}(h_{it}))]$, where $var(\hat{\nu}(h_{it}))$ is the sample variance of the estimated log medical spending residuals by health type in t.

IV.3 Estimates of Earning and Medical Spending

In this section, I present empirical estimates for the earning determination equation (equation 7) and the medical spending prediction equation (equation 8). The first column of Table 3 reports part of the parameter estimates of the log earning equation. The reference group is females whose ages range from 26 to 30 and who are among the healthiest group in both periods $t-4$ and t. They are also stayers in firms with the lowest earning level type. Moreover, they belong to the category of the lowest person earning type and work in firms with the lowest earning risk type. The estimates reveal that workers with higher person pay premiums are predicted to earn more on average. Moreover, age group dummies, gender, and firm risk types are also important predictors of log earnings. The second column of Table 3 shows how

[13] Because only the annual health risk score is observed, I assume individuals first predict the annual medical spending for the year that contains the quarter of prediction interest. Then the quarterly medical spending prediction is obtained by multiplying the annual prediction by $\frac{1}{4}$.

[14] I add \$1 to the zero medical spending observations because the Logarithm function is undefined at zero.

[15] If I ever observe this individual receiving treatment or diagnosis of diabetes or hypertension from 2013 to 2015 in the data, I consider that this person always has diabetes or hypertension during the 3 years.

[16] An individual can change health insurance during the year of interest, so I use the quarterly average of the actuarial value of the insurances he enrolled in a year.

age, medical spending during the last year, plan characteristics, and whether the worker has diabetes or hypertension help predict the annual log medical spending. On average, people with chronic conditions like diabetes or hypertension are predicted to pay more for medical services. Lagged medical spending is also a good predictor of this year's medical spending, reflecting the persistence of medical spending. On average, individuals enrolled in plans with higher medical coverage are also predicted to spend more on medical services.

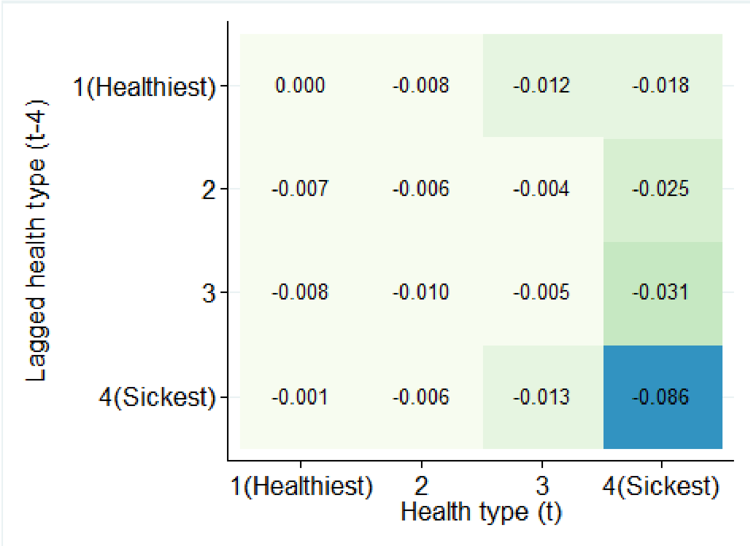
Figure 4 reveals a negative correlation between earnings and medical spending. Panel (a) reports the differences in the mean of log earning relative to the reference state, which is one of the healthiest groups in periods $t-4$ and period t. Panel (b) shows how health type transitions influence medical spending prediction. Sicker individuals are expected to face higher log medical spending.

Table 3: Some parameter estimates for log earning and medical spending equations

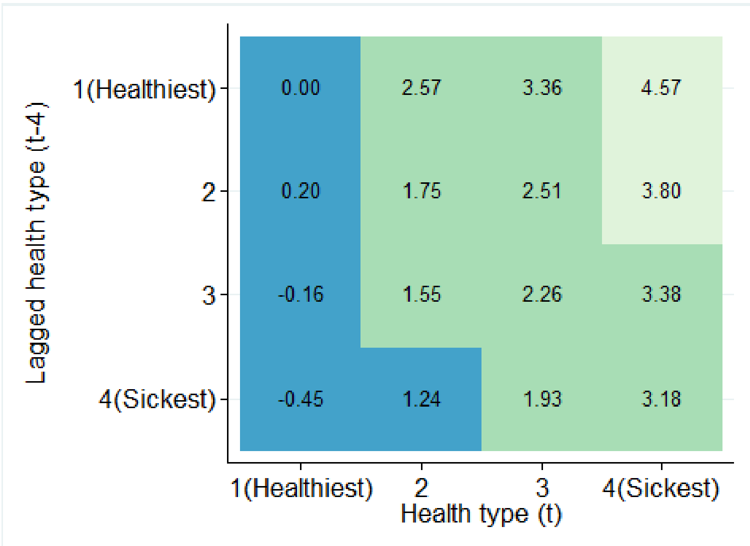
	Log Quarterly Earning	Log Annual Medical Spending
Person type 2	0.795*	
	[0.002]	
Person type 3	1.112*	
	[0.001]	
Person type 4	1.385*	
	[0.001]	
Person type 5	1.682*	
	[0.002]	
Person type 6	2.219*	
	[0.002]	
Male	0.050*	–0.431*
	[0.001]	[0.005]
Age 31-35	0.109*	0.022*
	[0.001]	[0.011]
Age 36-40	0.230*	0.051*
	[0.001]	[0.011]
Age 41-45	0.311*	0.064*
	[0.001]	[0.011]
Age 46-50	0.349*	0.048*
	[0.001]	[0.011]
Age 51-55	0.365*	–0.017
	[0.001]	[0.011]
Age 56-60	0.357*	–0.068*
	[0.001]	[0.011]
Age 61-64	0.314*	–0.030*
	[0.002]	[0.013]
Firm risk type 2	0.025*	
	[0.001]	
Firm risk type 3	–0.005*	
	[0.001]	
Firm risk type 4	–0.033*	
	[0.001]	
Last year medical spending		0.340*
		[0.002]
Plan characteristics		0.604*
		[0.026]
Has Diabetes		0.214*
		[0.006]
Has Hypertension		0.199*
		[0.004]
Constant	7.074*	2.879*
	[0.002]	[0.024]
N	3,224,032	629,370
R-Sq	0.724	0.520

Notes: Column 1 reports part of the parameter estimates of the log earning equation. The reference group is females whose ages range from 26 to 30 and whose health type transits from 1 to 1. They are also stayers in firms with earning level type 1 — the firm type with the lowest compensation level. Moreover, their person earning level is type 1 (the lowest person earning type), and they work in firms with risk type 1 (the firms with the lowest risk). Column 2 shows how age, medical spending of last year, plan characteristics, and whether the worker has diabetes or hypertension help predict the annual log medical spending.

Figure 4: Impact of health type transitions on the mean of log earnings and medical spending



(a) Log quarterly earning



(b) Log annual medical spending

Note: This figure reports the parameters of health type transitions in the earning equation (equation 7) and the medical spending equation (equation 8). Panel (a) reports the parameters of health-type transitions in the log earning equation: β_h. The reported numbers can be interpreted as the differences in the mean of log earnings relative to the reference type, which is one of the healthiest groups in both periods $t-4$ and period t. Panel (b) shows how health type transitions influence the prediction of log annual medical spending, which is expressed in the parameters γ_h in equation 8. The reported numbers can be interpreted as the differences in the mean of log medical spending relative to the reference type, which is one of the healthiest groups in both periods $t-4$ and period t.

Figure 5 reports the impact of job mobility status on the mean of log earnings. The reference type is the job stayers who continue to work in firms with the lowest firm earning level. On average, and conditional on the same job mobility status, the workers who end up moving firms with a higher earning level type earn more. Moreover, conditional on moving to firms with the same earning level type, stayers are predicted to earn more than movers and newly-employed workers. One potential reason for this is that the movers and the newly employed do not work the entire job transition quarter. Alternatively, the movers and the newly-employed workers are new to the destination firm, and, thus, they earn less because of their shorter tenure in the firm. Unfortunately, I cannot distinguish between the two possibilities because hours of working and separation reasons are unobservable in the dataset.

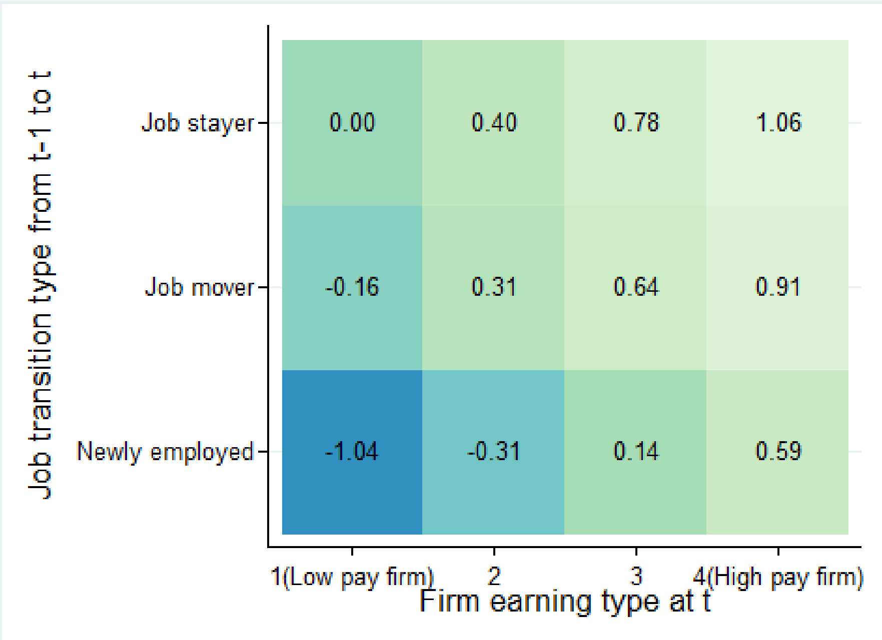
Table 4 reports the sample standard deviation of the log earning residuals and log medical spending residuals. The standard deviation of the log earning residual is higher for firms with higher earning risk types. The standard deviation of the log medical spending residuals is stable across health types except for the group predicted to face the lowest medical risk. The deviation shows that the uncertainty over the log medical spending is exceptionally high for the lowest medical spending group. Perhaps costly medical treatment for accidents, which are unpredictable with past medical utilization and diagnosis, explains the high uncertainty.

Table 4: Sample Standard deviation of the log earnings and medical spending residuals

Panel A: log earnings residuals	
Firm earning risk type	Standard deviations
1(Lowest risk)	0.35
2	0.46
3	0.58
4(Highest risk)	0.75
Average	0.49
Panel B: log medical spending residuals	
Health type	Standard deviations
1(Healthiest)	2.36
2	0.91
3	0.87
4(Sickest)	1
Average	1.95

Notes: This table reports the sample standard deviation of the log earning residuals ($\sqrt{var(\hat{\epsilon}(k^{\sigma}_{it}))}$ in equation 7) by firm earning risk type k^{σ}_{it} and log medical spending residuals ($\sqrt{var(\hat{\nu}(h_{it}))}$ in equation 8) by health type h_{it}.

Figure 5: Impact of Job Transition Types on the mean of log earning



Note: This figure reports the parameters of the interaction term between job mobility transitions from $t-1$ to t and destination firm earning level types in t ($d_{it} \times k^{\mu}_{it}$) in equation 7. The reported numbers can be interpreted as the differences in the mean of log earnings relative to the reference type: stayers who remain in firms with the lowest firm earning level type. There are three job mobility types. Stayers are those who do not switch employers from $t-1$ to t. Movers are those who change employers. New earners are those who are not employed in $t-1$ and receive paychecks in t. Firm earning level types reflect the firm compensation, which is estimated in Section 4. The firm earning level types take values of 1 to 4, with 1 standing for the lowest earning level type and 4 standing for the highest earning level type.

IV.4 Health and Employment Transitions

In this section, I discuss how individual i predicts the probability of realizing each possible combination of earning and medical spending. I assume that the individual predicts the probabilities in two steps. First, he predicts the probability of realizing the health and employment types. Second, conditional on the types realized, he predicts the probability of the random draws of log earning and log medical spending residuals.

As mentioned in Section IV.1 and Section IV.2, I assume that log earning residuals ϵ_{it} and log medical spending residuals ν_{it} are independently drawn from $N[0, var(\hat{\epsilon}(k_{it}^{\sigma}))]$ and $N[0, var(\hat{\nu}(h_{it}))]$, respectively. In practice, I discretize them using quadrature methods.

Prediction Process of Types. — The individual i first predicts health types for period t, based on his health type in period $t-4$, gender, and age. I denote the probability as $Pr(h_{it})$. Second, conditional on each possible realization of health type in period t, he predicts the job mobility types d_{it}. The probability is denoted as $Pr(d_{it}|h_{it})$. Third, conditional on health type h_{it} and job mobility types d_{it}, he predicts the probability of working in different types of firms in period t, which is denoted as $Pr(k_{it}|d_{it}, h_{it})$. The probability of realizing the types $Pr(h_{it}, d_{it}, k_{it}) = Pr(k_{it}|d_{it}, h_{it})Pr(d_{it}|h_{it})Pr(h_{it})$. In this section, I introduce how each step of this prediction process is modeled and estimated.

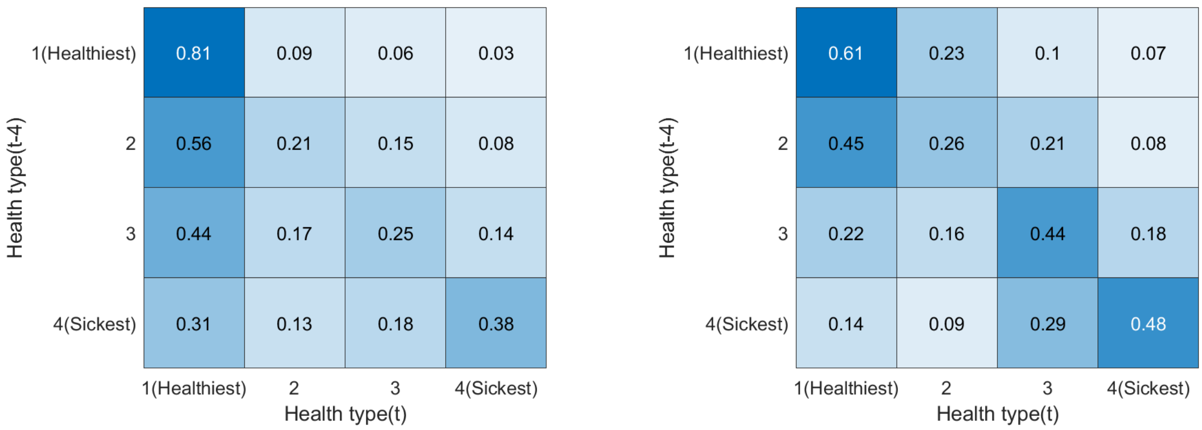
Health Type Transitions. — I allow the transition probabilities for health to depend on health type four quarters ago $h_{i,t-4}$, gender, and age group (in 5-year bins). Following Atal et al. (2021), I estimate the probabilities by fitting the transitions observed in the data from 2013-2015 into a multinomial logit model specified as:

$$\pi_{it}^{j} = \gamma_j X_{it} + \omega_j L_{i,t-4} + \lambda_j X_{it}^{age} \times L_{i,t-4} + \eta_{it}^{j} \tag{9}$$

where π_{it}^{j} represents the log odds for $h_{it} = j$. X_{it} includes the indicators for the age group individual i belongs to in period t as well as his gender. $L_{i,t-4}$ is a set of indicators for the categories of health types four quarters ago. $X_{it}^{age} \times L_{i,t-4}$ is the interactions of age groups and $L_{i,t-4}$. γ_j, ω_j, λ_j are the associated parameter vectors.

