Health Insurance Exchange Design in an Empirical Equilibrium Labor Market Model*†

JOB MARKET PAPER

Naoki Aizawa‡
University of Pennsylvania
January 10, 2014

Abstract

The Affordable Care Act (ACA) requires that all U.S. states establish regulated individual insurance markets, known as health insurance exchanges (HIX), to facilitate the purchase of health insurance. In this paper, I evaluate the current HIX system and examine its optimal design, accounting for adverse selection and equilibrium labor market interactions. I first develop and empirically implement a life cycle equilibrium labor market search model integrated with the pre-ACA health insurance market. Various forms of individual heterogeneity are incorporated to understand the welfare consequences of introducing HIX. I estimate the model by the method of simulated moments using a combination of micro data sources. Through counterfactual experiments, I find that the ACA substantially reduces the uninsured rate. However, it also decreases aggregate labor productivity because more workers choose employment in less productive firms. Next, I examine the optimal design of HIX by choosing the values of three major design components—individual mandates (tax penalties on the uninsured), premium subsidies and age-based rating regulations. I find that the optimal combination of these components makes it less beneficial for older workers relative to younger workers to purchase health insurance from HIX. Implementing the optimal structure leads to a substantial welfare gain relative to the HIX implemented under the ACA while achieving higher labor productivity and a slightly lower uninsured rate.

*Please check my website for the most recent version: https://economics.sas.upenn.edu/graduate-program/candidates/naoki-aizawa
†I am grateful to my advisors Andrew Shephard, Holger Sieg, Kenneth I. Wolpin, and especially Hanming Fang for their guidance and support at all stages of this research project. I also thank Mike Abito, Cecilia Fieler, Camilo García-Jimeno, Jonathan Kolstad, Dirk Krueger, Iourii Manovskii, Petra Todd, Bob Town, Rakesh Vohra, Yuichi Yamamoto, an extensive list of Penn Econ graduate students, and seminar participants at the University of Pennsylvania and 2013 North American Summer Meeting of the Econometric Society for comments and discussions. Financial support from the Center for Retirement Research at Boston College, Grant #5001537 from the Social Security Administration, is gratefully acknowledged. Any errors are my own.
‡E-mail: naizawa@sas.upenn.edu. Department of Economics, University of Pennsylvania, 3718 Locust Walk, Philadelphia, PA 19104.
1 Introduction

The Patient Protection and Affordable Care Act of 2010 (ACA) represents the most significant health care reform in the United States in the past 40 years. The proposal and passage of the ACA was driven by the fact that close to 20% of the U.S. population does not have health insurance. Under the pre-ACA system, individuals tend to be uninsured if their family heads do not work at firms offering health insurance through employer sponsored health insurance (ESHI). Among the uninsured, 80% are employed, but most of them are not offered health insurance by their employer; the remaining 20% are not employed. Although many provisions of the ACA will take effect starting in 2014, one of the most important provisions is the establishment of health insurance exchanges (HIX). HIX are individual insurance markets where insurance plans cannot price or deny coverage based on preexisting conditions. A core idea of the ACA is to let the uninsured purchase health insurance from HIX. To provide individuals with incentives to participate in HIX, the ACA includes features such as individual mandates (a tax penalty on the uninsured), premium subsidies and an age-based rating. Although the desirability of these features has been discussed extensively in policy debates, there have been few studies evaluating the current HIX system or examining the possibility of alternative designs that improve welfare.

To examine the welfare impact of HIX design, a number of issues need to be addressed. First, HIX may be subject to adverse selection due to the prohibition of pricing based on preexisting conditions. This prohibition, by attracting disproportionately more unhealthy individuals, may cause healthy individuals to remain uninsured and incur the tax penalty rather than pay an actuarially unfair (higher) premium. Second, HIX interacts with the labor market due to the presence of ESHI. Individual health insurance purchase decisions in HIX, the key decisions affecting the extent of adverse selection, depend on the availability of ESHI, an outcome endogenously determined in the labor market. HIX may also affect worker labor supply and job mobility decisions, as well as firm decisions about offering ESHI.

In this paper, I evaluate the current HIX system and consider the question of its optimal design, accounting for adverse selection and equilibrium labor market interactions. First, I develop and empirically implement a life cycle equilibrium labor market search model integrated with the pre-ACA health insurance market. Various forms of individual heterogeneity are incorporated to understand the heterogeneous welfare consequences of HIX. Second, using the estimated model, I investigate the impact of HIX, as implemented under the ACA, on individuals with different characteristics and the aggregate consequences. Understanding these responses is necessary to conduct the welfare analysis. Third, I conduct a normative analysis. I examine the optimal design of HIX.

---

1 Roughly 60% of the non-elderly have ESHI. Also, 10% of worker compensation consists of ESHI premia (Kaiser Family Foundation and Health Research and Educational Trust, 2009).
2 Health insurance exchanges are also known as health insurance marketplaces.
3 Other important components of the ACA affecting the uninsured rate are Medicaid expansion and employer mandate (a tax penalty to large firms not providing coverage); these will be incorporated in my analysis as well.
4 In the text of the ACA, individual (employer) mandates are formally referred to as individual (employer) shared responsibility.
by finding, subject to the ACA’s government revenue constraint, the values of three major design components that maximize a utilitarian social welfare function—tax penalties on the uninsured, premium subsidies and age-based rating regulations.  

The benchmark model is designed to explain key patterns among health, health insurance and labor market outcomes observed in the pre-ACA economy. It builds on several strands of literature. First, it builds on a growing literature of empirical life cycle models of labor supply and health (for example, Rust and Phelan (1997); Blau and Gilleskie (2008); Khwaja (2010); French and Jones (2011); Papageorge (2012), among others). Second, it builds on the small literature on equilibrium search models with endogenous ESHI provision (Dey and Flinn (2005); Bruegemann and Manovskii (2010); Aizawa and Fang (2013)).

In the model, individuals make health care utilization, labor supply and job mobility decisions over the life cycle. In each period, health status affects current period utility, the distribution of latent medical expenditure shocks and labor productivity. Health status is influenced by health care utilization, which takes into account its impact on future health status and its cost. Medical expenditure risk is insured if individuals are covered by health insurance. Health insurance may be offered by firms as part of compensation. In the labor market, both non-employed and employed workers meet firms randomly and then decide whether to accept a job offer based on the compensation package, which consists of a wage and an ESHI offering. Employed workers accumulate work experience, which increases their stock of skills. Finally, individuals differ by education and unoberved type, the latter being a determinant of risk preference, initial labor market skill, and the evolution of health status.

Firms, which differ by their productivity, determine compensation packages to maximize their steady state flow profit. To account for various anti-discrimination laws restricting their choice of compensation package, I assume that firms set their wage offers by choosing a skill price subject to the constraint that the health cannot be priced, and decide whether to offer health insurance to all employees or not. Therefore, their coverage decisions are made to take into account the impact on the composition of workers having different characteristics within the firm. Moreover, I assume that health insurance costs are tax-deductible. I characterize a steady-state equilibrium where economic decisions by workers and firms are simultaneously determined.

The model features various forms of individual heterogeneity, both observed and unobserved, to understand the welfare consequences of HIX. I incorporate life cycle decision-making and skill characteristics because the major components of HIX design (individual mandates, premium subsidies and age-based rating regulation) vary by age and income. I also incorporate heterogeneity in health

---

5The rationale of these policies in insurance markets with adverse selection has been extensively studied in public economics literature. See Einav, Finkelstein, and Levin (2010) and Einav and Finkelstein (2011) for the survey.

6Existing anti-discrimination laws in the U.S. prohibit compensation packages from being based on age and health. Starting in 2014, firms are also prohibited from offering different sets of health plans to full-time employees with different income levels. Also, if ESHI is offered, workers cannot obtain premium subsidies from HIX unless the premium contribution of the ESHI plan for singles exceeds 9.5% of annual income. While part-time workers can obtain premium subsidies regardless of the ESHI offering, this paper considers a demographic group that consists of few part-time workers.
risk and risk preference, because those characteristics are known to be important determinants of health insurance purchasing decisions. Moreover, all of these characteristics impact the worker’s and the firm’s responses to the design of HIX. Individuals of different ages and skills change their labor supply and job mobility decisions differently because premia and subsidies in HIX vary by a worker’s age and income. Also, individual heterogeneity impacts the firm’s decision about offering ESHI given that it is affected by the composition of the firm’s workforce. Although incorporating various life cycle dimensions of individuals into an equilibrium labor market search model substantially complicates the analysis, the model is still numerically tractable under my approach to solve the firm’s compensation package, using techniques that are an extension of Barlevy (2008), Burdett, Carrillo-Tudela, and Coles (2011) and Bagger, Fontaine, Postel-Vinay, and Robin (2013).

I estimate the model using three data sources: the 2004 Survey of Income and Program Participation (SIPP), the 2004-2007 Medical Expenditure Panel Survey (MEPS), and the Kaiser Family 2004-2007 Employer Health Insurance Benefit Survey (Kaiser). The first two are panel data sets on worker-side labor market, health, health insurance, and medical expenditure, while the third is a cross-sectional firm-level data set containing information about firms’ characteristics and health insurance coverage. Worker-side data shows that individual health insurance coverage status is positively correlated with wages, education status, and age. Firm-side data shows that large firms tend to offer ESHI. Estimation is carried out via the method of simulated moments. The model fits worker-side moments such as health insurance, wage, employment, health, medical expenditure and their transitions over age profiles and education status, as well as firm-side moments such as the coverage rate and the size distribution.

The model estimates show that high-productivity firms are more likely to offer ESHI, which can be understood as follows. In the model, although firms want to attract more productive workers, high productivity firms are more likely to attract them because they can offer greater compensation. These more productive workers are typically experienced workers, who tend to be older and thus have a higher demand for health insurance. They also are more educated workers. From my estimates, educated workers are more risk-averse and have a higher demand for health insurance than less educated workers. The consequence of sorting of workers with high health insurance demand leads high-productivity firms to offer ESHI. This mechanism, although intuitive but generally ignored in the literature, simultaneously explains positive correlations among worker age, wage, education status, and health insurance coverage and also between firm size and health insurance coverage.

In counterfactual experiments, I introduce HIX to the pre-ACA economy as a competitive individual insurance market where individuals can purchase health insurance. Moreover, I incorporate other important features of the ACA, including employer mandate and Medicaid expansion. I first investigate the impact of the ACA. The key features of the ACA’s design for HIX are as follows. First, the tax penalty imposed on the uninsured increases with their income. Second, premium subsidies decrease with income. Third, the age-based rating regulation is such that the maximum

\[ \text{See Chetty and Finkelstein (2012) for an extensive review of the findings in the literature of empirical insurance markets.} \]
allowable premium ratio (MPR) between the oldest and the youngest is 3.

I find that the ACA decreases the uninsured rate in my estimation sample from 23.6% to 7.8%, where the remaining uninsured are mainly young employed workers who are healthy. Compared with them, the population of individuals purchasing health insurance from HIX consists of the sicker and older individuals. It therefore indicates the presence of adverse selection. The decrease in the uninsured rate leads to an increase in the fraction of healthy individuals, by 1 percentage point, as insured individuals tend to take health care utilization whenever they are hit by medical expenditure shocks relative to the uninsured.

Although the fraction of healthy individuals increases, and healthy individuals are more productive, I find that the steady state level of aggregate labor productivity decreases by 0.6%. This result arises because more workers are allocated to low productivity firms. The main channel is the inability of firms to make ESHI offerings separately for each employee and to provide the opportunity for employees to choose the source of coverage (i.e., from HIX or ESHI). If firms offer ESHI, their employees lose the opportunity to purchase subsidized health insurance from HIX. This makes accepting a job offer from a firm offering ESHI less attractive to older and less skilled workers, who can purchase health insurance at the lower cost given the ACA’s design of premium subsidies and rating regulations. Those workers prefer not to move to high productivity jobs offering ESHI unless these jobs offer a high wage. As a result, more workers are allocated to low productivity firms. This decline of labor productivity contributes to a decrease in output per capita. Although the ACA increases the employment rate by 0.4 percentage point due to the increase in the fraction of healthy individuals, output per capita decreases by 0.2%.

Next, I investigate the optimal design of HIX, which maximizes social welfare subject to the ACA’s government revenue constraint. I allow the government to optimally set the age-based rating regulation, which determines the maximum premium ratio (MPR) between the oldest and the youngest, and the premium subsidies and tax penalties to the uninsured as nonlinear functions of age and income. I find that the optimal combination of these policies increases aggregate labor productivity by 0.5% relative to the ACA while achieving a slightly lower level of the uninsured rate, 7.6% under the optimal HIX. To achieve the same welfare under HIX implemented by the ACA, the government needs to provide an annual lump-sum transfer to individuals amounting to $195 per year, which corresponds to 7.6% of medical expenditure. The optimal structure makes it less beneficial for old workers relative to young workers to purchase health insurance from HIX by setting larger MPR and subsidies that decrease with age, rather than age-independent subsidies as in the ACA. In this structure, the adverse selection problem among the young is partially resolved. Moreover, it gives older workers an incentive to work at firms offering ESHI. This increases the allocation of workers from low to high productivity firms, raising aggregate labor productivity.

Finally, I assess the importance of modeling equilibrium labor market interactions to evaluate the design of HIX. I assume that firm’s compensation package is exogenously determined and the same as the one under the pre-ACA economy. Then, I evaluate the impact of each component of

8The exact specification of HIX designs and other components in the ACA (employer mandate and Medicaid expansion) modeled in the paper is in section.
HIX, as well as the optimal design of HIX. I find that both are qualitatively and quantitatively very different from the one with endogenous compensation packages. This finding suggests the importance of modeling equilibrium labor market interactions to evaluate the design of HIX.

The related literature. This study is related to several strands of literature. First of all, it belongs to a new and growing literature evaluating the impact of the ACA. Ericson and Stær (2012), Hackmann, Kolstad, and Kowalski (2013) and Handel, Hendel, and Whinston (2013) develop an equilibrium model of HIX with adverse selection and examine HIX designs. Ericson and Stær (2012) study the role of imperfect competition in the Massachusetts (MA) unsubsidized health insurance exchanges established due to the 2006 MA health care reform which has many features in common with the ACA, while Hackmann, Kolstad, and Kowalski (2013) investigate the extent of adverse selection and the optimal individual mandate in those exchanges. Handel, Hendel, and Whinston (2013) examine equilibria when multiple insurance products can be traded in the competitive insurance market. As in this study, these studies investigate the efficacy of individual mandates, premium subsidies, and age-based rating regulations. The common assumption of these studies is that individuals have access to HIX only and therefore they do not consider equilibrium labor market interactions.

Aizawa and Fang (2013) study the labor market effects of the ACA, arguing that explicit modeling of the labor market equilibrium is crucial to evaluate the ACA. They examine the impact of the ACA under the setting where workers are infinitely lived and homogeneous except in health. I advance this research agenda in a number of directions. First and most importantly, I shift the focus to the normative analysis, especially the optimal design of HIX, which requires modeling individual life cycle decisions and heterogeneity. Although incorporating these features into an equilibrium labor market search model substantially complicates the framework, I show how to maintain the tractability of analyses. Second, I show how equilibrium effects of HIX interact with individual heterogeneity. In particular, because the current HIX system forces a partial pooling between young and old individuals, the impact of the ACA on individuals with different ages differs substantially. Third, I show that the optimal structure of HIX consists of age-dependent policies.

The welfare analysis conducted in this paper is also related to Kolstad and Kowalski (2012), who evaluate the welfare impact of the 2006 MA health care reform by focusing on labor market distortions. This paper has a number of important differences from theirs: First, theirs is based on the sufficient statistic approach which allows to conduct a welfare analysis using difference-in-difference identification strategy. This approach requires both pre- and post-reform data, the latter being not available for the ACA at present. Second, while they consider labor market distortions in terms of the level of employment and wages, my paper also takes into account distortions arising

---

9 See Bundorf, Levin, and Mahoney (2012) and Geruso (2012) for studying premium rating within employers.

10 See Fohll (2012) and Hai (2013) for an evaluation of Medicaid expansion.

11 See also Finkelstein, Taubman, Wright, Bernstein, Gruber, Newhouse, Allen, Baicker, and the Oregon Health Study Group (2012), Baicker, Finkelstein, Song, and Taubman (2013) and Garthwaite, Gross, and Notowidigdo (2013) for an evaluation of the impact of Medicaid on health and labor supply by focusing on recent Oregon Medicaid experiments and Tennessee Medicaid reform.
from the misallocation of workers. Third, while they examine welfare costs from the labor market distortion alone, my welfare analysis considers both the labor market and HIX together. My focus on the general equilibrium effects of HIX design is also related to studies assessing the macroeconomic impacts of ACA (Bruegemann and Manovskii (2010), Ozkan (2011), Cole, Kim, and Krueger (2012), Hansen, Hsu, and Lee (2012), Pashchenko and Porapakkarm (2013), and others). Although their models are richer than mine in certain dimensions (e.g., saving and capital accumulation), my paper complements theirs by endogenizing firm coverage decisions and allowing both individual and firm heterogeneity, which play an important role in characterizing the optimal design of HIX.

Second, this paper is also related to a large literature investigating the link between health insurance systems and labor markets, early contributions of which are reviewed by Currie and Madrian (1999) and Gruber (2000). A main topic in this literature has been whether the existing employer based health insurance system leads to an inefficient sorting of workers across firms, known as job lock and job push problems. Among them, Dey and Flinn (2005) is the closest to my study. They develop a search-matching-bargaining model with endogenous ESHI provisions and quantify the extent of job lock and job push in the pre-ACA economy. This paper extends their approach and shows that HIX causes additional inefficiency in worker allocation.

Finally, this paper is methodologically related to several branches of the labor and public finance literature. The empirical approach is closely related to the literature on structural estimation of equilibrium search allowing worker heterogeneity. Recently, Bagger, Fontaine, Postel-Vinay, and Robin (2013) estimate an equilibrium search model with two sided heterogeneity, allowing for individual human capital accumulation. This paper adds to the literature by fitting various additional worker life cycle economic events and decisions within an empirical equilibrium labor market search model. My focus on how a certain institution affects labor productivity through worker reallocation is also related to Gourio and Roys (2012) and Garicano, LeLarge, and Van Reenen (2013), who empirically study the impact of regulations in France that are dependent on firm-size on worker reallocation and labor productivity. My empirical welfare analysis is related to the literature studying optimal insurance and taxation policies, including Einav, Finkelstein, and Schrimpf (2010) for optimal mandates in annuity markets, and also Conesa, Kitao, and Krueger (2009) and Blundell and Shephard (2012) for optimal taxation. Finally, my welfare analysis about age- and income-based pricing and subsidies is related to the optimal taxation literature studying the role of tagging (e.g., Akerlof (1978); Michelacci and Ruffo (2011); Weinzierl (2011); Farhi and Werning (2013)).

The rest of the paper is organized as follows. Section 2 presents the pre-ACA model of this paper; Section 3 describes the data sets; Section 4 explains my estimation strategy; Section 5...
presents my estimation results; Section 6 describes the main results from evaluating the ACA and the components of HIX; Section 7 shows the main results from evaluating the ACA and the current HIX system; Section 8 concludes and discuss directions for future research.

2 Model

2.1 Environment

In this section, I first lay out an economic environment for the benchmark model, i.e., the model of the pre-ACA economy, which I use for estimation. Time is discrete and measured in periods of four months. Consider an economy populated with a continuum of workers with a measure $M > 0$ and a continuum of employers with a measure normalized to 1. They are randomly matched in a frictional labor market. Each worker lives for a finite horizon $t = t_0, \ldots, T$, while employers exist infinitely. I use $\beta \in (0, 1)$ to denote the discount factor. Each worker makes health care, labor supply and job mobility decisions up to the period $T$. Then, they exogenously retire from the labor market and are replaced by newborn workers. Upon entering the labor market, the new workers are initially heterogeneous with respect to their education status $ed$ which is either college graduate ($C$) or non-college graduate ($NC$) and with respect to their time-invariant type $type \in \{type1, type2\}$, the latter of which is a determinant of individual preference, labor market skills, and health transitions.

2.1.1 Individual preference

Each individual has time-separable, expected-utility preferences, which are defined over consumption $C_{it}$; employment status $P_{it} \in \{0, 1\}$, which takes a value of 1 if he is employed and 0 if he is not employed; and health status $h_{it} \in \{H, U\}$, which takes on value $H$ if he is healthy and $U$ if he is unhealthy. Let $U_t(C_{it}, P_{it}, h_{it}; type)$ be the period utility for individuals with time-invariant type $type$, which takes the following functional form:

$$U_t(C_{it}, P_{it}, h_{it}; type) = \exp(-\gamma_{type} C_{it}) - \eta_{pt} P_{it} - \eta_{ht} I(h_{it} = U) - \eta_{htp} P_{it} I(h_{it} = U)$$

where $\gamma_{type}$ is the CARA coefficient, $\eta_{pt}$ is the disutility from working which varies with individual age; $\eta_{ht}$ is the disutility from being unhealthy $U$; and $\eta_{htp}$ is the disutility of work for an unhealthy individual which varies with individual age. The last term is incorporated to fit the relationship between health and employment status. I assume that individuals can neither save nor borrow. The budget constraint of the individual is then given by

$$C_t = \tau_m(w_t) P_t + (1 - P_t) b - OOP^{INS}(x_t, m_t)$$

15 My choice of four months as a unit of time is motivated by the construction of data sets. See Section 3 for details.

16 As is common with many empirical life cycle labor supply models, $type$ will be treated as heterogeneous and unobserved by the econometrician in the empirical part of this paper.
where $\tau_w(w_t)$ is after-tax labor income, $b$ is non-employed income, and $OOP^{INS}(x_t m_t)$ is out-of-pocket expenditure. Out-of-pocket expenditure is a function of period $t$‘s medical expenditure and health insurance status: $x_t \in \{0, 1\}$ is health care choice, $m_t$ is latent medical expenditure shocks, and $INS \in \{0, 1\}$ is health insurance status, where $INS = 0$ if the individual is uninsured and $INS = 1$ if the individual is insured through ESHI. Although it is plausible to allow heterogeneous characteristics of insurance plans, given the limitation of the data, I assume that insurance is a homogeneous product in the pre-ACA model. Moreover, I assume that health insurance provides full insurance to individuals.

In addition, I specify that after period $T$, individuals receive the terminal value, which is merely a function of health status. The formal specification of the terminal value is described in Section 2.2.1.

