# Fragility: A Quantitative Analysis of the US Health Insurance System*

Björn Brügemann
Yale University
bjoern.bruegemann@yale.edu

Iourii Manovskii
University of Pennsylvania
manovski@econ.upenn.edu

**Preliminary and Incomplete.**[1]

**Abstract**

We develop a framework for the quantitative study of several aspects of the US health insurance system. The system is largely employer based. Employer purchased insurance is tax deductible, while individually purchased insurance is not. Employers are prohibited from discrimination based on health status when offering coverage. Insurance contracts have a typical duration of one year and insurance companies are allowed to change the premiums with few binding restrictions.

One feature of the data is that smaller employers are less likely to provide coverage than larger employers. This might be due to the existence of fixed cost of obtaining and maintaining coverage. Alternatively, a more volatile health composition associated with being small may make coverage less sustainable.

We develop a quantitative equilibrium model that features tax deductibility of employer-provided coverage, non-discrimination restrictions, fixed costs of coverage and employers that hire discrete numbers of workers in frictional labor markets. We use the calibrated model to understand what drives the patterns of insurance provision with employer size and to evaluate the effects of this system on the flows of workers across health insurance coverage status, labor market flows, as well as on the size distribution of establishments, and aggregate productivity. We also evaluate the effects of various proposed reforms of the system.

JEL Classification: E20, E24, E25, E62, I10, J32, J63, J78.

Keywords: Health, Health Insurance, Tax Policy, Labor Markets, Labor Mobility, Discrimination.

*We would like to thank conference participants at the 2009 NBER Summer Institute, 2009 "Health and the Macroeconomy" conference at the Laboratory for Aggregate Economics and Finance at the University of California – Santa Barbara, and 2009 Yale Health Economics Workshop for their comments. We would also like to thank Susan Hoffman and Martha Keon of Littler Mendelson, P.C. for helping us interpret the US health insurance laws and regulations. This research has been supported by the National Science Foundation Grant No. SES-0617876.

[1]First version: April 28, 2008. This version: February 10, 2010. Please check authors' websites `http://www.econ.yale.edu/~bb338/` and `http://www.econ.upenn.edu/~manovski/` for the most recent version.

# 1 Introduction

In this paper we develop a basic equilibrium model that can be used to quantitatively study the decisions of employers to provide health insurance coverage to their employees and the aggregate implications of these decisions. In particular, we quantitatively assess the implications of the current design of the US health insurance system, the effects of various attempts to reform the system at the state level and the likely consequences of implementing the proposed federal legislation currently debated in Congress.

The US health insurance system for those younger than 65 is largely employer-based. Among those with private coverage from any source, 94.6 percent of adults and 94.0 percent of children held employment-related coverage in 2006 (Selden and Gray (2006)). The predominance of employer-provided insurance is often attributed to its tax deductibility, which is not available to employees purchasing insurance individually. This provides employers with strong incentives to provide coverage. At the same time, employer-provided insurance is subject to numerous regulations. By law health insurance options that an employer offers to its employees cannot be conditioned on employees' health. Moreover, most severe medical conditions make employees eligible for protection against discrimination in wages, benefits, hiring, and firing. An important feature of the system is that employers assume much of the risk. Large employers often self-insure. And while smaller employer contract with insurance carriers, the typical contract duration is annual, and (in most states) premiums can be adjusted with the claims experience of the employer, leaving employers exposed to long-term risk.

Despite the tax subsidy, almost half of employers do not provide insurance. A prominent feature of the data is that large employers are much more likely to offer health insurance than small employers. Only 46% percent of establishments with up to 10 employees offer health insurance, compared with 96% of establishments with more than 50 employees. This difference is often attributed to much higher administrative loads that insurance carriers apply to small groups. The dynamics of coverage also vary with employer size, with smaller employers more likely to discontinue coverage. This may be a consequence of the more volatile health composition of small employers. If a single employee, or dependent, develops a costly persistent medical condition, then experience rating together with non-discrimination provisions generates a large increase in premiums for all employees upon contract renewal. The tax subsidy may then be insufficient to make employer-provided insurance attractive to healthy employees, prompting the

employer to drop coverage. We refer to this mechanisms as the "fragility channel", and assessing its quantitative implications is one of the main objectives of the paper.

A concern about experience rating and fragility of coverage in small businesses were the key motivation for proponents of health insurance reform from the 1990s to the present day. Several states adopted some form of community rating, either pure community rating where all groups within a given insurance pool pay the same premium, or modified community rating, which allows premiums to vary with a few specific characteristics such as age, gender, or industry. But most states adopted a more conservative approach, imposing rating bands that require premiums to be within a certain range (e.g. 30%) of an average premium for a particular class of business. The argument for proceeding cautiously was that community rating, or rating bands which are to tight, might induce a severe adverse selection in the small group market.

Both the concern about experience rating and the concern about the possibility of an adverse selection death spiral are driven by the same feature that distinguishes small employers from large employers, namely that the health status composition of the workforce is more variable in small employers. Given a stable health composition of large employers, experience rating would induce little variation in premiums from year to year. Less dispersion in health composition also reduces the scope for adverse selection. Thus the link between employer size and variability of health composition lies at the heart of the trade-off that shaped state-level small group reform in the 1990s.

While qualitatively it is clear that health composition is more variable for small employers, it is less clear whether this link is sufficiently strong quantitatively to generate substantial fragility of coverage under experience rating, and substantial adverse selection under community rating. Statistical evidence (discussed below) on the prevalence and effects of experience rating is scarce because of the lack of data that links employers' coverage decisions to changes in the health composition of their workforce. Moreover, we show below that such statistical evidence would be hard to interpret given the presence of strong equilibrium effects. Using a quantitative equilibrium model allows us to bring to bear additional evidence on these questions. Specifically, we utilize evidence on the dynamics of medical expenditures at the individual level, and evidence on the employer-size distribution.

Our model combines elements from two literatures. To generate an employer-size distribution we build on the establishment-size dynamics literature, specifically the model

of Hopenhayn and Rogerson (1993). The latter has frictionless labor markets. For our purposes it is not innocuous to abstract from labor market frictions, since the strength of adverse selection in a system of employer-provided health insurance depends on the ease of changing employers. This leads us to introduce search frictions, building on Mortensen and Pissarides (1994).