Figure 6 presents the health type transition matrices from $t-1$ to t for females aged 26 to 30 and 61 to 64. Two facts emerge from Figure 6. First, the health risk is highly persistent. If an individual transits into a sicker state, he faces a low probability of transiting out of the bad health status. Second, transition rates are highly dependent on age. The probability of remaining in the healthiest state decreases from 81% among individuals aged 25-30 to 61% among individuals above 61. Also, the probability of recovering from a sicker state to a healthier state declines with age.

Figure 6: Health Type Transitions by Age group



(a) Age: 26-30 **(b)** Age: 61-64

Note: This figure presents the health type transitions for females in two age groups: 26-30 and 61-64. The health types are constructed using ACG risk scores. People with higher risk scores are predicted to have higher medical spending. Health type takes values of 1, 2, 3, and 4, which represent people with risk scores below 1, between 1 to 1.5, between 1.5 and 2.5, and above 2.5, respectively.

Job Mobility Type Transitions. — I assume that the transition probabilities of job mobility types depend on current health types h_{it}, past job mobility types $d_{i,t-1}$, person types a_i, age groups, and gender. The transitions are estimated using a multinomial logit model, with details given in Appendix C.

Firm type transitions. — Conditional on the possible health type h_{it} and job mobility type d_{it}, individuals further predict the types of their new employers in period t if they change employers from $t-1$ to t. There is no need to predict the firm-type transitions if the individuals are stayers because the firm types will remain unchanged. I assume that for movers, the transition probabilities of firm types depend on gender, age group, person earning types, health type, past firm earning level, and risk types. Newly-employed workers predict the probabilities of their new employers by gender, age groups, and person earning types. The transitions are estimated using a multinomial logit model, with details given in Appendix C.

V Life Cycle Model with Precautionary Savings

As discussed in Section II.1, the assets that individuals hold before they make insurance decisions for period t are an important object for estimating the willingness to pay for health insurance. However, one empirical difficulty is that I do not observe assets directly in the data. In this section, I introduce how I estimate the assets if individuals face the heterogenous earnings dynamics estimated in Section IV. Earnings dynamics can cause wealth-holding inequality (De Nardi and Fella, 2017). Higher uncertainty in labor market outcomes may motivate people to save more to smooth consumption between today and tomorrow. But negative earning shocks, too, can lead to wealth reduction. In particular, the persistent attribute of negative earning shocks makes it even harder to accumulate wealth afterward.

V.1 Life Cycle Model

Individual i seeks to maximize his expected lifetime utility at tth quarter of his life after birth until the last quarter of age 100. The individuals maximize the lifetime expected utility by choosing consumption c. Each quarter, the individual's utility depends only on consumption — the flow utility from consumption is the CRRA utility function with a risk aversion parameter of 2: $u(c) = -c^{-1}$.[17]

Because assets are not directly observed in the data, I assume individuals save according to a life-cycle model and enter the labor market at age 26 with zero assets. All individuals die at age 100 and derive no utility from left-over assets after death. Therefore, there is no bequest motive in this saving model. All individuals retire at the first quarter of age 65 and begin to receive constant paychecks from social security each quarter until death.[18] Moreover, I assume individuals always follow the earning determination equation (equation 7) and the transition matrices estimated in Section IV when they predict earnings for each quarter during the life cycle.[19] I further assume individuals save as if they face no medical risk. Under these assumptions, the different saving incentives only arise from individuals' heterogenous earnings dynamics.

The next period's assets are given by:

$$A_{t+1} = A_t + \tau_t(rA_t + w_t) + b_t - c_t \tag{10}$$

Where w_t stands for earning at period t, and A_t is the asset holding at the beginning of period t. $\tau_t(rA_t + w_t)$ denotes the post-tax income, with $\tau_t(.)$ standing for a function that maps pre-tax income with post-tax income. Assets have to satisfy a borrowing constraint: $A_t \geq 0$. b_t denotes government transfers. I also assume that government transfers b_t to individuals to

[17]Individuals' utility does not depend on health status.

[18]Women receive \$3293 and men receive \$4589 per quarter. The numbers are calculated using Table 5.J3 from the Annual Statistical Supplement, 2014:
https://www.ssa.gov/policy/docs/statcomps/supplement/2014/5j.html#table5.j3

[19]This assumption indicates that people make current saving decisions without adjusting their beliefs about the following over time: (1) the person earning level types and the firm's earning level and risk types; (2) the transition matrices of job mobility and health status remain unchanged.

provide a consumption floor at $\underline{c}$.

$$b_t = \max\{0, \underline{c} - [A_t + \tau_t(rA_t + w_t)]\} \tag{11}$$

The value function for a single individual of type δ_t is given by

$$V_t(A_t, \delta_t, w_t) = \max_{c_t, A_{t+1}} \{u(c_t) + \beta s_t E_t V_{t+1}(A_{t+1}, \delta_{t+1}, w_{t+1})\} \tag{12}$$

subject to equations 10 and 11. s_t stands for the probability that an individual is alive at period $t+1$, conditional on gender and being alive at period t. w_{t+1} is the predicted earning in $t+1$ that is associated with possible type realization δ_{t+1} and random draws of log earning residuals.

When estimating, the problem is redefined in terms of cash on hand x_t to save on state variables. The problem is rewritten as follows. The value function for a single agent is given by

$$V_t(x_t, \delta_t, w_t) = \max_{c_t, x_{t+1}} \{u(c_t) + \beta s_t E_t V_{t+1}(x_{t+1}, \delta_{t+1}, w_{t+1})\} \tag{13}$$

subject to:

$$x_t = A_t + \tau(rA_t + w_t) + b_t \tag{14}$$

$$A_{t+1} = x_t - c_t \tag{15}$$

$$x_{t+1} = x_t - c_t + \tau(r(x_t - c_t) + w_{t+1}) + b_{t+1} \tag{16}$$

To enforce the consumption floor, I impose that for all t:

$$x_t \geq \underline{c} \tag{17}$$

The nonnegative assets require that:

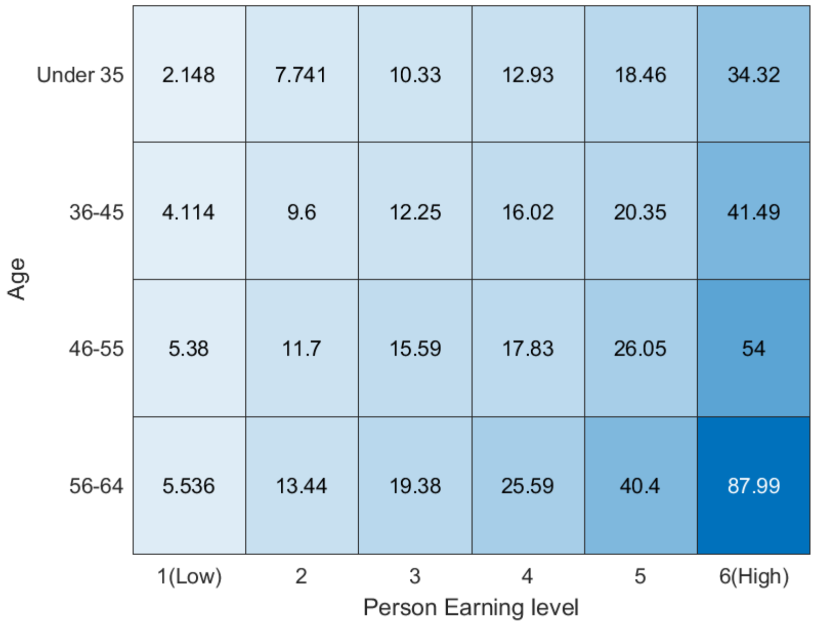
$$c_t \leq x_t \tag{18}$$

The estimation of an individual's assets in the sample follows two steps. First, I simulate for this individual one potential past path of type realizations. Second, I simulate the assets by combining the consumption-saving strategies estimated from life-cycle models and his simulated past life path.

V.2 Estimated wealth distribution

Figure 7 reports the mean of simulated assets by person-earning-level types and age groups. People who have a higher person-earning-level type tend to accumulate more assets than those with a lower person-earning-level type. Moreover, the assets accumulated tend to grow with age. This is consistent with a pattern observed in actual US wealth distribution: higher-income people and older people tend to accumulate more assets.

Figure 7: Mean Asset(000s of 2013 dollars) by the person earning level type and age group



Note: This figure reports the mean of simulated assets by the person-earning-level types and age groups. The statistics are bootstrapped 50 times with replacement and re-simulation of assets for the randomly selected 1 percent of the original sample.

Several reasons explain why the estimated wealth distribution differs from the observed net worth data in the US. First, the saving model does not allow borrowing or intergeneration transfers from parents, making it hard for individuals in the model to own homes. However, equity in home ownership is a huge component of net worth in US data. Second, this single-agent saving model does not capture the effect of marriage and children on wealth. Third, medical expenses are not considered in the saving model. Medical expenses can motivate high-income people to save (De Nardi, French, and Jones, 2010). At the same time, a huge medical bill can destroy accumulated wealth.

VI Market Aggregation and Adverse Selection

This Section aggregates the individual-level willingness to pay for health insurance to market demand to study how the incoporation of earnings dynamics affects adverse selection. In Section IV, I model and estimate how individuals predict joint earnings and medical spending dynamics for period t. Then, in Section V, I use a life-cycle saving model to estimate the assets that individuals hold at the beginning of period t. As mentioned in Section II.1, individuals' willingness to pay can be calculated using these two components.

In this Section, I focus on a randomly selected approximated 1% subsample and end up with $N = 3219$ individuals. I study how they calculate their willingness to pay for a health insurance plan that can fully cover their medical cost in the next quarter t under four models.[20] I start with the baseline model, which assumes that the only heterogeneity among the consumers is the medical spending risk. This can be considered as the textbook model with a consumption floor. I then sequentially add different sources of heterogeneity among consumers. Section VI.1 introduces the details of the four models. Section VI.2 introduces market equilibrium and social efficiency, which are applied to compare the welfare and adverse selection of the four models. Finally, in Section VI.3 and Section VI.4, I compare the four models and discuss the impact of earnings dynamics on adverse selection.

VI.1 Models of earnings dynamics

This section introduces the four models with different levels of heterogeneity in earnings dynamics among consumers. I assume that individual i predicts the same medical spending distribution for period t across the four models of interest. However, his asset level and prediction of joint dynamics of earnings and medical spending differ across models. I further assume that all individuals have CRRA utility with a risk aversion parameter of 2 and they face a consumption floor at \$2000.

I begin by introducing the model that considers all sources of heterogeneity in earnings dynamics among individuals, which I also call the "Full-heterogeneity" model (or Model 4) in the discussion that follows. I first introduce this model because it is easier to define the other models in relationship to it. To make insurance choices, the individual i, who holds A_i, predicts the joint distribution of earnings and medical spending for period t as $f(w_{it}, m_{it})$. The estimation of assets is introduced in Section V, and how the individual predicts the joint distribution of earnings and medical spending is presented in Section IV.

The baseline model, which I also call the "No-heterogeneity" model (or Model 1) in the discussion that follows, considers a sample of individuals who differ only in terms of medical risk. In this "No-heterogeneity" model, individual i holds the average asset of the sample $\bar{A}$. Moreover, he predicts that for all possible states of the world, he obtains the sample average of the mean earning $\bar{\mu}_{wt}$, which is calculated in two steps. First, individuals calculate the mean of the earning distribution they face.

$$\mu_{iwt} = \int w_{it} f(w_{it}) dw_{it} \tag{19}$$

[20] Empirically, I focus on the insurance choice problem for the second quarter year 2014.

where w_{it} is the variable that documents the earning realizations in different states. $f(w_{it})$ is the PDF of his predicted earning distribution. Second, the sample average is given as $\bar{\mu}_{wt} = \frac{1}{N}\sum_i \mu_{iwt}$.

Based on the baseline model, I add differences in assets. In this model with different assets (also called Model 2 in the following discussion), individuals now hold A_i at the beginning of period t. I also add differences in the mean of earning distribution. In this model (Model 3), individual i faces μ_{iwt} in every state of the world.

VI.2 Market Equilibrium and Social Efficiency

In this section, I introduce how to aggregate individual-level willingness to pay for health insurance to market-level analysis, including market equilibrium and social efficiency outcomes. I also introduce some statistics of interest in order to compare adverse selection and welfare across the four models discussed in Section VI.1. The statistics of interest include take-up rates for health insurance, insurance premiums, consumer surplus, and deadweight loss.

VI.2.1 Aggregation to the Market

In the health insurance market, N individuals, indexed by $i \in \{1...I\}$, must choose between a full-coverage health insurance plan and being uninsured. Individual i calculates her willingness to pay for the health insurance plan as g_i. I further assume perfect competition among insurers.

VI.2.2 Market Equilibrium

The equilibrium premium p^* is thus the price under which insurers earn zero expected profits. The equilibrium take-up rate is the share of the people enrolled in the full-coverage plans at the market equilibrium, which is given as

$$q^* = \frac{1}{N}\sum_i \mathbf{1}(g_i \geq p^*) \tag{20}$$

where $\mathbf{1}(g_i \geq p^*)$ equals 1 if $g_i \geq p^*$. The consumer surplus at the equilibrium is

$$CS^* = \frac{1}{N}\sum_i [(g_i - p^*)\mathbf{1}(g_i \geq p^*)] \tag{21}$$

Because of the zero expected profits assumption, the producer surplus is 0. Therefore, the total surplus is simply the consumer surplus.

VI.2.3 Social Efficiency

The textbook adverse selection model calculates the socially efficient take-up rate and premiums by finding the intersection between demand and marginal cost curves. However, when earnings dynamics are considered, the marginal cost curves may not be monotonic. Willingness to pay for health insurance is no longer solely determined by medical risk; it is also affected by earnings dynamics and its correlation with medical risk. Thus, an individual with a higher willingness to pay does not necessarily face higher medical risk. This non-monotonicity creates difficulty when the analyst attempts to accurately find the intersection. Consequently, I introduce two measures of socially efficient outcomes.