2.1.2 Health shock and health transition

In each period, a worker is hit by latent medical expenditure shock $m_t$ which is a function of health $h_t$, age (measured in four-month intervals) $t$, and idiosyncratic shock $\epsilon_t$. I specify its functional form as follows:

\[
\begin{align*}
m_t &= \max\{m_t^* - \kappa_{h_t}, 0\} \\
\mu_t &= \exp(\omega_1^t + \omega_2^t h_t + \omega_3^t t^2 + \epsilon_t^m), \\
\epsilon_t^m | h_t &\sim i.i.d. N(0, \sigma_{h_t}^2), \\
h_t &\in \{H, U\}
\end{align*}
\]

where $m_t^*$ is a latent health shock which is distributed over the log normal distribution; and $\epsilon_t$ is i.i.d. idiosyncratic shock which is conditionally heteroskedastic with respect to health status $h_t$. In this specification, I incorporate a parameter $\kappa_{h_t} > 0$ which is used to capture the possibility that individuals do not report any positive medical expense because they are not hit by a positive amount of latent medical expenditure shocks, i.e., they are not hit by health shocks at all. Moreover, note that all of the parameters in the latent medical expenditure distribution differ according to the beginning of period health status $h_t$.

Conditional on latent medical expenditure shocks $m_t$, the worker chooses health care utilization $x_t \in \{0, 1\}$, which affects the realization of the next period health status. The transition to next period health status is determined by

\[
\begin{align*}
\Pr[h_{t+1} = k|x_t, h_t, \epsilon_t^m, t, type] &= \frac{\exp(\phi_1 x_t m_t + \phi_2 x_t + \phi_3 x_t m_t + \phi_4 t + \sum \phi_5^i (h_t = i) + \sum \phi_6^i (type = j))}{\sum_{k'} \exp(\phi_1 x_t m_t + \phi_2 x_t + \phi_3 x_t m_t + \phi_4 t + \sum \phi_5^i (h_t = i) + \sum \phi_6^i (type = j))}
\end{align*}
\]

where $1(type = j)$ is the dummy variable for individual unobserved type $j$. It is important to allow an interaction term between $x_t$ and $m_t$. Otherwise, to avoid reducing consumption, individuals tend to choose no health care utilization when they are hit by large latent medical expenditure shocks.
2.1.3 Health insurance market

In the baseline model intended to capture the pre-ACA U.S. health insurance market, I assume that workers can obtain health insurance only if their employers offer it. This is a simplifying assumption meant to capture the fact that the individual private insurance market is very small in the U.S.\footnote{Indeed, the fraction of individuals with individual insurance among the whole sample is 2%, and the fraction of such individuals among the sum of individuals with individual insurance and the uninsured is just 10% in my data set. Moreover, even for those who are covered by individual insurance, insurance products typically do not cover pre-existing conditions. In addition, one in seven applicants for health insurance are rejected (Hendren (2013)). \footnote{Hendren (2013) provides both theoretical and empirical analyses showing that adverse selection leads to a collapse of many other individual insurance markets.}} In my counterfactual experiment, I will introduce a competitive individual insurance market which I call the health insurance exchange (HIX).

2.1.4 Individual labor productivity

Each individual possesses labor productivity which affects the size of their compensation. I assume that an individual produces output $e_X(p)$ as a function of (1) a vector of individual characteristics $X = (ed, type, E_t, h_t)$ where $ed$ is education, $type$ is individual permanent type, $E_t$ is labor market experience, and $h_t$ is health status, and (2) the permanent productivity of the firm the individual is currently matched with, denoted by $p$. The log of output is specified as

$$
\ln(e_X(p)) = e_w^*(ed, type, E_t) + e_h^*(h_t) + p
$$

where $e_w^*(ed, type, E_t)$ is the worker skill explained by $(ed, type, E_t)$ and $e_h^*(h_t)$ is the worker skill explained by $h_t$. I assume that output is multiplicatively separable in $(ed, type, E_t)$ and $h_t$. The separability of health is assumed to maintain tractability when I characterize firm’s optimal wage policy.

Individual labor market experience is accumulated as long as the individual is employed. That is,

$$
E_{t+1} = \begin{cases} 
E_t + 1 & \text{if } P_t = 1 \\
E_t & \text{if } P_t = 0
\end{cases}
$$

2.1.5 Firm

Firms are heterogeneous with respect to their permanent productivity. In the population of firms, the distribution of productivity is denoted by $\Gamma(\cdot)$ which has a density function $d\Gamma$ that is continuous and positive everywhere. In my empirical application, I specify $\Gamma$ to be lognormal with mean $\mu_p$ and variance $\sigma^2_p$, i.e., $p \sim \ln N(\mu_p, \sigma^2_p)$. 

$$
\ln(e_X(p)) = e_w^*(ed, type, E_t) + e_h^*(h_t) + p
$$
Firms have access to a constant return to scale production function. In each period, they offer a package of wage and health insurance provision to maximize their steady state profit flow. If they offer health insurance, they incur the cost of health insurance provision, which is equal to the sum of the total expected medical expenditure of their workforce and a fixed administrative cost $\xi_{ESHI}$. The health insurance costs are tax-exempt.

### 2.1.6 Compensation package

Health insurance provision and the wage offer are determined as a solution of the firm’s profit maximization problem. Given the laws prohibiting discrimination in compensation packages based upon health or age, firms cannot offer separate compensation packages to each individual. Moreover, given this constraint and dimensionality of individual characteristics, it is hard to solve the optimal compensation contract for each worker. Therefore, I reduce the dimension of the potential contract space so that the model is still tractable, but captures the most important patterns of the data.

Specifically, I assume that firms post a skill price $\theta^{ed}$ for each skill group $ed$ subject to the constraint that health $h_t$ cannot be priced. Moreover, I assume that firms decide whether to offer health insurance to all of their workforce $INS \in \{0, 1\}$. As a result, a worker with characteristics $\tilde{X} = (ed, type, E_t, h_t)$ in a firm offering a compensation package $(\theta, INS)$ receives a wage offer which is equal to

$$w^{INS}_{\tilde{X}}(\theta) = \theta^{INS}_{INS} \exp(e_w^*(ed, type, E_t)).$$

(3)

Given the wage offer, the flow profit of a firm with productivity $p$ from hiring a worker with $\tilde{X}$ is $\exp(e_w^*(ed, type, E_t))(\exp(p)e_h^*(h_t) - \theta)$ where $g_h^*(h_t)$ is the productivity effect of health.\footnote{I implicitly assume that employer contribution of insurance premium is 100% in this baseline model. That is, if workers want to opt out the coverage, they cannot receive an additional wage to compensate. This is not an unrealistic choice because the current U.S. average is 85% for single workers’ premiums.}

Finally, I assume that wages are subject to classical measurement error, with errors following a log normal distribution.\footnote{One could instead model these errors as transitory skill shock. My treatment of measurement error follows the work of Low, Meghir, and Pistaferri (2010) which estimates the wage process of a life-cycle search model using the same data source (SIPP).}

### 2.1.7 Labor market

The labor market is frictional and workers and employers randomly meet. A non-employed worker receives a new job offer from an employer with probability $\lambda_u^{ed}$ and an employed worker receives an offer from an employer with probability $\lambda_e^{ed}$ where $ed$ is individual education status. The compensation is drawn from the offer distribution $F^{ed}(\theta, INS)$. Upon receiving the job offer, the worker decides whether to accept it.

In addition to changing jobs, employed workers are allowed to quit and become non-employed. Furthermore, they are hit by an exogenous job destruction shock with probability $\delta^{ed}$, upon which workers lose their current jobs. Because the model period is a relatively long, I allow the exogenous
job destruction shock and the arrival of a new job offer to occur within the same period with probability $\delta^{ed}\lambda^{ed} > 0$. Moreover, I allow an additively separable preference shock to being non-employed $\epsilon_t^n$ which follows a Type-I extreme value distribution with scale parameter $\sigma_t$.\(^{20}\)

2.1.8 Timing in a period

At the beginning of each period, individuals, who are heterogeneous in their education, unobserved type, age, and health status, are either unemployed or working for employers offering different wage and health insurance packages. I now describe the explicit timing assumptions for a period that I use in the derivation of the value functions in Section 2.2. These particular timing assumptions simplify our derivation, but they are not crucial.

1. An employed individual produces output and accumulates labor market experience.
2. Idiosyncratic health shock $\epsilon_t^m$ is realized.
3. An individual makes health care decision $x_t$.
4. The next period health status is realized.
5. An employed worker is hit by an exogenous job destruction shock with probability $\delta$.
6. An individual draws a preference shock for being non-employed.
7. A non-employed worker receives a job offer with probability $\lambda_u$ and decides whether to accept the job offer. An employed worker receives a job offer with probability $\lambda_e$ and chooses to accept the offer, to stay at the current job, or to quit and become non-employed. The employed worker who does not receive the offer decides whether to stay at the current job or quit and become non-employed.

2.1.9 Initial condition

I assume that non-college graduates enter the labor market at age 18, while college graduates enter at age 22. I assume that upon entry into the labor market, they have no labor market experience and start their career as non-employed. Notice that the return to education is captured by labor productivity. Workers are assigned an unobserved type, the distribution of which is specific to education. Moreover, I assume that all newborn workers are healthy.\(^{21}\) Finally, I assume that the population of individuals and firms grows at the constant rate $n$ in each period. Let $\mu^r_t$ denote

\(^{20}\)This preference shock is incorporated to smooth the labor supply function with respect to wage offers, which is useful when I solve the model numerically. An alternative is to add time-invariant continuous heterogeneous flow utility of being non-employed, as done by [Bontemps, Robin, and Van den Berg 1999]. Because of the nonstationary nature of the individual problem in my model, I need to allow for dynamics of unobserved flow utility, which substantially complicates analysis.

\(^{21}\)Note, however, that their health evolution is affected by their permanent type type and therefore they face different health risks.
the fraction of a cohort size of age \( t \) in the total population of individuals at period \( \tau \). I consider a steady state economy in which \( \mu^\tau \) is constant over time, i.e., \( \mu^\tau = \mu \).

### 2.2 Analysis of the Model

In this section, I define the steady state equilibrium of the model. To do so, I first consider the individual life cycle optimization problem. The solution to the individual optimization problem is then used to determine the steady state distribution of individuals. Next, I formulate the firm’s optimization problem and characterize their optimal decisions for posting compensation package.

#### 2.2.1 Individual optimization problem over the life cycle

At the beginning of a period, the state space of an individual at age \( t \) consists of a six dimensional vectors: \( (\tilde{X}_t, \theta, INS) \): four dimensional vector of individual characteristics \( \tilde{X}_t = (ed, type, E_t, h_t) \) and the compensation package \( (\theta, INS) \). Note that \( INS_t = 0 \) for any non-employed worker. Consider that all individuals face offer distribution of compensation package, which is denoted by \( F^{ed}(\theta, INS) \). Denote \( V_0^t(\tilde{X}_t) \) as the value function of a non-employed worker with age \( t \) who is in a state \( \tilde{X}_t \) and \( V_1^t(\tilde{X}_t, \theta, INS) \) is the value function of an employed worker with age \( t \) who is in a state \( (\tilde{X}_t, \theta, INS) \).

Consider a non-employed worker having characteristics \( \tilde{X}_t = (ed, type, E_t, h_t) \). His value function is defined by

\[
V_0^t(\tilde{X}_t) = \mathbb{E}_{\epsilon^m} \left[ \max_{x_t} U_t(C_t(x_t, \epsilon_t^m, \tilde{X}_t), 0, h_t) + \beta \sum_h \Pr(h_{t+1} = \hat{h}|\tilde{X}_t, x_t, \epsilon_t^m) \lambda_u^{ed} V_0^{t+1}(\tilde{X}_{t+1}) \right. \\
+ \left. \beta \sum_{\hat{h}} \Pr(h_{t+1} = \hat{h}|\tilde{X}_t, x_t, \epsilon_t^m)(1 - \lambda_u^{ed}) V_0^{t+1}(\tilde{X}_{t+1}) \right]
\]

subject to budget constraint (1) where \( \tilde{X}_{t+1} \) is the next period’s individual characteristics, \( \tilde{X}_{t+1} = (ed, type, E_t, \hat{h}) \). \( \tilde{V}_0^t(\tilde{X}_{t+1}) \) is defined as

\[
\tilde{V}_0^t(\tilde{X}_{t+1}) = \int \mathbb{E}_{\epsilon^m} \left[ \max \{ V_0^{t+1}(\tilde{X}_{t+1}), \epsilon_t^m, V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS) \} \right] dF^{ed}(\theta, INS)
\]

The first term is the flow utility from the current period’s consumption which is affected by the health care utilization choice \( x_t \) and the realization of an i.i.d. latent medical expenditure shock \( \epsilon_t^m \). The second term is the expected value from receiving a job offer, denoted by \( \tilde{V}_0^t(\tilde{X}_{t+1}) \), multiplied by the health transition probability which is affected by \( x_t \) and the realization of the i.i.d. latent medical expenditure shock \( \epsilon_t^m \). Conditional on receiving the offer, the individual decides whether to accept it, which depends on the value from staying as unemployed \( V_0^{t+1}(\tilde{X}_{t+1}) \) and the value from accepting the job offer with compensation package \( (\theta, INS) \) denoted \( V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS) \). Note that the decision is subject to the preference shock of working, denoted by \( \epsilon_t^p \). The third term is the expected value from not receiving a job offer.

Similarly, consider an employed worker who is in a state \( (\tilde{X}_t, \theta, INS) \). The value function of such an individual is defined by

\[
\tilde{V}_0^t(\tilde{X}_{t+1}) = \int \mathbb{E}_{\epsilon^m} \left[ \max \{ V_0^{t+1}(\tilde{X}_{t+1}), \epsilon_t^m, V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS) \} \right] dF^{ed}(\theta, INS)
\]
\[ V_t^1(\tilde{X}_t, \theta, INS) = \mathbb{E}_{x_t}^m \left[ \max_{x_t} U_t(C_t(x_t, \epsilon_t^m, \tilde{X}_t, \theta, INS), 1, h_t) + \beta \sum_{\tilde{h}} \Pr(h_{t+1} = \tilde{h}|\tilde{X}_t, x_t, \epsilon_t^m) \delta^{ed} V_{t+1}^1(\tilde{X}_{t+1}) \right] \]

subject to budget constraint \[ (\text{6}) \] where \( \tilde{X}_{t+1} = (ed, type, E_t + 1, \tilde{h}) \) is the next period individual characteristics and

\[ \tilde{V}_1^1(\tilde{X}_{t+1}) = (1 - \lambda_e^{ed}) V_{t+1}^1(\tilde{X}_{t+1}) + \lambda_e^{ed} \mathbb{E}_{x_t}^m \left[ \max\{V_0^t+1(\tilde{X}_{t+1}) + \epsilon_t^m, V_1^t+1(\tilde{X}_{t+1}, \theta, INS')\}\right] dF^{ed}(\theta', INS') \]

and

\[ \hat{V}_1^1(\tilde{X}_{t+1}, \theta, INS) = (1 - \lambda_e^{ed}) \mathbb{E}_{x_t}^m \left[ \max\{V_0^t+1(\tilde{X}_{t+1}) + \epsilon_t^m, V_1^t+1(\tilde{X}_{t+1}, \theta, INS')\}\right] \\
+ \lambda_e^{ed} \mathbb{E}_{x_t}^m \left[ \max\{V_0^t+1(\tilde{X}_{t+1}) + \epsilon_t^m, V_1^t+1(\tilde{X}_{t+1}, \theta, INS), V_1^t+1(\tilde{X}_{t+1}, \theta', INS')\}\right] dF^{ed}(\theta', INS'). \]

As before, the first term is the flow utility from the current period’s consumption, which is affected by health care utilization choice \( x_t \) and the realization of the i.i.d. latent medical expenditure shock \( \epsilon_t^m \). The second term is the expected value from being hit by an exogenous job destruction shock, \( \tilde{V}_1^1(\tilde{X}_{t+1}) \), which consists of two parts, expressed in the following square bracket. The first part is the gain from not receiving a job offer. In this case, the worker is forced to be non-employed in the next period. The second part is the gain from receiving a job offer. In this case, the worker decides whether to accept the job offer or remain non-employed. The third term is the expected value from being hit by an exogenous job destruction shock, denoted by \( \hat{V}_1^1(\tilde{X}_{t+1}, \theta, INS) \). Again, it consists of two parts, expressed in the following square bracket. The first part is the gain from not receiving a job offer from another firm. In this case, the worker decides whether to stay in the current job or to become non-employed. The second part is the case in which the worker receives a job offer. If the worker receives a job offer with compensation package \( (\theta', INS') \), he decides whether to move to a new firm, stay in his current job, or quit and become non-employed.

Because individuals live for a finite period, we need to specify the value of the terminal period. The terminal value function of an individual with state \( S_T \) is simply given by

\[ V_T^i(S_T) = v_T^i h_{T_T} U, \]

where \( i \in \{0, 1\} \) is his employment status at period \( T \).

It is easy to see that the decisions of non-employed individuals about accepting a job offer follow the standard reservation rule strategy. Here, notice that the strategy is nonstationary because my model uses a life cycle environment. It is easy to see that for the employed worker working at a firm offering a compensation package of skill price and health insurance provision status \( (\theta, INS) \) and receiving a job offer from a firm offering a compensation package \( (\theta', INS') \), his job acceptance decision is purely determined by comparing these two compensation packages and is not affected.
by firm productivity per se. This is because individual wages are fully determined by skill prices given the skill level, as seen from (3).

Moreover, it is clear that individual decisions to choose health care utilization \( x_t \) take into account its impact on the current period’s consumption and the dynamics of health status. Because health enters into an individual’s direct utility, which allows interactions with disutility from working, an individual’s health care decisions affect his ability to work in the future.

### 2.2.2 Steady state worker distribution

The steady state distribution of workers is given by workers’ optimal choices. I assume that at \( t = 1 \), new workers enter the population as non-employed with exogenously determined health, education, and type. Because workers live for a finite period, the steady state distribution of individuals includes age as state variables. Denote \( g_t(\tilde{X}, \theta, INS) \) as a steady state measure of employed workers with age \( t \) and characteristics \( \tilde{X}_t = (ed, type, E, \hat{h}) \) receiving compensation packages \( (\theta, INS) \). Similarly, \( u_t(\tilde{X}) \) is defined as a steady state measure of the non-employed with age \( t \) and characteristics \( \tilde{X} \). Note that \( g_t \) and \( u_t \) are fully determined by the inflow from the distribution of \( g_{t-1} \) and \( u_{t-1} \). The determinants of \( g_t(\tilde{X}, INS, \theta^{ed}_{INS}) \) and \( u_t(\tilde{X}) \) are described as follows. Given the offer distribution of compensation package \( F^{cd}(\theta, INS) \):

\[
g_t(\tilde{X}_t, \theta, INS) \frac{1}{1 + n} = \sum_{h_{t-1}} g_{t-1} \left( \tilde{X}_{t-1}^{A}, \theta, INS \right) \mathbb{E} e^m_t \left[ \Pr(h_t = \hat{h} | \tilde{X}_{t-1}^{A}, \theta, INS, e^m_t) \right] \Psi_{t-1}(\tilde{X}_t, \theta, INS) + \sum_{h_{t-1}} u_{t-1} \left( \tilde{X}_{t-1}^{B} \right) \mathbb{E} e^m_t \left[ \Pr(h_t = \hat{h} | \tilde{X}_{t-1}^{B}, e^m_t) \right] \Pr(\Omega_1(\tilde{X}_t, \theta, INS)) \times \lambda^{ed}_{c} F^{ed}(\theta, INS) + \sum_{h_{t-1}} \sum_{INS'} \int g_{t-1} \left( \tilde{X}_{t-1}^{A}, \theta', INS' \right) \mathbb{E} e^m_t \left[ \Pr(h_t = \hat{h} | \tilde{X}_{t-1}^{A}, \theta', INS', e^m_t) \right] 1(\theta, INS, \theta', INS') d\theta' \times (1 - \delta^{cd}) \lambda^{ed}_{c} F^{ed}(\theta, INS) + \sum_{h_{t-1}} \sum_{INS'} \int g_{t-1} \left( \tilde{X}_{t-1}^{A}, \theta', INS' \right) \mathbb{E} e^m_t \left[ \Pr(h_t = \hat{h} | \tilde{X}_{t-1}^{A}, \theta', INS', e^m_t) \right] \Pr(\Omega_1(\tilde{X}, \theta, INS)) d\theta' \times \delta^{cd} \lambda^{ed}_{c} F^{ed}(\theta, INS)
\]

where \( \tilde{X}_{t-1}^{A} = (ed, type, E_{t-1}, h_{t-1}) \) and \( \tilde{X}_{t-1}^{B} = (ed, type, E_t, h_{t-1}) \) are individual characteristics in the last period for the employed and non-employed, respectively, which can turn into \( \tilde{X} \) in this period. \( \Psi_{t-1}(\tilde{X}_t, \theta, INS) \) is the probability of staying at the same firm

\[
\Psi_{t-1}(\tilde{X}_t, \theta, INS) = (1 - \delta^{cd}) \left( 1 - \lambda^{ed}_{c} \right) \Pr(\Omega_1(\tilde{X}_t, \theta, INS)) + (1 - \delta^{cd}) \lambda^{ed}_{c} \Pr(\Omega_1(\tilde{X}, \theta, INS)) \Pr(\Omega_1(\tilde{X}_t, \theta, INS)).
\]

(Note that I will provide an intuition for each term of \( \Psi_{t-1} \) in the next paragraph). \( \Pr(\Omega_1(\tilde{X}_t, \theta, INS)) \) is the probability that individuals with \( \tilde{X} \) prefer to work a job with compensation package \( (\theta, INS) \) over being non-employed, where the set \( \Omega_1(\tilde{X}_t, \theta, INS) \) is formally defined as

\[
\Omega_1^{ed}(\tilde{X}_t, \theta, INS) = V^0_t(\tilde{X}_t) + e^m_t < V^1_t(\tilde{X}_t, \theta, INS)
\]
where $\epsilon_t^I$ is a preference shock from not working. Because it follows a type-I extreme value distribution with scale parameter $\sigma_n$, \( \text{Pr}(\Omega_t^E(X_t, \theta, INS)) \) is characterized as

\[
\text{Pr}(\Omega_t^E(X_t, \theta, INS)) = \frac{\exp(V_1^I(X_t, \theta, INS))}{\exp(V_1^I(X_t, \theta, INS)) + \exp(V_1^I(X_t, \theta, INS))}.
\]

Next, \( \text{Pr}(\Omega^E_2(X_t, \theta, INS)) \) is the probability that individuals who receive a job offer from other firms decide to stay in the current job:

\[
\text{Pr}(\Omega^E_2(X_t, \theta, INS)) = \text{Fed}(\theta, INS) + \text{Fed}(\hat{\theta}_{INS}(\tilde{X}_t, \theta, INS)).
\]

for $INS \neq INS$ where $\hat{\theta}_{INS}(\tilde{X}_t)$ is the threshold skill price which can be defined as

\[
V_1^I(\tilde{X}_t, \theta, INS)) = V_1^I(\tilde{X}_t, \hat{\theta}_{INS}(\tilde{X}_t, \theta, INS)).
\]

$1(\theta, INS, \theta', INS')$ is the indicator function such that individuals prefer to take an offer from $(\theta, INS)$ over $(\theta', INS')$:

\[
1(\theta, INS, \theta', INS') = \begin{cases} 1 & \text{if } V_1^I(\tilde{X}_t, \theta, INS)) > V_1^I(\tilde{X}_t, \theta', INS')) \\
0 & \text{otherwise} \end{cases}
\]

While the expression of how $g_t$ is determined looks rather complicated, it can be understood fairly easily. The first term is the inflow from the workers who work in firms offering the same contract $(\theta, INS)$, transition to the health status $\tilde{h}$ this period, and decide to stay at the same firm. Notice that the health transition probability is denoted by \( \text{Pr}(h_t = \tilde{h}|X_{A1}, \theta, INS', \epsilon_t^m) \) which does not include health care decision $x_t$, as the optimal health care decision is a function of a vector of $(X_{A1}, \theta, INS', \epsilon_t^m)$. The probability of staying at the same firm, denoted by \( \Psi_{t-1}(\theta, INS, \tilde{X}) \), consists of two terms. The first term is the probability that the worker does not receive any offer at all and prefers to stay at the same firm over quitting and becoming non-employed. The second term is the probability that the worker receives an offer from another firm but turns it down and stays at the same firm. The second term of the right-hand side of (7) is the inflow from the non-employed. This happens if a non-employed individual gets an offer with probability $\lambda_u^f \text{Fed}(\theta, INS)$ and decides to accept it. The third and fourth terms of the right-hand side of (7) are the inflow from the currently employed working in other firms but receiving and accepting offers with compensation $(\theta, INS)$. The difference between the third and fourth terms derives from whether they are hit by an exogenous job destruction shock, which affects the probability of accepting a job offer. Finally, in order to determine the size of $g_t$, I need to take into account that population size grows at a constant rate $n$. Therefore, the population at age $t$, $g_t$, should be divided by $1 + n$.