The key new element of our model is that employer size is discrete rather than continuous. This is necessary to capture that health-expenditure shocks at the worker level translate into health-composition shocks at the establishment level. This also introduces a strategic element, as employees take into account the effect of their own mobility decisions on the future health composition and coverage decisions of their employers.

The model incorporates the institutional environment in which employers make coverage decisions, specifically tax deductability and the non-discrimination provisions discussed above.

The individual health expenditure process is calibrated using data from the Medical Expenditure Panel Survey (MEPS). Through the establishment size distribution, this maps into health composition shocks at the employer level. The shape of this distribution, in particular the share of employment in small establishments, is another important determinant of adverse selection. Thus the US establishment-size distribution is a key target in our calibration.

In models of employer-provided insurance, health and otherwise, an important modeling choice concerns employers' ability to commit to employment contracts. We show that if employers have no ability to commit, then the model has the robust counterfactual implication that both the level and the stability of coverage are *decreasing* in employer size. We show that a simple commitment technology allows the model to match observed patterns of coverage. Interestingly, large employers endogenously choose to commit to provide insurance, while small employers do not. We also show that costs of starting or stopping coverage act similarly to commitment, but that the levels of costs needed to account for observed coverage are implausibly large.

We use the calibrated model to analyze the type of reforms implemented at the state level in the 1990s, as well as stylized versions of reform proposals currently debated in Congress. Concerning state-level reforms, we ask whether the model is consistent with the empirical finding that community rating or the imposition of rating bands had relatively small effects on aggregate coverage rates in the small group market. We find

that whether community rating increases or decreases aggregate coverage indeed depends on how this type of scheme is implemented. Coverage increases if the threshold below which employers are subject to community rating is sufficiently high, and if the scheme pools not only medical expenditures but also administrative expenses. But coverage may fall if the scheme does not include sufficiently many large establishments, or if pooling of administrative expenses is limited. In future drafts of this paper we will also include an analysis of the rating bands imposed in many states.

Current reform proposals combine elements of community rating among small employers, subsidies to coverage for small employers, penalties for large employers that do not provide, as well as an individual mandate. In this draft we analyze the first three elements, both separately as well as jointly in order to shed light on their interaction. We find that community rating among small employers leads to a collapse in coverage if it is not complemented by subsidies. Penalties imposed on large employers have small effects on aggregate coverage. Future drafts will incorporate the proposed individual mandate into the analysis.

In future work we plan to use the model to address several additional questions. In particular, we will study the distortions associated with the current system of employer-provided insurance. Are their non-trivial effects on worker flows across establishments, the size distribution of establishments and establishments dynamics? What are the costs (or benefits) of the system when the consequences of these distortions are taken into account? We also plan to study implications for the dynamics of coverage at the individual level. In 2007 19.7 percent of adults between the ages of 19 and 65 were not covered by health insurance (Henry J. Kaiser Family Foundation (1998)) and over 80 percent of the uninsured were wage earners or members of working families. However, this is not a static pool of people, there are substantial flows in and out of it. To what extent does the fragility of coverage by small employers together with worker flows across employment and non-employment states and across employers that do and do not offer health insurance play a role in accounting for these flows?

Two other papers studied related issues using equilibrium models. Dey and Flinn (2005) present an equilibrium model of health insurance provision by firms and wage determination. They investigate the effect of employer-provided health insurance on job mobility rates and economic welfare using an on-the-job search and bargaining framework. They find that the employer-provided health insurance system does not lead to any serious

inefficiencies in mobility decisions. However, they assume that wages and health coverage are negotiated between workers and firms on an individualistic basis, that is, without reference to the composition of the firm's current workforce. Moreover, they do not model the favorable tax treatment of the employer-provided health insurance coverage. Thus, they abstract from the key mechanisms evaluated in this paper.

Jeske and Kitao (2008) also study the effects of tax deductibility of group coverage using a dynamic general equilibrium model with heterogeneous agents. The key difference is that they do not model firms. Instead, it is assumed that workers either have a stochastic opportunity to join a group plan (or opt out) or they do not have such an opportunity and have a choice of whether to purchase individual coverage. Thus, all workers with the opportunity to purchase health insurance belong to one large group. In contrast, in our model, such groups are defined at the employer level. As in our model, healthy workers that participate in the group health plan subsidize unhealthy workers. The extent of their willingness to do so depends on the amount of tax savings they obtain. Without the tax subsidy, group insurance unravels. The model in Jeske and Kitao (2008) is not designed to address the fragility of coverage of small employers, the effects of the health insurance system on labor market flows, and its heterogeneous effects on employers of different sizes.

The paper is organized as follows. In Section 2 we describe the key facts motivating our analysis. In Section 3 we develop the model of the employer-provided health insurance coverage that features the fragility channel. In Section 4 we define equilibrium. In Section 5 we calibrate the model and perform the quantitative analysis of the current system and evaluate the effects of several proposed reforms. Section 6 concludes.

# 2 Facts

## 2.1 Data Sources

### 2.1.1 1997 Robert Wood Johnson Foundation Employer Health Insurance Survey

The 1997 Robert Wood Johnson Foundation Employer Health Insurance Survey (Long and Marquis (1997)) is a nationally representative survey of public and private employers conducted in 1996 and 1997. Data were collected on employers' offers of health insurance coverage, employees' eligibility and enrollment in health plans, and, for each plan offered, the plan type (HMO, POS, PPO, conventional), premiums (employer and em-

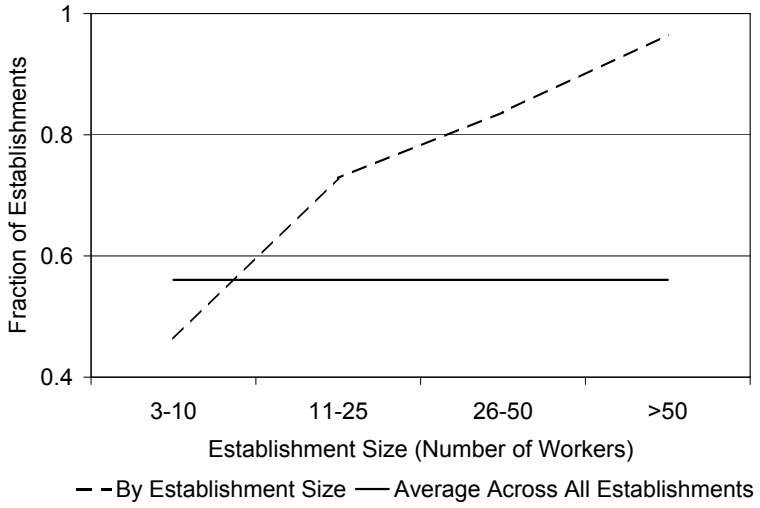


Figure 1: Fraction of Establishments Providing Insurance by Establishment Size.

ployee contributions), benefits, cost-sharing, and employer self-insurance status. The study also collected information on the characteristics of employers and workers, including the number of employees at the establishment, the number of employees statewide and nationwide, and the distribution of workers by hours worked, age, sex, and earnings. Our analysis is based on a sample of $21,545$ private sector employers. All results are weighted by the sampling weights provided by the Survey.