Social Planner Measure. — In the first of these I measure social efficiency from the perspective of a social planner who holds full information about the distribution of willingness to pay and medical costs. I consider it socially efficient to cover individuals who are willing to pay more than their expected medical costs. This measure of social efficiency cannot be obtained in general under a single pricing assumption. The socially efficient take-up rate is given as

$$q^o = \frac{1}{N}\sum_i [\mathbf{1}(g_i \geq z_i)] \tag{22}$$

The socially efficient level of consumer surplus is given as

$$CS^o = \frac{1}{N}\sum_i [(g_i - z_i)\mathbf{1}(g_i \geq z_i)] \tag{23}$$

The deadweight loss under this measure is thus the consumer surplus of those who should be efficiently covered but who remain uninsured in the competitive equilibrium.

$$DWL^o = \frac{1}{N}\sum_i [(g_i - z_i)\mathbf{1}(g_i \geq z_i)] - \frac{1}{N}\sum_i [(g_i - z_i)\mathbf{1}(g_i \geq p^*)] \tag{24}$$

Smoothed Marginal Cost Measure. — In the second measure, I assume that policymakers smooth the non-monotonic marginal cost curves and regard the intersection between the smoothed marginal cost curve with the demand curve as the socially efficient premium for welfare evaluation. I denote the socially efficient premium as p^{mo}.

The take-up rate under this measure is given by

$$q^{mo} = \frac{1}{N}\sum_i [\mathbf{1}(g_i \geq p^{mo}) \tag{25}$$

The consumer surplus under this measure is thus

$$CS^{mo} = \frac{1}{N} \sum_i [(g_i - z_i)\mathbf{1}(g_i \geq p^{mo})] \tag{26}$$

The deadweight loss under this measure is thus the consumer surplus of those willing to pay above p^{mo} but below p^*.

$$DWL^{mo} = \frac{1}{N} \sum_i [(g_i - z_i)\mathbf{1}(g_i \geq p^{mo})] - \frac{1}{N} \sum_i [(g_i - z_i)\mathbf{1}(g_i \geq p^*)] \tag{27}$$

VI.3 Results and Discussion

In this section, I present and discuss how adverse selection differs in the four models introduced in Section VI.1. I begin by examining the changes in the willingness-to-pay distribution. Then I discuss the changes in market equilibrium and social efficiency, which are introduced in Section VI.2.

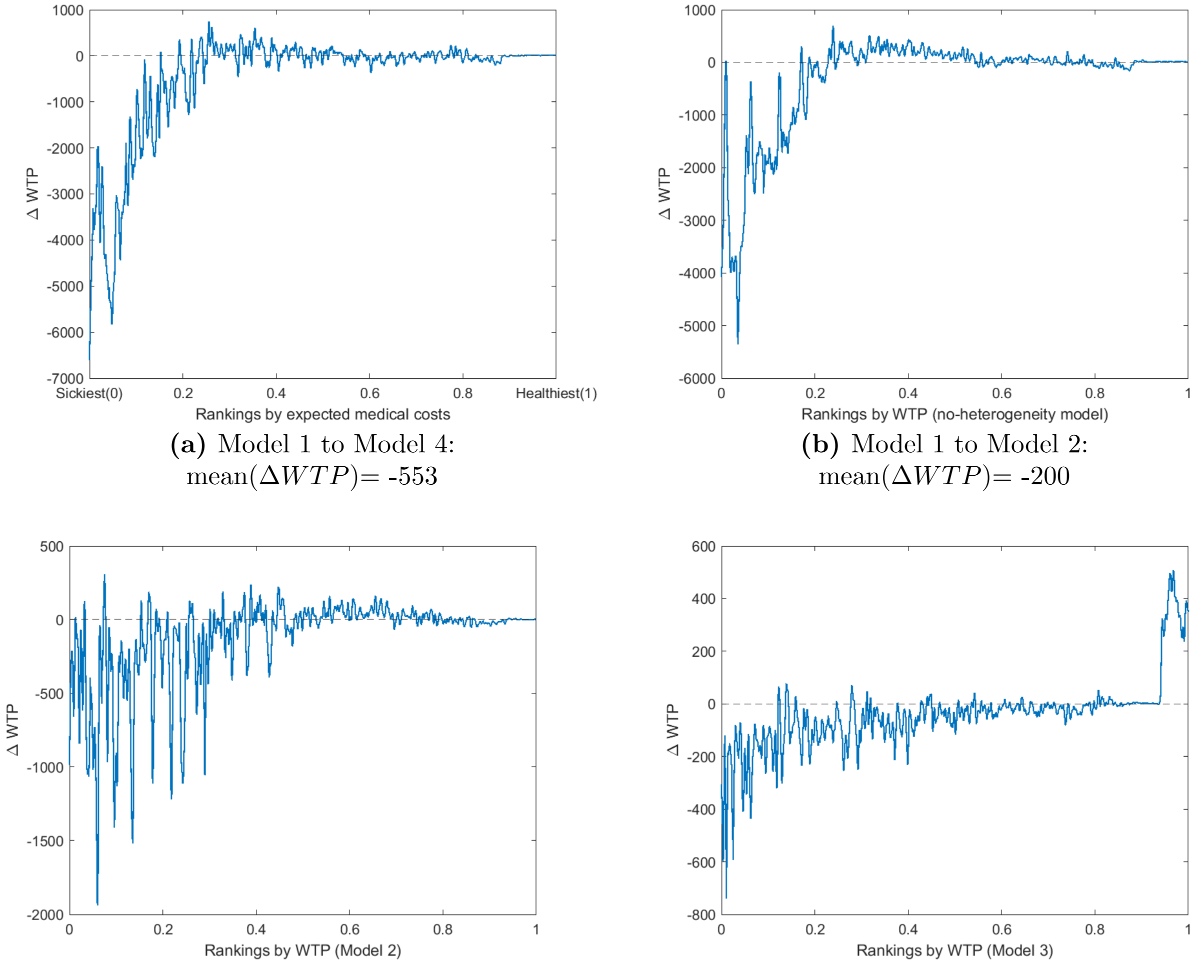
VI.3.1 Willingness-to-pay distribution

Heterogeneous changes in willingness to pay. — The changes in willingness-to-pay distribution can lead to shifts in the demand curve because, at each possible market price, the share of individuals who are willing to buy changes with the willingness-to-pay distribution. Moreover, when the correlation between willingness to pay and expected medical cost changes, the marginal cost curve may also change because individuals with higher expected medical costs may no longer be the ones with a higher willingness to pay. In Figure 8, I plot the smoothed changes in the willingness to pay when earnings dynamics are incorporated.[21]I find that, on average, ignoring earnings dynamics overestimates willingness to pay. Figure 8 also reveals heterogeneous changes in willingness to pay. Panel (a) reports that for the people who are ranked at around the first 20% of the expected medical cost distribution, their willingness to pay tends to drop after allowing different earnings dynamics. However, the willingness to pay tends to increase for people ranked from around 20% to 40%. Finally, the changes in the willingness to pay are small for people with low expected medical costs (ranked after around 60%). Therefore, Figure 8 (a) provides evidence that the willingness to pay changes could be correlated with expected medical costs.

A potential explanation of this finding is that individuals who have higher medical costs tend to receive more protection from the consumption floor, significantly reducing their willingness to pay. The negative correlation between earnings and medical risks causes some unhealthy people to earn less. Moreover, they face higher difficulty in accumulating assets

[21]The curves are smoothed with robust linear regression over each window of 20 points.

Figure 8: Willingness-to-pay changes with earnings dynamics



(a) Model 1 to Model 4: mean(ΔWTP)= -553

(b) Model 1 to Model 2: mean(ΔWTP)= -200

(c) Model 2 to Model 3: mean(ΔWTP)= -293

(d) Model 3 to Model 4: mean(ΔWTP)= -61

Note: This figure compares how the willingness to pay changes across models with different levels of heterogeneity in earnings dynamics. Four models are considered, with details introduced in Section VI.1. In Model 1, individuals only differ in terms of medical risk. In Model 2, differences in assets are added. In Model 3, individuals are assumed to face the mean of individual-level predicted earning distribution in all states of the world. In Model 4, individuals are assumed to face the predicted earning distribution. Panel (a) plots the differences in willingness to pay estimated in Model 1 and Model 4. People are ranked based on the expected medical costs in Panel (a). In Panel (b), (c), and (d), people are ranked based on willingness to pay, which is estimated in Model 1, Model 2, and Model 3, respectively. People are ranked from higher willingness to pay to lower willingness to pay.

because bad earning shocks can destroy their wealth. The persistent earning shocks, such as unemployment, make it difficult for them to re-accumulate assets. Given their bad financial situation, they are more likely to hit the consumption floor than healthier individuals who can successfully accumulate assets.

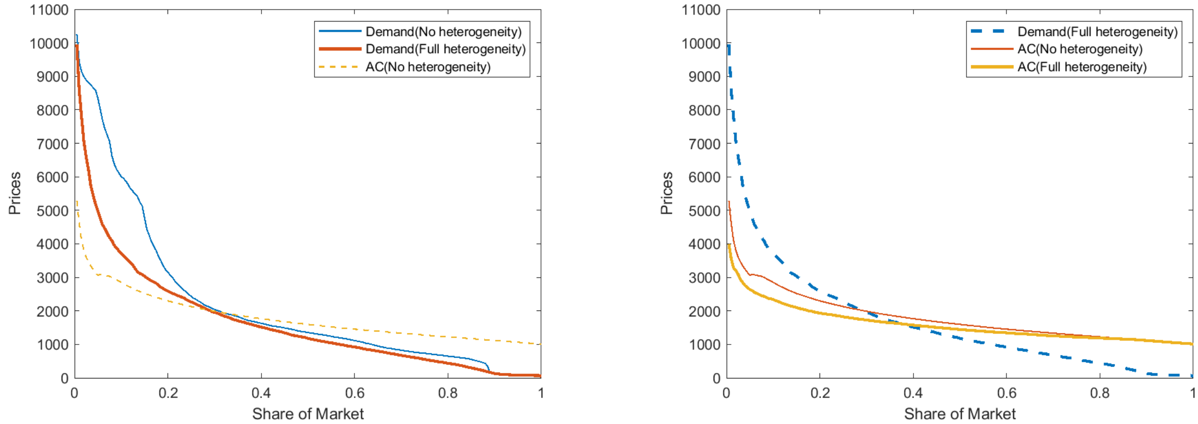
The impact of adding earning uncertainty. — The "Full-heterogeneity" model (Model 4) adds earnings uncertainties to Model 3, in which individuals certainly face the mean of the earning distribution. In Figure 8 (d), I plot the changes in willingness to pay for people ranked by their willingness to pay as estimated in Model 3. This figure reveals two interesting findings. First, ignoring the earnings uncertainty could overestimate the willingness to pay of the individuals who have a high willingness to pay in Model 3. This is so because their protection from the consumption floor is underestimated. The higher earning uncertainties increase their probability of receiving transfers from the consumption floor. Second, we could underestimate the willingness to pay for those who are considered very unwilling to purchase the plan in Model 3 (ranked after around 95%). These people have an extremely low mean of earnings and almost zero wealth. In Model 3, health insurance is of almost zero value because the probability of obtaining resources above the consumption floor is extremely low. However, in the "Full-heterogeneity" model they face a positive probability of earning enough to have resources above the consumption floor. Therefore, they consider health insurance valuable in these states.

VI.3.2 Market Equilibrium

Table 5 Panel A reports the changes in the market equilibrium relative to the baseline "No-heterogeneity" Model. When I sequentially add differences in assets, the mean of earnings, and earnings uncertainties, the equilibrium take-up rate rises and the price of the insurance drops. Compared to the "No-heterogeneity model," 5.2% more individuals enroll in the full-insurance plans, and the equilibrium premiums decrease by around \$333. Moreover, the consumer surplus per person estimated for models with earnings dynamics decreases dramatically.

The changes in the outcomes are the net impact of both changes in the demand curves and average cost curves. Figure 9 shows how demand and average cost curves change when one shifts from the "No-heterogeneity" model to the "Full heterogeneity" model. In Panel (a), I compare the changes in the demand curve while keeping the average cost curve unchanged. The demand curve shifts downward when earnings dynamics are considered: the equilibrium take-up rate decreases by around 5.5% and the price increases by \$148. Panel (b) shows that the average cost curve is steeper when earnings dynamics are incorporated: the equilibrium take-up increases by 10.6%, and the premium goes down by \$481. The impact of the changes in the average cost curve on the market equilibrium dominates the shifts in demand curves.

Figure 9: Changes in demand and average cost curves when earnings dynamics are incorporated



(a) Demand:
s^*: -0.055[0.015]]
p^*: +148[44.2]

(b) Average cost:
s^*: +0.106[0.015]]
p^*: -481[56.5]

Note: This figure shows how demand and average cost curve change when we compare the "No-heterogeneity" model, in which individuals only differ in medical risk, and the "Full-heterogeneity" model, which considers differences in predicted earning distribution and assets. Panel (a) compares the changes in demand curves. Panel (b) shows the changes in average cost curves. The changes in equilibrium take-up and premiums are reported for each step.

Steeper average cost curves. — The adverse selection implied by the textbook model is severe because the willingness to pay is a perfect predictor of individuals' medical risks. Though insurers cannot distinguish among the buyers, they understand that those with a higher willingness to pay are always more expensive to cover. However, incorporating earnings dynamics changes people's willingness to pay, making those who have higher incentives to enroll no longer always more costly. As a result, and inconsistent with the predictions of the textbook model, the marginal cost curve is no longer monotonically decreasing. The non-monotonically decreasing marginal cost curves affect the average cost curves insurers use to price the plans. At each possible price, insurers face a group of relatively healthier consumers, leading to a steeper average cost curve. Steeper average cost curves can result in a higher equilibrium take-up rate and a lower price.