Similarly, one can express the determinant of $u_t(X_t)$ as
\[
\frac{u_t(\tilde{X}_t)}{1 + \eta} = \sum_{h_{t-1}} u_{t-1} \left( \tilde{X}_{t-1}^B \right) \mathbb{E}_{\tilde{t}^n} \left[ \Pr(h_t = \hat{h}_t | \tilde{X}_{t-1}^B, c_t^m) \right] (1 - \lambda_{u}^{cd}) \\
+ \sum_{h_{t-1}} u_{t-1} \left( \tilde{X}_{t-1}^B \right) \mathbb{E}_{\tilde{t}^n} \left[ \Pr(h_t = \hat{h}_t | \tilde{X}_{t-1}^B, c_t^m) \right] \Pr(\Omega_t^U(\tilde{X}_t)) \lambda_{u}^{cd} \\
+ \sum_{h_{t-1}} \sum_{INS} \int g_{t-1} \left( \tilde{X}_{t-1}^A, \theta, INS \right) \mathbb{E}_{\tilde{t}^n} \left[ \Pr(h_t = \hat{h}_t | \tilde{X}_{t-1}^A, c_t^m) \right] \Pr(\Omega_t^U(\tilde{X}_t, \theta, INS)) dF(\theta, INS)(1 - \delta^{cd})(1 - \lambda_{e}^{cd}) \\
+ \sum_{h_{t-1}} \sum_{INS} \int g_{t-1} \left( \tilde{X}_{t-1}^A, \theta, INS \right) \mathbb{E}_{\tilde{t}^n} \left[ \Pr(h_t = \hat{h}_t | \tilde{X}_{t-1}^A, c_t^m) \right] \Pr(\Omega_t^U(\tilde{X}_t, \theta, INS)) dF(\theta, INS)(1 - \delta^{cd}) \lambda_{e}^{cd} \\
+ \sum_{h_{t-1}} \sum_{INS} \int g_{t-1} \left( \tilde{X}_{t-1}^A, \theta, INS \right) \mathbb{E}_{\tilde{t}^n} \left[ \Pr(h_t = \hat{h}_t | \tilde{X}_{t-1}^A, c_t^m) \right] \Pr(\Omega_t^U(\tilde{X}_t)) d\theta \delta^{cd} \lambda_{e}^{cd} \\
\tag{8}
\]

where \( \tilde{X}_{t-1}^A \) and \( \tilde{X}_{t-1}^B \) are defined as above, and \( \Pr(\Omega_t^U(\tilde{X}_t)) \) is the probability that the unemployed with characteristics \( \tilde{X} \) decides to turn down the offer:

\[
\Pr(\Omega_t^U(\tilde{X}_t)) = \int \Pr(V_0^l(\tilde{X}_t) + \epsilon_t^n > V_1^l(\tilde{X}_t, \theta, INS)) dF^{cd}(\theta, INS).
\]

\( \Pr(\Omega_t^U(\tilde{X}_t, \theta, INS)) \) is the probability that an employed individual with characteristics \( \tilde{X}_t \) who has a job with compensation package \( (\theta, INS) \) decides to quit and become non-employed:

\[
\Pr(\Omega_t^U(\tilde{X}_t, \theta, INS)) = \Pr(V_0^l(\tilde{X}_t) + \epsilon_t^n > V_1^l(\tilde{X}_t, \theta, INS)).
\]

Finally, \( \Pr(\Omega_t^U(\tilde{X}_t, \theta, INS)) \) is the probability that an employed individual with characteristics \( \tilde{X}_t \) who has a job with compensation package \( (\theta, INS) \) receives a job offer from another firm but decides to quit and become non-employed:

\[
\Pr(\Omega_t^U(\tilde{X}_t, \theta, INS)) = \int 1 \left( V_0^l(\tilde{X}_t) + \epsilon_t^n = \max \left\{ V_0^l(\tilde{X}_t) + \epsilon_t^n, V_1^l(\tilde{X}_t, \theta, INS), V_1^l(\tilde{X}_t, \theta', INS') \right\} \right) dF^{ed}(\theta', INS').
\]

Again, one can interpret (8) very intuitively. The first two terms of the right-hand side of (8) are the inflow from the currently non-employed workers with state variable \( \tilde{X}_{t-1}^B \). They stay non-employed and have state variable \( \tilde{X}_t \) in the following period if they are not offered health insurance, or if they receive a job offer but turn it down. The third and the fourth terms are the inflow from currently employed workers with state variable \( \tilde{X}_{t-1}^A \) who choose to quit and become non-employed: the third term is the case where they do not receive a job offer from another firm; the fourth term is the case where they receive a job offer from another firm but turn it down. The fifth and sixth terms are the inflow from the currently employed who are hit by job destruction shock. The fifth term depicts those individuals who are forced to be unemployed and the sixth term depicts individuals who also receive job offers from other firms, but choose to quit and become non-employed.
Finally, the steady state condition requires that the sum of all measures of workers is equal to $M$, that is
\[ \sum_t \sum_{X_t} u_t(X_t) + \sum_t \sum_{X_t} \sum_{INS} g_t(X_t, \theta, INS) d\theta = M. \] (9)

Therefore, by using (7), (8) and (9), one can characterize the steady state worker distribution. Although the expression looks rather complicated, it can be derived fairly easily by using forward induction. That is, we can analytically calculate the whole distribution once we know the period 1 distribution $g_1$ and $u_1$, the worker’s value function, health transition function, and the offer distribution $F^{ed}(\theta, INS)$.

From the steady state employment measure $g_t(X_t, \theta, INS)$, one can define the terms related to firm size, by following the same spirit as Burdett and Mortensen (1998) and Postel-Vinay and Robin (2002). Specifically, the density of employees with age $t$ and characteristics $X_t$ for firms offering compensation package ($\theta, INS$) is given by
\[ l_t(X_t, \theta, INS) = \frac{g_t(X_t, \theta, INS)}{f^{ed}(\theta, INS)} \]
where $f^{ed}(\theta, INS)$ is density of firms offering compensation package ($\theta, INS$).

### 2.2.3 Firm’s optimization problem

Next, I formalize the firm’s problem. Firms choose wage offers and health insurance offerings to maximize steady state profit flow. I assume that the firm draws a shock, $\epsilon^{ESH1}$, in each period, which is specific to its choice of whether to offer health insurance. The shock is additively separable from the steady state profit flow, but affects the payoff of health insurance provisions. I incorporate the choice-specific shock to smooth the insurance provision decisions of the employers. This problem can be formulated as
\[ \Pi(p, \epsilon^{HI}) = \max \{ \Pi_1(p) + \epsilon^{ESH1}, \Pi_0(p) \} \]
where $\Pi_{INS}$ is the conditional profit under the health insurance offering status $INS \in \{0, 1\}$. It is a solution of the following problems:
\[ \Pi_1(p) = \max_{\theta_1^{ed}} \sum_t \sum_{X} \pi_1(X, t, p, \theta_1^{ed}) l_t(\bar{X}, \theta_1^{ed}, 1) - \xi_{ESH1} \] (10)
\[ \Pi_0(p) = \max_{\theta_0^{ed}} \sum_t \sum_{X} \pi_0(X, t, p, \theta_0^{ed}) l_t(\bar{X}, \theta_0^{ed}, 0). \] (11)

where $\pi_1(X, t, p, \theta_1^{ed})$ is the flow net profit of firms with productivity $p$ offering skill price $\theta_1^{ed}$ and health insurance when hiring a worker with characteristics $(\bar{X}, t)$
\[ \pi_1(X, t, p, \theta_1^{ed}) = \exp(\epsilon^*_w(ed, type, E_t)) \left( \exp(p) \exp(\epsilon^*_h(h_t)) - \theta_1^{ed} \right) - E[m^*_t|INS(p) = 1]. \] (12)

The second term of the right-hand side is the expected medical expenditure of the employed worker having characteristics $X$ and age $t$ at firms with productivity $p$ offering health insurance.
\( \pi_0(\bar{X}, t, p, \theta_{1}^{ed}) \) is the flow of net profit of firms with productivity \( p \) offering skill price \( \theta_{1}^{ed} \) but not offering health insurance by hiring a worker with characteristics \((\bar{X}, t)\):

\[
\pi_0(\bar{X}, t, p, \theta_0^{ed}) = \exp(e^*_w(ed, type, E_t)) \left( \exp(p) \exp(e^*_r(h_t)) - \theta_0^{ed} \right).
\]  

(13)

I assume that \( \epsilon^{ESH1} \) follows an i.i.d. Type-I extremum value distribution with scale parameter \( \sigma_f \). As a result, the fraction of firms with productivity \( p \) offering health insurance is characterized by

\[
\Delta(p) = \frac{\exp\left(\frac{\Pi_1(p)}{\sigma_f}\right)}{\exp\left(\frac{\Pi_0(p)}{\sigma_f}\right) + \exp\left(\frac{\Pi_1(p)}{\sigma_f}\right)}.
\]

(14)

### 2.2.4 Steady state equilibrium

Now, I am in a position to define an equilibrium.

**Definition 1.** A steady state equilibrium consists of workers’ value functions \( \{V_0^t, V_1^t\} \) and corresponding policy functions, the steady state measure of workers with characteristics of \( \bar{X} = (ed, type, E, h) \), \( g_t(\bar{X}, \theta, INS) \) and \( u_t(\bar{X}) \), the firms’ compensation packages consisting of skill price \( \{\theta_1^{ed}(p), \theta_0^{ed}(p)\} \) and insurance offer \( \{\Delta(p)\} \) for all \( p \), and offer distribution \( F^{ed}(\theta, INS) \) such that

1. Given offer distribution \( F^{ed}(\theta, INS) \), the value functions at age \( t \) \( \{V_0^t, V_1^t\} \) and corresponding policy functions solve (4), (5), and (6).

2. Given worker’s optimization behavior described by \( \{V_0^t, V_1^t\} \) and corresponding policy functions and offer distribution \( F^{ed}(\theta, x) \), \( g_t(\bar{X}, \theta, INS) \) and \( u_t(\bar{X}) \) must satisfy (7), (8) and (9).

3. Given \( F^{ed}(\theta, INS) \) and the steady state employee sizes implied by \( g_t(\bar{X}, \theta, INS) \) and \( u_t(\bar{X}) \), a firm with productivity \( p \) chooses to offer health insurance with probability \( \Delta(p) \), where \( \Delta(p) \) is given by (14). Moreover, conditional on insurance choice \( x \), the firm offers skill price \( \{\theta_1^{ed}(p)\} \) that solves (10) and (11) respectively for \( INS \in \{0, 1\} \).

4. The postulated distributions of offered compensation packages are consistent with the firms’ optimization behavior \( \{\theta_1^{ed}(p), \theta_0^{ed}(p), \Delta(p)\} \). Specifically, \( F^{ed}(\theta, INS) \) must satisfy

\[
F^{ed}(\theta, 1) = \int_0^\infty \mathbf{1}(\theta_1^{ed}(p) < \theta^{ed}) \Delta(p) d\Gamma(p),
\]

(15)

\[
F^{ed}(\theta, 0) = \int_0^\infty \mathbf{1}(\theta_0^{ed}(p) < \theta^{ed}) [1 - \Delta(p)] d\Gamma(p).
\]

(16)

for each education group \( ed \in \{NC, C\} \).

### 2.2.5 Characterization of equilibrium

I can characterize the firm’s optimal skill price \( \theta_{INS}^{ed}(p) \) by extending an approach by Bontemps, Robin, and Van den Berg (1999) and Bontemps, Robin, and Van den Berg (2000). By applying an envelope condition to (10), we obtain that

\[
\Pi_1(p) = \sum_t \sum_{\bar{X}} \left( e^{\bar{X}}(p) - \frac{\partial E[m^{\bar{X}}_r|HI = 1]}{\partial p} \right) l_t(\bar{X}, \theta_1^{ed}(p), 1).
\]
Integrating over \([p_1, p]\), where \(p_1\) represents the lowest productivity firms that hire a positive number of workers, we obtain that

\[
\Pi_1(p) = \Pi_1(p_1) + \int_{p_1}^{p} \sum_t \sum_{\tilde{X}} \left( e_{\tilde{X}}(p') - \frac{\partial E[m^t_{X_1}\text{INS} = 1]}{\partial p'} \right) l_t \left( \tilde{X}, \theta_1(p'), 1 \right) dp'.
\]

By equating this with (10), one can characterize \(\theta^e_1(p)\) and \(\theta^e_0(p)\) as follows:

**Proposition 2.** For \(p > p_1\), \(\theta^e_1(p)\) and \(\theta^e_0(p)\) satisfy

\[
\theta^e_1(p) = \frac{\sum_t \sum_{\tilde{X}} \left( e_{\tilde{X}}(p) - E[m^t_{X_1}] \right) l_t \left( \tilde{X}, \theta^e_1(p), 1 \right) - \Pi_1(p_1)}{\sum_t \sum_{\tilde{X}} \exp(e^*_t(\text{ed}, \text{type}, E_t)) l_t \left( \tilde{X}, \theta^e_1(p), 1 \right)} - \frac{\int_{p_1}^{p} \sum_t \sum_{\tilde{X}} \left( e_{\tilde{X}}(p') - \frac{\partial E[m^t_{X_1}\text{INS} = 1]}{\partial p'} \right) l_t \left( \tilde{X}, \theta^e_1(p'), 1 \right) dp'}{\sum_t \sum_{\tilde{X}} \exp(e^*_t(\text{ed}, \text{type}, E_t)) l_t \left( \tilde{X}, \theta^e_1(p), 1 \right)}.
\]

\[
\theta^e_0(p) = \left( \frac{\sum_t \sum_{\tilde{X}} e_{\tilde{X}}(p) l_t \left( \tilde{X}, \theta^e_1(p), 0 \right) - \Pi_0(p_0) - \int_{p_0}^{p} \sum_t \sum_{\tilde{X}} e_{\tilde{X}}(p') l_t \left( \tilde{X}, \theta^e_0(p'), 0 \right) dp'}{\sum_t \sum_{\tilde{X}} \exp(e^*_t(\text{ed}, \text{type}, E_t)) l_t \left( \tilde{X}, \theta^e_1(p), 0 \right)} \right). \tag{18}
\]

This form can be utilized when we numerically solve the equilibrium of the model. \(\theta^e_{\text{INS}}(p_1)\) and \(\theta^e_{\text{INS}}(p_0)\) must be solved by maximizing (10) and (11) without relying on (17) and (18).

Due to the complexity of the model, I cannot solve the equilibrium analytically. I instead solve the equilibrium numerically using the algorithm described in Appendix A. The complexity of the model also prevents me from establishing a proof of the existence and uniqueness of the equilibrium, but, by conducting extensive numerical simulations, I always find a unique equilibrium based on this algorithm.

### 3 Data Set

In this section, I describe my data set and its sample selection. The model describes rich individual-level dynamics over the life cycle regarding labor market outcome, health status, health insurance coverage, and medical expenditure. Moreover, it describes how firms with different productivity decide wage and health insurance provisions which determine the size of the firm. Therefore, an ideal data set for estimating the model is to use single employee-employer matched data which contains this information. However, such a data set is not available in the U.S. Instead, I combine three separate data sets for the estimation: (1) 2004 Survey of Income and Program Participation; (2) 2004-2007 Medical Expenditure Panel Survey; and (3) Kaiser Family 2004-2007 Employer Health Benefit Survey. I choose the data period 2004-2007 because estimating the model using the data after 2008 is not ideal for the following two reasons. First, the Great Recession has generated dramatic changes in the labor market. Second, possibly due to the policy announcement effect, there was a sharp jump in the health insurance offering rate in 2010, which disappears after
2011. These short-term dramatic changes are difficult to capture with my model. Instead, I choose 2004-2007, when the economic environment was relatively stable.

3.1 Survey of Income and Program Participation

I obtain an individual-level labor market outcome and associated health status, health insurance coverage status, and demographic information from the 2004 Panel of Survey of Income Program Participation (hereafter, SIPP 2004). The SIPP 2004 interviews individuals every four months for up to twelve times, so that an individual may be interviewed over a four-year period. It consists of two parts: (1) the core module, and (2) the topical module. The core module, which is based on the interviews from each wave, contains detailed monthly information regarding individuals’ demographic characteristics and labor force activity, including earnings, number of weeks worked, average hours worked, and employment status, as well as whether the individual changed jobs during any month in the survey period. In addition, at each interview date the SIPP 2004 gather a variety of health insurance variables. It specifies the source of insurance so we know whether it is ESHI, private individual insurance, or Medicaid, and we also know whether it is obtained through the individual’s own name or the spouse’s name. The topical module, which is based on annual interviews (i.e., at interview waves 3, 6, 9 and 12), contains yearly information about the worker and his/her family members’ health status and out-of-pocket medical expenditure. To link individual health status with individual labor market outcomes and health insurance coverage status information, I match the core module with the topical module.

I construct the estimation sample as follows. The total sample size after matching the topical module and the core module is 131,532. I restrict my sample to men (dropping 71,283 female individuals) whose age are between 25-59 (dropping an additional 33,652 individuals). In addition, I only keep individuals who are not in school, are not self-employed, do not work in the public sector, do not engage in military service and do not participate in any government welfare program (dropping an additional 11,433 individuals). I also limit our sample to individuals who are either uninsured or covered by ESHI in their own name (dropping an additional 1,949 individuals). Finally, I exclude individuals receiving Social Security income (dropping an additional 95 individuals). The sample size for the estimation is a total of 11,797 individuals.

3.2 Medical Expenditure Panel Survey (MEPS)

The SIPP data set allows me to capture the dynamics of health insurance coverage driven by the labor market mobility, one of the main drivers determining individual insurance status under the pre-ACA health insurance system. However, a problem with using SIPP data for my estimation is the lack of information about total medical expenditure. To obtain the information, I use the Medical Expenditure Panel Survey (hereafter, MEPS) 2004-2007. MEPS is a set of

\[ \text{In both SIPP and MEPS, I use the self-reported health status to construct whether the individual is healthy or unhealthy. The self-reported health status has five categories. I categorize “Excellent”, “Very Good” and “Good” as Healthy (H) and “Fair” and “Poor” as Unhealthy (U).} \]
large-scale annual rotating panel surveys. I use its Household Component (HC), which surveys households in two consecutive years, collecting detailed information for each person in the household on demographic characteristics, health conditions, health status, use of medical services, charges and source of payments, access to care, satisfaction with care, health insurance coverage, income, and employment. To construct the estimation sample, I use the same criteria as SIPP 2004. The sample size for the estimation is a total of 17,536 individuals.

3.3 Kaiser Family Employer Health Benefits Survey (Kaiser)

Finally, I obtain firm-side information about the health insurance offering status and associated firm characteristics from the Kaiser Family 2004-2007 Employer Health Benefit Survey.\(^{23}\) It is an annual survey of the nation’s private and public firms having three or more workers. It contains information about firms’ characteristics (such as industry and size) and categorical information about employees’ demographics (such as age and annual wage), as well as information about health insurance (such as whether the employer offers health insurance, the type of plan offered, employees’ eligibility and enrollment, and whether the employer, the employee, or both contribute to the purchase of insurance.) I restrict the sample to firms which belong to the private sector and have at least three employees. The estimation sample size is 18,593\(^{24}\).

3.4 Descriptive Statistics

Table 1 reports the descriptive statistics of key variables in SIPP 2004. It is shown that the employed who receive health insurance receive a higher wage than the employed who do not. Moreover, they are slightly healthier than individuals who do not have health insurance.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraction of workers who are college graduates</td>
<td>0.4521</td>
<td>0.4977</td>
</tr>
<tr>
<td>Average worker age</td>
<td>40.5462</td>
<td>8.8822</td>
</tr>
<tr>
<td>Fraction of insured among employed workers</td>
<td>0.8409</td>
<td>0.3658</td>
</tr>
<tr>
<td>Average 4-month wage for employed workers, in $10,000</td>
<td>1.9383</td>
<td>2.0832</td>
</tr>
<tr>
<td>Average 4-month wage for insured Employees, in $10,000</td>
<td>2.1445</td>
<td>2.1896</td>
</tr>
<tr>
<td>Average 4-month wage for uninsured Employees, in $10,000</td>
<td>0.8483</td>
<td>0.7230</td>
</tr>
<tr>
<td>Employment rate</td>
<td>0.9111</td>
<td>0.2847</td>
</tr>
<tr>
<td>Fraction of healthy workers</td>
<td>0.9338</td>
<td>0.2487</td>
</tr>
<tr>
<td>Fraction of healthy workers among insured</td>
<td>0.9501</td>
<td>0.2178</td>
</tr>
<tr>
<td>Fraction of healthy workers among uninsured</td>
<td>0.8771</td>
<td>0.3284</td>
</tr>
</tbody>
</table>

Table 1: Summary Statistics: SIPP 2004.

\(^{23}\)Note that other studies investigating ESHI often use Robert Wood Johnson Foundation Employer Health Insurance Survey (e.g., Cebul, Rebitzer, Taylor, and Votruba\(^{21}\)) and Aizawa and Fang\(^{21}\)). The unit of observation is the establishment. However, the survey has not conducted after 1997.