## 2.2 Insurance Provision by Establishment Size

It is well-known that insurance provision varies substantially with establishment size. Figure 1 plots the fraction of establishments providing insurance as a function of establishment size. Overall, 56% of all employers provide health insurance to their workers. However, only 46% of employers with 10 or fewer employees provide coverage. Insurance provision increases rapidly with establishment size to 73% of employers with 11 to 25 employees, 84% of employers with 26 to 50 employees, and 96% of employers with more than 50 employees.

Figure 2 illustrates that this pattern is robust to controlling for the average wage received by workers. Establishments that pay higher wages are indeed more likely to provide health insurance but the gradient of the probability of coverage with respect to

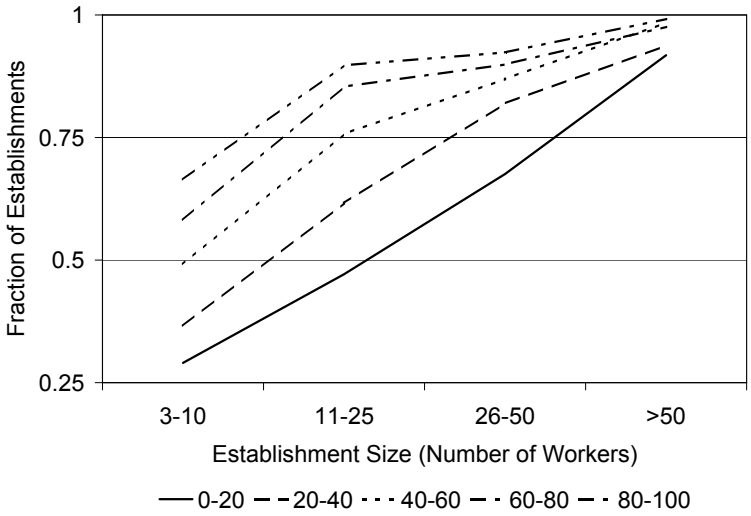


Figure 2: Fraction of Establishments Providing Insurance by Establishment Size and Average Payroll.

establishment size is almost independent of the average payroll. Strikingly, the lowest paying establishments with more than 50 employees are 40% more likely to offer health insurance than the highest paying establishments with less than 10 employees.[2]

## 2.3 Discontinuance of Insurance Provision by Establishment Size

One issue that has received little attention in the literature (Long and Marquis (1998) is one exception) is that establishments occasionally discontinue offering insurance coverage to their workers. The probability of discontinuing coverage varies systematically with establishment size. In particular, Figure 3 illustrates that 11% of establishments that offered insurance within two years prior to the survey date were not offering at the time of the survey. This fraction is 15% for establishments with less than 10 employees and declines to less than 1% of establishments with more than 50 employees.[3]

One drawback of the statistic above is that it might be affected by time-aggregation.

[2]The establishment were ranked by their percentile in the overall establishment distribution, not the distribution within a size class, making such a comparison meaningful.

[3]The fraction of establishments starting coverage by establishment size is nearly identical to the fraction of establishments discontinuing coverage so that the fraction of establishments providing coverage remains nearly constant across the size categories.

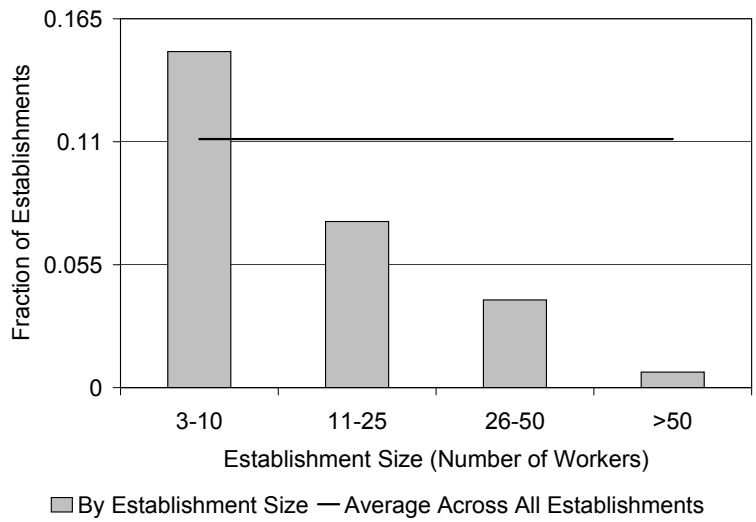


Figure 3: Fraction of Establishments Discontinuing Insurance over a 2 Year Period by Establishment Size.

The RWJ survey asks employers that do not provide insurance at the time of the interview whether they provided within the preceding two years. The survey does not ask whether they provided exactly two years ago (as we assumed to be the case when computing the preceding statistic). Similarly, the establishments that do provide insurance at the time of the interview are only asked how long they have been providing, not whether they were providing exactly two years ago (as we assumed to be the case when computing the preceding statistic). One way to address the time aggregation problem is to postulate a simple reduced form model of turnover in insurance provision.[4] Suppose there is a unit mass of employers that start coverage at rate $\alpha$ and stop coverage at rate $\sigma$. Then, the mass of employers providing coverage in steady state is $c = \frac{\alpha}{\alpha+\sigma}$. Using that the employers are asked whether they provided within the last two years, $\alpha = -\frac{1}{2}\log[1 - x]$, where x denotes the fraction of employers that are not providing coverage at the time of the interview. Since we observe the fractions of employers providing and not providing coverage at the time of the interview, we can solve for the instantaneous rate of discontinuing insurance. The results of this exercise are plotted in Figure 4. The

[4]We do this for the expositional purposes only. When we perform the quantitative analysis using the model developed below, we will exactly replicate the design of the RWJ survey on the model generated data.