VI.3.3 Social Efficiency

Significant changes happen to the estimated socially optimal outcomes. Table 5 Panel B reports the changes in socially efficient outcomes under the Social Planner Measure. This measure assumes that insurance is assigned by a social planner with full information and that individual willingness to pay above the expected medical cost characterizes those who should

Table 5: Impact of Earnings Dynamics on Market Equilibrium and Social Efficiency

Model	Description	Take-up	Price	Consumer surplus	Deadweight loss
Panel A: Equilibrium					
Model	Description	Take-up	Price	Consumer surplus	
1	**Only differ in medical risk**	**0.320**	**1941.8**	**944.9**	
	"No heterogeneity"	**(0.02)**	**(51.22)**	**(26.84)**	
2	Add different assets	+0.172	-346.8	-106.0	
		(0.01)	(37.38)	(25.24)	
3	Add different mean of earnings	+0.082	-359.6	-266.1	
		(0.02)	(40.06)	(23.59)	
4	Add earning uncertainty	+0.052	-333.1	-343.5	
	"Full heterogeneity"	(0.01)	(38.64)	(22.99)	
Panel B: Social Efficiency: Social Planner Measure					
Model	Description	Take-up	Price	Consumer surplus	Deadweight loss
1	**Only differ in medical risk**	**1.000**		**1190.9**	**246.0**
	"No heterogeneity"	**(0.00)**		**(27.88)**	**(12.29)**
2	Add different assets	+0.000		-200.0	-94.1
		(0.00)		(21.49)	(11.97)
3	Add different mean of earnings	-0.136		-370.9	-104.8
		(0.01)		(20.70)	(12.14)
4	Add earning uncertainty	-0.170		-442.5	-99.0
	"Full heterogeneity"	(0.01)		(21.86)	(10.59)
Panel C: Social Efficiency: Smoothed Marginal Cost Measures					
Model	Description	Take-up	Price	Consumer surplus	Deadweight loss
1	**Only differ in medical risk**	**0.905**	**172.0**	**1189.8**	**244.8**
	"No heterogeneity"	**(0.01)**	**(102.01)**	**(27.96)**	**(12.37)**
2	Add different assets	+0.095	-119.6	-199.0	-93.0
		(0.01)	(102.27)	(21.72)	(12.05)
3	Add different mean of earnings	-0.021	-48.9	-417.2	-151.1
		(0.04)	(120.32)	(22.69)	(12.58)
4	Add earning uncertainty	-0.088	+224.7	-534.6	-191.1
	"Full heterogeneity"	(0.06)	(173.51)	(27.12)	(15.64)

Notes: This table compares models that vary in the degree of heterogeneity in earnings dynamics. The numbers reported are the differences between Model 2-4 and the baseline model (Model 1). Details of the models are introduced in Section VI.1. Model 1 allows only heterogeneity in medical risk. Individuals are holding the average asset level of the sample $\bar{A}$ and expect to earn the sample average mean earning $\bar{\mu}_{wt}$. Model 2 add heterogeneity in assets, and individual i holds assets at A_i. Model 3 adds the heterogeneity in the mean of earnings based on Model 3. Individual i earns at μ_{iwt} in all possible states. In Model 4, the full-heterogeneity model, people are uncertain about earning, and the correlation between earning and medical spending is allowed. The medical risk distribution is kept unchanged in all models. All individuals are assumed to have the same constant relative risk aversion utility function with risk aversion at 2. The consumption floor is set at \$2000. Panel A reports the equilibrium take-up rate, premiums, and consumer surplus. To calculate the equilibrium outcomes, I assume that insurers obtain zero expected profits. Column 4 in Panel A reports the cost of public funds to raise taxes to pay for the wealth transfers when consumption hits the consumption floor. Panel B reports the socially efficient take-up rate, consumer surplus, and deadweight loss under the Social Planner Measure—people with a willingness to pay that is higher than expected medical cost are regarded as those who should be socially optimal to cover. Panel C smooths non-monotonic marginal cost curves and considers the interaction point between the smoothed marginal cost curve and demand curves as the socially efficient price. The marginal cost curves are smoothed with robust linear regression over each window of 20 points. More details can be found in Section VI.2. The reported values are the mean of 50 bootstrapped samples.

be socially efficient to cover. The socially efficient take-up rate in the "Full-heterogeneity" Model (Model 4) is 17% lower than in the "No-heterogeneity" Model (Model 1), in which consumers only differ in terms of their medical risk. Incorporating earnings dynamics into insurance choices leads to a lower estimation of socially optimal consumer surplus and deadweight loss.

The approximation of socially efficient outcomes. —In the textbook model, the intersection between the demand and the marginal cost curve gives the price that can achieve socially efficient allocations. Individuals with a willingness to pay above this price are socially efficient to cover. However, if insurance is assigned according to the interaction point between the smoothed marginal cost curve and demand curves, the outcome will differ from the outcome under the Social Planner Measure. Table 5 Panel C reports the socially efficient outcomes under the Smoothed Marginal Cost Measure. The socially efficient price p^{mo} is the intersection between smoothed marginal cost and demand curves. Individuals willing to pay more than p^{mo} are regarded as those who are efficient to cover. In models with earnings dynamics, this approximation results in an inaccurate estimate of deadweight loss — a factor that should be considered in policy designs.

VII Subsidies in models with earnings dynamics

Enormous money were spent on subsidizing low-income individuals in health insurance markets. Often subsidies were offered in order to reduce adverse selection. This section discusses how subsidy performance changes when earnings dynamics are incorporated. Intuitively, in models with earnings dynamics, individuals with a lower willingness to pay may not in all cases be healthier. Indeed, they might be unwilling to buy insurance because of their financial situation. Therefore, subsidies designed only on the basis of earnings may not reduce adverse selection in the amount that the textbook model predicts.

In this section, I investigate the changes in how much deadweight loss subsidies decline when earnings dynamics are considered. I begin by introducing how subsidies contribute to reductions in adverse selection. Then, I discuss the counterfactual analysis of subsidies.

How subsidies work. — I start by introducing how subsidies work step by step to reduce adverse selection. When the equilibrium price in the market is p^*, if an individual receives a subsidy of k_i, his price is adjusted to

$$\hat{p}_i^* = \max(p^* - k_i, 0) \tag{28}$$

Therefore, individuals face a price that is lower than or equal to the equilibrium price p^* at the status quo (before subsidies are implemented).

Through several steps subsidies affect adverse selection. First, the lower individualized prices motivate some uninsured individuals to buy insurance. Second, the average insurance cost changes because some uninsured individuals switch to being insured. According to the textbook model, the average cost goes down because the switchers are unambiguously cheaper to cover than those insured at the status quo. However, for models with earnings dynamics, the direction and the size of the changes in average cost are ambiguous. Third, the equilibrium price is adjusted according to the changes in average cost. More consumers are attracted to buying insurance if the equilibrium price drops. The process stops at a new equilibrium when insurers earn zero expected profits.

I compare the changes in equilibrium take-up rates, prices, consumer surplus, and deadweight loss before and after the subsidies. I assume that willingness to pay is the welfare-relevant metric for evaluating the welfare of the subsidy recipients. I define the new equilibrium price after the policy as p^*_{after}. The change in equilibrium price because of subsidies is given by $\Delta p^* = p^*_{after} - p^*$. Individuals now face individualized prices, which are adjusted by subsidies on the basis of this new equilibrium premium p^*_{after}. The individualized prices are given by:

$$\hat{p}^*_{i,after} = \max(p^*_{after} - k_i, 0) \tag{29}$$

The changes in equilibrium take-up rate are given by

$$\Delta q = \underbrace{\frac{1}{N}\sum_i \mathbf{1}(g_i \geq \hat{p}^*_{i,after})}_{\text{Equilibrium take-up after subsidies}(q_{after})} - \underbrace{\frac{1}{N}\sum_i \mathbf{1}(g_i \geq p^*)}_{\text{Equilibrium take-up before subsidies}(q_{before})} \tag{30}$$

The changes in consumer surplus are given by

$$\Delta CS = \underbrace{\frac{1}{N}\sum_i (g_i - \hat{p}^*_{i,after})\mathbf{1}(g_i \geq \hat{p}^*_{i,after})}_{\text{Consumer surplus after subsidies}(CS_{after})} - \underbrace{\frac{1}{N}\sum_i (g_i - p^*)\mathbf{1}(g_i \geq p^*)}_{\text{Consumer surplus before subsidies}(CS_{before})} \tag{31}$$

The deadweight loss at the status quo (before the subsidies are implemented) is given by:

$$DWL_{before} = \frac{1}{N}\sum_i (g_i - z_i)\mathbf{1}(g_i \geq z_i) - \frac{1}{N}\sum_i (g_i - z_i)\mathbf{1}(g_i \geq p^*) \tag{32}$$

The deadweight loss after the implementation of subsidies:

$$DWL_{after} = \frac{1}{N}\sum_i (g_i - z_i)\mathbf{1}(g_i \geq z_i) - \frac{1}{N}\sum_i (g_i - z_i)\mathbf{1}(g_i \geq \hat{p}^*_{i,after}) \tag{33}$$

The changes in deadweight loss are given by:

$$\Delta DWL = DWL_{after} - DWL_{before} \tag{34}$$

I further define the share of changes in deadweight loss as:

$$\Delta DWL\% = \frac{DWL_{after} - DWL_{before}}{DWL_{before}} \tag{35}$$

I also consider the social cost of raising taxes to fund the subsidies. Only individuals who are both eligible for subsidies and purchase insurance will receive subsidy payments from the policymakers. The social cost of the subsidies is given by

$$\lambda \frac{1}{N} \sum_i (p^*_{after} - \hat{p}^*_{i,after}) \mathbf{1}(g_i \geq \hat{p}^*_{i,after}) \tag{36}$$

where λ is the marginal cost of public funds. I use $\lambda = 0.3$ as the (standard estimate of) the marginal cost of public funds (e.g., Einav et al., 2010).

Subsidies reduce deadweight loss less in models with earnings dynamics than in a model that abstracts from earnings dynamics. — Table 6 reports the different impacts that uniform subsidies[22] have on adverse selection across the models that vary in heterogeneity in earnings dynamics, which I introduce in Section VII.1. The changes in equilibrium take-up rate, premiums, and consumer surplus are lower in models that incorporate earnings dynamics. The share of deadweight loss reduced is also lower. Perhaps this is so because subsidies attract sicker consumers in models that include earnings dynamics. Table 7 shows that approximately 8% of the switchers have higher expected medical costs than the average cost at the status quo. This feature suggests that when designing subsidies, it is important to consider non-monotonic marginal cost curves.

[22] My focus on uniform subsidies rather than ACA subsidies is reasonable because I study a hypothetical market that does not allow intensive margin choices between plans that provide different coverages. ACA subsidies are based on the prices of the second-lowest-cost silver plan. More details about ACA subsidies are provided in Appendix Section E.

Table 6: Impact of subsidies in models with earnings dynamics

Subsidy	Model	Description	ΔTake -up	ΔPremium	ΔConsumer Surplus	Social cost of subsidies	ΔWelfare	Δ% Deadweight Loss
400	1	Only differ in medical risk "No heterogeneity"	0.306	-521	422	75	347	−69.5%
	2	Add different assets	0.181	-237	369	81	288	−66.2%
	3	Add different mean of earnings	0.198	-250	321	72	249	−57.2%
	4	Add earning uncertainty "Full heterogeneity"	0.200	-241	297	68	228	−46.2%
800	1	Only differ in medical risk "No heterogeneity"	0.558	-804	946	211	735	−99.2%
	2	Add different assets	0.371	-440	836	207	629	−96.5%
	3	Add different mean of earnings	0.393	-440	736	191	546	−71.3%
	4	Add earning uncertainty "Full heterogeneity"	0.447	-430	705	196	509	−34.6%

Notes: This table reports the changes in equilibrium outcomes after implementing two levels of uniform subsidies: \$400 and \$800. Four models are considered. In Model 1, the individuals only differ in medical risk. In Model 2, differences in assets are added. In Model 3, individuals are assumed to face the mean of individual-level predicted earning distribution in all states of the world. In Model 4, individuals are assumed to face the predicted earning distribution. Details of these models are introduced in Section VI.1. When the equilibrium price in the market is p^*, an individual's price is adjusted to $\hat{p}_i^* = \max(p^* - k_i, 0)$ if he receives a subsidy of k_i. The consumer surplus per person is calculated as $\frac{1}{N}\sum_i(g_i - \hat{p}^*_{i,after})\mathbf{1}(g_i \geq \hat{p}^*_{i,after})$, where g_i is the willingness to pay, and $\hat{p}^*_{i,after}$ is the individualized premium at the new equilibrium price p^*_{after}. The social cost of subsidies is calculated using $\lambda = 0.3$ as the estimate of the marginal cost of public funds, which is given by $\lambda\frac{1}{N}\sum_i(p^*_{after} - \hat{p}^*_{i,after})\mathbf{1}(g_i \geq \hat{p}^*_{i,after})$. Welfare per person is the difference between consumer surplus and the social cost of the subsidy. Deadweight loss for the status quo (before subsidy) is calculated as $DWL_{before} = \frac{1}{N}\sum_i(g_i - z_i)\mathbf{1}(g_i \geq z_i) - \frac{1}{N}\sum_i(g_i - z_i)\mathbf{1}(g_i \geq p^*)$, where z_i is the expected medical costs. The deadweight loss after the subsidy is calculated as $DWL_{after} = \frac{1}{N}\sum_i(g_i - z_i)\mathbf{1}(g_i \geq z_i) - \frac{1}{N}\sum_i(g_i - z_i)\mathbf{1}(g_i \geq \hat{p}^*_{i,after})$. The changes in deadweight loss due to subsidy is: $\Delta DWL = DWL_{after} - DWL_{before}$. The share of deadweight loss change is thus defined as $\frac{\Delta DWL}{DWL_{before}}$. The reported values are the mean of 50 bootstrapped samples.

Table 7: Share of Newly Insured with Higher than Average Medical Costs

		Subsidy (dollars)			
Model	Description	400	800	1000	1200
1	Only differ in medical risk No-heterogeneity	0	0	0	0
2	Add different assets	0	0	0	0
3	Add different mean of earnings	0.046 (0.012)	0.031 (0.006)	0.031 (0.006)	0.032 (0.006)
4	Add earning uncertainty Full-heterogeneity	0.076 (0.014)	0.078 (0.010)	0.078 (0.010)	0.079 (0.009)

Notes: This table reports the share of switchers whose expected costs are higher than the average cost in the status quo (before subsidy) after implementing uniform subsidies. When the equilibrium price in the market is p^*, an individual's price is adjusted to $\hat{p}_i^* = \max(p^* - k_i, 0)$ if he receives a subsidy of k_i. The switchers are those who are willing to pay less than the status quo equilibrium premium but are willing to buy the insurance and face the same premium after each uniform subsidy design. The average cost in the status quo is estimated as the mean of those insured before subsidies are implemented. The reported values are the mean of 50 bootstrapped samples.

Offering more subsidies for low earners reduces deadweight loss less than targeting median earners. — In Table 8, I compare the welfare impact of three subsidy designs on the "Full-heterogeneity" Model, in which individuals differ in terms of their joint dynamics of earnings and medical spending. The three subsidies offer the same average subsidy to the sample but target different groups through their earning levels. Individuals are divided into three equal groups on the basis of their expected earnings: high earners, median earners, and low earners. In the uniform subsidy design, all individuals are offered a subsidy of \$400 if they purchase the plan. The second subsidy design offers zero subsidies to the highest-earning group and increases the subsidy to the low-earning group to \$800. The final subsidy offers \$800 to the median-earning group, while the individuals who belong to the low-earning group receive only \$400.