\(^{24}\)The data also shows whether firms consist of a single establishment or not. In my data selection, 90.15% of firms consist of a single establishment.
In Table 2, I report the descriptive statistics for the individuals in MEPS 2004-2007. Note that the proportion of healthy workers is similar to that reported in SIPP. The proportion of individuals who are insured is somewhat under-reported relative to SIPP.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average worker age</td>
<td>41.0386</td>
<td>9.7948</td>
</tr>
<tr>
<td>Fraction of healthy workers</td>
<td>0.9116</td>
<td>0.2839</td>
</tr>
<tr>
<td>Fraction of insured among employed workers</td>
<td>0.7244</td>
<td>0.4468</td>
</tr>
<tr>
<td>Annual medical expenditure, in $10,000</td>
<td>0.2478</td>
<td>0.9583</td>
</tr>
</tbody>
</table>


I provide the descriptive statistics for firm-side data based on Kaiser 2004-2007 in Table 3. In general, firms that offer health insurance tend to have many employees. Moreover, the composition of employees systematically differs by firm’s health insurance offering status. First, firms offering health insurance tend to employ a larger share of high income employees. Moreover, firms offering health insurance also consist of a larger share of older workers.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Mean</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average firm Size</td>
<td>20.8198</td>
<td>53.8675</td>
</tr>
<tr>
<td>... for those that offer health insurance</td>
<td>29.6444</td>
<td>67.7277</td>
</tr>
<tr>
<td>... for those that do not offer health insurance</td>
<td>7.8903</td>
<td>12.3044</td>
</tr>
<tr>
<td>Health insurance coverage rate</td>
<td>0.5943</td>
<td>0.4911</td>
</tr>
<tr>
<td>... for those with less than 50 workers</td>
<td>0.5686</td>
<td>0.4954</td>
</tr>
<tr>
<td>... for those with 50 or more workers</td>
<td>0.9338</td>
<td>0.2486</td>
</tr>
<tr>
<td>Average fraction of employees more than age 26</td>
<td>0.8364</td>
<td>0.2093</td>
</tr>
<tr>
<td>... for those that offer health insurance</td>
<td>0.8636</td>
<td>0.1519</td>
</tr>
<tr>
<td>... for those that do not offer health insurance</td>
<td>0.7991</td>
<td>0.2650</td>
</tr>
<tr>
<td>Average fraction of employees annual salaries $21,000 or less</td>
<td>0.2108</td>
<td>0.3107</td>
</tr>
<tr>
<td>... for those that offer health insurance</td>
<td>0.1244</td>
<td>0.2328</td>
</tr>
<tr>
<td>... for those that do not offer health insurance</td>
<td>0.3288</td>
<td>0.3621</td>
</tr>
<tr>
<td>Average fraction of employees annual salaries $50,000 or more</td>
<td>0.2345</td>
<td>0.2833</td>
</tr>
<tr>
<td>... for those that offer health insurance</td>
<td>0.2739</td>
<td>0.2935</td>
</tr>
<tr>
<td>... for those that do not offer health insurance</td>
<td>0.1807</td>
<td>0.2601</td>
</tr>
</tbody>
</table>

4 Estimation

4.1 Identification

In this section, I first discuss identification of several key parameters of the model. Due to the complexity of the model, the argument is mainly heuristic. The key parameters related to health and health insurance are risk aversion, \( \gamma_{\text{type}} \), the disutility from bad health, \( \eta_{h}^{\text{type}} \), productivity loss due to bad health, \( \alpha_{h} \), a scale parameter for firms offering ESHI, \( \sigma_{f} \), and the fixed cost of offering ESHI, \( \xi_{\text{ESHI}} \). These parameters all affect firm’s health insurance coverage and it may be unclear how each parameter is separately identified.

While it is true that higher \( \gamma_{\text{type}} \) and \( \eta_{h}^{\text{type}} \), and lower \( \alpha_{h} \) increase firm’s coverage rate, \( \gamma_{\text{type}} \) and \( \eta_{h}^{\text{type}} \) affect worker-side moments such as life cycle patterns of coverage rate and associated job-to-job transitions and medical expenditure. Specifically, if \( \gamma_{\text{type}} \) and \( \eta_{h}^{\text{type}} \) are larger, we would expect to observe more frequent transitions of workers from jobs without health insurance to jobs with health insurance, even if the transition involves a reduction in wages. Moreover, the magnitude of the wage cut a worker is willing to tolerate in order to switch from a job without health insurance to a job with health insurance increases with \( \gamma_{\text{type}} \) and \( \eta_{h}^{\text{type}} \). Then, \( \eta_{h}^{\text{type}} \) is disciplined by fitting medical expenditure: because medical expenditure is a choice variable in the model and the benefit of medical expenditure is to improve future health status, I can identify the disutility from bad health through the variation of the fraction of positive medical expenditure by individual health insurance status conditional on individual characteristics, e.g., age. Finally, heterogeneity of \( \gamma_{\text{type}} \) and \( \eta_{h}^{\text{type}} \) by unobserved type can be identified by differential patterns of the above moments by education status, as the type distribution differs across education.

The remaining parameters are then identified in the following way. The scale parameter \( \sigma_{f} \) affects the relationship between the probability of offering health insurance and firm productivity (and thus firm size). The parameter \( \alpha_{h} \) has an additional effect on the differences in wages offered by firms, depending on whether they offer health insurance. This therefore affects the wage difference between the insured and the uninsured. Finally, the administrative cost \( \xi_{\text{ESHI}} \) is identified from the the probability (in the level) of small firms offering health insurance.

Finally, I discuss additional identification assumptions related to labor market parameters. As in standard labor market equilibrium models, it is impossible to identify intercepts in individual skill function described by \( [2] \) from skill prices. Therefore, I assume that the mean of log normal distribution for firm productivity distribution \( \mu_{p} \) is set to be 0. Furthermore, I assume that the intercept for type 2 individual dummy \( \alpha_{2}^{2} \) is equal to 0.

4.2 Estimation Strategy

The solution of the model serves as an input to the estimation procedure. Estimation is by the method of simulated moments (MSM). Specifically, a weighted average distance between sample moments and simulated moments is minimized with respect to the model’s parameters. The weights are the inverses of the estimated variances of the moments. The procedure requires a choice of
moments.

The following is a list of moments used in the estimation. Each moment on the worker-side is a conditional moment by education and age cohort.

1. Labor market status and its dynamics

   (a) employment rate
   (b) transition rate from non-employment to employment
   (c) job-to-job (JJ) transition rate
   (d) transition rate from employment to non-employment

2. Wage and health insurance

   (a) the fraction of the uninsured among employed
   (b) wage change through JJ conditional on before-after health insurance coverage status
   (c) the distribution of wage conditional on health insurance coverage status
   (d) wage change of job stayers conditional on health insurance coverage status
   (e) the distribution of wage among previously unemployed workers, conditional on health insurance coverage status

3. Health and medical expenditure

   (a) health status conditional on employment and health insurance status
   (b) annual health transition conditional on health
   (c) annual health transition conditional on health and health insurance status
   (d) annual medical expenditure conditional on health
   (e) annual medical expenditure conditional on health and health insurance
   (f) the fraction of zero medical expenditure conditional on health and health insurance

4. Firm characteristics

   (a) the fraction of firms with less than 50 workers
   (b) health insurance coverage rate by whether firm size is less than 50 workers
   (c) mean firm size conditional on health insurance offering coverage

---

25 One could incorporate moments regarding within-firm worker distribution. However, the available demographic information (e.g., the fraction of employees whose age is less than 26, the fraction of employees earning more than $15 per hour, etc.) is very coarse. Therefore, I decide not to include them as a target. However, the model prediction of these moments (specifically conditional moments) is qualitatively consistent. The prediction is available from the author upon request.
The estimation is done in the standard nested fixed point algorithm and described as follows:

1. Guess a vector of parameters \( q \).
2. Given the parameters \( q \), solve the equilibrium of the model and then simulate the data.
3. Evaluate the objective function using the simulated moments:

\[
\min_{\{q\}} G(q)'\Omega G(q),
\]

where \( G(q) \) is the vector of the value of each moment \( j \), \( G_j(q) \). Each moment \( G_j(q) \) is constructed as

\[
G_j(q) = \tilde{G}_j - \mu_j(q)
\]

where \( \tilde{G}_j \) is the sample moment of \( j \) and \( \mu_j(q) \) is the simulated moment.
4. Repeat steps 1-3 and find \( q \) to minimize the objective value.

I compute asymptotic standard errors following Gourieroux, Monfort, and Renault (1993). Certain parameters of the model are calibrated without using the model. In this paper, I do not try to estimate \( \beta \) but set \( \beta = 0.99 \) so that the annual interest rate is about 3%.\(^{26}\) Moreover, the population growth rate \( n \) is estimated using the SIPP 2004 sample by running a regression of cohort size on age. Estimates are \( n = 1.0005 \) per four-month period. Finally, the after-tax income schedule is specified by following Kaplan (2012) who approximates the U.S. income tax code as

\[
T(w) = \tau_0 + \tau_1 w(1+\tau_2)^{1+\tau_2}. \]

Kaplan (2012) estimates the parameters as \( \tau_0 = 0.0056, \tau_1 = 0.6377 \) and \( \tau_2 = -0.1362 \).

5 Estimation Result

The parameter estimates are summarized in the Tables 4-8. Table 4 reports the parameter estimates for individual preferences. Table 4 shows that the CARA coefficient is very heterogeneous across different types. In my empirical analysis, I consider the case where type 1 individuals are more risk averse.\(^ {27} \) If we convert this into relative risk aversion, we get 2.99 for the very high income group of type 1 workers, while it is 1.67 for type 2 workers. These estimates are much lower than the standard estimates in consumption/saving literature, but consistent with the literature of labor supply elasticity.\(^ {28} \) On the other hand, the disutility from bad health is relatively homogenous.

\(^{26}\) It is known from Flinn and Heckman (1982) that it is difficult to separately identify the discount factor \( \beta \) from the flow unemployed income \( b \) in standard search models.

\(^{27}\) Note that I do not impose any restrictions in terms of the correlation of risk type with other characteristics such as labor market skills, health transition, and disutility from bad health. These correlations are estimated to fit the data.

\(^{28}\) For example, French and Jones (2011) estimate the CRRA coefficient as greater than 8 in his life-cycle model of labor supply and saving for older workers. On the other hand, Chetty (2006) shows that the upper bound of the relative risk aversion coefficient for the average worker should be 2 if one estimates it from labor supply behavior, under the assumption that both college graduates and non-college graduates share a common parameter value.
across different types. I also find that the disutility of working is increasing over ages. Moreover, the coefficient on the interaction between disutility from working and bad health is positive, implying that old unhealthy individuals suffer higher disutility from working relative to young unhealthy individuals.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimates</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>CARA Coefficient for type 1 workers (γ₁)</td>
<td>1.2973</td>
<td>(0.038)</td>
</tr>
<tr>
<td>CARA Coefficient for type 2 workers (γ₂)</td>
<td>0.7259</td>
<td>(0.033)</td>
</tr>
<tr>
<td>disutility from bad health for type 1 workers (η₁)</td>
<td>0.0360</td>
<td>(0.001)</td>
</tr>
<tr>
<td>disutility from bad health for type 2 workers (η₂)</td>
<td>0.0380</td>
<td>(0.001)</td>
</tr>
<tr>
<td>disutility from working interacted with age (ηₚ)</td>
<td>0.0005</td>
<td>(0.00001)</td>
</tr>
<tr>
<td>interaction between disutility from working and bad health (ηₚ)</td>
<td>0.0223</td>
<td>(0.00003)</td>
</tr>
<tr>
<td>non-employment income (b)</td>
<td>0.0891</td>
<td>(0.001)</td>
</tr>
<tr>
<td>terminal value from bad health</td>
<td>-4.0711</td>
<td>(2.003)</td>
</tr>
</tbody>
</table>

Table 4: Parameter Estimate for Individual Preference

Next, Table 5 shows the parameter estimates for latent medical expenditure shocks and health transition processes. I find that the distribution of latent medical expenditure shocks differs substantially between healthy and unhealthy individuals. For example, although the constant term for healthy individuals in health shock process is much smaller than the constant term for unhealthy individuals, the age coefficient in the health shock process for healthy individuals is much larger than the age coefficient for unhealthy individuals. Moreover, the health transition process differs substantially between type 1 and type 2 workers. I find that type 1 workers are more likely to transition to being healthy in the next period relative to type 2 workers.

Table 6 shows the parameter estimates for the individual labor market environment. An important finding is that type 1 workers, who are more risk averse and are more likely to be healthy, are more productive than type 2 workers. Another important finding from Table 6 is that the productivity loss of being unhealthy, αₕ, is -0.337, which means that unhealthy workers produce 71.39% of the output of healthy workers.

Table 7 shows the parameter estimates for the distribution of workers. It shows that the fraction of type 1 workers among the college graduates is larger than that among the non-college graduates. Therefore, by combining results from Tables 4, 5, and 6, parameter estimates show that the college graduates are on average more risk averse, more productive, and more healthy than non-college graduates in the model. Moreover, the total measure of workers (relative to the measure of firms) is estimated to be 22.2050 to fit the mean of firm size observed in the data.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimates</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant term for healthy in health shock process ($\omega_H^H$)</td>
<td>-2.8652</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Age coefficient in health shock process for healthy individuals ($\omega_2^H$)</td>
<td>0.0097</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>Age square coefficient in health shock process for healthy individuals ($\omega_3^H$)</td>
<td>1.7001E-07</td>
<td>(1.0E-08)</td>
</tr>
<tr>
<td>Standard deviation in health shock process for healthy individuals ($\sigma_H$)</td>
<td>0.5613</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Scale parameter in health shock process for healthy individuals ($\kappa_H$)</td>
<td>0.0684</td>
<td>(0.0004)</td>
</tr>
<tr>
<td>Constant term for unhealthy in health shock process ($\omega_U^H$)</td>
<td>-1.9286</td>
<td>(0.008)</td>
</tr>
<tr>
<td>Age coefficient in health shock process for unhealthy individuals ($\omega_2^U$)</td>
<td>0.0001</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>Age square coefficient in health shock process for unhealthy individuals ($\omega_3^U$)</td>
<td>0.0000</td>
<td>(1.0E-08)</td>
</tr>
<tr>
<td>Standard deviation in health shock process for unhealthy individuals ($\sigma_U$)</td>
<td>1.7218</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Scale parameter in health shock process for unhealthy individuals ($\kappa_U$)</td>
<td>0.0000</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Dummy for today’s good health on tomorrow’s good health ($\phi_{1H}$)</td>
<td>3.2557</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Dummy for today’s bad health on tomorrow’s bad health ($\phi_{1U}$)</td>
<td>0.4002</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Dummy for $x_t$ conditional on positive latent shock on tomorrow’s bad health ($\phi_{2U}$)</td>
<td>-0.3895</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Coeff. for latent health shock for tomorrow’s bad health ($\phi_{3U}$)</td>
<td>0.000015</td>
<td>(0.000002)</td>
</tr>
<tr>
<td>Coeff. for age on tomorrow’s bad health ($\phi_{4U}$)</td>
<td>0.004</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Coeff. for dummy for type 2 for tomorrow’s bad health ($\phi_{5U}$)</td>
<td>0.1530</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Coeff. for interaction between $x_t$ and $m_t$ for ($\phi_{6U}$)</td>
<td>-0.00016</td>
<td>(0.00002)</td>
</tr>
<tr>
<td>Coeff. for interaction between age and bad health for tomorrow’s bad health ($\phi_{7U}$)</td>
<td>0.0025</td>
<td>(0.0013)</td>
</tr>
</tbody>
</table>

Table 5: Parameter Estimate for Latent Medical Expenditure Shocks and Health Transition Process

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimates</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dummy for college graduate in labor productivity ($\alpha_{1C}^C$)</td>
<td>-0.8008</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Dummy for non-college graduate in labor productivity ($\alpha_{1}^{NC}$)</td>
<td>-1.1200</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Dummy for type 1 in labor productivity ($\alpha_{1}^1$)</td>
<td>0.3668</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Experience coeff. for college graduate in labor productivity ($\alpha_{4C}^C$)</td>
<td>0.00083</td>
<td>(0.0003)</td>
</tr>
<tr>
<td>Experience coeff. for non-college graduate in labor productivity ($\alpha_{4}^{NC}$)</td>
<td>3.91E-05</td>
<td>(1.5E-05)</td>
</tr>
<tr>
<td>Experience coeff. for type 1 in labor productivity ($\alpha_{4}^1$)</td>
<td>2.16E-06</td>
<td>(5.5E-06)</td>
</tr>
<tr>
<td>Experience square coeff. for college graduate in labor productivity ($\alpha_{5C}^C$)</td>
<td>1.0E-09</td>
<td>(1.16E-08)</td>
</tr>
<tr>
<td>Experience square coeff. for non-college graduate in labor productivity ($\alpha_{5}^{NC}$)</td>
<td>-1.5E-06</td>
<td>(1.90E-07)</td>
</tr>
<tr>
<td>Experience square coeff. for type 1 in labor productivity ($\alpha_{5}^1$)</td>
<td>1.0E-08</td>
<td>(2.1E-07)</td>
</tr>
<tr>
<td>Productivity loss due to bad health ($\alpha_{b}$)</td>
<td>-0.3370</td>
<td>(0.004)</td>
</tr>
<tr>
<td>Scale parameter for preference shock not to work ($\sigma_n$)</td>
<td>0.0126</td>
<td>(0.0003)</td>
</tr>
<tr>
<td>Offer arrival rate for the non-employed college graduate ($\lambda_{u}^C$)</td>
<td>0.4340</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Offer arrival rate for the employed college graduate ($\lambda_{e}^C$)</td>
<td>0.1360</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Probability of exogenous match destruction for college graduate ($\delta^C$)</td>
<td>0.0101</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>Offer arrival rate for the non-employed non-college graduate ($\lambda_{u}^{NC}$)</td>
<td>0.2048</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Offer arrival rate for the employed non-college graduate ($\lambda_{e}^{NC}$)</td>
<td>0.0874</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Probability of exogenous match destruction for non-college graduate ($\delta^{NC}$)</td>
<td>0.0244</td>
<td>(0.0001)</td>
</tr>
<tr>
<td>Standard deviation of measurement error ($\sigma_w$)</td>
<td>0.0100</td>
<td>(0.0002)</td>
</tr>
</tbody>
</table>

Table 6: Parameter Estimate for Individual Labor Market Activities
The fraction of type 1 workers among the college graduates 0.4045 (0.004)
The fraction of type 1 workers among the non-college graduates 0.2715 (0.003)
Total measure of workers ($M$) 22.2050 (1.26)

Table 7: Parameter Estimate for the Distribution of Workers

Finally, Table 8 shows parameter estimates for firm-side characteristics. Note that the scale parameter of firm productivity distribution is normalized to 0 and therefore only the shape parameter is estimated. The key finding here is that the fixed cost of offering ESHI is estimated to be $552 per four-month, which is equivalent to $1,656 annually.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimates</th>
<th>Std. Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shape parameter of firms' log normal productivity distribution ($\sigma_p$)</td>
<td>0.8517</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Fixed cost of offering employer sponsored health insurance (\xi_{ESHI})</td>
<td>0.0552</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Scale parameter of choice-specific shock to EHI offering ($\sigma_f$)</td>
<td>0.3587</td>
<td>(0.0008)</td>
</tr>
</tbody>
</table>

Table 8: Parameter Estimate for Firm-side Characteristics

5.1 Model Fit

In this section, I first provide tables about model fits to show that the model can fit the most salient features of the data.

<table>
<thead>
<tr>
<th>Age</th>
<th>College graduate with $HI$</th>
<th>Non-college graduate with $HI$</th>
<th>College graduate w/o $HI$</th>
<th>Non-college graduate w/o $HI$</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-31</td>
<td>1.8974 1.8695 0.8753 0.8680</td>
<td>1.2624 1.4014 0.7500 0.7772</td>
<td>2.5425 2.3713 0.9752 1.0004</td>
<td>1.4231 1.4741 0.8915 0.8012</td>
</tr>
<tr>
<td>32-38</td>
<td>2.5425 2.3713 0.9752 1.0004</td>
<td>1.4231 1.4741 0.8915 0.8012</td>
<td>3.9-45</td>
<td>2.7911 2.7291 0.9603 0.9815</td>
</tr>
<tr>
<td>46-52</td>
<td>2.9782 3.0157 1.0895 0.9501</td>
<td>1.6467 1.5203 0.8252 0.8461</td>
<td>53-59</td>
<td>2.8987 2.9818 0.9935 0.9065</td>
</tr>
</tbody>
</table>

Table 9: Model fit: mean 4-month wage conditional on coverage, age and education.
Notes: (a) the unit is $10,000. (b) with indicates employed individuals with ESHI; w/o indicates employed individuals are are uninsured.

Table shows the model fit for the average wage conditional on health insurance status, age, and education group. In the data, there is a strong positive correlation between wage and age for
college graduates for individuals who have health insurance. On the other hand, wage-age slopes in other groups are rather flat. The model is able to fit these moments well.

Next, Table 10 shows the pattern of health insurance coverage status among the employed workers over ages. The data shows a positive correlation between coverage rate and age, regardless of education group. The model captures the insurance gain over the life cycle quite well. The model somewhat underpredicts the coverage rate among college graduates in the age group 25-31, while it overpredicts the coverage rate among non-college graduates in the age group 25-31. Table 11 shows the model fit for the employment rate across each of the age groups. The model can quantitatively explain the age profile of the employment rate for both education groups.

<table>
<thead>
<tr>
<th>Age</th>
<th>Data College graduate</th>
<th>Data Non-college graduate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model</td>
<td>Model</td>
</tr>
<tr>
<td>25-31</td>
<td>0.9045 0.8516</td>
<td>0.6557 0.7494</td>
</tr>
<tr>
<td>32-38</td>
<td>0.9494 0.9417</td>
<td>0.7621 0.8005</td>
</tr>
<tr>
<td>39-45</td>
<td>0.9516 0.9653</td>
<td>0.8235 0.8478</td>
</tr>
<tr>
<td>46-52</td>
<td>0.9527 0.9766</td>
<td>0.8843 0.9028</td>
</tr>
<tr>
<td>53-59</td>
<td>0.9607 0.9786</td>
<td>0.8911 0.9185</td>
</tr>
</tbody>
</table>

Table 10: Model fit: coverage rate among the employed conditional on age and education.

<table>
<thead>
<tr>
<th>Age</th>
<th>Data College graduate</th>
<th>Data Non-college graduate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model</td>
<td>Model</td>
</tr>
<tr>
<td>25-31</td>
<td>0.9788 0.9692</td>
<td>0.9175 0.9282</td>
</tr>
<tr>
<td>32-38</td>
<td>0.9787 0.9795</td>
<td>0.9380 0.9278</td>
</tr>
<tr>
<td>39-45</td>
<td>0.9693 0.9809</td>
<td>0.9262 0.9268</td>
</tr>
<tr>
<td>46-52</td>
<td>0.9624 0.9757</td>
<td>0.9207 0.9129</td>
</tr>
<tr>
<td>53-59</td>
<td>0.9504 0.9678</td>
<td>0.9129 0.9094</td>
</tr>
</tbody>
</table>

Table 11: Model fit: employment rate conditional on age and education.