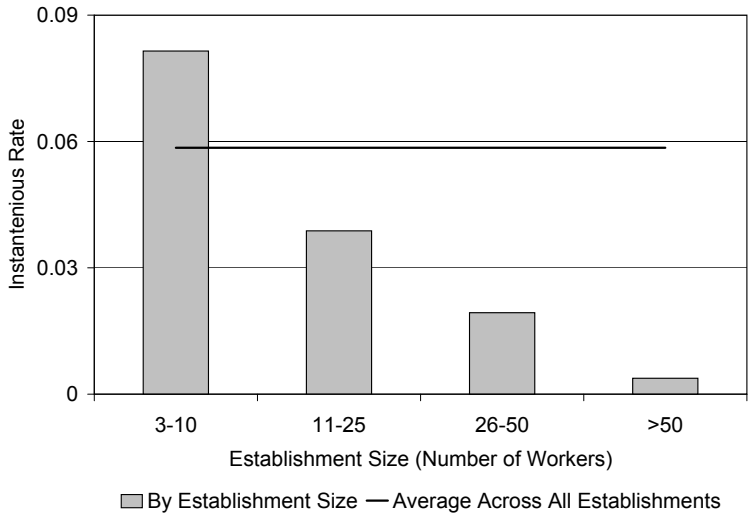


Figure 4: Instantaneous Rates of Discontinuing Insurance by Establishment Size.

instantaneous rate of discontinuing insurance equals 0.6 for establishments of all sizes and declines from 0.08 for establishments with 10 or fewer workers to 0.004 for establishments with more than 50 workers.

## 2.4 Variability of Insurance Premiums by Establishment Size

Health insurance premiums faced by establishments are quite volatile over time. Moreover, this volatility is also systematically related to establishment size.

The average increase in premium in the RWJS sample of private sector establishments providing coverage in 1996 and 1997 was around 2.5% irrespective of the establishment size. This relatively small increase in private health insurance premiums in 1997 accords well with other sources of data.[5]

The standard deviation of the premium change across establishments of all sizes was considerably higher at 9.9. This indicates that establishments experience substantial changes (positive and negative) to premiums from one year to the next. The standard deviation of the premium change declines from 10.4% for establishments employing 10 or fewer workers to 8.3% for establishments with more than 50 employees. As an alternative

[5]E.g., The U.S. Department of Health and Human Services Fact Sheet available at http://www.hhs.gov/news/press/2000pres/20000110.html.

Figure 5: Variability of Insurance Premiums.

statistic, consider the difference between the 90th and 10th percentile of the premium change distribution. This statistic equals 18% for establishments of all sizes and declines from 20% for establishments with less than 10 employees to 15% for establishments with more than 50 employees.[6]

## 2.5 Features of Observed Insurance Contracts

In practice health insurance contracts have a typical length of one year. Long-term health insurance contracts are virtually non-existent. Moreover, premiums adjust almost freely upon renewal (subject to some restrictions imposed in several states). Insurance premiums are typically based at least in part on the expected medical costs of the employer

[6]Cutler (1994) finds qualitatively similar patterns in the 1991 Health Insurance Association of America (HIAA) survey. He finds that the spread between the 90th and 10th percentile of the costs of comparable plans is 174% for firms with less than 50 workers, and it declines monotonically to 71% for firms with 501 to 1000 workers. The spread between the 90th and 10th percentile of the change in costs from one year to the next is 45% for firms with less than 50 workers, and it declines monotonically to 23% for firms with 501 to 1000 workers. Thus, he finds a larger spread of the distribution of cost changes. The difference may be attributable to the difference in survey years (1991 being a recession year and exhibiting insurance premium increase of 14% compared to 2.5% in 1997). Another difference between our studies is that we estimate the percent change in health insurance cost between the two years as directly reported by the respondent. Cutler's study is based on the data about the actual premiums paid which he attempts to adjusts for the different generosity of the plans. Finally, while our unit of analysis is an establishment, it is a firm in Cutler's study.

purchasing insurance, a method termed "experience rating" in the literature. Thus, establishments that experience adverse events will pay more for insurance than establishments that do not.

These features imply that the typical insurance contract insures the establishment against the risk that the medical costs within a year will exceed their expected value at the beginning of the year. However, if some workers in an establishment learn that they are permanently less healthy than they thought[7], future premiums will be higher than they expected. Current system does not provide insurance against such intertemporal risk, except to the extent that large establishments can achieve a relatively stable overall health composition of their workforce.

There is no consensus in the literature as to what forces have shaped the system to have these features. Some of the hypothesis offered in the literature include the presence of adverse selection and moral hazard, whereby had the insurers not experience rated, they would attract establishments with higher expected health costs or encourage less healthy behavior on the part of those insured. Another possibility involves inability of establishments and insurers to commit to long-term contracts. If employers can walk away from a contract, the establishments that have learned that their workers are healthier than they thought will do so. Establishments might be reluctant to agree to sizable pre-payment required to overcome this selection problem given the uncertainties they face about the future. Cochrane (1995) suggests that the competition for healthy groups among health insurers is a relatively recent phenomenon to which the regulatory and legal systems have not adapted yet to enforce potentially feasible long-term contracts. In this paper we do not attempt to contribute to this debate. Instead, we assume that the nature of the contracts offered is a feature of the environment that we take as given.

## 2.6 Legal Framework

Health insurance in the US is primarily regulated at the state level. Federal laws also apply and typically establish minimum requirements on, e.g., the standards for the availability of coverage. These requirements can be and often are made more stringent by state laws. Three federal laws with the most significant impact on health insurance regulation are the Employee Retirement Income Security Act (ERISA), American with Disabilities

[7]For example, having a heart attack today signals a greater risk of a heart attack in the future. Babies born with birth defects often require medical care throughout their life.

Act (ADA), Family and Medical Leave Act (FMLA), and Health Insurance Portability and Accountability Act (HIPAA).

### 2.6.1 No Mandate

There is no legal requirement for employers to offer health insurance. Individuals are also not required to purchase health insurance.