Before the equilibrium price adjustment, we see similar changes in the subsidy take-up rates that target low and median earners. The take-up rate changes in the uniform subsidy design are around 1% lower than the subsidies that target groups by earnings. However, the average cost of the insured is significantly lower if the subsidy targets the median-earning group. As mentioned earlier, lowering the average cost by attracting healthier new enrollees is an important step that leads to a lower equilibrium price. Panel B reports the outcomes after adjusting the price to the new equilibrium. Results show the importance of considering whom to target in subsidy designs. Offering more subsidies to the median earners reduces subsidies more than offering subsidies that target the lower-earning group: it leads to a higher equilibrium take-up rate, consumer surplus, and a more significant price reduction. The

Table 8: Subsidy Design Targeted by Earnings

Panel A: Equilibrium Price not Adjusted			
	Uniform	Low-earning	Median-earning
Take-up	0.485	0.494	0.497
Average cost	1460	1508	1435
Panel B: Equilibrium Price adjusted			
	Uniform	Low-earning	Median-earning
Take-up	0.572	0.557	0.591
Premium	1368	1445	1337
Consumer surplus	898	848	931
Social cost of subsidy	68	71	81
Welfare per person	829	777	849
Δ Deadweight loss	-68	-10	-58

Notes: This table compares three counterfactual subsidy designs on the "Full-heterogeneity" Model, which considers differences in medical risk, assets, mean of earnings, and earning uncertainties. When the equilibrium price in the market is p^*, the price is adjusted to $\hat{p}_i^* = \max(p^* - k_i, 0)$ for an individual who receives a subsidy of k_i. Individuals are divided into three groups by their expected earnings: high-earning, median-earning, and low-earning. Column 1 reports results for a uniform subsidy design at $400. Column 2 reports the results for a subsidy that targets the low-earning group with a subsidy at $800 and offers a subsidy to the median-earning group at $400. Column 3 reports the results for a subsidy that targets the median-earning group with a subsidy at $800 and offers a subsidy to the low-earning group at $400. Panel (A) reports the outcomes when the equilibrium price is not adjusted. Panel (B) reports the equilibrium outcomes. The consumer surplus per person is calculated as $\frac{1}{N}\sum_i (g_i - \hat{p}^*_{i,after})\mathbf{1}(g_i \geq \hat{p}^*_{i,after})$, where g_i is the willingness to pay, and $\hat{p}^*_{i,after}$ is the individualized premium at the new equilibrium price p^*_{after}. The social cost of subsidies is calculated using $\lambda = 0.3$ as the estimate of the marginal cost of public funds, which is given by $\lambda \frac{1}{N}\sum_i (p^*_{after} - \hat{p}^*_{i,after})\mathbf{1}(g_i \geq \hat{p}^*_{i,after})$. Welfare per person is the difference between consumer surplus and the social cost of the subsidy. Deadweight loss for the status quo (before subsidy) is calculated as $DWL_{before} = \frac{1}{N}\sum_i (g_i - z_i)\mathbf{1}(g_i \geq z_i) - \frac{1}{N}\sum_i (g_i - z_i)\mathbf{1}(g_i \geq p^*)$, where z_i is the expected medical costs. The deadweight loss after the subsidy is calculated as $DWL_{after} = \frac{1}{N}\sum_i (g_i - z_i)\mathbf{1}(g_i \geq z_i) - \frac{1}{N}\sum_i (g_i - z_i)\mathbf{1}(g_i \geq \hat{p}^*_{i,after})$. The changes in deadweight loss due to subsidies are $\Delta DWL = DWL_{after} - DWL_{before}$. The reported values are the mean of 50 bootstrapped samples.

welfare per person, calculated by subtracting the social cost of subsidies from the consumer surplus, is also higher. Surprisingly, subsidies that target the lower-earning group reduce deadweight loss less than the other two subsidy designs do. Because bad health can lead to earning reductions, individuals in the low-earning group are more likely to be unhealthy. The subsidies that target the low-earning group can attract many individuals who value the insurance more than their expected costs.

VIII Conclusion

In this paper, I incorporate joint dynamics of earnings and medical spending into the modeling of individuals' insurance choices. I study how individuals who face uncertainty over earnings and medical spending decide between being uninsured and fully insured. I empirically estimate the individual-level willingness to pay using a dataset that links Utah All-payer Claims Data to earnings records derived from the UI database. I first estimate how individuals predict the joint distribution of earnings and medical spending for a nearly population-level sample. Second, I estimate the assets via a life-cycle model in which individuals face earnings dynamics estimated in the first step.

I aggregate the individual willingness to pay for health insurance to the market level and study how earnings dynamics affect adverse selection. I document significant heterogeneous changes in the willingness-to-pay distribution. Moreover, the willingness to pay is no longer a straightforward predictor for medical costs as in textbook models. By reducing the correlation between willingness to pay and expected medical cost, I find that my model, which incorporates earnings dynamics, predicts higher equilibrium take-up rates, lower equilibrium premiums, and lower deadweight loss than a model abstracting from earnings dynamics.

Moreover, ignoring earnings dynamics in the demand model leads to overestimation of the effectiveness of subsidies as a policy tool for combatting adverse selection. Counterfactual simulations also show that offering more subsidies to lower earners reduces deadweight loss less than with uniform subsidies or when targeting the median earners. These results reveal the importance of who to target when designing income-based policies like subsidies.

My findings point to several directions for future research. The first is to incorporate marriage into the model. The impact of including marriage and family in the model is theoretically ambiguous. Marriage is one form of implicit insurance for some people. However, the assortative mating feature of the marriage market may also cause increases in earnings dynamics heterogeneity. Second, my paper focuses on the extensive margins (the choice between being uninsured and insured). Allowing choices among plans with different levels of coverage could affect selection into different health insurance plans and corresponding optimal subsidy designs. Finally, studies can be done on the optimal adjustment of health insurance policies when individuals face financial crises, industry-specific shocks, or changes in other labor market policies.

IX Reference

Abowd, J.M., Kramarz, F. and Margolis, D.N., 1999. High-wage workers and high-wage firms. *Econometrica*, 67(2), pp.251-333.

Abowd, J.M., McKinney, K.L. and Schmutte, I.M., 2019. Modeling endogenous mobility in earnings determination. *Journal of Business & Economic Statistics*, 37(3), pp.405-418.

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

Aizawa, N. and Fang, H., 2020. Equilibrium labor market search and health insurance reform. *Journal of Political Economy*, 128(11), pp.4258-4336.

Akerlof, G.A., 1978. The market for “lemons”: Quality uncertainty and the market mechanism. *In Uncertainty in economics* (pp. 235-251). Academic Press.

Altonji, J.G., Smith Jr, A.A. and Vidangos, I., 2013. Modeling earnings dynamics. *Econometrica*, 81(4), pp.1395-1454.

Atal, J.P., Fang, H., Karlsson, M. and Ziebarth, N.R., 2020. *Long-term health insurance: Theory meets evidence* (No. w26870). National Bureau of Economic Research.

Bagger, J., Fontaine, F., Postel-Vinay, F. and Robin, J.M., 2014. Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics. *American Economic Review*, 104(6), pp.1551-96.

Bewley, T., 1976. *The permanent income hypothesis: A theoretical formulation.* HARVARD UNIV CAMBRIDGE MASS.

Blundell, R., Borella, M., Commault, J. and De Nardi, M., 2020. *Why does consumption fluctuate in old age and how should the government insure it?* (No. w27348). National Bureau of Economic Research.

Bonhomme, S., Lamadon, T. and Manresa, E., 2019. A distributional framework for matched employer employee data. *Econometrica*, 87(3), pp.699-739.

Burdett, K. and Mortensen, D.T., 1998. Wage differentials, employer size, and unemployment. *International Economic Review*, pp.257-273.

Brevoort, K., Grodzicki, D. and Hackmann, M.B., 2020. The credit consequences of unpaid medical bills. *Journal of Public Economics*, 187, p.104203.

Brot-Goldberg, Z.C., Chandra, A., Handel, B.R. and Kolstad, J.T., 2017. What does a deductible do? The impact of cost-sharing on health care prices, quantities, and spending dynamics. *The Quarterly Journal of Economics*, 132(3), pp.1261-1318.

Cahuc, P., Postel-Vinay, F. and Robin, J.M., 2006. Wage bargaining with on-the-job search: Theory and evidence. *Econometrica*, 74(2), pp.323-364.

Card, D., Heining, J. and Kline, P., 2013. Workplace heterogeneity and the rise of West German wage inequality. *The Quarterly journal of economics*, 128(3), pp.967-1015.

Carroll, C.D., 2006. The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics letters*, 91(3), pp.312-320.
Charles, K.K., 2003. The longitudinal structure of earnings losses among work-limited disabled workers. *Journal of Human Resources*, 38(3), pp.618-646.
Chung, YoonKyung. 2013. "Chronic Health Conditions and Economic Outcomes." http://www.solejole.org/14225.pdf.
Cutler, D.M., Meara, E. and Richards-Shubik, S., 2011. *Health Shocks and Disability Transitions Among Near-elderly Workers* (No. onb11-08). National Bureau of Economic Research.
Cohen, A. and Einav, L., 2003. The effects of mandatory seat belt laws on driving behavior and traffic fatalities. *Review of Economics and Statistics*, 85(4), pp.828-843.
Decarolis, F., 2015. Medicare part d: Are insurers gaming the low income subsidy design?. *American Economic Review*, 105(4), pp.1547-80.
Decarolis, F., Polyakova, M. and Ryan, S.P., 2020. Subsidy design in privately provided social insurance: Lessons from Medicare Part D. *Journal of Political Economy*, 128(5), pp.1712-1752.
De Nardi, M. and Fella, G., 2017. Saving and wealth inequality. *Review of Economic Dynamics*, 26, pp.280-300.
De Nardi, M., French, E. and Jones, J.B., 2010. Why do the elderly save? The role of medical expenses. *Journal of political economy*, 118(1), pp.39-75.
De Nardi, M., French, E. and Jones, J.B., 2016. Medicaid insurance in old age. *American Economic Review*, 106(11), pp.3480-3520.
De Meza, D. and Webb, D.C., 2001. Advantageous selection in insurance markets. *RAND Journal of Economics*, pp.249-262.
Dey, M.S. and Flinn, C.J., 2005. An equilibrium model of health insurance provision and wage determination. *Econometrica*, 73(2), pp.571-627.
Doherty, N.A. and Schlesinger, H., 1983. Optimal insurance in incomplete markets. *Journal of political economy*, 91(6), pp.1045-1054.
Dobkin, C., Finkelstein, A., Kluender, R. and Notowidigdo, M.J., 2018. The economic consequences of hospital admissions. *American Economic Review*, 108(2), pp.308-52.
Di Addario, S., Kline, P., Saggio, R. and Sølvsten, M., 2022. It ain't where you're from, it's where you're at: hiring origins, firm heterogeneity, and wages. *Journal of Econometrics.*
Einav, L. and Finkelstein, A., 2011. Selection in insurance markets: Theory and empirics in pictures. *Journal of Economic perspectives*, 25(1), pp.115-38.
Einav, L., Finkelstein, A. and Cullen, M.R., 2010. Estimating welfare in insurance markets using variation in prices. *The quarterly journal of economics*, 125(3), pp.877-921.
Einav, L., Finkelstein, A., Ryan, S.P., Schrimpf, P. and Cullen, M.R., 2013. Selection on moral hazard in health insurance. *American Economic Review*, 103(1), pp.178-219.

Ericson, K.M. and Sydnor, J.R., 2018. *Liquidity constraints and the value of insurance* (No. w24993). National Bureau of Economic Research.
Fang, H., Keane, M.P. and Silverman, D., 2008. Sources of advantageous selection: Evidence from the Medigap insurance market. *Journal of Political Economy*, 116(2), pp.303-350.
Farber, H.S. and Levy, H., 2000. Recent trends in employer-sponsored health insurance coverage: are bad jobs getting worse?. *Journal of Health Economics*, 19(1), pp.93-119.
Finkelstein, A. and McGarry, K., 2006. Multiple dimensions of private information: evidence from the long-term care insurance market. *American Economic Review*, 96(4), pp.938-958.
Finkelstein, A., Mahoney, N. and Notowidigdo, M.J., 2018. What does (formal) health insurance do, and for whom?. *Annual Review of Economics*, 10, pp.261-286.
Finkelstein, A., Hendren, N. and Shepard, M., 2019. Subsidizing health insurance for low-income adults: Evidence from Massachusetts. *American Economic Review*, 109(4), pp.1530-67.
Finkelstein, A., Hendren, N. and Luttmer, E.F., 2019. The value of Medicaid: Interpreting results from the Oregon health insurance experiment. *Journal of Political Economy*, 127(6), pp.2836-2874.
French, E. and Jones, J.B., 2004. On the distribution and dynamics of health care costs. *Journal of Applied Econometrics*, 19(6), pp.705-721.
Garthwaite, C., Gross, T. and Notowidigdo, M.J., 2018. Hospitals as insurers of last resort. *American Economic Journal: Applied Economics*, 10(1), pp.1-39.
Ghosh, A., Simon, K. and Sommers, B.D., 2019. The effect of health insurance on prescription drug use among low-income adults: evidence from recent Medicaid expansions. *Journal of health economics*, 63, pp.64-80.
Handel, B.R., 2013. Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, 103(7), pp.2643-82.
Handel, B.R. and Kolstad, J.T., 2015. Health insurance for" humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105(8), pp.2449-2500.
Handel, B., Hendel, I. and Whinston, M.D., 2015. Equilibria in health exchanges: Adverse selection versus reclassification risk. *Econometrica*, 83(4), pp.1261-1313.
Hendren, N., 2021. Measuring ex-ante welfare in insurance markets. *The Review of Economic Studies*, 88(3), pp.1193-1223.
Hendren, N., 2013. Private information and insurance rejections. *Econometrica*, 81(5), pp.1713-1762.
Hansen, G.D. and Imrohoroğlu, A., 1992. The role of unemployment insurance in an economy with liquidity constraints and moral hazard. *Journal of political economy*, 100(1), pp.118-142.
Jaffe, S.P. and Shepard, M., 2017. *Price-linked subsidies and health insurance markups.*

Cambridge (MA): National Bureau of Economic Research.
Kaufmann, C., Schmid, C. and Boes, S., 2017. Health insurance subsidies and deductible choice: Evidence from regional variation in subsidy schemes. *Journal of health economics*, 55, pp.262-273.
Lakdawalla, D. and Sood, N., 2013. Health insurance as a two-part pricing contract. *Journal of public economics*, 102, pp.1-12.
Lockwood, L. 2022. Health Insurance Increases Financial Risk.
Mahoney, N., 2015. Bankruptcy as implicit health insurance. *American Economic Review*, 105(2), pp.710-46.
Mahoney, N. and Weyl, E.G., 2017. Imperfect competition in selection markets. *Review of Economics and Statistics*, 99(4), pp.637-651.
Manning, W.G., 1998. The logged dependent variable, heteroscedasticity, and the retransformation problem. *Journal of health economics*, 17(3), pp.283-295.
Meyer, B.D. and Mok, W.K., 2019. Disability, earnings, income and consumption. *Journal of Public Economics*, 171, pp.51-69.
Meghir, C. and Pistaferri, L., 2004. Income variance dynamics and heterogeneity. *Econometrica*, 72(1), pp.1-32.
Poterba, J.M., Venti, S.F. and Wise, D.A., 2017. The asset cost of poor health. *The Journal of the Economics of Ageing*, 9, pp.172-184.
Postel–Vinay, F. and Robin, J.M., 2002. Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica*, 70(6), pp.2295-2350.
Polyakova, M. and Ryan, S.P., 2019. *Subsidy targeting with market power* (No. w26367). National Bureau of Economic Research.
Pratt, J.W., 1978. Risk aversion in the small and in the large. *In Uncertainty in economics* (pp. 59-79). Academic Press.
Rothschild, M. and J. E. Stiglitz (1976): *Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information* The Quarterly Journal of Economics, 90, 630-49.
Sasso, A.T.L. and Lurie, I.Z., 2009. Community rating and the market for private non-group health insurance. Journal of Public Economics, 93(1-2), pp.264-279.
Shepard, Mark. 2022. Hospital Network Competition and Adverse Selection: Evidence from the Massachusetts Health Insurance Exchange. *American Economic Review*, 112 (2): 578-615.
Slusky, D.J. and Ginther, D.K., 2021. Did Medicaid expansion reduce medical divorce? *Review of Economics of the Household*, 19(4), pp.1139-1174.
Smith, J.P., 1999. Healthy bodies and thick wallets: the dual relation between health and economic status. *Journal of Economic perspectives*, 13(2), pp.145-166.