Next, Table 12 shows the model fit for health status. It reports the fraction of healthy individuals across age groups and education status. One striking pattern observed in the data is that the difference between the fraction of college graduates and non-college graduates that are healthy increases over time. The model quantitatively accounts for this pattern well.

<table>
<thead>
<tr>
<th>Age</th>
<th>Data College graduate</th>
<th>Data Non-college graduate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Model</td>
<td>Model</td>
</tr>
<tr>
<td>25-31</td>
<td>0.9788 0.9692</td>
<td>0.9175 0.9282</td>
</tr>
<tr>
<td>32-38</td>
<td>0.9787 0.9795</td>
<td>0.9380 0.9278</td>
</tr>
<tr>
<td>39-45</td>
<td>0.9693 0.9809</td>
<td>0.9262 0.9268</td>
</tr>
<tr>
<td>46-52</td>
<td>0.9624 0.9757</td>
<td>0.9207 0.9129</td>
</tr>
<tr>
<td>53-59</td>
<td>0.9504 0.9678</td>
<td>0.9129 0.9094</td>
</tr>
</tbody>
</table>

Table 12: Model fit: health status conditional on age and education.

Next, Table 13 reports the model fit for the pattern of health status conditional on insurance and employment status. It is shown that the employed workers with ESHI are the most healthy; the employed who do not have health insurance are less healthy and the non-employed are the least healthy. The model somewhat over-predicts the fraction of the healthy among the non-employed, but it accounts for these qualitative patterns.

Table 14 shows the model fit for the health transition rate. The data demonstrates increasing persistence of transitions from unhealthy to unhealthy. Moreover, there is a stark difference in health transitions between college graduates and non-college graduates. The table shows that the
Table 12: Model fit: health status conditional on age and education.
Note: the number in the table is the fraction of individuals who are healthy.

<table>
<thead>
<tr>
<th>Age</th>
<th>Data</th>
<th>Model</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-31</td>
<td>0.9889</td>
<td>0.9613</td>
<td>0.9585</td>
<td>0.9278</td>
</tr>
<tr>
<td>32-38</td>
<td>0.9784</td>
<td>0.9523</td>
<td>0.9397</td>
<td>0.8968</td>
</tr>
<tr>
<td>39-45</td>
<td>0.9590</td>
<td>0.9362</td>
<td>0.9229</td>
<td>0.8692</td>
</tr>
<tr>
<td>46-52</td>
<td>0.9564</td>
<td>0.9185</td>
<td>0.8820</td>
<td>0.8857</td>
</tr>
<tr>
<td>53-59</td>
<td>0.9182</td>
<td>0.8576</td>
<td>0.8223</td>
<td>0.7822</td>
</tr>
</tbody>
</table>

Table 13: Model fit: mean health status conditional on insurance, employment and education.

<table>
<thead>
<tr>
<th></th>
<th>Emp. with HI</th>
<th>Emp. w/o HI</th>
<th>Non-employed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Data</td>
<td>Model</td>
<td>Data</td>
</tr>
<tr>
<td>College graduate</td>
<td>0.9354</td>
<td>0.9507</td>
<td>0.8796</td>
</tr>
<tr>
<td>Non-College graduate</td>
<td>0.9009</td>
<td>0.9222</td>
<td>0.8365</td>
</tr>
</tbody>
</table>

Table 14: Model fit: Annual health transition

Table 15 shows how well the model fits medical expenditure patterns conditional on health status. One interesting pattern in the data is that the variation of medical expenditure due to health status is much larger for older individuals than for younger. While the model tends to produce a steeper relationship between age and medical expenditure relative to the data, it capture the overall pattern of the data reasonably well.

Finally, Table 16 shows the model fit of firm-side moments. It demonstrates that the model fits reasonably well with the data. Specifically, it fits remarkably well for coverage rate and firm size distributions.

Overall, the model does a good job of quantitatively explaining most salient features of health, health insurance, and labor market outcomes. While the model contains various mechanisms generating these outcomes, one of the most important mechanisms is as follows. First of all, the model estimates predict that high productivity firms tend to offer health insurance. In the model, firms
want to attract more productive workers. High productivity firms are more likely to attract such workers because they can offer higher compensation. These more productive workers are typically experienced workers, who tend to be older and thus have a higher demand for health insurance. College graduates are another type of productive workers. From my estimates, college graduates are more risk-averse and have a higher demand for health insurance than less educated workers. The consequence of sorting of workers with high health insurance demand leads high-productivity firms to offer ESHI. This then leads to a positive correlation between firm size and the rate of firms offering ESHI, as high productivity firms attract more workers and become larger. Furthermore, the model also explains why the coverage rate increases over age. By moving from low to high productivity firms through job-to-job transitions over their life cycles, workers gain health insurance. Overall, this mechanism simultaneously explains positive correlations among workers’ age, wage, education status, and health insurance coverage and correlations among firm size and ESHI offering rate.\footnote{In addition to this mechanism, the model has several additional mechanisms that contribute toward generating a higher rate of firms offering ESHI by large firms. These include tax-deductibility of ESHI and an improvement of health, and therefore labor productivity, through more usage of health care (conditional on health shocks). Given my estimates, these mechanisms quantitatively play a minor role relative to the mechanism described in the main text. The quantitative comparison is available on request.}

Table 15: Model fit: mean annual medical expenditure conditional on age and education and health.
Note: The unit is $10,000.

<table>
<thead>
<tr>
<th>Variable Name</th>
<th>Data</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average establishment Size</td>
<td>20.8198</td>
<td>20.9779</td>
</tr>
<tr>
<td>... for those that offer health insurance</td>
<td>29.6444</td>
<td>31.6429</td>
</tr>
<tr>
<td>... for those that do not offer health insurance</td>
<td>7.8903</td>
<td>6.1163</td>
</tr>
<tr>
<td>The frac. of firms having less than 50 workers</td>
<td>0.9295</td>
<td>0.8858</td>
</tr>
<tr>
<td>Health insurance Coverage Rate</td>
<td>0.5943</td>
<td>0.5957</td>
</tr>
<tr>
<td>... for those having less than 50 workers</td>
<td>0.5686</td>
<td>0.5486</td>
</tr>
<tr>
<td>... for those having 50 or more workers</td>
<td>0.9338</td>
<td>0.9612</td>
</tr>
</tbody>
</table>

Table 16: Model fit for firm-side moment
6 Counterfactual Experiments: Evaluating the ACA and the components of HIX

In this section, I use the estimated model to evaluate the impact of ACA and each component of HIX and then study the optimal design of HIX. To do so, I first consider a stylized version of the ACA as a benchmark counterfactual environment, and then evaluate the role of HIX in such a context. I consider a stylized version of the ACA which incorporates its five main components: first, all individuals are required to have health insurance or pay a penalty (individual mandate); second, all firms with more than 50 workers are required to offer health insurance or pay a penalty (employer mandate); third, HIX are established where individuals can purchase health insurance at a modified community rated premium; fourth, the individuals purchasing health insurance from HIX can obtain income-based subsidies; fifth, Medicaid is expanded.

The introduction of HIX requires a substantial departure from the pre-ACA model because the premium in HIX will be endogenously determined. As a result, I will first describe how I extend and analyze the pre-ACA model to incorporate HIX.

6.1 The Model for evaluating the ACA

I provide a brief explanation of the main changes in the economic environment, as well as the definition of equilibrium, for the model used to evaluate the ACA and components of HIX.

6.1.1 The Main change in individuals’ environment

I assume that individuals who are not offered health insurance by their employers and those who are non-employed can purchase individual insurance from HIX. I also assume that the insurance product offered from HIX is the same as that offered by the employers, in that it also fully insures medical expenditure risk. Moreover, as in the pre-ACA economy, health insurance premia for ESHI are pre-paid by firms before an individual accepts a job offer from the firm. Therefore, there is no incentive for the workers not to take the offer. Thus in the extended model, an individual’s insurance status \( INS \) is defined as:

\[
INS = \begin{cases} 
0 & \text{if uninsured} \\
1 & \text{if insured through ESHI} \\
2 & \text{if insured through HIX.}
\end{cases}
\]

I also incorporate the premium subsidies to the individuals and penalties if uninsured into the model. Let \( S(y, R^{HIX}(t)) \) denote income-based subsidies to an individual with income \( y \) who purchases health insurance from HIX at premium \( R^{HIX}(t) \); note that the subsidy amount does not explicitly depend on individual age \( t \) under the specification of the ACA. Similarly, let \( IM(y) \) denote the penalty to individuals who remain uninsured, which is merely function of income under the ACA.

\[
C_t = \tau_w(w_t)P_t + (1 - P_t)b - OOP^{HI}(x, m_t) - 1(x = 0)IM(y) - 1(x = 2) \left( R^{HIX}(t) - S\left(y, R^{HIX}(t)\right) \right)
\]

(19)
where $y = w_t$ if employed ($P_t = 1$) at firms offering $(\theta, INS)$ and $y = b$ otherwise ($P_t = 0$).

Modeling individual decisions of health insurance purchases from HIX requires me to modify the timing in a period specified in Section 2.1.8. I assume that the decision is made at the end of the period: after making a working decision in the labor market, an employed individual not offered health insurance can decide whether to purchase health insurance from HIX.

Of course, the introduction of HIX into the individual decision problem makes the expression of individual value functions and steady state worker distributions rather complicated. However, the derivation itself is a straightforward extension of the pre-ACA version. Therefore, I introduce value functions and steady state worker distribution in Appendix [B].

6.1.2 The main change in firms’ environment

Because of the employer mandate, firms pay a penalty if they do not offer health insurance. This changes the determination of profits obtained by firms not offering ESHI. The flow profit $\Pi_0(p)$ is specified as

$$\Pi_0(p) = \max_{\theta_0^e} \sum_t \sum_X \pi_0(\bar{X}, t, p, \theta_0^e) l_t(\bar{X}, \theta_0^e, 0) - EM \left( \sum_t \sum_X \left( l_t(\theta_0^e, 0) + l_t(\theta_0^e, 2) \right) l_t(\bar{X}, \theta_0^e, 0) \right)$$

where $\pi_0(\bar{X}, t, p, \theta_0^e)$ is defined in the pre-ACA economy, which is expressed in (13), and $EM(l)$ is the tax penalty amount, which depends on firm size $l = \sum_t \sum_X \left( l(\theta_0^e, 0) + l(\theta_0^e, 2) \right)$. The flow profit of firms offering ESHI, $\Pi_1(p)$, as well as the decision of whether to offer ESHI, $\Delta(p)$, are determined as before, and are expressed in (10) and (14) respectively.

6.1.3 HIX

I assume that HIX is a competitive insurance market, as in Hackmann, Kolstad, and Kowalski (2013) and Handel, Hendel, and Whinston (2013)

The premium in HIX is regulated as a modified community rating. It can vary based on individual age. In equilibrium, it must satisfy

$$\sum_t \sum_X \int g_t(\bar{X}, \theta, 2) d\theta R_{HIX}(t) = (1 + \xi_{HIX}) \sum_t \sum_X \int E[m_x(\bar{X}, \theta, 2) INS_t = 2] g_t(\bar{X}, \theta, 2) d\theta$$

where the left-hand side is the total premium paid by individuals purchasing health insurance from HIX, and the right-hand side is the total expected medical expenditure by those participants multiplied by the loading factor $\xi_{HIX}$.\footnote{Ericson and Starc (2012), in their analysis of the HIX in Massachusetts (MA), instead assume that HIX is an imperfect insurance market. As I will explain, one of the main differences between the MA reform and the ACA is that the MA reform does not have medical loss ratio regulation while the ACA has. This limits insurance companies from charging a higher markup for insurance premia. While allowing imperfect competition in HIX allows us to study the optimal choice of medical loss ratio, such an assumption substantially complicates the whole analysis and requires the actual data on national samples of individuals purchasing health insurance from HIX, which is not available at present. Therefore, I leave this issue for future work.}
Note that the health insurance premium $R_{HIX}(t)$ is subject to government regulation such that it cannot depend on individual health status. The variation of premia due to age is partially limited due to the age-based rating regulation. Specifically, the regulation determines the maximum allowable premium ratio between the oldest and the youngest $\omega_{\text{AGE}}$:

$$\omega_{\text{AGE}} R_{HIX}(1) \geq R_{HIX}(T).$$

6.1.4 The definition of equilibrium

The definition of steady state equilibrium is a straightforward extension of the equilibrium under the pre-ACA economy. Specifically, it now includes premiums in HIX $\{R_{HIX}(t)\}$ as equilibrium objects. Insurance premia must satisfy a break even condition defined in (20).

6.1.5 Parameterization of policy components in the ACA

In order to evaluate the impact of ACA and the components of HIX, I need to parameterize each component. It requires me to address several issues regarding how to introduce the specifics of the ACA provisions into my model, such as the penalty associated with the individual mandate and employer mandate, the premium subsidies, age based rating regulation, and Medicaid. First, the final picture of the reform regarding Medicaid expansion is still unclear. As of the end of August 2013, 25 states plan to expand Medicaid as the ACA requires, while the others plan not to. Because the model does not describe heterogeneity in terms of states, I can only consider extreme cases: (a) all states expand Medicaid; (b) no states expand Medicaid. In this analysis, I first show the result for case (a), and then discuss how results change in case (b). Second, I need to decide on the magnitude of the loading factor $\xi_{HIX}$ that appeared in (20), which is applicable in HIX. I calibrate $\xi_{HIX}$ based on the ACA requirement that all insurance sold in the exchange must satisfy the newly-imposed regulation by the ACA that the medical loss ratio must be at least 80%. This implies that $\xi_{HIX} = 0.25$. Third, the amount of penalties and subsidies are defined as an annual level, while my model uses four-month as a model period. I simply divide all monetary units by 3 to obtain the applicable number for a four-month period. Given these assumptions, I specify the details of the components of the ACA, which is in Appendix C.
### Panel A: Effects on the Firm Side

<table>
<thead>
<tr>
<th></th>
<th>Pre-ACA (1)</th>
<th>ACA (2)</th>
<th>ACA w/o IM (3)</th>
<th>ACA w/o Subs (4)</th>
<th>ACA + more Subs (5)</th>
<th>ACA w/o AGE (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frac. of firms offering ESHI</td>
<td>0.5957</td>
<td>0.6268</td>
<td>0.5394</td>
<td>0.6768</td>
<td>0.4093</td>
<td>0.5993</td>
</tr>
<tr>
<td>...if firm size is less than 50</td>
<td>0.5486</td>
<td>0.5797</td>
<td>0.4814</td>
<td>0.6359</td>
<td>0.3382</td>
<td>0.5487</td>
</tr>
<tr>
<td>...if firm size is 50 or more</td>
<td>0.9612</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9999</td>
<td>0.9999</td>
</tr>
<tr>
<td>Average labor productivity</td>
<td>2.4883</td>
<td>2.4736</td>
<td>2.4725</td>
<td>2.4862</td>
<td>2.4664</td>
<td>2.4730</td>
</tr>
<tr>
<td>Output per capita</td>
<td>2.2483</td>
<td>2.2439</td>
<td>2.2427</td>
<td>2.2476</td>
<td>2.2430</td>
<td>2.2436</td>
</tr>
</tbody>
</table>

### Panel B: Effects on the Worker Side

<table>
<thead>
<tr>
<th></th>
<th>Pre-ACA (1)</th>
<th>ACA (2)</th>
<th>ACA w/o IM (3)</th>
<th>ACA w/o Subs (4)</th>
<th>ACA + more Subs (5)</th>
<th>ACA w/o AGE (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uninsured rate</td>
<td>0.2359</td>
<td>0.0788</td>
<td>0.1656</td>
<td>0.0934</td>
<td>0.0200</td>
<td>0.0842</td>
</tr>
<tr>
<td>The frac. of ind. with ESHI</td>
<td>0.7641</td>
<td>0.7687</td>
<td>0.7057</td>
<td>0.7992</td>
<td>0.5886</td>
<td>0.7686</td>
</tr>
<tr>
<td>Non-employment rate</td>
<td>0.0958</td>
<td>0.0918</td>
<td>0.0918</td>
<td>0.0941</td>
<td>0.0908</td>
<td>0.0921</td>
</tr>
<tr>
<td>Average wage</td>
<td>1.7859</td>
<td>1.7717</td>
<td>1.7756</td>
<td>1.7778</td>
<td>1.7802</td>
<td>1.7667</td>
</tr>
<tr>
<td>The fraction of employed in the top 50% productivity rank firms</td>
<td>0.7412</td>
<td>0.7332</td>
<td>0.7331</td>
<td>0.7402</td>
<td>0.7301</td>
<td>0.7335</td>
</tr>
<tr>
<td>Medical expenditure</td>
<td>0.0724</td>
<td>0.0847</td>
<td>0.0823</td>
<td>0.0838</td>
<td>0.0861</td>
<td>0.0844</td>
</tr>
<tr>
<td>The frac. of healthy workers</td>
<td>0.9501</td>
<td>0.9589</td>
<td>0.9572</td>
<td>0.9581</td>
<td>0.9598</td>
<td>0.9587</td>
</tr>
</tbody>
</table>

### Panel C: Effects on Government Revenue

<table>
<thead>
<tr>
<th></th>
<th>Pre-ACA (1)</th>
<th>ACA (2)</th>
<th>ACA w/o IM (3)</th>
<th>ACA w/o Subs (4)</th>
<th>ACA + more Subs (5)</th>
<th>ACA w/o AGE (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revenue from Income Tax</td>
<td>0.6853</td>
<td>0.6814</td>
<td>0.6832</td>
<td>0.6726</td>
<td>0.6860</td>
<td>0.6792</td>
</tr>
<tr>
<td>Subsidies to HIX &amp; Medicaid</td>
<td>-</td>
<td>0.0139</td>
<td>0.0150</td>
<td>0.0115</td>
<td>0.0226</td>
<td>0.0166</td>
</tr>
<tr>
<td>Revenue from penalties</td>
<td>-</td>
<td>0.0016</td>
<td>1.63e-07</td>
<td>0.0034</td>
<td>0.0003</td>
<td>0.0020</td>
</tr>
<tr>
<td>Total Revenue</td>
<td>0.6853</td>
<td>0.6690</td>
<td>0.6682</td>
<td>0.6726</td>
<td>0.6639</td>
<td>0.6655</td>
</tr>
</tbody>
</table>

Table 17: Counterfactual Policy Experiments: Evaluation of the ACA and its Variations.

Notes: (a) Column (1) reports the statistics generated under the pre-ACA economy. Note that all the statistics are calculated by including individuals less than age 25. The main pattern is unchanged even if I exclude them in this table. (b) Column (2) reports the statistics generated under the ACA. (c) Column (3) reports the statistics generated under the ACA without individual mandate. (d) Column (4) reports the statistics generated under the ACA without premium subsidies in the HIX. (e) Column (5) reports the statistics generated under the ACA with additional premium subsidies which are set that individuals who are eligible for premium subsidies in HIX obtain the full premium subsidies, subsidies which are equal to premiua. (f) Column (6) reports the statistics generated under the ACA without maximum premium ratio regulation between the oldest and youngest. out premium subsidies in the HIX.
6.2 The Main Results

6.2.1 The aggregate impact of the ACA

Table 17 shows the main results from simulating the ACA and a variety of its combinations. The results for the ACA are reported in Column (2). It reports the outcomes of several important aggregate variables determined in the model. Panel A reports the main outcomes regarding the firm side. First, I find that the ACA substantially increases firms’ coverage rate from 59.6% to 62.7%. The increase in firm’s coverage rate comes not only from firms with more than 50 workers which are subject to the penalty of employer mandate, but also from firms with less than 50 workers, which are not. Secondly, and more interestingly, I find that the ACA reduces labor productivity and output per worker. The aggregate labor productivity, defined as the output per employed worker, decreases by 0.6%, while output per capita decreases by 0.2%, which are substantial.

Panel B reports the main outcomes regarding the worker side. It shows that the uninsured rate in the estimated sample decreases from 23.6% to 7.8%, which is a substantial reduction. Half of the reduction is due to the Medicaid expansion, which covers 9.2% of non-employed workers. Moreover, 10.2% of workers, who are employed workers at firms not offering ESHI, have health insurance from HIX. The remaining increase is explained by the fact that more workers are covered by ESHI. It also shows that the non-employment rate decreases by 0.4 percentage point. Furthermore, the fraction of healthy workers increases by 1 percentage point.

Combining the results from Panel A and B, it is somewhat surprising to observe the decline of aggregate labor productivity and output. It is important to stress that the impact of the ACA on aggregate labor productivity and output are theoretically ambiguous. There are several mechanisms counteracting the decline of these variables. First, the fraction of healthy workers increases as more workers are insured and utilize more health care when they are hit by health shocks. Because the unhealthy individuals produce 30% less than the healthy individuals, this should increase aggregate labor productivity. Second, the non-employment rate decreases relative to the pre-ACA economy, which is a consequence of the increase in the fraction of healthy individuals, as the unhealthy individuals face higher disutility of working. This channel has a positive effect on the output per capita. However, the main reason for observing the decrease in aggregate labor productivity and output is that more workers are allocated to low productivity firms, reported in Panel B. The

On June 28, 2012, the U.S. Supreme Court’s ruled unconstitutional the law’s provision that, if a State does not comply with the ACA’s new coverage requirements, it will lose not only the federal funding for those requirements, but all of its federal Medicaid fund. This ruling allows states to opt out of ACA’s Medicaid expansion, leaving each state’s decision to participate in the hands of the nation’s governors and state leaders.

The medical loss ratio is the ratio of the total claim costs that the insurance company incurs to total insurance premium collected from participants.

The medical loss ratio implied by (20) is simply $1/(1 + \xi)$, thus an 80% medical loss ratio corresponds to $\xi = 0.25$. The ACA requires that $\xi \leq 0.25$.

Note that an increase in the employment rate is not a trivial result. Indeed, given the expansion of Medicaid which provides free coverage to non-employed workers, the ACA gives a strong disincentive to work, as argued by Mulligan (2013b) and Mulligan (2013a). However, I find that such a disincentive effect is smaller than the effect from improved health which leads to more labor force participation.
fraction of employed workers in the top 50% of firms in productivity ranking decreases by 1.11%.

Panel C reports the main outcome regarding the government budget. It shows that government revenues decrease by 3%. The source of the declines is an expansion of subsidized coverage in HIX and an expansion of Medicaid.

6.2.2 The heterogeneous impact of the ACA

While Table 17 is mainly about aggregate outcome, it is also useful to know whether the ACA has different effects on individuals with different characteristics.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>Pre-ACA</th>
<th>ACA</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-35</td>
<td>0.18</td>
<td>0.17</td>
</tr>
<tr>
<td>35-45</td>
<td>0.12</td>
<td>0.12</td>
</tr>
<tr>
<td>45-55</td>
<td>0.09</td>
<td>0.10</td>
</tr>
</tbody>
</table>

Table 18: Counterfactual Policy Experiments: the Impact of the ACA by Age Groups.
Notes: (a) Employed not offered ESHI is the measure of employed workers who are not offered ESHI. (b) Uninsured who are employed is the measure of individuals who are uninsured and employed.