### 2.6.2 Tax Treatment

Health insurance premium paid by employers are fully tax deductible as a business expense. Individual purchase of health insurance is done with after-tax income.[8]

### 2.6.3 Non-Discrimination

ERISA is the major element of federal legislation that regulates employee benefit plans including employer-sponsored health insurance plans. ERISA's requirements are largely procedural. While ERISA states that plan participants cannot be discriminated against for asserting ERISA rights, these protections are relatively weak. Based on the premise that ERISA does not require employers to offer any plan at all, a sequence of court decisions establishes that ERISA does not prohibit employers from capping insurance benefits for particular disabilities, or indeed, from making any plan changes, even though those changes may completely, and even intentionally, exclude only a single employee from the plan's benefits. The restrictions on this are established in other laws discussed below.

ADA prohibits discrimination against qualified employees with disabilities (or even based on a (mis)perception of a disability) in "job application procedures, the hiring, advancement, or discharge of employees, employee compensation, job training, and other terms, conditions, and privileges of employment." The ADA makes it unlawful for an

[8]Under current law, employer-provided health insurance coverage is excluded from employees' income for determining their federal income taxes. Exclusions also apply to Social Security, Medicare, and unemployment taxes (both employer and employee portions) and to state income and payroll taxes as well. Considering the average cost of employment-based insurance, now around $4,750 a year for single coverage and $12,700 for family coverage (Lyke (2008)), these exclusions result in significant costs to the government. Joint Committee on Taxation (2008) estimated calendar year 2007 tax expenditures for the employer coverage exclusion to have been between $105 and $145.3 billion for the federal income tax and $100.7 billion for Social Security and Medicare taxes. The federal income tax component alone represents the single largest source of revenue loss in the U.S. Budget and is, e.g., 60 percent larger than the revenue loss from the federal income tax deduction of mortgage interest and 3 times larger than the current revenue loss form the tax-deductibility of contributions and earnings in 401(k) retirement plans (Table 19-1 in Office of Management and Budget (2008)).

employer to refuse to hire a disabled employee because of concerns about insurance premiums, or to discharge an employee because he or she has in fact caused an increase in health insurance premiums.

ADA also contains provisions that make it more difficult to learn about medical conditions of employees and prospective employees. Section 102(c) states that prohibition of discrimination includes prohibition of medical examinations and inquiries about medical conditions. The only acceptable pre-employment inquiries must concern the ability of an applicant to perform job-related functions. Employment entrance medical examinations are permitted only after an offer of employment is made and accepted, and only if all entering employees are subject to such an examination, and the results are used only in accordance with the anti-discrimination provisions of the ADA.

ADA has certain limitations. First, it applies only to employers with 15 or more employees. Most states have extended its reach to the smaller employers through the state law. Second, while case law associated with the Rehabilitation Act of 1973 (a precursor to the ADA that applied only to Federal employees and contractors) held conditions such as cancer, diabetes, multiple sclerosis, and HIV to be disabilities, a series of Supreme Court decisions in 1999 and early 2000s meant that the presence of a disability had to be determined with reference to any mitigating or corrective measures the individual uses to offset the effects of impairment, and that courts should only consider the present state of the individual. In response many states introduced changes to their statutes clarifying that their definition is meant to be broader than the Supreme Courts interpretation of the ADA. The U.S. Congress also clarified its intent in ADA Amendment Act of 2008 (ADAAA) which states that courts must determine whether a disability is present without regard to mitigating measures, that impairments which are episodic or in remission are disabilities if they substantially limit a major life activity when active, and provides an expanded list of "major live activities" which appear to include most serious health conditions.

FMLA allows employees who have been with an employer for more than a year to take up to twelve weeks of unpaid leave each year in connection with the birth or illness of a family member, or for the employee's own medical needs. Although the leave is unpaid, the FMLA requires the employer to maintain the employee in the group health insurance plan, and to continue to pay the same portion of the premium as if the employee were working. While FMLA is restricted to employers with more than 50 employees, most

states adopted similar legislation, often applying to smaller employers but requiring that employee pays the entire cost of their premiums during the leave.

HIPAA explicitly prohibits all group health plans from applying different eligibility rules, offering different benefits, or charging a different premium to any individual within a group on the basis of "health factors" including, among others, health status, medical condition, claims experience, medical history, and genetic information. HIPAA applies to employers of all sizes and provides most stringent protections at the federal level against discriminating individual workers on the basis of health, both with respect to coverage and premiums. In addition, HIPAA contains privacy provisions that prohibit group health plans from disclosing health information of individual workers to employers for any employment-related actions or decisions.

Finally, HIPAA contain portability restrictions that specify that pre-existing conditions can be excluded for a maximum of 12 months. However, previous coverage, say from a previous job or continuing coverage in between jobs, can be credited against the pre-existing conditions exclusion. Thus, individuals with a pre-existing conditions who maintained continuous coverage for 12 months cannot be discriminated against in hiring and will become immediately eligible to participate in the health plan of the new employer. While HIPAA guarantees access to coverage, it does not restrict insurers in setting premiums (beyond providing that rates had to be the same for all employees). Thus, insurers are free to raise rates for the whole group when an individual with a costly medical condition joins the group.

# 3 Model

At the beginning of every period there is a mass one of workers. Workers permanently leave the labor market with probability $1-\rho$ per period and leavers are replaced by labor market entrants. Preferences of workers are

$$\mathbb{E}_t \sum_{t'=t}^{\infty} \beta^{t'-t} U(c_{t'})$$

with $U' > 0$ and $U'' \leq 0$.

Workers are subject to idiosyncratic shocks to their health. Their health status follows a two state Markov process: healthy $h$ or unhealthy $u$ with transition probabilities $q^H_{ii'}$ for $i, i' \in \{h, u\}$. Health status in turn determines medical expenditures $e_u > e_h$. Let $q^0_i$ denote the fraction of labor market entrants in health status $i$.

The model is designed to study the link between shocks to the health composition at the establishment level and coverage decisions. It is computationally infeasible to apply our approach of modeling health composition shocks to very large establishment, those with more than a few hundred workers. At the same time, it is critical to have very large establishments in the model. Since they account for a large share of employment and almost always provide coverage their presence is needed to accurately capture the chance of workers to find a job that provides insurance. We choose to model very large establishments in a simplified way, described in Section 4.7. Unless explicitly stated otherwise, the discussion until then refers to establishments which are not very large.