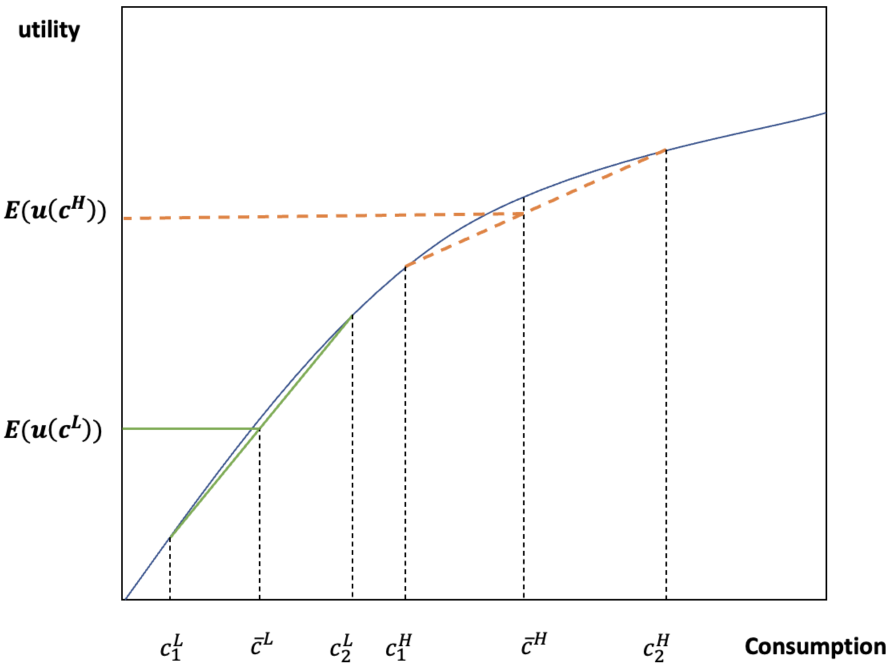
Smith, J., 2005. Consequences and predictors of new health events. *In Analyses in the Economics of Aging* (pp. 213-240). University of Chicago Press.

Tebaldi, P., 2022. *Estimating equilibrium in health insurance exchanges: Price competition and subsidy design under the aca* (No. w29869). National Bureau of Economic Research.

Appendix

A. More Theoretical Discussion

Figure A1. Lower earning mean leads to the lower expected utility of uninsured



Note: This figure illustrates that individuals with a lower mean of earnings tend to have a lower expected utility of being uninsured. Facing the same level of earning uncertainty, if the average consumption equals $\bar{c}^L$, the expected utility is $E(u(c^L))$. This is lower than $E(u(c^H))$, which is the expected utility when the average consumption is $\bar{c}^H$.

B. Sample Selection Details

In this section, I present the details in the summary statistics change of the sample when I apply sample selection criteria in Table 1. When I restrict the sample to people insured in relatively large plans with estimated plan actuarial value, the average earning increases, and the probability of not being employed or changing employers decreases. This evidence reveals that people with higher earnings may be more likely to access health insurance. The potential full-time workers also earn more quarterly and are less liked not to be employed or change employers.

B1. Summary Statistics by Sample section criteria

	(1) Starting Sample mean	(2) 1. Age(25-64) mean	(3) 2. Insured mean	(4) 3. With plan characteristics mean	(5) 4. no missing types mean	(6) 5. potential stayers in labor force mean	(7) 6. potential full-time workers mean
Age	35.88	42.45	42.98	43.62	43.59	43.61	43.80
Male	0.52	0.53	0.52	0.52	0.52	0.54	0.60
Quarterly Earning	11591.71	13947.54	14650.62	15792.11	15917.53	15946.03	18025.77
Quarterly Earning(Imputed)	7661.68	10576.08	11917.18	13264.50	13458.91	14851.94	17542.20
Not employed	33.90	24.17	18.66	16.01	15.45	6.86	2.68
Stay in the same firm	88.05	91.54	93.56	95.61	96.24	96.52	97.98
Change employer	6.18	4.29	3.39	2.10	1.86	1.86	1.32
Change to earn zero amount	5.77	4.17	3.05	2.29	1.90	1.62	0.70
Newly hired	12.43	13.09	14.18	11.69	9.92	23.48	25.61
Continue to earn zero amount	87.57	86.91	85.82	88.31	90.08	76.52	74.39
Inpatient Spending	161.41	183.80	209.84	190.73	192.00	180.41	145.76
Outpatient Spending	214.73	252.55	308.26	306.76	307.59	299.06	273.85
Office visits spending	273.18	308.90	375.01	388.09	391.41	375.91	324.12
Pharmacy spending	151.85	181.68	221.54	229.45	231.05	221.50	211.27
Total quarterly medical spending	801.17	926.93	1114.66	1115.03	1122.05	1076.88	955.00
Total yearly medical spending	3191.25	3690.89	4354.66	4432.45	4460.77	4281.73	3796.54
Acg risk score	1.04	1.17	1.21	1.25	1.25	1.23	1.16
Number of chronic conditions	0.04	0.04	0.04	0.04	0.04	0.04	0.04
Has Diabetes	4.35	5.30	5.47	5.72	5.80	5.77	5.63
Has hypertension	10.50	12.89	13.34	14.30	14.39	14.41	14.54
N	14118929	8675205	7008764	4203342	3829606	3461535	2778325

Notes: This table reports the summary statistics for the sample with sample selection criteria, and is a detailed version of Table 1. I begin with the whole sample in Column 1. In Column 2, I restrict the sample to people between 26 to 64. In Column 3, I restrict to people who are insured in all quarters in 2013-2015. In Column 4, I restrict to people who are enrolled in plans with a large enough number of enrollees. in Column 5, I restrict the sample to people who are linked to employers so that we could construct firm earning level types and volatility types. in Column 6, I consider the sample size change if we focus on the people who were employed from 2014-2015 so that we could limit the impact of including people who have exited the labor force after 2013. In Column 7, I consider the changes if people who are potentially part-time workers are excluded. I consider workers whose quarterly earnings are above $3480 to be more likely to be full-time workers. $3480 is the quarterly total earning if the individual works at the minimum wage $7.25 for 8 hours a day, 5 days a week, 48 weeks annually.

C. Transition Matrices Estimations

In this section, I describe the details of how to estimate job mobility and the firm-type transition matrices.[23]

Job mobility Type transitions. — I first estimate the job mobility transitions if the individual i is not employed in period $t-1$. The transition probabilities for job mobility types d_{it} from $t-1$ to t are assumed to depend on health type h_{it}, gender, age groups (in 5-year bins), and person earning type a_i. I fit the transitions into a multinomial logit model given by

$$\pi_{it}^{j} = \gamma_j X_{it} + \omega_j L_{it} + \lambda_j a_i + \eta_{it}^{j} \tag{37}$$

where π_{it}^{j} represents the log odds for $d_{it} = j$. X^{it} includes the indicators for the age group individual i belongs to in period t, and his gender. L_{it} is a set of indicators for the categories of health types in period t. a_i is 1×6 vector that describes the category of the six person earning types individual i belongs to. γ_j, ω_j, λ_j are the associated parameter vectors.

Second, I estimate the job mobility transitions if the individual i is employed in period $t-1$. The transition probabilities for job mobility types d_{it} from $t-1$ to t are assumed to depend on health type h_{it}, gender, age groups (in 5-year bins), and person earning type a_i. I fit the transitions into a multinomial logit model given by

$$\pi_{it}^{j} = \gamma_j X_{it} + \omega_j L_{it} + \lambda_j a_i + \alpha_j D_{i,t-1} + \eta_{it}^{j} \tag{38}$$

where π_{it}^{j} represents the log odds for $d_{it} = j$. X^{it} includes the indicators for the age group individual i belongs to in period t, and his gender. L_{it} is a set of indicators for the categories of health types in period t. a_i is 1×6 vector that describes the category of the six person earning types individual i belongs to. $D_{i,t-1}$ is a set of indicators for the category of the job mobility type from period $t-2$ to $t-1$. γ_j, ω_j, λ_j, and α_j are the associated parameter vectors.

Figure C1 shows the probability of remaining unemployed in period t if currently not employed in period $t-1$. We can see that individuals with lower earning types are more likely to remain unemployed, and bad health can increase the probability of remaining unemployed.

Figure C2 shows the job mobility transitions if currently employed. We can also observe that individuals with lower earning types are less likely to stay in their current firm. Moreover, bad health increases the probability of moving or losing jobs slightly.

Firm Type Transitions. — If the individuals are stayers from $t-1$ to t, there is no need to estimate the probability of the types of the firms they work in period t because the type

[23] Only part of the transition matrices are shown, to meet disclosure requirements.

Figure C1. Probability to remain not employed

Note: This figure presents the probability of remaining unemployed in t if currently not employed in $t-1$. The health types are constructed using health risk scores. People with higher risk scores are predicted to have higher medical spending. Health type takes values of 1, 2, 3, and 4, representing the people with risk scores below 1, between 1 to 1.5, between 1.5 and 2.5, and above 2.5, respectively. Person fixed earning types a_i are constructed from a two-way fixed effects model as explained in Section 4.1.1. The lower earning types reports are the average probabilities over people with person earning types 1, 2, and 3. The higher earning types estimates are the average probabilities over people with person earning types 4, 5, and 6.

of their employers does not change. I first estimate the probability of firm types when the individuals are not employed in period $t-1$. I assume that the type of firms they work for in period t, conditional on being newly employed, depends on their gender, age group, and person earning types. I estimate the transitions for firm earning level type and uncertainty type separately. I fit the transitions into a multinomial logit model given by

$$\pi_{it}^{j} = \gamma_j X_{it} + \lambda_j a_i + \eta_{it}^{j} \tag{39}$$

where π_{it}^{j} represents the log odds for firm earning level type $k_{it}^{\mu} = j$ (or for firm earning uncertainty type $k_{it}^{\sigma} = j$). X^{it} includes the indicators for the age group individual i belongs to in period t, and his gender. a_i is 1×6 vector that describes the category of the six person earning types individual i belongs to. γ_j and λ_j are the associated parameter vectors.

Second, I estimate the probability of their destination firm types in period t when individuals are employed in period $t-1$. These people are classified as movers. I assume that the type of firms they work for in period t, conditional on being movers, depends on their gender, age group, health types, past firm types, and person earning types. For the transitions for firm types, the multinomial logit model is given by

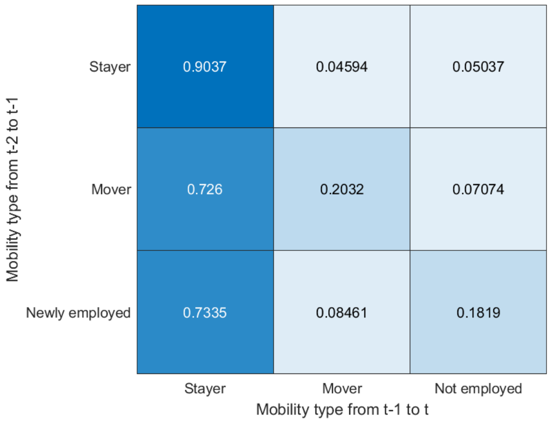
$$\pi_{it}^{j} = \gamma_j X_{it} + \omega_j L_{it} + \lambda_j a_i + \beta_j k_{i,t-1} + \eta_{it}^{j} \tag{40}$$

where π_{it}^{j} represents the log odds for $k_{it}^{\mu} = j$(or for firm earning uncertainty type $k_{it}^{\sigma} = j$). X^{it} includes the indicators for the age group individual i belongs to in period t, and his gender. L_{it} is a set of indicators for the categories of health types in period t. a_i is 1×6 vector that describes the category of the six person earning types individual i belongs to. $k_{i,t-1}$ is a set of indicators for the category of the type of firms they work for (including both firm earning level types and uncertainty types) in $t-1$. γ_j, ω_j, λ_j, and β_j are the associated parameter vectors.

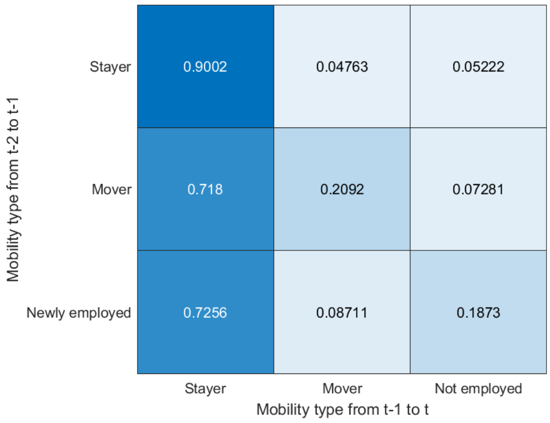
Figure C3 shows the firm-type transitions if the individuals are newly employed in period t. People with lower earning types are more likely to move to firms with lower earning level types and higher risk types.

Figure C4 and C5 show the firm-type transitions if the individuals are movers. The figures reveal that people with lower earning types are more likely to move to firms with lower earning level types and higher risk types. Moreover, bad health slightly reduces the probability of transitioning into higher-earning, lower-risk firms.