Table 18 shows the uninsured rate by age cohorts. Interestingly, the third column in Table 18 shows that individuals older than age 45 achieve almost universal coverage. Most remaining uninsured are younger individuals. The third column in Table 18 shows that 10% of individuals between age 45 and 55 are working for firms not offering ESHI, while it is 17% for age 25-35. Therefore, this age difference is caused by the low participation to HIX by younger individuals who are working for firms not offering ESHI. This pattern is explained by the following two reasons. First, the maximum premium ratio between the youngest and oldest is binding in my counterfactual. Therefore, younger individuals need to pay higher premia relative to their expected medical expenditures. Second, the distribution of medical expenditure among the young is more skewed relative to the old in my estimates. As a result, HIX for the young pool suffers from more of an adverse selection problem. These effects lead to a lower proportion of the young participating in HIX. Another interesting finding is that the fraction of older workers in firms not offering ESHI is higher in the ACA than the fraction in the pre-ACA economy. Because firms not offering ESHI tend to be low productivity firms, a higher proportion of older workers are allocated to low productivity firms. Such a pattern is documented in the Figure 1, which shows that the fraction of older individuals in more productive firms is lower in the ACA than the fraction in the pre-ACA economy.

Table 19 shows the change in the distribution of health status among the uninsured. As shown in the table, the fraction of healthy workers who are working at firms not offering ESHI is 0.94 in
Figure 1: The fraction of employed workers in top 50% of firms (by productivity)

<table>
<thead>
<tr>
<th>Pre-ACA (1)</th>
<th>ACA (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>The frac. healthy in firms w/o ESHI</td>
<td>The frac. healthy in firms w/o ESHI</td>
</tr>
<tr>
<td>0.94</td>
<td>0.96</td>
</tr>
<tr>
<td>The frac. healthy uninsured</td>
<td>0.99</td>
</tr>
</tbody>
</table>

Table 19: Counterfactual Policy Experiments: the Impact of the ACA by Health Status.
Notes: (a) the frac. healthy in firms w/o ESHI is the fraction of healthy workers among those who are working in firms not offering ESHI. (b) the frac. healthy uninsured is the fraction of healthy workers among those who choose to stay uninsured.

The frac. healthy in firms w/o ESHI is 0.94. The ACA increases it to 0.96. Moreover, those who remain uninsured are healthy individuals. Therefore, almost all unhealthy individuals are covered by health insurance.

### 6.2.3 The impact of the components of HIX

To understand the main driving forces and mechanisms generating the impact of the ACA, I show the importance of each component of HIX in the ACA, which is summarized in Table 17.

First, Column (3) shows the simulation result under the ACA without individual mandate. In this case, the uninsured rate more than doubles. This increases in the uninsured rate is explained by the decline in both the participation rate in HIX and the rate of firms offering ESHI. This indicates the importance of the individual mandate in this economy for achieving a lower uninsured rate. The increase in the uninsured rate comes from a lower proportion of firms offering ESHI and lower take-up in HIX. Mainly due to an increase in the uninsured rate without individual mandates, the fraction of unhealthy individuals increases, contributing to the decline of labor productivity and

---

35 An additional set of the results showing each component of the ACA, in addition to the HIX, is available on request.
output. However, the magnitude of the decline is quite small.

Column (4) shows the simulation result under the ACA without premium subsidies. Relative to the case under the ACA without individual mandate, an increase in the uninsured rate from the ACA is relatively small: it is 9.34% while it is 16.56% under the ACA without individual mandate. One reason for the increase in the uninsured rate is the lower participation in HIX, which is the effect we observed in the case of the removal of the individual mandate. However, this small response is due to the fact that the fraction of firms offering ESHI in the absence of premium subsidies is 5 percentage points higher relative to the ACA with the premium subsidies. This is due to the fact that premium subsidies decrease the ESHI offering rate. Therefore, while both an individual mandate and premium subsidies increase participation in HIX, each of the components of the ACA has a qualitatively different effect on the rate of firms offering ESHI.

An interesting finding from Column (4) is that the decline in labor productivity and the output per capita in the ACA, reported in Column (2) is accounted for by the premium subsidies. Column (2) also shows that premium subsidies account for the most of the decline in the fraction of employed workers in firms ranked in the top 50% productivity ranking.

One can see a similar pattern from Column (5). Here, I consider the case where individuals whose income is less than 400% obtain full premium subsidies so that they can obtain free health insurance from HIX. Both output and labor productivity decline, while the fraction of healthy workers and employment rate increase.

The main reason why more workers are allocated to low productivity firms is that firms cannot allow individual workers to choose between ESHI and HIX as their source of coverage. Specifically, if firms offer ESHI, their workers lose the chance to buy subsidized health insurance from HIX. This will be a more important problem for older and low skilled workers, who obtain health insurance at lower costs from the HIX relative to high skilled workers. Therefore, more older and low skilled workers show their preference for staying at low productivity firms not offering ESHI by turning down offers from high productivity firms offering ESHI, unless these high productivity firms offer wages which are high enough. On the other hand, high skilled workers, who gain less from purchasing health insurance from HIX prefer to work at jobs offering ESHI. Such heterogeneous preferences can arise from not only skill heterogeneities, but also other characteristics such as age. These preference heterogeneities then lead to an inefficient sorting of workers across different firms.

This result is very different from the standard perception that HIX can lead to more efficient sorting of workers across firms. Indeed, there has been a large literature investigating how the pre-ACA, employer based health insurance system leads to misallocation of workers by hampering reallocation of workers from low to high productivity firms, which are labeled as job-lock and job-push problems. However, the existing evidence about the quantitative importance of these channels is mixed. Indeed, Dey and Flinn (2005), who estimate a search, matching, and bargaining model of wage and health insurance provisions, find that the existing employer based health insurance system leads to a negligible amount of inefficient job mobility decisions (i.e., decisions to

---

36See Currie and Madrian (1999) and Gruber (2000) for an excellent survey of this literature.
turn down job offers from more productive jobs.) My paper, rather, points out that the particular design of HIX can lead to less efficient allocation of workers across firms.

In terms of the quantitative significance of this result, it is comparable to other studies investigating the impact of institution on labor productivity. For example, Gourio and Roys (2012) find that size dependent regulations on firms in France lead to a decline in labor productivity of 0.27%. By using a different model, Garicano, LeLarge, and Van Reenen (2013) evaluate the impact of the same institution in France and find that it reduces labor productivity by 0.02% if wages are endogenously adjusted while it reduces labor productivity by 4% if wages are not adjusted.

Finally, column (6) shows the simulation result under the ACA without age based rating regulation. Here, I assume that each age group consists of different risk pools and that premia for each group is determined as a competitive price. Interestingly, the uninsured rate increases slightly, which reflects the fact that given the individual mandate and premium subsidies, perfect age based rating causes more adverse selection in each group. More interestingly, the premium subsidies to HIX substantially increase relative to the original ACA.

6.2.4 The partial equilibrium analysis: the role of equilibrium labor market interactions

To understand the role of equilibrium labor market interactions in Table 17, I conduct counterfactual experiments where the distribution of offers of compensation packages is assumed to be exogenous and the same as in the pre-ACA economy. As I find in Table 17, the rate of firm offering ESHI varies substantially depending on the design of HIX. The key question is how important is this channel to understanding the impact on the uninsured rate or other labor market outcomes. Note that this analysis is not directly comparable to Ericson and Starc (2012), Handel, Hendel, and Whinston (2013) and Hackmann, Kolstad, and Kowalski (2013). In these models, individuals can choose either to stay uninsured or to purchase health insurance from HIX. Therefore, their models do not account for the possibility that individuals change labor supply and job mobility decisions to take advantage of the benefit of purchasing health insurance from HIX. In this analysis, I still allow the individual-level responses in the labor market.

The result is reported in Table 20. In Column (2), I show the results under the case where all the components of the ACA are implemented. Compared with Column (2) in Table 17, where the offer distribution is endogenously determined, the labor productivity decline is much smaller. A similar result is obtained in Column (4) and (5), where the amount of premium subsidies is changed from the ACA.

To understand this result, it is important to recognize how the offer distribution is adjusted. As I explain in the last section, low skilled workers prefer to stay with low productivity firms if they have premium subsides from HIX. This makes low productivity firms that do not offer ESHI more attractive relative to more productive firms offering ESHI, as these low productivity firms can hire and retain workers for longer. As a result, low productivity firms are less likely to offer ESHI, which increases the likelihood that these workers work at low productivity firms. Consequently,
this channel becomes an additional mechanism that lower aggregate labor productivity.

Another interesting pattern involves the uninsured rate. In Column (2) and (3), the reported uninsured rate is much lower relative to the one in Table 17. The main difference is the rate of firms offering ESHI. In Table 17, firms’ ESHI offering rate is much higher, particularly for large firms. This causes workers to have less incentive to obtain health insurance and utilize health care when they are young, as they expect to be covered in the future.

Column (6) is the case where the premium is perfectly rated on the basis of age. Interestingly, in this case we find much more decline in output and labor productivity relative to the case where the offer distribution is endogenously adjusted. Here, young workers, who are more likely to experience job-to-job quits, face much lower premia relative to older workers in HIX. These workers have more incentive to stay at firms not offering ESHI. In the case where offer distribution of compensation packages is endogenously adjusted, this effect is partially muted because firms increase wage offer to take into account premium differences in HIX.

Overall, these results indicate that the simulated outcomes substantially differ between partial and general equilibrium settings. This indicates that an explicit modeling of labor market interactions is crucial to understand the welfare impact of HIX.

7 Normative Analysis: Optimal Design of Health Insurance Exchanges

In this section, I study whether the government can increase the welfare by altering the major design components of HIX. The previous section has shown mechanisms whereby each component of HIX leads to changes in the uninsured rate and labor market outcomes. I find that the current design of HIX system decreases the aggregate labor productivity by allocating more workers to relatively low productivity firms. This is mainly due to the premium subsidies, which contributes to lower uninsured rate in HIX. Moreover, I also find that an adverse selection problem is much more severe for young individuals under the current structure of HIX. Given these findings, the natural question is whether there is a room for improving the welfare by altering the HIX system. To answer the question, I study the optimal design problem of HIX system. First, I specify policy tools for designing HIX. Next, I define the government problem of designing HIX. Then, I show the main results obtained from the optimal design problem.

7.1 The Government Problem

7.1.1 Policy instruments for HIX

In my welfare analysis, I consider the following three policy instruments for HIX: (1) age-based pricing regulation $\omega_{AGE}$, (2) premium subsidies $S$, and (3) tax penalties on the uninsured $IM$. I choose these policy instruments because these policies are regarded as the major instruments

\[\text{Note that the specification under the ACA is explained in Appendix C.}\]
### Panel A: Effects on the Firm Side

<table>
<thead>
<tr>
<th></th>
<th>Pre-ACA (1)</th>
<th>ACA (2)</th>
<th>ACA w/o IM (3)</th>
<th>ACA w/o Subs (4)</th>
<th>ACA + more Subs (5)</th>
<th>ACA w/o AGE (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frac. of firms offering ESHI</td>
<td>0.5957</td>
<td>0.5957</td>
<td>0.5957</td>
<td>0.5957</td>
<td>0.5957</td>
<td>0.5957</td>
</tr>
<tr>
<td>...if firm size is less than 50</td>
<td>0.5486</td>
<td>0.5486</td>
<td>0.5486</td>
<td>0.5486</td>
<td>0.5486</td>
<td>0.5486</td>
</tr>
<tr>
<td>...if firm size is 50 or more</td>
<td>0.9612</td>
<td>0.9612</td>
<td>0.9612</td>
<td>0.9612</td>
<td>0.9612</td>
<td>0.9612</td>
</tr>
<tr>
<td>Average labor productivity</td>
<td>2.4883</td>
<td>2.4777</td>
<td>2.4771</td>
<td>2.4820</td>
<td>2.4742</td>
<td>2.4742</td>
</tr>
<tr>
<td>Output per capita</td>
<td>2.2483</td>
<td>2.2461</td>
<td>2.2474</td>
<td>2.2477</td>
<td>2.2452</td>
<td>2.2448</td>
</tr>
</tbody>
</table>

### Panel B: Effects on the Worker Side

<table>
<thead>
<tr>
<th></th>
<th>Pre-ACA (1)</th>
<th>ACA (2)</th>
<th>ACA w/o IM (3)</th>
<th>ACA w/o Subs (4)</th>
<th>ACA + more Subs (5)</th>
<th>ACA w/o AGE (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uninsured rate</td>
<td>0.2359</td>
<td>0.0665</td>
<td>0.1114</td>
<td>0.0972</td>
<td>0.0200</td>
<td>0.0977</td>
</tr>
<tr>
<td>The frac. of ind. with ESHI</td>
<td>0.7641</td>
<td>0.7794</td>
<td>0.7748</td>
<td>0.7834</td>
<td>0.5886</td>
<td>0.7776</td>
</tr>
<tr>
<td>Non-employment rate</td>
<td>0.0958</td>
<td>0.0924</td>
<td>0.0925</td>
<td>0.0942</td>
<td>0.0908</td>
<td>0.0921</td>
</tr>
<tr>
<td>Average wage</td>
<td>1.7859</td>
<td>1.7817</td>
<td>1.7821</td>
<td>1.7854</td>
<td>1.7798</td>
<td>1.7794</td>
</tr>
<tr>
<td>Medical expenditure</td>
<td>0.0724</td>
<td>0.0850</td>
<td>0.0838</td>
<td>0.0836</td>
<td>0.0864</td>
<td>0.0839</td>
</tr>
<tr>
<td>The frac. of the healthy workers</td>
<td>0.9501</td>
<td>0.9591</td>
<td>0.9581</td>
<td>0.9579</td>
<td>0.9599</td>
<td>0.9583</td>
</tr>
</tbody>
</table>

### Panel C: Effects on Government Revenue:

<table>
<thead>
<tr>
<th></th>
<th>Pre-ACA (1)</th>
<th>ACA (2)</th>
<th>ACA w/o IM (3)</th>
<th>ACA w/o Subs (4)</th>
<th>ACA + more Subs (5)</th>
<th>ACA w/o AGE (6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Revenue from Income Tax</td>
<td>0.6853</td>
<td>0.6854</td>
<td>0.6856</td>
<td>0.6857</td>
<td>0.6851</td>
<td>0.6848</td>
</tr>
<tr>
<td>Subsidies to HIX &amp; Medicaid</td>
<td>-</td>
<td>0.0138</td>
<td>0.0132</td>
<td>0.0117</td>
<td>0.0166</td>
<td>0.0162</td>
</tr>
<tr>
<td>Revenue from penalties</td>
<td>-</td>
<td>0.0013</td>
<td>1.63e-05</td>
<td>0.0019</td>
<td>1.57e-04</td>
<td>0.0019</td>
</tr>
<tr>
<td>Total Revenue</td>
<td>0.6853</td>
<td>0.6729</td>
<td>0.6724</td>
<td>0.6757</td>
<td>0.6687</td>
<td>0.6686</td>
</tr>
</tbody>
</table>

Table 20: Counterfactual Policy Experiments: Partial Equilibrium.

Notes: (a) Results between Column (1)-(6) are obtained under the case where the offer distribution of compensation package is exogenous and fixed as the distribution under the pre-ACA economy. (b) Column (1) reports the statistics generated under the pre-ACA economy. Note that all the statistics are calculated by including individuals less than age 25. The main pattern is unchanged even if I exclude them in this table. (c) Column (2) reports the statistics generated under the ACA. (d) Column (3) reports the statistics generated under the ACA without individual mandate. (e) Column (4) reports the statistics generated under the ACA without premium subsidies in the HIX. (f) Column (5) reports the statistics generated under the ACA with additional premium subsidies which are set that individuals who are eligible for premium subsidies in HIX obtain the full premium subsidies, subsidies which are equal to premiums. (g) Column (6) reports the statistics generated under the ACA without maximum premium ratio regulation between the oldest and youngest. out premium subsidies in the HIX.
affecting the uninsured rate in HIX. Moreover, in public economics, these policy instruments are also regarded as the major instruments to correct welfare loss from adverse selection in insurance markets.\textsuperscript{38} The key novelty of my analysis is to take into account equilibrium labor market interactions. To define the government problem, I specify their policy constraints in the optimal design problem. I proceed with analysis in the following three cases:

Case 1. $S$ and $IM$ follow the same functional form as the ACA. For example, tax penalties are specified as $(\omega_0^{im}, \omega_1^{im})$ such that $IM(y) = \max \{\omega_0^{im} \times y, \omega_1^{im}\}$. In the ACA, $\omega_0^{im} = 0.025$ and $\omega_1^{im} = $695. I parameterize $S$ in the same way:

$$S(y, R_{HIX}(t)) = \begin{cases} \max \left\{R_{HIX}(t) - \left[\omega_0^{s} + \omega_1^{s} \frac{3y}{\omega_2^{s}}\right], 0\right\} & \text{if } y < \frac{\omega_2^{s}}{3} \\ 0, & \text{otherwise,} \end{cases}$$

In the ACA, $\omega_0^{s} = 0.035$, $\omega_1^{s} = 0.060$, and $\omega_2^{s}$ is the income at 400% federal poverty level $FPL_{400\%}$.

Case 2. $S$ and $IM$ take a more flexible functional form. Specifically, they are specified so that

$$S(y, R_{HIX}(t)) = \frac{\exp(\omega_0^{a} + \omega_1^{b} y_t + \omega_2^{c} y_t^2)}{1 + \exp(\omega_0^{a} + \omega_1^{b} y_t + \omega_2^{c} y_t^2)} R_{HIX}(t)$$

$$IM(y, R_{HIX}(t)) = \frac{\exp(\omega_0^{im} + \omega_1^{im} y_t + \omega_2^{im} y_t^2)}{1 + \exp(\omega_0^{im} + \omega_1^{im} y_t + \omega_2^{im} y_t^2)} R_{HIX}(t)$$

Case 3. $S$ and $IM$ are explicitly age-dependent:

$$S(y, t, R_{HIX}(t)) = \frac{\exp(\omega_0^{a} + \omega_1^{a} y_t + \omega_2^{a} y_t^2 + \omega_3^{a} t + \omega_4^{a} t^2)}{1 + \exp(\omega_0^{a} + \omega_1^{a} y_t + \omega_2^{a} y_t^2 + \omega_3^{a} t + \omega_4^{a} t^2)} R_{HIX}(t)$$

$$IM(y, t, R_{HIX}(t)) = \frac{\exp(\omega_0^{im} + \omega_1^{im} y_t + \omega_2^{im} y_t^2 + \omega_3^{im} t + \omega_4^{im} t^2)}{1 + \exp(\omega_0^{im} + \omega_1^{im} y_t + \omega_2^{im} y_t^2 + \omega_3^{im} t + \omega_4^{im} t^2)} R_{HIX}(t)$$

I study the possibility that the government can condition premium subsidies and tax penalties on the uninsured by individual age for the following reasons. First, in HIX, insurance premia are, to some extent, pooled across individuals with different ages due to the age based pricing regulation. Second, I find in Section 6 that there is a sharp difference in terms of the impact of the ACA across different age groups. Third, it is known from recent optimal taxation and social insurance literature that age dependent policies are welfare improving (see, for example, Akerlof (1978); Michelacci and Ruffo (2011); Weinzierl (2011); Farhi and Werning (2013)). Motivated by these reasons, I examine the welfare impact of age dependent premium subsidies and tax penalties, in addition to age based rating regulation.

7.1.2 Optimal design of HIX

The government chooses a combination of three policy instruments for HIX, \( T_{HIX} = \{\omega_{AGE}, S, IM\} \), to maximize the social welfare subject to the revenue constraint. There are several ways to specify the social welfare function. By following the standard approach used in the social insurance and optimal taxation literature, I assume that the government is utilitarian. Moreover, I define the social welfare function as the ex-ante lifetime utility of newborn individuals:

\[
SW(T_{HIX}) = \sum_{\tilde{X}_0} V_0(\tilde{X}_0) g_0(\tilde{X}_0, 0, 0)
\]

where \( V_0(\tilde{X}) \) is lifetime utility of newborn individuals. The revenue constraint is given by

\[
RV_{tax}(T_{HIX}) + RV_p(T_{HIX}) - EXP_{sub}(T_{HIX}) \geq R.
\]

The first term in the left hand side is the revenue from the income tax in the equilibrium under the policy parameters \( T_{HIX} \), which is given as

\[
RV_{tax}(T_{HIX}) = \sum_t \sum_{\tilde{X}} \sum_{INS} \int T_l(\tilde{X}, \theta, INS) g_t(\tilde{X}, \theta, INS) d\theta,
\]

and \( RV_p(T_{HIX}) \) is the revenue from tax penalties imposed on the uninsured and on large firms not offering ESHI, given as

\[
RV_p(T_{HIX}) = \sum_t \sum_{\tilde{X}} \int IM \left( y_t, R^{HIX}(t), t \right) g_t(\tilde{X}, \theta, 0) d\theta
+ \sum_t \sum_{\tilde{X}} IM(\tilde{X}, \theta, 0) u_t(\tilde{X}, 0)
+ \int_p EM \left( \sum_t \sum_{\tilde{X}} l_t(\tilde{X}, \theta_0^d, 0) + l_t(\tilde{X}, \theta_0^d, 2) \right) (1 - \Delta(p)) d\Gamma(p)
\]

where the first line is the tax penalty on the uninsured who are employed, the second line is the penalty on the uninsured among non-employed, and the third term is the penalty on the large firms not offering ESHI. \( EXP_{sub}(T_{HIX}) \) is government subsidies for health insurance, which consist of the expenditure on the premium subsidies to HIX and Medicaid:

\[
EXP_{sub}(T_{HIX}) = \sum_t \sum_{\tilde{X}} \int S_y \left( y_t, R^{HIX}(t), t, \tilde{X} \right) g_t(\tilde{X}, \theta, 2) d\theta
- \sum_t \sum_{\tilde{X}} S_u \left( R^{HIX}(t), t, \tilde{X} \right) u_t(\tilde{X}, 1)
\]

Finally, I specify that \( R \) is the government revenue obtained at the ACA, \( T_{HIX} = T_{HIX}^{ACA} \). Therefore, the solution of this government problem tells us the maximum welfare gain achieved given the policy constraints relative to the ACA. In order to solve the planner’s problem numerically, I use KNITRO,
which is a solver for nonlinear optimization allowing nonlinear inequality constraints. Moreover, I fix revenue constraint $R$ as the revenue obtained if $T_{HIX}$ takes the parameter values adopted by the ACA.

### 7.2 Main Results

#### 7.2.1 The general equilibrium analysis (with endogenous offer distribution of compensation)

First, I investigate the optimal structure of HIX in a general equilibrium setting, i.e., the setting where the offer distribution of compensation package is endogenously determined by firm’s optimization problem. I begin by reporting the optimal policy parameters for Case 1, which is the case where the government’s policy set is restricted within the functional form implemented under the ACA.