There is a large mass of potential establishments. Employers share workers' discount factor $\beta$ but are risk neutral. Establishments are subject to idiosyncratic productivity shocks: $z$ indexes the productivity of an establishment, which follows a discrete Markov process with transition probabilities $q^Z_{zz'}$. Output of an establishment with productivity $z$ is given by

$$zF(g_e),$$

where $g_e$ is the number of workers employed, and $F(0) = 0$, $F' > 0$, $F'' \leq 0$. Active establishments are subject to a fixed operating cost $c_f$, which can only be avoided through exit. New establishments can enter by paying an entry cost $c_e$, and draw an initial productivity from the invariant distribution associated with $q^Z$.

The labor market is not competitive due to search frictions, to be described below. Compensation consists of wage payments and health insurance. Due to regulation, compensation cannot discriminate between healthy and unhealthy workers. To keep the determination of compensation as simple as possible given this restriction, we assume that establishments have all the bargaining power.

Non-discrimination regulations also apply to dismissal. To capture these constraints, we model the dismissal decision of an establishment in two steps, with the following timing. First, the establishment decides how many workers of each health status to retain. In this step, the establishment is subject to the constraint that workers asked to leave must not prefer to stay, given the compensation package. For example, if an establishment offers health insurance and unhealthy workers prefer to stay, the establishment cannot dismiss them in this step. In the second step, the establishment can dismiss workers at random without considering health status, not facing the constraint that workers must not prefer to stay. In addition to endogenous dismissal, an employed worker separates

exogenously with probability $\delta$ every period.

An establishment can recruit new workers by posting vacancies. It posts $g_v \in \{0, 1, \ldots\}$ vacancies at cost $c_v$ per vacancy.

The probability that a vacancy contacts a worker and the probability that a searching worker contacts a vacancy are given by

$$q(\theta) = M\left(\frac{1}{\theta}, 1\right) \quad \text{and} \quad f(\theta) = M\left(1, \theta\right), \tag{1}$$

respectively. Here $M$ is a constant returns to scale matching function and $\theta = \frac{v}{m}$ is the ratio of vacancies to searchers.

Establishments can purchase health insurance contracts that cover the medical expenditures of their employees. While wage income is taxed at rate $\tau \in [0, 1)$, the provision of health insurance is not subject to taxes. Health insurers charge an administrative load which may depend on the number of employees $\kappa(g_e)$. Apart from these administrative costs, insurance is offered on actuarially fair terms.

We want to use the model to understand whether the extent of employers' ability to commit to providing insurance is an important determinant of coverage level. We assume that employers can provide coverage in two different ways. First, they can provide coverage today without restricting future coverage decisions. Second, they can commit to provide coverage today and in the future. We parametrize the strength of this commitment by assuming that commitment lapses with probability $q^I$ every period.

We allow that setting up coverage is associated with a possibly size-dependent starting cost $b_I(g_e)$. In conjunction with the commitment technology, this implies that an employer can be in one of three insurance provision states, denoted as follows: $C$ indicates commitment to coverage; $Y$ indicates that the establishment has provided coverage in the past but is not committed; $N$ indicates that no coverage has been provided in the preceding period, hence setting up coverage is subject the cost $b_I(g_e)$.

Workers who do not receive coverage through their employer pay out of pocket for their medical expenditures.

Events within a period unfold as follows. New establishments enter. Establishments decide on exit, health insurance, wages, retainment and recruitment of workers. Production takes place. Health shocks are realized and insurance payments are made. Consumption takes place. Some workers exit the labor market and some employed workers separate exogenously. Idiosyncratic shocks to establishment productivity are realized.

Commitment to coverage lapses stochastically. Searching workers are matched with vacancies.

# 4 Stationary Equilibrium

The state of an establishment at the beginning of the period $s = (z, \psi_h, \psi_u, I)$ is given by its productivity $z \in \mathcal{Z}$, its workforce $(\psi_h, \psi_u) \in \Psi$, and its insurance coverage status $I \in \{C, Y, N\}$. Let $\mathcal{S} \equiv \mathcal{Z} \times \Psi \times \{C, Y, N\}$ denote the state space. Let $\mathcal{S}^L \equiv S \cup \{s^L\}$ denote the enlarged state space that includes very large establishments. For $s \in \mathcal{S}$ we write $z(s)$ for establishment productivity in that state, using analogous notation for the other state variables. In a given period an establishment make seven choices, which are collected in a policy vector $g = (g_h, g_u, g_e, g_v, g_I, g_w, g_x)$. The first three entries concern dismissal of incumbent employees. In the first dismissal step, the establishment chooses the number of healthy workers $g_h$ and unhealthy workers $g_u$ to retain. In the second step, the establishment dismisses workers at random, and $g_e$ denotes the total number of workers retained after random dismissal. The fourth choice concerning the size of the workforce is the number of vacancies $g_v$. The next two choices describe the compensation package offered by the establishment. The insurance decision is between three alternatives $g_I \in \{c, y, n\}$, indicating commitment, coverage without commitment, and no coverage, respectively. The decision concerning wages is binary: since the establishment has all the bargaining power it always makes one of the two worker types indifferent between staying and leaving, hence the wage decision reduces to the decision which type to make indifferent $g_w \in \{h, u\}$. The final entry is the decision of the establishment whether to exit $g_x \in \{0, 1\}$, where 1 indicates the choice to exit.

Let $\mathcal{G}$ denote the set of all pure policies

$$\mathcal{G} \equiv \{(g_h, g_u, g_e, g_v, g_I, g_w, g_x) \in \mathbb{N}_0^4 \times \{c, y, n\} \times \{h, u\} \times \{0, 1\}\}.$$

The equilibrium concept allows employers to use mixed policies, i.e. each establishment chooses an element of $\Gamma \equiv \Delta(\mathcal{G})$, the set of all probability distributions over the set $\mathcal{G}$. Having defined the state space and the policy space, next we define a list of objects that make up a stationary equilibrium:

1. An establishment value function $J(\cdot) : \mathcal{S} \to \mathbb{R}$, with $J(s)$ giving the value of an establishment in state $s$.