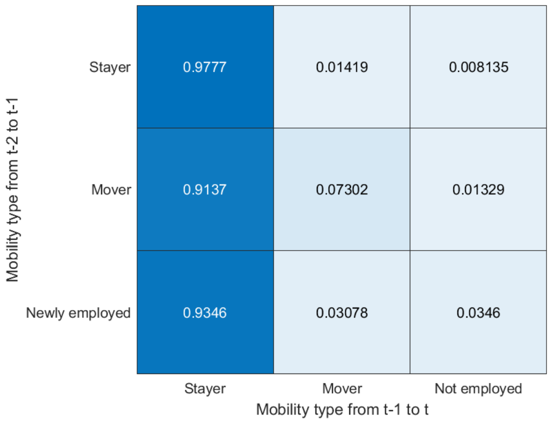
Figure C2. Job mobility type transitions for the currently employed



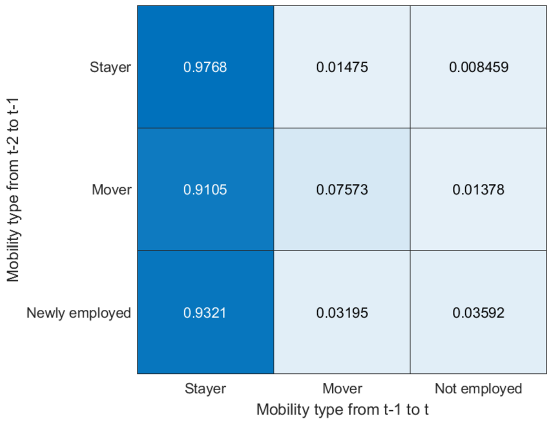
(a) $h_t = 1$, $\alpha = 1$



(b) $h_t = 4$, $\alpha = 1$



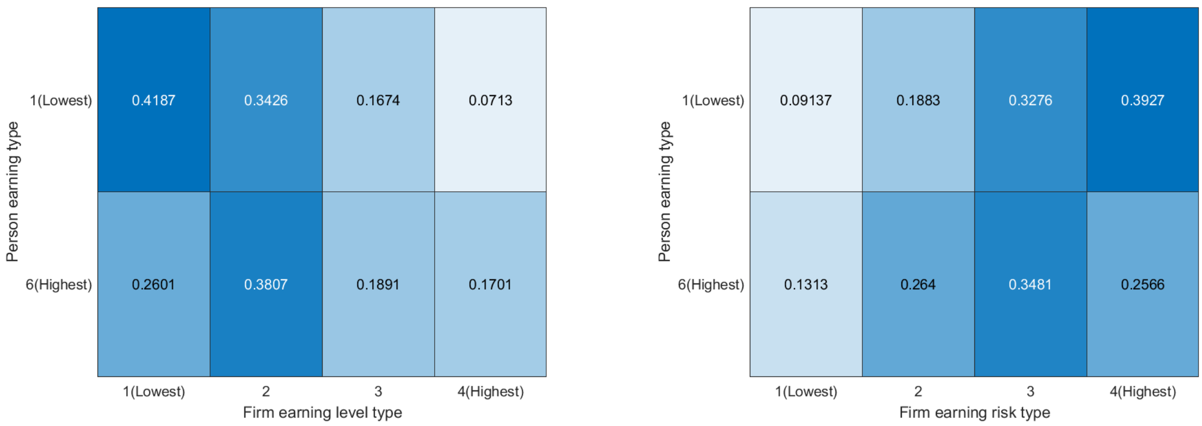
(c) $h_t = 1$, $\alpha = 6$



(d) $h_t = 4$, $\alpha = 6$

Note: This figure presents the job mobility type transitions from $t-1$ to t by health type if currently employed in $t-1$. The stayers are those who stay in the same firm. Movers are those who are continuously employed but change employers. Newly employed means not employed last quarter. The health types are constructed using ACG risk scores. People with higher risk scores are predicted to have higher medical spending. Health type takes values of 1, 2, 3, and 4, representing the people with risk scores below 1, between 1 to 1.5, between 1.5 and 2.5, and above 2.5, respectively.

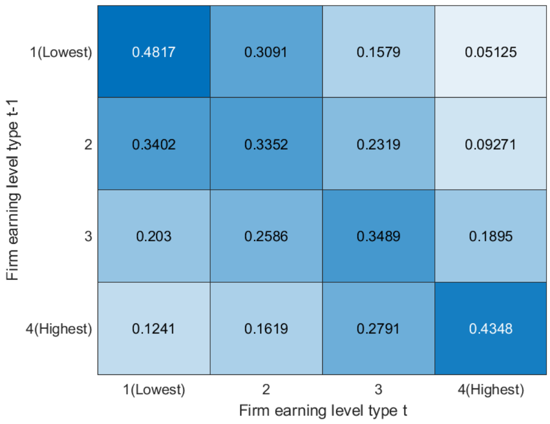
Figure C3. Firm type transitions if newly employed



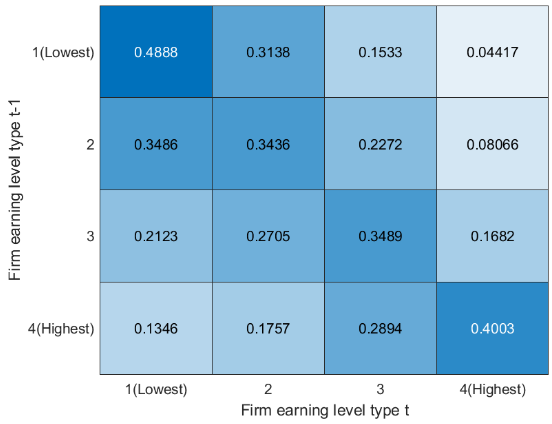
(a) Firm earning level type **(b)** Firm earning risk type

Note: This figure presents the probability of transiting into different firm-level type k_j^{μ} and firm risk type k_j^{σ} in time t if newly employed from $t-1$ to t. Newly employed means not employed last quarter. The health types are constructed using ACG risk scores. People with higher risk scores are predicted to have higher medical spending. Health type takes values of 1, 2, 3, and 4, representing the people with risk scores below 1, between 1 to 1.5, between 1.5 and 2.5, and above 2.5, respectively.

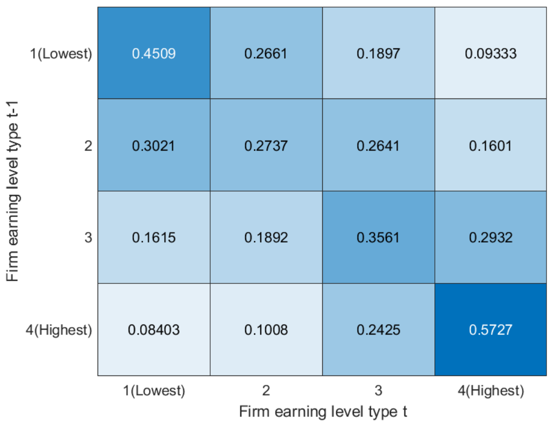
Figure C4. Firm earning level type transitions for movers



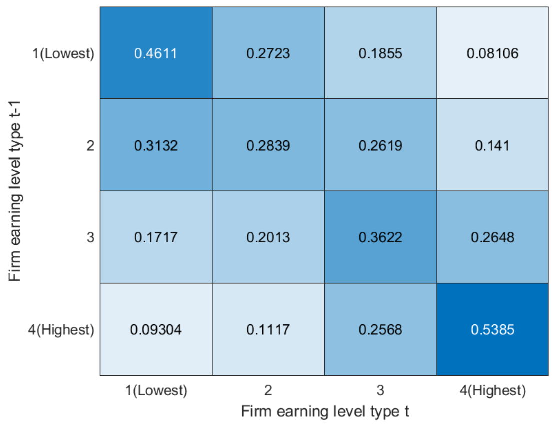
(a) $h_t = 1$, $a_i = 1$



(b) $h_t = 4$, $a_i = 1$



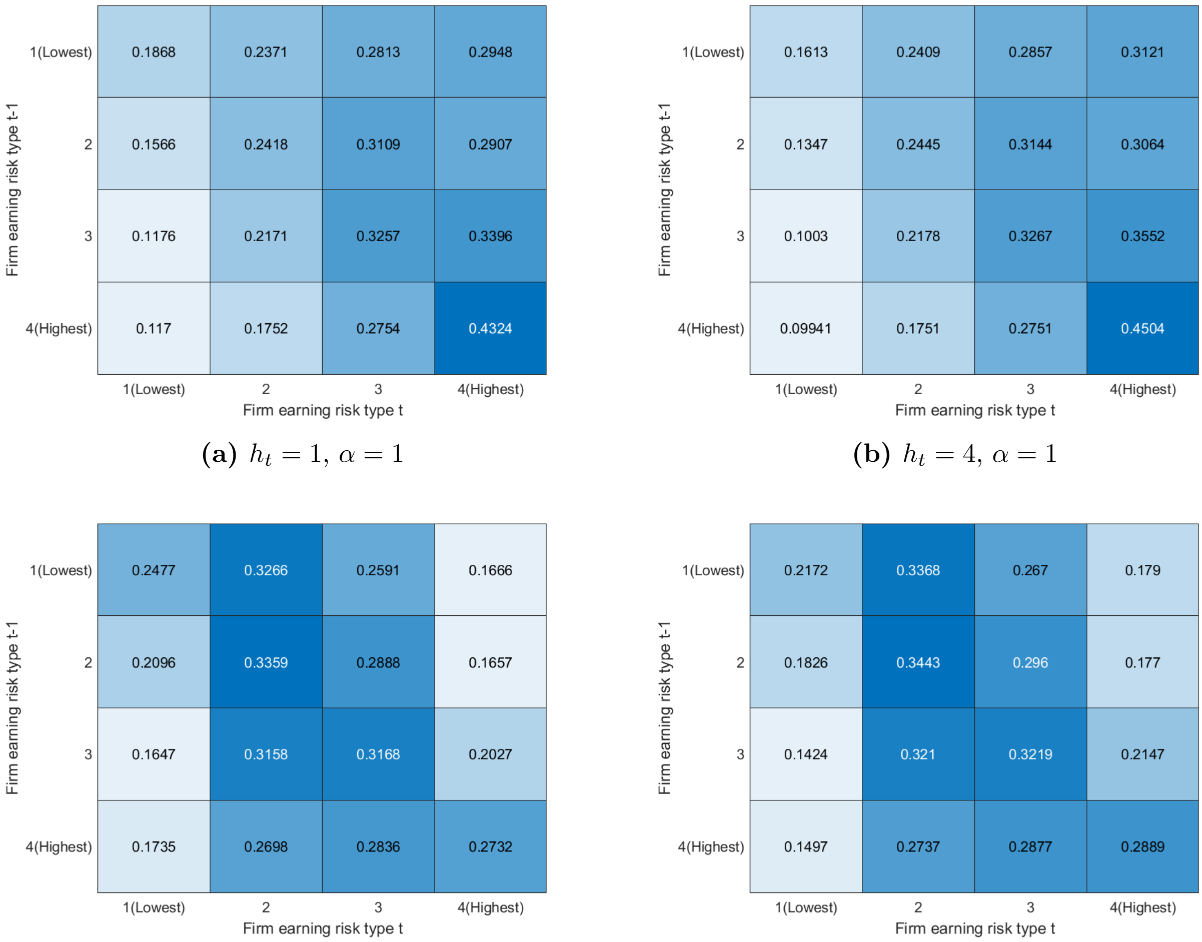
(c) $h_t = 1$, $a_i = 6$



(d) $h_t = 4$, $a_i = 6$

Note: This figure presents the probability of transiting into different firm-level type k_j^{μ} in time t if change employer from $t-1$ to t. The health types h_t are constructed using ACG risk scores. People with higher risk scores are predicted to have higher medical spending. Health type takes values of 1, 2, 3, and 4, representing the people with risk scores below 1, between 1 to 1.5, between 1.5 and 2.5, and above 2.5, respectively. Person fixed earning types a_i are constructed from a two-way fixed effects model as explained in Section 4, with 1 standing for the lowest earning type and 6 standing for the highest earning type.

Figure C5. Firm earning risk type transitions for movers



(a) $h_t = 1$, $\alpha = 1$ **(b)** $h_t = 4$, $\alpha = 1$

(c) $h_t = 1$, $\alpha = 6$ **(d)** $h_t = 4$, $\alpha = 6$

Note: This figure presents the probability of transiting into different firm risk type k_j^σ in time t if change employer from $t-1$ to t. The health types h_t are constructed using ACG risk scores. People with higher risk scores are predicted to have higher medical spending. Health type takes values of 1, 2, 3, and 4, representing the people with risk scores below 1, between 1 to 1.5, between 1.5 and 2.5, and above 2.5, respectively. Person fixed earning types a_i are constructed from a two-way fixed effects model as explained in Section 4, with 1 standing for the lowest earning type and 6 standing for the highest earning type.

D. Decompose Demand curve and Cost curves changes

In this section, I discuss the outcome differences across the four models with the textbook model, which are introduced in detail in Section VI.1. In Model 1, the individuals only differ in medical risk. In Model 2, differences in assets are added. In Model 3, individuals are assumed to face the mean of individual-level predicted earning distribution in all states of the world. In Model 4, individuals are assumed to face the predicted earning distribution.

I first discuss the impact of demand curve changes across models while keeping the cost curves in the textbook model. In this way, the changes in the equilibrium and social efficient outcomes reflect only the changes in demand curves. Second, I compare the impact of cost curves with the textbook model cost curves while keeping the demand curve estimated in the "Full-heterogeneity" model (Model 4). Keeping the demand curve unchanged across the models can separate out the impact of cost curve changes.

Demand Curve Changes. — To separate the impact of demand curve shifts, I keep the cost curves of the textbook curves and report the equilibrium and social efficient outcomes relative to the "Full-heterogeneity" model. The reason to compare with "Full-heterogeneity" model is that when studying the impact of changes in cost curves, I keep the demand curve in the "Full-heterogeneity" model. Table D1 Panel A reports the equilibrium take-up rates, price, and consumer surplus relative to Model 4. The intersection between the demand curves and the textbook average cost curves is the equilibrium price. It shows that the take-up rate and consumer surplus are lower, price is higher when considering earnings dynamics. Panel B and Panel C report the socially efficient outcomes. We observe significant differences between Model 3 and Model 4 in the socially efficient take-up rates.

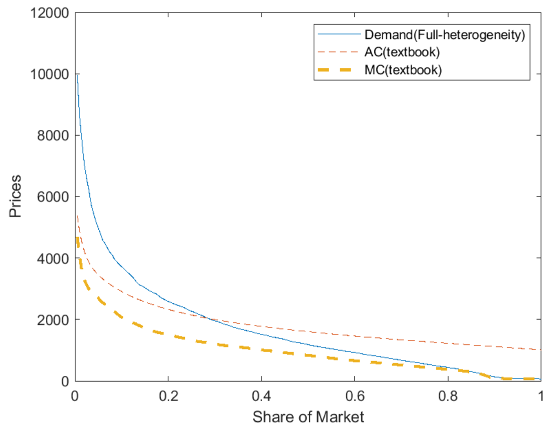
Table D1. The impact of Demand curves changes relative to the "Full-heterogeneity" model
Conditional on textbook cost curve

Model	Description	Take-up	Price	Consumer surplus	Deadweight loss
Panel A: Equilibrium					
Model	Description	Take-up	Price	Consumer surplus	
4	**Full-heterogeneity**	0.259	2123.4	441.4	
		(0.02)	(78.36)	(27.51)	
1	Only differ in medical risk	+0.057	-161.3	+497.1	
		(0.02)	(51.86)	(29.49)	
2	Add different assets	+0.222	-489.6	+378.7	
		(0.01)	(58.19)	(16.09)	
3	Add different expected earning	+0.051	-146.6	+97.9	
		(0.01)	(35.45)	(10.39)	
Panel B: Social Efficiency: Social Planner Measure					
Model	Description	Take-up	Price	Consumer surplus	Deadweight loss
4	**Full-heterogeneity**	0.995		637.7	196.3
		(0.01)		(24.05)	(14.29)
1	Only differ in medical risk	0.005		+553.2	+56.1
		(0.01)		(25.22)	(12.19)
2	Add different assets	0.005		+353.1	-25.6
		(0.01)		(10.76)	(11.42)
3	Add different expected earning	-0.180		+74.8	-23.1
		(0.01)		(5.93)	(7.87)
Panel C: Social Efficiency: Smoothed Marginal Cost Measures					
Model	Description	Take-up	Price	Consumer surplus	Deadweight loss
4	**Full-heterogeneity**	0.965	105.0	637.2	195.8
		(0.06)	(101.46)	(24.07)	(14.41)
1	Only differ in medical risk	+0.035	-53.9	+553.8	+56.7
		(0.06)	(101.31)	(25.13)	(12.36)
2	Add different assets	+0.035	-52.7	+353.6	-25.1
		(0.06)	(101.26)	(10.73)	(11.59)
3	Add different expected earning	-0.174	+280.1	+75.3	-22.7
		(0.06)	(99.35)	(6.21)	(8.04)