<table>
<thead>
<tr>
<th>Policy Parameter</th>
<th>ACA</th>
<th>optimal (Case 1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: policy for age based rating</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the maximum premium ratio between the youngest and the oldest ($\omega_{AGE}$)</td>
<td>3.00</td>
<td>4.044</td>
</tr>
<tr>
<td><strong>Panel B: premium subsidies to HIX</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>constant term ($\omega_0^s$)</td>
<td>0.035</td>
<td>0.005</td>
</tr>
<tr>
<td>the coeff. for income ($\omega_1^s$)</td>
<td>0.060</td>
<td>0.052</td>
</tr>
<tr>
<td>the maximum 4-month income eligible for subsidies</td>
<td>400% of FPL</td>
<td>242% of FPL</td>
</tr>
<tr>
<td><strong>Panel C: individual mandates (tax penalty to the uninsured)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>the minimum amount of annual tax penalty ($\omega_0^{im}$)</td>
<td>$695</td>
<td>$40.7</td>
</tr>
<tr>
<td>the percentage of tax penalty as a function of income ($\omega_1^{im}$)</td>
<td>2.5%</td>
<td>2.99%</td>
</tr>
</tbody>
</table>

Table 21: Optimal Policy Parameters under Case 1

Table 21 shows the optimal structure of HIX under Case 1 and its comparison from the ACA. The major differences are as follows: (1) the maximum premium ratio between the oldest and the youngest $\omega_{AGE}$ is larger; (2) the premium subsidies becomes more progressive: available up to 242% of FPL, more generous subsidies to low income; (3) individual mandates are set max \{0.03 \times y, $29.5\}$, which indicates a larger tax penalty for high income individuals and a lower penalty to low income individuals. However, the welfare gain is modest. I measure the welfare gain by the annual lump-sum transfer to individuals to have the same utility under the ACA environment as the one in the optimal. For Case 1, it is merely $45.

Next, I examine the optimal design under Case 2, which allows more nonlinearity in terms of the choice of premium subsidies and tax penalties. The result is shown in Table 22. I find that

---

39 KNITRO is a derivative based optimization toolbox, and thus requires smoothness of the objective function. To guarantee the smoothness, I add preference shock for insurance purchasing decisions for HIX. I specify that it follows a Type-I extreme value distribution, where the scale parameter is set very small: 0.005.
Optimal Policy Parameter

Panel A: policy for age based rating

The maximum premium ratio between the youngest and the oldest ($\omega_{AGE}$) 4.70

Panel B: premium subsidies to HIX

 Constant term ($\omega_{a}$) 0.05
 The coeff. for income ($\omega_{b}$) -0.50
 The coeff. for income squared ($\omega_{c}$) 0.002

Panel C: individual mandates (tax penalty to the uninsured)

 Constant term ($\omega_{im}^{a}$) -3.63
 The coeff. for income ($\omega_{im}^{b}$) 0.16
 The coeff. for income squared ($\omega_{im}^{c}$) 0.01

Table 22: Optimal Policy Parameters under Case 2

$\omega_{AGE}$ is even larger and the premium subsidies are still progressive. To calculate the welfare gain, I use the same measure as used in Case 1. The welfare gain is equivalent to an annual lump-sum transfer to each individual of $111. Overall, I find that progressive subsidies and more transfers from the old to the young are welfare improving.

Now, I consider Case 3, which allows age dependent premium subsidies and individual mandates. The optimal policy parameter is reported in Table 23. The general feature of the optimal policy relative to the ACA scheme is: (a) it allows a larger maximum premium ratio between the youngest and the oldest; (b) the ratio of premium subsidies to the premium is substantially decreasing in income and age; (c) the ratio of tax penalties over premium is very slightly increasing in income and age. Specifically, an important pattern is that premium subsidies are decreasing in age, while they are set independently in the ACA. The shape of optimal subsidies is summarized in Figure 2 and 3. As is clear from Figure 2, individuals with income at 300% of the Federal Poverty Level (FPL 300%) have very small subsidies over age, while those at FPL 133% receive large premium subsidies when they are young, decreasing gradually over time. Moreover, from Figure 3, one can see that premium subsidies become almost zero around a 4-month income of $15,000, which is close to FPL 400% regardless of age. This pattern is similar to the ACA.
Optimal Policy Parameter

Panel A: policy for age based rating
The maximum premium ratio between the youngest and the oldest ($\omega_{AGE}$) 3.74

Panel B: premium subsidies to HIX
- Constant term ($\omega_a^s$) 3.95
- The coeff. for income ($\omega_i^s$) -0.46
- The coeff. for income squared ($\omega_c^s$) 0.06
- The coeff. for age ($t = \ldots, 132$) ($\omega_d^s$) -0.02
- The coeff. for age squared ($\omega_e^s$) -0.02

Panel C: individual mandates (tax penalty to the uninsured)
- Constant term ($\omega_a^{im}$) -3.60
- The coeff. for income ($\omega_i^{im}$) 0.005
- The coeff. for income squared ($\omega_c^{im}$) 0.000
- The coeff. for age ($t = \ldots, 132$) ($\omega_d^{im}$) 0.003
- The coeff. for age squared ($\omega_e^{im}$) 0.000

Table 23: Optimal Policy Parameters under Case 3: Allowance of Age-Dependent Subsidies and Individual Mandate

Figure 2: Optimal Subsidies Rate over Ages
Column (2) in Table 24 reports the outcome under optimal HIX. First, it shows the substantial increase in output and labor productivity: output per worker increases by 0.1% and aggregate labor productivity increases by 0.6%. The increase in output is achieved despite the fact that the nonemployment rate increases. The uninsured rate is 7.53% which is somewhat lower relative to the ACA, while the magnitude of the changes is quite small compared with the pre-ACA economy where the uninsured rate is 23.6%.

In order to measure the welfare gain relative to the ACA, I assume that the government provides a lump-sum transfer to all individuals in the ACA economy. I find that the ACA will achieve the same level of utility as the optimal HIX if the government provide the transfer which is $195 annually. The transfer corresponds to 7.6% of medical expenditure. This eventually contributes to an increase in government expenditure of 1.1%, which is substantial.

Columns (3) to (5) report what happens when each component of optimal HIX is replaced by the component of HIX implemented under the ACA. Column (3) reports the major outcomes if the premium subsidies scheme is replaced by the scheme implemented under the ACA. I find that the uninsured rate is much higher while labor productivity is lower. The rise in the uninsured rate is mainly explained by the fact that more young and healthy individuals are uninsured. Column (3) reports the major outcomes if the individual mandate scheme is replaced by the scheme implemented under the ACA. I find that the uninsured rate is much smaller, as it induces more participation into HIX. While labor productivity is still higher than the ACA, it is lower relative to the optimal HIX reported in Column (2). Finally, Column (5) reports the major outcomes if the age based rating regulation is replaced by the regulation under the ACA. I find that the uninsured rate is higher while most labor market outcomes are similar to those reported in Column (2). These results demonstrate that all of the components contribute to the changes in outcomes, but redesigning premium subsidies has a crucial effect on labor productivity.
To understand the mechanism generating the welfare gain, it is important to recognize that under the optimal structure of HIX, relatively old workers gain less benefit than young workers from purchasing health insurance from HIX, because the optimal structure sets higher maximum premium ratio (MPR) and also allows premium subsidies to decrease with age. By giving more subsidies to the young, the optimal structure can ameliorate the adverse selection problems explained in Section 6.2. Because young workers are given more premium subsidies, they have more incentive to participate in HIX. Moreover, because the gain from participating in HIX is smaller for relatively old workers, this structure gives old workers more incentives to work at firms offering ESHI. Because high productivity firms are more likely to offer ESHI, this optimal structure can give more incentives to workers to move from low to high productivity firms over the life cycle. This intuition is confirmed by Figure 4, which shows that older individuals are allocated more on more productive firms.

While these economic forces give the benefit of redistribution from old to young workers, its cost is to make old workers worse off, in particular for those that lose their job offering ESHI due to exogenous job destruction shocks and start working at firms not offering ESHI. Indeed, Table 25 shows that the measure of the uninsured among the age group 45-55 increases due to the decrease in premium subsidies. The optimal structure of HIX is therefore determined to take into account such a cost as well as the benefit described above.

Finally, an interesting feature of the optimal structure is while premium subsidies are substantially decreasing in age, tax penalties are not. This difference reflects the differential effect of subsidies and tax penalties, which arise in an economy where both HIX and ESHI are modeled. Both larger subsidies and higher tax penalties give individuals more incentive to obtain health insurance. However, while larger subsidies give individuals incentive to purchase heath insurance
<table>
<thead>
<tr>
<th></th>
<th>ACA (1)</th>
<th>Optimal HIX (2)</th>
<th>Optimal w. ACA subs (3)</th>
<th>Optimal w. ACA IM (4)</th>
<th>Optimal w. ACA AGE (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Effects on the Firm Side</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frac. of firms offering ESHI</td>
<td>0.6268</td>
<td>0.6073</td>
<td>0.5672</td>
<td>0.6201</td>
<td>0.6079</td>
</tr>
<tr>
<td>Average labor productivity</td>
<td>2.4777</td>
<td>2.4875</td>
<td>2.4701</td>
<td>2.4814</td>
<td>2.4870</td>
</tr>
<tr>
<td>Output per capita</td>
<td>2.2461</td>
<td>2.2472</td>
<td>2.2464</td>
<td>2.2492</td>
<td>2.2471</td>
</tr>
<tr>
<td>Panel B: Effects on the Worker Side</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uninsured rate</td>
<td>0.0788</td>
<td>0.0753</td>
<td>0.1159</td>
<td>0.0357</td>
<td>0.0907</td>
</tr>
<tr>
<td>Non-employment rate</td>
<td>0.0918</td>
<td>0.0941</td>
<td>0.0910</td>
<td>0.0925</td>
<td>0.0935</td>
</tr>
<tr>
<td>Average wage</td>
<td>1.7717</td>
<td>1.7840</td>
<td>1.7743</td>
<td>1.7799</td>
<td>1.7836</td>
</tr>
</tbody>
</table>


Notes: (a) Column (1) reports the main aggregate outcomes under the ACA. (b) Column (2) reports the main aggregate outcomes under the optimal structure of the HIX. (c) Column (3) reports the main aggregate outcomes under the optimal structure of the HIX but the premium subsidies are replaced by the premium subsidies which are implemented under the ACA. (d) Column (4) reports the main aggregate outcomes under the optimal structure of the HIX but individual mandates (tax penalties on the uninsured) are replaced by the individual mandates which are implemented under the ACA. (e) Column (5) reports the main aggregate outcomes under the optimal structure of the HIX but an age based rating regulation (maximum premium ratio between the oldest and youngest) is replaced by the regulation which is implemented under the ACA.
Table 25: Comparison of the uninsured rate by different age groups under the ACA and under the optimal structure of HIX in Case 3.

<table>
<thead>
<tr>
<th>Age Group</th>
<th>ACA (1)</th>
<th>Optimal HIX (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>25-35</td>
<td>0.12</td>
<td>0.07</td>
</tr>
<tr>
<td>35-45</td>
<td>0.07</td>
<td>0.07</td>
</tr>
<tr>
<td>45-55</td>
<td>0.004</td>
<td>0.04</td>
</tr>
</tbody>
</table>

[Note: Uninsured who are employed is the measure of individuals who are uninsured and employed.

from HIX rather than to obtain health insurance from employers, higher penalties do not directly give such incentive to individuals. Therefore, in order to give incentive the old individuals to obtain health insurance from employers, premium subsidies are set to decrease in age. This differential impact of the premium subsidies and tax penalties has not pointed out in the existing works that study HIX designs, as they do not consider ESHI. Therefore, this finding indicates the importance of modeling ESHI and the labor market to understand the optimal design of HIX.

7.2.2 The partial equilibrium analysis

Next, I investigate how the optimal structure differs if one ignores general equilibrium effects of labor markets. To this end, I assume that the offer distribution of compensation package is fixed and the same as the one in the pre-ACA economy. Then, I solve the worker’s problem taking into account the equilibrium determination of health insurance premia in HIX.

The optimal policy parameters under Case 3 are summarized in Table 26. The key features of the optimal policies relative to the optimal HIX under the general equilibrium case are as follows. First, average premium subsidies are, in general, larger, particularly for low income individuals. This reflects the fact that the offer distribution is not adjusted, and therefore does not contribute a reduction of labor productivity. Second, the tax penalty on the uninsured is higher, particularly for high income individuals. Because tax penalties do not increase the ESHI offering rate, higher penalties are needed to improve the pool of HIX. Finally, the maximum premium ratio (MPR) between the oldest and the youngest is slightly higher than the MPR under the general equilibrium case.

To understand how the implications differ if we use the optimal policies obtained under the partial equilibrium setting, I report the comparison of outcomes simulated under the general equilibrium setting between optimal HIX policies under general equilibrium (in Column (2)) and optimal HIX policies under partial equilibrium (in Column (3)) in Table 27. First, by providing more subsidies, aggregate labor productivity is lower under optimal HIX policies under partial equilibrium (in Column (3)) than the productivity under general equilibrium (in Column (2)). Moreover, due to a larger expenditure on premium subsidies, the fraction of firms offering ESHI decreases and the total government expenditure is slightly higher under the partial equilibrium. These results show the importance of taking into account equilibrium effects when designing the optimal HIX.
### Optimal Policy Parameters

#### Panel A: policy for age based rating

<table>
<thead>
<tr>
<th>Parameter Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>The maximum premium ratio between the youngest and the oldest ($\omega_{\text{AGE}}$)</td>
<td>3.91</td>
</tr>
</tbody>
</table>

#### Panel B: premium subsidies to HIX

<table>
<thead>
<tr>
<th>Parameter Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant term ($\omega_a$)</td>
<td>4.63</td>
</tr>
<tr>
<td>The coeff. for income ($\omega_b$)</td>
<td>-0.31</td>
</tr>
<tr>
<td>The coeff. for income squared ($\omega_c$)</td>
<td>0.22</td>
</tr>
<tr>
<td>The coeff. for age ($t = \ldots, 132$) ($\omega_d$)</td>
<td>-0.01</td>
</tr>
<tr>
<td>The coeff. for age squared ($\omega_e$)</td>
<td>-0.001</td>
</tr>
</tbody>
</table>

#### Panel C: individual mandates (tax penalty to the uninsured)

<table>
<thead>
<tr>
<th>Parameter Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant term ($\omega_{\text{im}}$)</td>
<td>-3.62</td>
</tr>
<tr>
<td>The coeff. for income ($\omega_{\text{im}}$)</td>
<td>0.16</td>
</tr>
<tr>
<td>The coeff. for income squared ($\omega_{\text{im}}$)</td>
<td>0.01</td>
</tr>
<tr>
<td>The coeff. for age ($t = \ldots, 132$) ($\omega_{\text{im}}$)</td>
<td>0.00</td>
</tr>
<tr>
<td>The coeff. for age squared ($\omega_{\text{im}}$)</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Table 26: Optimal Policy Parameters under Case 3 and Exogenous Offer Distribution

<table>
<thead>
<tr>
<th></th>
<th>Optimal HIX: GE</th>
<th>Optimal HIX: partial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frac. of firms offering ESHI</td>
<td>0.6073</td>
<td>0.5646</td>
</tr>
<tr>
<td>Average labor productivity</td>
<td>2.4875</td>
<td>2.4775</td>
</tr>
<tr>
<td>Uninsured rate</td>
<td>0.0753</td>
<td>0.0853</td>
</tr>
<tr>
<td>Non-employment rate</td>
<td>0.0941</td>
<td>0.0938</td>
</tr>
</tbody>
</table>

Table 27: Counterfactual Policy Experiments: Comparison: Outcomes in general equilibrium under (1) optimal policies obtained with endogenous offer distribution of compensation packages and (2) optimal policies obtained with exogenous offer distribution of compensation packages.

### 7.3 Discussion: the role of Medicaid expansion

So far, all the exercises are conducted under the assumption that Medicaid is expanded as projected. However, as of August 2013, 25 states do not plan to expand Medicaid. In this section, I investigate the interaction between Medicaid expansion and HIX. Medicaid expansion affects HIX as it affects the risk pools in HIX. Specifically, if Medicaid is not expanded, individuals who would have been eligible for Medicaid will be absorbed by HIX. Because these individuals are less healthy relative to the average population in the economy, their absorption into HIX may increase health insurance premium in HIX, worsening the adverse selection problem.\(^{40}\) Therefore, the question is how much effect Medicaid expansion will have on risk pooling in HIX. At this stage, it is not clear how state governments that decide not expand Medicaid will choose premium subsidies for those who are eligible for Medicaid but ineligible for federal premium subsidies. However, for simplicity, I assume that individuals who were eligible for Medicaid obtain full premium subsidies from HIX.

\(^{40}\)The link between Medicaid and the risk pools in individual insurance markets is investigated by Clemens (2013).
That is, those individuals will be guaranteed to be insured through HIX. I call this situation "no-Medicaid expansion" in the remainder of the paper. Therefore, the impact of no-Medicaid expansion is realized purely through its impact on risk pools in HIX.

First, I evaluate the impact of the ACA with no-Medicaid expansion. I find that the uninsured rate is 8.3%, which is higher than the rate under the ACA with Medicaid expansion, which was 7.8%. The main reason is that the risk pool in HIX is worse due to the entry of pre-Medicaid pools who are less healthy. This negative externality is large enough that more employed workers choose to be uninsured. Next, I examine the optimal structure of HIX, given the ACA budget constraint. To make the comparison transparent, I assume that the pre-Medicaid eligible population still obtain health insurance for free from HIX. I find that the optimal structure of HIX is qualitatively similar to the case with Medicaid expansion, including the age dependence of premium subsidies. More interestingly, I find that the welfare gain from the optimal design of HIX is rather small: it amounts to $80, much smaller than the welfare gain with Medicaid expansion, which amounts to $195. These results highlight the importance of interactions between Medicaid and HIX.

8 Conclusion

In this paper, I evaluate the current HIX system and investigate its optimal design, accounting for adverse selection and equilibrium labor market interactions. I first develop and estimate a life cycle equilibrium labor market search model integrated with the pre-ACA health insurance market. Various forms of individual heterogeneity are incorporated to understand the welfare consequences of HIX. Through counterfactual experiments, I find that the ACA substantially reduces the uninsured rate. However, the ACA also decreases aggregate labor productivity by allocating more workers to less productive firms.

Next, I examine the optimal design of HIX by choosing the values of three major design components – individual mandates, premium subsidies and age-based rating regulations. I found that the optimal combination of these components makes it less beneficial for older workers to purchase health insurance from HIX, relative to young workers. Implementing the optimal structure leads to a substantial welfare gain relative to HIX implemented under the ACA, while achieving higher labor productivity and a slightly lower uninsured rate. In this structure, the adverse selection problem in HIX among the young is reduced. Moreover, it gives older workers an incentive to work at firms offering ESHI. This increases the allocation of workers from low to high productivity firms, raising aggregate labor productivity.

Finally, I assess the role of equilibrium labor market interactions by assuming that the distribution of offers of compensation package is fixed and evaluating the impact of each component of HIX as well as the optimal design of HIX. I find that both the impact of HIX components and their optimal design are qualitatively and quantitatively very different from the equilibrium with endogenous compensation packages, indicating the importance of modeling equilibrium labor market interactions to evaluate HIX design.

There are a number of dimensions in which my stylized model could be extended to capture
other important features. One of the most important extensions is to allow multiple insurance products offered in HIX to understand the importance of adverse selection across insurance plans within HIX. Such an extension allows us to study the optimal regulation of insurance contracts, such as the choice of minimum creditable coverage. However, there are a number of difficulties. First, it requires careful choices about the notion of equilibrium, as certain types of equilibria (e.g., Nash equilibrium) may not exist in such environment. Handel, Hendel, and Whinston (2013) make an important contribution in this regard. Second, an additional important challenge is to model the situation where firms offer a menu of ESHI contracts to workers, and to obtain data about the set of insurance plans that workers face and firms offer. At this stage, it is very difficult to obtain such data. Third, it also requires a more sophisticated modeling of individual health care utilization decisions. While each of these issues is very challenging and requires a substantial departure from this framework, these extensions are exciting opportunities for future work.
References


Appendix

A Numerical Algorithm

I describe the numerical algorithm which is used to solve the equilibrium of pre-ACA model described in section 2. I discretize the support of $\Gamma, [p, \bar{p}]$, into finite points. Then, I solve the equilibrium by the following fixed-point algorithm.

1. First, I provide an initial guess of the skill price and the fraction of firms offering ESHI $(\langle \theta_{0,0}^e(p), \theta_{1,0}^e(p) \rangle^e, \Delta_0(p))$ for all $p$ on support $[p, \bar{p}]$.

2. At iteration $\iota = 0, 1, \ldots$, I do the following sequentially, where I index the objects in iteration $\iota$ by superscript $\iota$:

   (a) Given the current guess of the health insurance costs and the health insurance offer probability $(\langle \theta_{0,0}^e(p), \theta_{1,0}^e(p) \rangle^e, \Delta_\iota(p))$, I construct an offer distribution of compensation package $F^e(\theta, INS)$.

   (b) Then, I numerically solve individual value functions backwards from the period $T - 1$. The main obstacle of solving individual life cycle problems is the large size of state spaces. There, to speed up the computation, I apply Keane and Wolpin (1994)'s interpolation method.

   (c) Given the value function, I solve the steady state distribution $\hat{\theta}_\iota \left( \tilde{X}, INS, \theta \right)$ and $u_\iota(\tilde{X})$ sequentially from age $t = 1$.

   (d) Using $\hat{\theta}_\iota \left( \tilde{X}, INS, \theta \right)$ and $u_\iota(\tilde{X})$, I solve $(\langle \hat{\theta}_{0}^e(p), \hat{\theta}_{1}^e(p) \rangle^e, \hat{\Delta}(p))$ for each $p$ using [17], [18], and [14].

3. After completing the step (d) at iteration $\iota$, I check whether the equilibrium object converges.

   (a) If $(\langle \theta_{0,\iota}^e(p), \theta_{1,\iota}^e(p) \rangle^e, \Delta(p))$ satisfies $d(\theta_{0,\iota}^e(p), \hat{\theta}_{0}^e(p)) < \epsilon_{tol}$, $d(\theta_{1,\iota}^e(p), \hat{\theta}_{1}^e(p)) < \epsilon_{tol}$ for $ed \in \{NC, C\}$ and $d(\Delta^e(p), \Delta_\iota(p)) < \epsilon_{tol}$ where $\epsilon_{tol}$ is a pre-specified tolerance level of convergence and $d(\cdot, \cdot)$ is a distance metric, then firm’s optimal policy converges and we have an equilibrium.