2. An establishment policy function $\gamma(\cdot|\cdot) : \mathcal{S} \to \Gamma$, where $\gamma(\cdot|s) \in \Gamma$ is the mixed policy for establishments in state $s$, $\Gamma$ is the set of all mixed policies, and $\gamma(g|s)$ is the probability that an establishment in state $s$ implements pure policy $g \in \mathcal{G}$.

3. A policy vector for very large establishments $g^L = \left(g_h^L, g_u^L, g_v^L\right) \in \mathbb{R}_+^3$, specifying healthy employment, unhealthy employment, and vacancies, respectively.

4. A wage function $w(\cdot,\cdot) : \mathcal{S} \times \mathcal{G} \to \mathbb{R}$ with $w(s,g)$ giving the wage an establishment in state $s$ would pay if it were to adopt policy $g$.

5. A correspondence $\mathcal{G}(s) \subset \mathcal{G}$ for $s \in \mathcal{S}$, where $\mathcal{G}(s)$ is the set of policies feasible for an establishment in state $s$.

6. Worker value functions $V_h(\cdot), V_u(\cdot) : \mathcal{S}^L \to \mathbb{R}$, with $V_i(s)$ giving the utility of a worker with health status $i \in \{h, u\}$ employed in an establishment in state $s$.

7. Worker continuation values $C_h(\cdot,\cdot), C_u(\cdot,\cdot) : \mathcal{S} \times \mathcal{G} \to \mathbb{R}$, where $C_i(s,g)$ is the continuation value of a worker with health status $i$ in an establishment in state $s$ pursuing policy $g$.

8. Values of searching $V_h^s, V_u^s \in \mathbb{R}$, for healthy and unhealthy workers, respectively.

9. An invariant distribution $\mu(\cdot) : \mathcal{S}^L \to [0,1]$ of the state of establishments at the beginning of the period.

10. A mass of workers $m \in [0,1]$ searching for employment, and the fraction of searching workers $\nu \in [0,1]$ which is healthy.

11. A mass of active establishments $\phi$.

12. Labor market tightness $\theta \in \mathbb{R}_+$.

In the following subsections we derive the conditions relating these objects, followed by a formal definition of stationary equilibrium.

## 4.1 Establishment Decision

Consider an establishment in state $s = (z, \psi_h, \psi_u, I) \in \mathcal{S}$. The expected flow utility of a worker that remains with the establishment at the time of production, as a function of

wage $w$ and insurance provision $g_I$, is given by

$$\begin{aligned} u_i(w, g_I) &= \iota(g_I)U((1-\tau)w) \\ &\quad + (1-\iota(g_I))\Big[q_{ih}^H U((1-\tau)w - e_h) + q_{iu}^H U((1-\tau)w - e_u)\Big], \end{aligned}$$

where $\iota(g_I) = 1$ if $g_I \in \{c, y\}$ and $\iota(g_I) = 0$ otherwise. The first term is the flow utility of the worker if insurance is provided by the employer. The second term gives expected flow utility of the worker if health expenditures must be paid out of pocket.

The lifetime expected utility at the time of production of a worker with health status $i$ is given by $u_i(w(s,g), g_I) + \beta\rho C_i(s,g)$. This must be equal to $V_{g_w}^s$ for type $g_w$ if the establishment chooses to make type $g_w$ indifferent. Thus the following equilibrium relationship implicitly determines the wage $w(s,g)$:

$$u_{g_w}(w(s,g), g_I) + \beta\rho C_{g_w}(s,g) = V_{g_w}^s. \tag{2}$$

Next we determine the set of pure policies that are feasible for an establishment in state $s$, denoted $\mathcal{G}(s) \subset \mathcal{G}$. For $g$ to be feasible it must be that workers asked to stay must not prefer to leave. The choice of the wage insures that this is true for type $g_w$. For the other type $\neg g_w$ it must be that

$$u_{\neg g_w}(w(s,g), g_I) + \beta\rho C_{\neg g_w}(s,g) \geq V_{\neg g_w}^s \quad \text{if} \quad g_{\neg g_w} > 0. \tag{3}$$

If $g_i < \psi_i(s)$, then the pure policy $g$ also calls on some workers of type $i$ to leave. A leaving worker can induce a deviation from the establishment policy by staying. Let $G_i^+(g)$ be the same pure policy as $g$ except that one additional worker of type $i$ stays:

$$G_h^+(g_h, g_u, g_e, g_v, g_I, g_w, g_x) \equiv (g_h + 1, g_u, g_e + 1, g_v, g_I, g_w, g_x)$$

for healthy workers, with $G_u^+$ defined analogously. For $g$ to be feasible it must be that

$$u_i(w(s,g), g_I) + \beta\rho C_i(s, G_i^+(g)) \leq V_i^s \quad \text{if} \quad g_i < \psi_i(s). \tag{4}$$

Finally, if an establishment is committed to provide coverage it is constrained to set $g_I = c$:

$$g \in \mathcal{G}(z, \psi_h, \psi_u, C) \Rightarrow g_I = c. \tag{5}$$

Thus the set of feasible policies for an establishment in state $s$ is

$$\mathcal{G}(s) = \{g \in \mathcal{G} | g_h \leq \psi_h \wedge g_u \leq \psi_u \wedge (3)\text{–}(5) \text{ hold}\}. \tag{6}$$

It is convenient to define net revenue for active establishments, deducting fixed cost of operating, health insurance premiums, recruiting costs, and applicable costs of starting health insurance coverage from output:

$$R(s,g) \equiv z(s)F(g_e) - c_f - \iota(g_I)p(g)g_e - c_v g_v - b_I(g_e)\iota(g_I)\mathcal{I}(I(s)=N).$$

The insurance premium $p(g)$ is given by

$$p(g) \equiv (1+\kappa(g_e))\frac{g_h(q^H_{hh}e_h + q^H_{hu}e_u) + g_u(q^H_{uh}e_h + q^H_{uu}e_u)}{g_h + g_u},$$

where $\kappa(g_e)$ is the administrative load.[9] The current state $s$, a pure policy $g$, tightness $\theta$, and the fraction of healthy workers among searchers $\nu$, together induce a distribution over the establishment's future state $\mu(s'|s,g,\theta,\nu)$, which is derived in Appendix I.1. The establishment value function $J(\cdot)$ must satisfy the Bellman equation

$$J(s) = \max_{g\in\mathcal{G}(s)} g_x \left[R(s,g) - w(s,g) + \beta\sum_{s'\in\mathcal{S}} J(s')\mu(s'|s,g,\theta,\nu)\right] \quad \text{for all} \quad s\in\mathcal{S}. \tag{7}$$

The policy function $\gamma(\cdot|\cdot)$ must satisfy $\gamma(\cdot|s)\in\Gamma$ for all $s\in\mathcal{S}$ and

$$\gamma(g|s) > 0 \quad \Rightarrow g \in \arg\max_{g\in\mathcal{G}(s)} g_x \left[R(s,g) - w(s,g) + \beta\sum_{s'\in\mathcal{S}} J(s')\mu(s'|s,g,\theta,\nu)\right] \tag{8}$$

for all $g\in\mathcal{G}$ and all $s\in\mathcal{S}$.