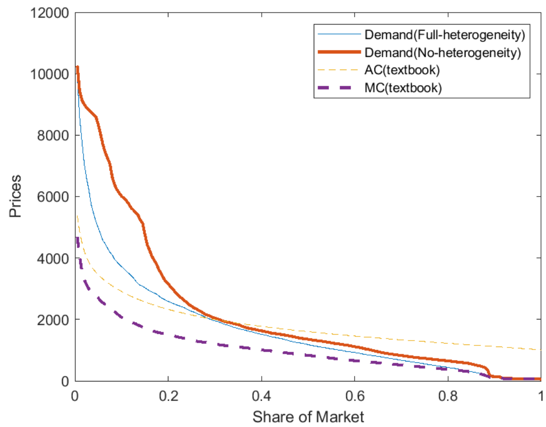
Notes: This table compares Model 1 to 3 with Model 4 — "Full-heterogeneity" model. Details of the models are introduced in Section VI.1. Model 1 is the model that allows only heterogeneity in medical risk. Individuals are holding the average asset level of the sample $\bar{A}$ and expect to earn the sample average mean earning $\bar{\mu}_{wt}$. Model 2 add heterogeneity in asset, and individual i holds assets at A_i. Model 3 adds the heterogeneity in the mean of earnings based on Model 3. Individual i gets the earning at μ_{iwt} in all possible states. Model 4 is the full-heterogeneity model, in which people are uncertain about earning, and the correlation between earning and medical spending is allowed. The medical risk distribution is kept unchanged in all models. All individuals are assumed to have the same constant relative risk aversion utility function with risk aversion at 2. The consumption floor is set at $2000. Panel A reports the equilibrium take-up rate, premiums, and consumer surplus. To calculate the equilibrium outcomes, I assume that insurers obtain zero expected profits. Column 4 in Panel A reports the cost of public funds to raise taxes to pay for the wealth transfers when consumption hits the consumption floor. Panel B reports the socially efficient take-up rate, consumer surplus, and deadweight loss under the Social Planner Measure—people with a higher willingness to pay than expected medical cost are considered as those who should be socially optimal to cover. Panel C smooths non-monotonic marginal cost curves and considers the interaction point between the smoothed marginal cost curve and demand curves as the socially efficient price. The marginal cost curves are smoothed with robust linear regression over each window of 20 points. More details can be found in Section VI.2.

Figure D1 plots the Demand curves relative to the demand in the Full-heterogeneity model, while keeping the cost curves in the textbook model. We can see that the changes in demand curves are not parallel shifts.

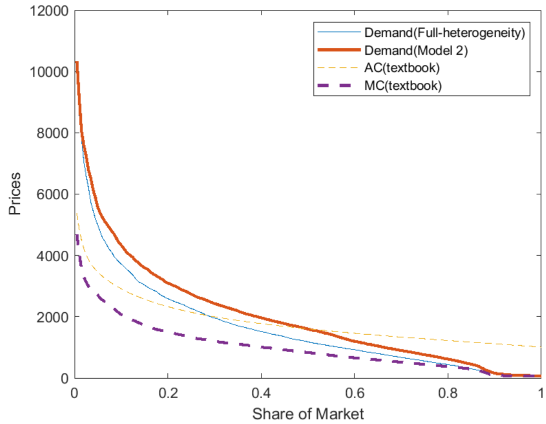
Figure D1. Compare Demand curves with the Full-heterogeneity model



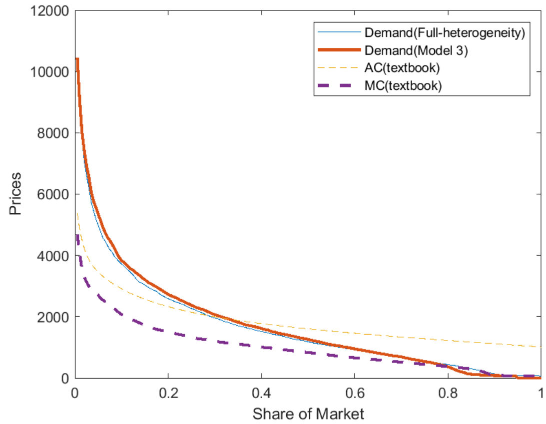
(a) Full-heterogeneity Model
s^*: 0.259[0.018]
s^o: 0.995[0.006]



(b) Model 1: only differ in medical risk
Δs^*: +0.057[0.015]
Δs^o: +0.0052[0.006]



(c) Model 2: add different asset
Δs^*: +0.222[0.014]
Δs^o: +0.0052[0.06]



(d) Model 3: add different mean of earnings
Δs^*: +0.051[0.010]
Δs^o: -0.18[0.012]

Note: This figure compares how the demand curve change across models with different level of heterogeneity in earning dynamics.

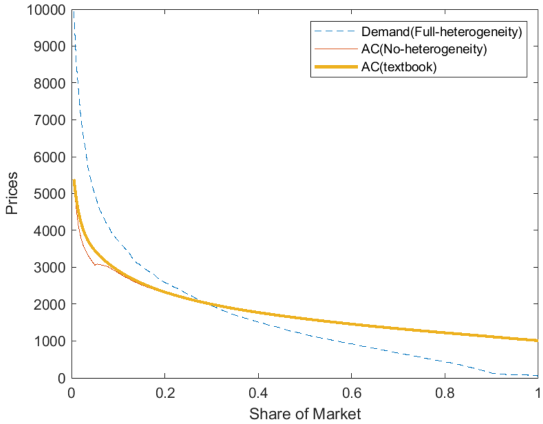
Cost Curve Changes. — In Table D2, I report the relative differences of outcomes with the textbook cost curves while keeping the demand curve estimated in Model 4, the "Full-heterogeneity" model. This table thus separates the impact of cost curves from demand curves. The numbers reported are the differences between cost curves in Models 1 to 4 and the textbook cost curves. We observe significant increases in the equilibrium take-up rate, consumer surplus, and price reductions when earnings dynamics are incorporated. We also see large decreases in the socially efficient take-up rate and deadweight loss.

Table D2. The impact of deviation from textbook marginal cost curve
Conditional on "Full-heterogeneity" Model Demand curve

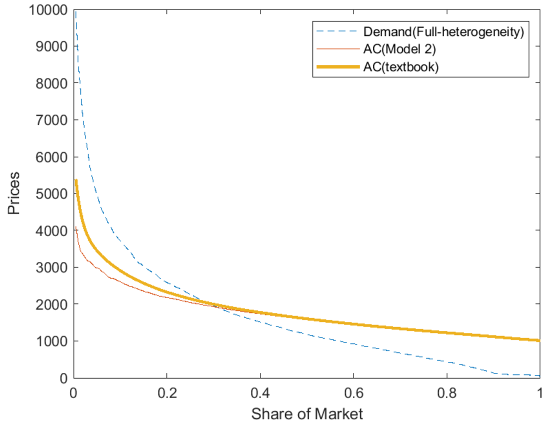
Panel A: Equilibrium					
Model	Description	Take-up	Price	Consumer surplus	
1	Only differ in medical risk	+0.006	-33.8	+8.7	
		(0.00)	(12.41)	(2.85)	
2	Add different assets	+0.030	-158.0	+42.9	
		(0.01)	(30.75)	(6.65)	
3	Add different expected earning	+0.105	-483.2	+148.4	
		(0.01)	(62.49)	(14.58)	
4	Full-heterogeneity	+0.112	-514.7	+160.0	
		(0.02)	(63.69)	(15.03)	
Panel B: Social Efficiency: Social Planner Measure					
Model	Description	Take-up	Price	Consumer surplus	Deadweight loss
1	Only differ in medical risk	-0.020		+1.6	-7.1
		(0.01)		(0.27)	(2.95)
2	Add different assets	-0.075		+3.7	-39.3
		(0.01)		(0.80)	(6.58)
3	Add different expected earning	-0.133		+118.5	-30.0
		(0.01)		(6.89)	(12.31)
4	Full-heterogeneity	-0.165		+110.7	-49.3
		(0.01)		(5.92)	(12.69)
Panel C: Social Efficiency: Smoothed Marginal Cost Measures					
Model	Description	Take-up	Price	Consumer surplus	Deadweight loss
1	Only differ in medical risk	+0.005	-28.5	+0.1	-8.7
		(0.07)	(106.94)	(1.21)	(3.25)
2	Add different assets	-0.053	+77.2	-0.7	-43.6
		(0.08)	(135.73)	(1.35)	(7.05)
3	Add different expected earning	-0.043	-5.9	+69.3	-79.2
		(0.06)	(99.92)	(5.93)	(14.04)
4	Full-heterogeneity	-0.148	+291.7	+18.0	-142.0
		(0.07)	(148.15)	(10.23)	(18.74)

Notes: This table aims to compare how the average cost curves and marginal cost curves deviate from the textbook cost curves across models. The demand curve is kept at the demand curve of model 4 across all models. The textbook version's marginal cost curve orders consumers by their expected medical costs. The textbook version's average cost curve is calculated based on the textbook version's marginal cost curves. Details of the models are introduced in Section VI.1. Model 1 is the model that allows only heterogeneity in medical risk. Individuals are holding the average asset level of the sample $\bar{A}$ and expect to earn the sample average mean earning $\bar{\mu}_{wt}$. Model 2 add heterogeneity in asset, and individual i holds assets at A_i. Model 3 adds the heterogeneity in the mean of earnings based on Model 3. Individual i gets the earning at μ_{iwt} in all possible states. Model 4 is the full-heterogeneity model, in which people are uncertain about earning, and the correlation between earning and medical spending is allowed. The medical risk distribution is kept unchanged in all models. All individuals are assumed to have the same constant relative risk aversion utility function with risk aversion at 2. The consumption floor is set at $2000. Panel A reports the equilibrium take-up rate, premiums, and consumer surplus. To calculate the equilibrium outcomes, I assume that insurers obtain zero expected profits. Column 4 in Panel A reports the cost of public funds to raise taxes to pay for the wealth transfers when consumption hits the consumption floor. Panel B reports the socially efficient take-up rate, consumer surplus, and deadweight loss under the Social Planner Measure—people with a higher willingness to pay than expected medical cost are considered as those who should be socially optimal to cover. Panel C smooths non-monotonic marginal cost curves and considers the interaction point between the smoothed marginal cost curve and demand curves as the socially efficient price. The marginal cost curves are smoothed with robust linear regression over each window of 20 points. More details can be found in Section VI.2.

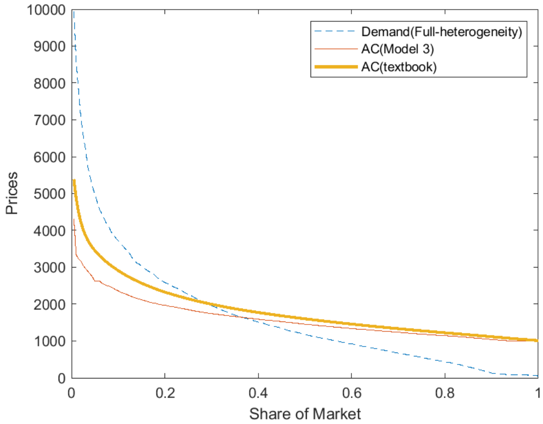
Figure D2. Compare the deviation of average cost curves from the textbook model



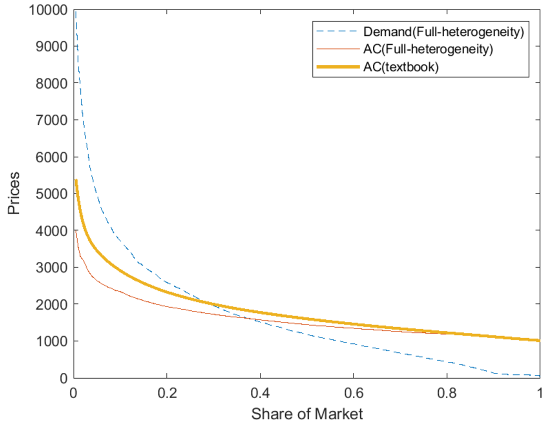
(a) Model 1: Only differ in Medical cost
Δs^{*}: +0.006[0.003]



(b) Model 2: add different asset
Δs^{*}: +0.03[0.007]



(c) Model 3: add different expected earning
Δs^{*}: +0.104[0.015]



(d) Model 4: add earning uncertainty
Δs^{*}: +0.112[0.016]

Note: This figure compares how the average curve change across models with different level of heterogeneity in earning dynamics.

E. ACA subsidy summary

This section summarizes the subsidy (Premium Tax Credit) in ACA. The Affordable Care Act caps the amount that individuals eligible for the tax credit to lower the monthly insurance payment in the Health Insurance Marketplace. How much tax credit to receive is based on income and household information. The eligibility and generosity of this subsidy in 2017 are given in Table E1. The final amount of the tax credit each person receives is based on the final yearly income. They will get the refund when they file the federal income tax return.

I then present an example to help illustrate how the tax credit is calculated. We consider a person with an income of $\$30,000$, which is 253% of poverty. This person's percentage of income is 8.28% of income, which means that the maximum premium this person will have to pay is $\$2,485 = \$30,000 \times 8.28\%$ annually for the second-lowest-cost silver plan. This person can receive a tax credit if that plan's premium is higher than $\$2,485$. If the premium is $\$4,485$, then this person receives $\$4,485 - \$2,485 = \$2,000$ tax credit annually.

Table E1. Affordable Care Act Tax Credit Premium Cap for single individuals, by income in 2017

Income %Poverty	Income $	Premium Cap
<100%	<11,880	No Cap
100% - 133%	11880 - 15800	2.04%
133% - 150%	15800 - 17820	3.06% - 4.08%
150% - 200%	17820 - 23760	4.08% - 6.43%
200% - 250%	23760 - 29700	6.43% - 8.21%
250% - 300%	29700 - 35640	8.21% - 9.69%
300% - 400%	35640 - 47520	8.21% - 9.69%
Over 400%	Over 47520	No Cap

Notes: This table presents the tax credit premium cap by income in 2017 under the Affordable Care Act. The premium cap is the maximum percent of the income one must pay for the second-lowest silver plan available to their area. Source: Kaiser Family Foundation

In summary, if a person's income is too high to be eligible for the subsidy, even if his willingness to pay for the plan is low, he will not receive the tax credit. However, if an individual's willingness to pay for the plan is higher than the equilibrium premium, but he is eligible to receive the tax credit, he would still benefit from a price reduction.

'Federal Subsidies for Health Insurance Coverage for People Under Age 65: 2019 to 2029 — Congressional Budget Office. 2019. 2 May 2019. https://www.cbo.gov/publication/55085.

Tebaldi, Pietro. 2022. 'Estimating Equilibrium in Health Insurance Exchanges: Price Competition and Subsidy Design under the ACA. Working Paper. Working Paper Series. National Bureau of Economic Research. https://doi.org/10.3386/w29869.