   (b) Otherwise, update $(\langle \theta_{0,\iota+1}^e(p), \theta_{1,\iota+1}^e(p) \rangle^e, \Delta_{\iota+1}(p))$ as follows:

$$
\theta_{0,\iota+1}^e(p) = w \theta_{0,\iota}^e(p) + (1 - w) \hat{\theta}_{0}^e(p),
\theta_{1,\iota+1}^e(p) = w \theta_{1,\iota}^e(p) + (1 - w) \hat{\theta}_{1}^e(p),
\Delta_{\iota+1}(p) = w \Delta_\iota(p) + (1 - w) \hat{\Delta}(p),
$$

for the pre-specified weight $w \in (0, 1)$ and continue Step 2 at iteration $\iota' = \iota + 1$.
B Omitted Derivations in the Counterfactual Experiments

B.1 Value Function

Consider a non-employed worker having characteristics $\tilde{X}_t = (ed, type, E_t, h_t)$ and insurance status $INS \in \{0, 2\}$. Then, his value function is defined by

$$V_0^t(\tilde{X}_t, INS) = \mathbb{E}_{e_t}^{m} \left[ \max_{x_t} U_t(C_t(x_t, \epsilon_t^m, \tilde{X}_t, INS), 0, h_t) + \beta \sum_{h} \Pr(h_t = h|\tilde{X}_t, x_t, \epsilon_t^m) \lambda^d_t V_0^{t+1}(\tilde{X}_{t+1}) \right], \quad (B1)$$

subject to budget constraint \([19]\) where $\tilde{X}_{t+1} = (ed, type, E_{t+1}, h)$ and

$$\check{V}_0^t(\tilde{X}_{t+1}) = \int \mathbb{E}_{e_t^{UP}} \left[ \max \left\{ V_0^{t+1}(\tilde{X}_{t+1}) + \epsilon_t^n, V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS) \right\} \right] dF^d(\theta, INS)$$

$$V_0^{t+1}(\tilde{X}_{t+1}) = \max \left\{ V_0^{t+1}(\tilde{X}_{t+1}, 0) + \epsilon_t^{HIX}, V_0^{t+1}(\tilde{X}_{t+1}, 2) \right\}.$$  

$\epsilon_t^{HIX}$ is i.i.d. preference shock to purchase health insurance from HIX. Note that the main change is a possibility that the unemployed can purchase health insurance from HIX.

Similarly, consider an employed worker having characteristics $\tilde{X}_t = (ed, type, E_t, h_t)$ and insurance status $INS \in \{0, 1, 2\}$ and who are offered skill price $\theta$. His value function is also defined by:

$$V_1^t(\tilde{X}_t, \theta, INS) = \mathbb{E}_{e_t}^{m} \left[ \max_{x_t} U_t(C_t(x_t, \epsilon_t^m, \tilde{X}_t, \theta, INS), 1, h_t) + \beta \sum_{h} \Pr(h_t = h|\tilde{X}_t, x_t, \epsilon_t^m) \delta^{d} \check{V}_1^t(\tilde{X}_{t+1}) \right], \quad (B2)$$

subject to budget constraint \([19]\) where $\tilde{X}_{t+1} = (ed, type, E_{t+1}, h)$ and $\check{V}_1^t(\tilde{X}_{t+1})$ and $\check{V}_1^t(\tilde{X}_{t+1}, INS, \theta)$ are defined as follows:

$$\check{V}_1^t(\tilde{X}_{t+1}) = (1 - \lambda^d_t) V_0^{t+1}(\tilde{X}_{t+1}) + \lambda^d_t \int \mathbb{E}_{e_t^{UP}} \left[ \max \left\{ V_0^{t+1}(\tilde{X}_{t+1}) + \epsilon_t^n, V_1^{t+1}(\tilde{X}_{t+1}, \theta', INS') \right\} \right] dF(\theta', INS')$$

where $V_0^{t+1}(\tilde{S}_{t+1}^0)$ is defined above; $V_1^{t+1}(\tilde{X}_{t+1}, \theta', INS')$ is defined as follows. If the health insurance is offered from the employer, then

$$V_1^{t+1}(\tilde{X}_{t+1}, \theta', INS') = V_1^{t+1}((\tilde{X}_{t+1}, \theta', 1)),$$

Otherwise, it takes the form

$$V_1^{t+1}(\tilde{X}_{t+1}, \theta', INS') = \max \left\{ V_1^{t+1}(\tilde{X}_{t+1}, \theta, 0) + \epsilon_t^{HIX}, V_1^{t+1}(\tilde{X}_{t+1}, \theta, 2) \right\}.$$  

Similarly,

$$\check{V}_1^t(\tilde{X}_{t+1}, \theta, INS) = (1 - \lambda^e) \mathbb{E}_{e_t^{UP}} \left[ \max \left\{ V_0^{t+1}(\tilde{X}_{t+1}) + \epsilon_t^n, V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS) \right\} \right]$$

$$+ \lambda^e \int \mathbb{E}_{e_t^{UP}} \left[ \max \left\{ V_0^{t+1}(\tilde{X}_{t+1}) + \epsilon_t^n, V_1^{t+1}(\tilde{X}_{t+1}, \theta, INS), V_1^{t+1}(\tilde{X}_{t+1}, \theta', INS') \right\} \right] dF(\theta', INS').$$
The difference between $V_{t+1}^t(\tilde{X}_{t+1}, \theta, INS)$ and $V_{t+1}^t(\tilde{X}_{t+1}, \theta', INS')$ is that the compensation package in the former is determined by the current employer, but by a potential employer in the latter. The terminal value is the same as [6].

### B.2 Steady State Worker Distribution

The steady state distribution of $g_t(\tilde{X}, \theta, INS)$ and $u_t(\tilde{X}, INS)$ are characterized as follows. First, define $\tilde{g}_t(\tilde{X}, \theta, NINS)$ as the measure of employed workers having characteristics $(t, \tilde{X})$ who are offered skill price $\theta$ but not offered health insurance:

$$\tilde{g}_t(\tilde{X}, \theta, NINS) = g_t(\tilde{X}, \theta^{ed}, 0) + g_t(\tilde{X}, \theta^{ed}, 2).$$

It is determined by

$$\frac{\tilde{g}_t(\tilde{X}, \theta, NINS)}{1 + n} = \sum_{NINS=0,2} \sum_{h_{t-1}} g_{t-1}(\tilde{X}^A_{t-1}, \theta, INS') \mathbb{E}_{\epsilon_{t-n}} \left[ \Pr(h_t = \hat{h} | \tilde{X}^A_{t-1}, \theta, INS', \epsilon_t) \right]$$

$$\times \left[ (1 - \delta^{ed}) (1 - \lambda^{ed}_e) \Pr(\Omega_1^E(\theta, 1, \tilde{X})) + (1 - \delta^{ed}) \lambda^{ed}_e \Pr(\Omega_2^E(\theta, 1, \tilde{X})) \Pr(\Omega_1(\theta, 1, \tilde{X})) \right]$$

$$+ \sum_{h_{t-1} \in INS'} \sum_{h_{t-1} \in INS'} u_{t-1}(\tilde{X}^B_{t-1}, INS') \mathbb{E}_{\epsilon_{t-n}} \left[ \Pr(h_t = \hat{h} | \tilde{X}^B_{t-1}, \epsilon_t) \right] \Pr(\Omega_1^E(\theta, 1, \tilde{X}))$$

$$\times \lambda_u f^{ed}(\theta, 1)$$

$$+ \sum_{h_{t-1} \in INS'} \sum_{h_{t-1} \in INS'} \int g_{t-1}(\tilde{X}^A_{t-1}, \theta', INS') \mathbb{E}_{\epsilon_{t-n}} \left[ \Pr(h_t = \hat{h} | \tilde{X}^A_{t-1}, \theta', INS', \epsilon_t) \right] 1(\theta, 1, \theta', INS') d\theta'$$

$$\times (1 - \delta^{ed}) \lambda^{ed}_e f^{ed}(\theta, 1)$$

$$+ \sum_{h_{t-1} \in INS'} \sum_{h_{t-1} \in INS'} \int g_{t-1}(\tilde{X}^A_{t-1}, \theta', INS') \mathbb{E}_{\epsilon_{t-n}} \left[ \Pr(h_t = \hat{h} | \tilde{X}^A_{t-1}, \theta', INS', \epsilon_t) \right] \Pr(\Omega_1^E(\theta, 1, \tilde{X})) d\theta'$$

$$\times \delta^{ed} \lambda^{ed}_e f^{ed}(\theta, 1)$$

(B3)

where $\tilde{X}^A_1 = (ed, type, E_{t-1}, \hat{h})$ and $\tilde{X}^B_1 = (ed, type, E_{t}, \hat{h})$ are individual characteristics in the last period for the employed and non-employed respectively which can turn into $\tilde{X}$ in this period, and $\Pr(\Omega_1^E(\tilde{X}, \theta, INS))$ is the probability that individuals with $\tilde{X}$ prefer to work a job with compensation package $(\theta^{ed}, INS)$ over being non-employed, which can formally be expressed as

$$\Pr(\Omega_1^E(\tilde{X}, \theta, INS)) = \Pr(\tilde{V}_0^t(\tilde{X}) + \epsilon_t < V_1^t(\tilde{X}, \theta^{ed}, INS))$$

where the value of choosing non-employed is

$$\tilde{V}_0^t(\tilde{X}) = \mathbb{E}[\max\{\tilde{V}_0^t(\tilde{X}, 0), \tilde{V}_0^t(\tilde{X}, 2) + \epsilon_t^{HIX}])$$

and $\Pr(\Omega_2^E(\theta^{ed}, INS))$ is the probability that individuals who receive a job offer from other firms decide to stay the current job:

$$\Pr(\Omega_2^E(\tilde{X}, INS, \tilde{X})) = F^{ed}(\theta, INS) + F^{ed}(\hat{\theta}_{INS}(\tilde{X}, \theta), INS)$$

where $F^{ed}(\theta, INS)$ is the cumulative distribution function of the employment distribution with parameters $(\theta, INS)$.

62
for $\overline{\text{INS}} \neq \text{INS}$ where $\tilde{\theta}_{\text{INS}}(\tilde{X}, \theta)$ is threshold skill price which can be defined as

$$V_t^1(\tilde{X}, \theta, \text{INS}) = V_t^1(\tilde{X}, \tilde{\theta}_{\text{INS}}(\tilde{X}, \theta), \overline{\text{INS}}).$$

$I(\theta, \text{INS}, \theta', \text{INS}')$ is the indicator function such that individuals prefer to take an offer from $(\theta, \text{INS})$ over $(\theta', \text{INS}')$:

$$I(\theta, \text{INS}, \theta', \text{INS}') = \begin{cases} 
1 & \text{if} \quad V_1^t(\tilde{X}, \theta, \text{INS}) > V_1^t(\tilde{X}, \theta', \text{INS}') \\
0 & \text{otherwise}
\end{cases}$$

Then, one can characterize $g_t(\tilde{X}, \theta, \text{INS})$ as

$$g_t(\tilde{X}, \theta, 2) = \mathbb{E}_{h|X} \left[ \Pr \left[ V_t^1(\tilde{X}, \theta, 2) + \epsilon_t^{HIX} > V_1^t(\tilde{X}, \theta, 0) \right] \right] g_t(\tilde{X}, \theta, \text{NINS})$$

and $g_t(\tilde{X}, \theta, 0) = \tilde{g}_t(\tilde{X}, \theta, \text{NINS}) - g_t(\tilde{X}, \theta, 2)$. $g_t(\tilde{X}, \theta, 1)$ is derived in the same way as in the main text. Therefore, I omit its derivation here.

One can characterize the determinants of steady state measures of non-employed at age $t$ with characteristics $\tilde{X}$, $\tilde{u}_t(\tilde{X})$, as follows. First, define $\tilde{u}_t(\tilde{X})$ as the measure of non-employed workers with characteristics $(t, \tilde{X})$ which satisfy

$$\tilde{u}_t(\tilde{X}) = u_t(\tilde{X}, 0) + u_t(\tilde{X}, 2)$$

Then, characterize $\tilde{u}_t(\tilde{X})$ as

$$\frac{\tilde{u}_t(\tilde{X})}{1 + n} = \sum_{I\text{INS}=0.2} \sum_{b_{t-1}} u_{t-1} \left( \tilde{X}_{b_{t-1}}^B, \text{INS}' \right) \mathbb{E}_{\epsilon_t^{m}} \left[ \Pr(h_t = \hat{h}_t|\tilde{X}_{b_{t-1}}^B, \epsilon_t^{m}) \right] \left( 1 - \lambda_{ed} \right) + \sum_{I\text{INS}=0.2} \sum_{b_{t-1}} u_{t-1} \left( \tilde{X}_{b_{t-1}}^A, \text{INS}' \right) \mathbb{E}_{\epsilon_t^{m}} \left[ \Pr(h_t = \hat{h}_t|\tilde{X}_{b_{t-1}}^A, \epsilon_t^{m}) \right] \Pr(\Omega_t^U(\tilde{X})) \lambda_{ed}^c$$

$$+ \sum_{b_{t-1}} \sum_{\text{INS}} \int g_{t-1} \left( \tilde{X}_{b_{t-1}}^A, \theta, \text{INS} \right) \mathbb{E}_{\epsilon_t^{m}} \left[ \Pr(h_t = \hat{h}_t|\tilde{X}_{b_{t-1}}^A, \epsilon_t^{m}) \right] \Pr(\Omega_t^U(\tilde{X}, \theta^{ed}, \text{INS})) dF(p) \left( 1 - \delta_{ed} \right)$$

$$+ \sum_{b_{t-1}} \sum_{\text{INS}} \int g_{t-1} \left( \tilde{X}_{b_{t-1}}^A, \theta, \text{INS} \right) \mathbb{E}_{\epsilon_t^{m}} \left[ \Pr(h_t = \hat{h}_t|\tilde{X}_{b_{t-1}}^A, \epsilon_t^{m}) \right] d\delta_{ed} \left( 1 - \lambda_{ed} \right)$$

where $\tilde{X}_{b_{t-1}}^A$ and $\tilde{X}_{b_{t-1}}^B$ are defined as above, and $\Pr(\Omega_t^U(\tilde{X}))$ is the probability that the non-employed with characteristics $\tilde{X}$ decides to turn down the offer:

$$\Pr(\Omega_t^U(\tilde{X})) = \int \Pr(\tilde{V}_t^0(\tilde{X}) + \epsilon_t^n > V_t^1(\tilde{X}, \theta^{ed}, \text{INS})) dF(\theta^{ed}, \text{INS}).$$

$\Pr(\Omega_t^U(\tilde{X}, \theta^{ed}, \text{INS}))$ is the probability that the employed with characteristics $\tilde{X}$ and a job with compensation package $(\theta^{ed}, \text{INS})$ decides to quit into the non-employed:

$$\Pr(\Omega_t^U(\tilde{X}, \theta^{ed}, \text{INS})) = \Pr(\tilde{V}_t^0(\tilde{X}) + \epsilon_t^n > V_t^1(\tilde{X}, \theta^{ed}, \text{INS})).$$

Then, one can characterize $u_t(\tilde{X}, \text{INS})$ as

$$u_t(\tilde{X}, 2) = \mathbb{E}_{h|X} \left[ \Pr \left[ V_0^t(\tilde{X}, 2) + \epsilon_t^{HIX} > V_0^t(\tilde{X}, 0) \right] \right] \tilde{u}_t(\tilde{X})$$

and $u_t(\tilde{X}, 0) = \tilde{u}_t(\tilde{X}) - u_t(\tilde{X}, 2)$. 63
C Parameterization of Policy Parameters

I describe the approach to parameterize the stylized version of the ACA in the model. The approach follows Aizawa and Fang (2013) which examines the impact of ACA on labor market outcomes, but several modifications are made to fit the model environment in this paper.

C.1 Penalties associated with individual mandate

The tax penalty on the uninsured in the ACA (from 2016 when the law is fully implemented) is set that the uninsured need to pay a tax penalty of the greater value of $695 per year or 2.5% of the taxable income above the Tax Filing Threshold (TFT), which can be written as:

\[
IM^{\text{ACA}}(y) = \max \left\{ 0.025 \times (y - \text{TFT}_{2011}), 695 \right\},
\]

where \(y\) is annual income.

I adjust the above formula in several dimensions. First, I adjust the scale of policy parameters to fit the 2007 economic environment. I estimated the model using data sets in 2004-2007 where the price level is normalized to 2007 value, while the ACA policy parameters are chosen to suit the economy in 2011. It is well known that the U.S. health care sector has a very different growth rate than that of overall GDP; in particular, there are substantial increases in medical care costs relative to GDP. Thus I need to appropriately adjust the policy parameters in the ACA to make them more in line with the U.S. economy in 2007. I implement the adjustment as follows: the $695 amount is adjusted by the ratio of the 2007 Medical Care CPI (CPI\text{Med}_{2007}) relative to the 2011 Medical Care CPI (CPI\text{Med}_{2011}); I choose this adjustment given the idea that the penalty amount $695 is chosen to be proportional to 2011 medical expenditures. I then multiply it by 1/3 to reflect the fact that the period-length in the model is four-month. Second, I adjust the TFT\text{2011} by the ratio of 2007 CPI of all goods (CPI\text{All}_{2007}) relative to the 2011 CPI of all goods (CPI\text{All}_{2011}). I also multiply it by 1/3 to reflect the choice of the four-month model period in this paper. Finally, I adjust the percentage 2.5% by the differential growth rate of medical care and GDP, i.e., multiply it by the relative ratio of \(\frac{\text{CPI\text{Med}_{2007}}}{\text{CPI\text{All}_{2007}}}\) and \(\frac{\text{CPI\text{Med}_{2007}}}{\text{CPI\text{All}_{2011}}}\). With these adjustments, the tax penalties on the uninsured are parameterized as:

\[
IM^{\text{ACA}}(y) = \max \left\{ \frac{0.025 \times (\text{CPI\text{Med}_{2007}} / \text{CPI\text{All}_{2007}})}{\left( \text{CPI\text{Med}_{2011}} / \text{CPI\text{All}_{2011}} \right) \times \left( y - \frac{1}{3} \text{TFT}_{2011} \times \frac{\text{CPI\text{Med}_{2007}}}{\text{CPI\text{All}_{2011}}} \right)} \times \frac{1}{3} \times 695 \times \frac{\text{CPI\text{Med}_{2007}}}{\text{CPI\text{All}_{2011}}} \right\},
\]

where \(w\) is four-month income in dollars.

C.2 Penalties associated with employer mandate

Tax penalties on employers in the ACA are set that firms with 50 or more full-time employees that do not offer coverage need to pay a tax penalty of $2,000 per full-time employee per year, \[I obtain CPI data for medical care and all goods both from Bureau of Labor Statistics website: http://www.bls.gov/cpi/data.htm.\]
excluding the first 30 employees from the assessment. That is,

\[ EM^{ACA}(l) = (l - 30) \times 2,000. \]  

As in the case of individual mandate, I first adjust the above formula by first scaling the $2,000 per-worker penalty using the ratio of the 2007 Medical Care CPI relative to the 2011 Medical Care CPI and by multiplying it by 1/3 to reflect our period-length of four months instead of a year, i.e., for \( l \geq 50 \),

\[ EM^{ACA}(l) = \frac{1}{3} \left( (l - 30) \times 2,000 \times \frac{CPI_{Med, 2007}}{CPI_{Med, 2011}} \right). \]

While it is ideal to apply this formula precisely in the model, to simplify the numerical algorithm, I approximate this penalty function as a differentiable function by removing the discontinuity, by following the approximation technique used by MaCurdy, Green, and Paarsch (1990).

C.3 Medicaid expansion

ACA stipulates that individuals with income below 133% of Federal Poverty Level (FPL) are able to enroll in the free public insurance Medicaid. While it is ideal to model this threshold carefully, given my sample selection those who are below 133% of Federal Poverty Level (FPL) tend to be non-employed. Moreover, there is a certain technical difficulty to model this threshold as described below. Therefore, to simplify the analysis, I assume that only non-employed individuals will be covered by Medicaid. As I mention in Section 6.1.5, I also consider the case that Medicaid expansion is not implemented.

C.4 Premium subsidies in HIX

In the ACA, federal premium subsidies are available to individuals who purchase health insurance from HIX if their incomes are less than 400% of the Federal Poverty Level (FPL), denoted by FPL400\(^3\). The premium subsidies will be set on a sliding scale such that the premium contributions are limited to a certain percentage of income for specified income levels. If an individual’s income is at 133% of the FPL, denoted by FPL133, premium subsidies will be provided so that the individual’s contribution to the premium is equal to 3.5% of his income; when an individual’s income is at FPL400, his premium contribution is set to be 9.5% of the income. If his income is above FPL400, he is no longer eligible for premium subsidies. Note that the premium support rule as described in ACA creates a discontinuity at FPL133: individuals with income below FPL133 receive free Medicaid, but those at or slightly above FPL133 have to contribute at least 2.3% of his income to purchase health insurance from HIX. To avoid this discontinuity issue, I instead adopt a slightly modified premium support formula as follows:

\[
S(y, R^{HIX}(t)) = \begin{cases} 
\max \left\{ R^{HIX}(t) - \left[ 0.0350 + 0.060 \frac{3y}{FPL400} \right] \times \frac{CPI_{Med, 2007}}{CPI_{Med, 2011}}, 0 \right\} & \text{if } y < \frac{FPL400}{3} \\
0, & \text{otherwise,}
\end{cases}
\]

\(2\)In July 2013, the government decided to postpone the implementation of the employer mandate until 2015.

\(3\)I assume that FPL is defined for a single person. In 2007, it is $11,200 annually.
where \( w \) is four-month income.

As I mention in Section 6.1.5 at this stage, it is unclear whether premium subsidies will be given to individuals with less than FPL133 who live in states where Medicaid is no expanded. In this paper, I assume that those individuals will obtain health insurance at zero premium from HIX.

### C.5 Age-based pricing regulation

In the ACA, they set the maximum premium ratio between the oldest and the youngest \( \omega_A = 3 \). Therefore, if it binds,

\[
\omega_{AGE} R^{HIX}(1) = R^{HIX}(T) .
\]

The issue is that it does not specify how the premium can vary over age. For simplicity, I assume that the premium is linearly increasing in \( t \). Given this restriction, in order to satisfy the market clearing condition in HIX, insurance premia in HIX must be determined as follows:

\[
R^{HIX}(t) = R^{HIX}(1) + \left( \frac{t - 1}{T - 1} \right) R^{HIX}(1)
\]

\[
R^{HIX}(1) = (1 + \xi_{HIX}) R^*_HIX \left( \frac{T - 1}{\omega_{AGE} - 1 + \frac{\omega_{AGE} - 1}{T + 1}} \right) \tag{C8}
\]

where \( R^*_HIX \) is the pooled premium among the all the participants in HIX:

\[
R^*_HIX = (1 + \xi_{HIX}) \sum_t \sum_X \int E[m_s^X_t|INS_t = 2] g_t \left( \bar{X}, \theta, 2 \right) d\theta .
\]

To understand the role of age based pricing regulation, I also consider the case where there is no regulation: in such a case, premia are determined as

\[
R^{HIX}(t) = (1 + \xi_{HIX}) \sum_X \int E[m_s^X_t|INS_t = 2] g_t \left( \bar{X}, \theta, 2 \right) d\theta . \tag{C9}
\]

That is, each age group consists of separate risk pool.