## 4.2 Worker Value Functions and Continuation Values

The probability that a worker in health status $i$ is retained if her employer is in state $s$ and pursues policy $g$ is

$$\sigma_i(s,g) \equiv \frac{g_i}{\psi_i(s)}\frac{g_e}{g_h+g_u}.$$

The first factor is the probability of being retained during selective dismissal, and the second factor is the probability of being retained during random dismissal. The worker value functions must then satisfy the equilibrium relationship

$$V_i(s) = \sum_{g\in\mathcal{G}(s)} \gamma(g|s)\left\{\sigma_i(s,g)\left\{u_i(w(s,g),g_I) + \beta\rho C_i(s,g)\right\} + \left[1-\sigma_i(s,g)\right]V_i^s\right\}. \tag{9}$$

[9] The formula implicitly assumes that premiums are paid before random dismissal. Thus $p(g)$ is based on the expected health expenditures given the health composition of the workforce before random dismissal. This is actuarially fair, since in expectation random dismissal does not change the health composition of the workforce. Since employers are risk neutral, nothing would change if premiums are based on the health composition after random dismissal, since expected premium cost associated with providing coverage are the same.

The current state $s$, a pure policy $g$, tightness $\theta$, and the fraction of healthy workers among searchers $\nu$, together induce a distribution over the establishment's future state. It is denoted $\mu_{ii'}\left(s'\,|s,g,\theta,\nu\right)$ where the subscripts indicate that this distribution is conditional on the worker transiting from health status $i$ to $i'$. This distribution is derived in Appendix I.2. Worker continuation values must satisfy the equilibrium relationship

$$\begin{aligned} C_i(s,g) = (1-\delta)\Big\{ & q_{ih}^H \sum V_h(s')\mu_{ih}\left(s'\,|s,g,\theta,\nu\right) \\ & + q_{iu}^H \sum V_u(s')\mu_{iu}\left(s'\,|s,g,\theta,\nu\right) \Big\} + \delta\left[q_{ih}^H V_h^s + q_{iu}^H V_u^s\right]. \end{aligned} \tag{10}$$

## 4.3 Value of Searching

The flow utility of a searching worker with health status $i$ is

$$u_i^s = q_{ih}^H U(b-e_h) + q_{iu}^H U(b-e_u),$$

where $b$ is the flow value of non-market activity, and searchers pay health expenditures out of pocket. The value of searching of a worker in health status $i$ must satisfy the equilibrium relationship

$$\begin{aligned} V_i^s = {} & u_i^s + \beta\rho\left(1-f(\theta)\right)\left[q_{ih}^H V_h^s + q_{iu}^H V_u^s\right] \\ & + \beta\rho f(\theta)\Big\{ q_{ih}^H \sum V_h(s')\mu_h^s\left[s'\,|\mu(\cdot),\gamma(\cdot|\cdot),\theta,\nu\right] + q_{iu}^H \sum V_u(s')\mu_u^s\left[s'\,|\mu(\cdot),\gamma(\cdot|\cdot),\theta,\nu\right] \Big\} \end{aligned} \tag{11}$$

The second part of the continuation utility corresponds to contacting an establishment, while the first part applies in the absence of a contact. Here $\mu_i^s\left(s'\,|\mu(\cdot),\gamma(\cdot|\cdot),\theta,\nu\right)$ for $i \in \{h,u\}$ is the distribution of the state of the worker's new establishment conditional on the worker having new health status $i$. It is derived in Appendix I.3.

## 4.4 Invariant Distribution of Establishments

In a stationary equilibrium the number of establishments must remain constant, thus entry must compensate for exit. This implies that we can compute the invariant distribution by taking the transition law $\mu(s'|s,g,\theta,\nu)$, modifying it by letting exiting establishments start over as new entrants. Using superscript $x$ to denote application of the operator that performs this modification, the invariant distribution $\mu(s)$ must satisfy the equilibrium relationship

$$\mu(s') = \sum_{s\in\mathcal{S}} \sum_{g\in\mathcal{G}(s)} \mu^x(s'|s,g,\theta,\nu)\gamma(g|s)\mu(s). \tag{12}$$

## 4.5 Mass of Searchers

Let $\mu_i$ denote the fraction of workers with health status $i$ in the stationary health status distribution. The mass of searchers in health status $i$ can be computed by deducting from $\mu_i$ the mass of workers which remain employed by an establishment at the time of matching:

$$m_{i'}[\gamma(\cdot|\cdot), g^L, \mu(\cdot)] \equiv \mu_{i'} - \rho\phi \sum_{i\in\{h,u\}} q^H_{ii'} \left[ \sum_{s\in\mathcal{S}} \sum_{g\in\mathcal{G}} \sigma_i(s,g)\psi_i(s)\gamma(g|s)\mu(s) + g^L_i \mu(s^L) \right]$$

for $i' \in \{h, u\}$. This yields the equilibrium relationships

$$\begin{aligned} m &= m_h[\gamma(\cdot|\cdot), g^L, \mu(\cdot)] + m_u[\gamma(\cdot|\cdot), g^L, \mu(\cdot)], \\ \nu &= \frac{m_h[\gamma(\cdot|\cdot), g^L, \mu(\cdot)]}{m}. \end{aligned} \tag{13}$$

## 4.6 Tightness

By definition tightness is the ratio of the number of vacancies and the mass of searcher, giving rise to the equilibrium relationship

$$\theta = \phi \frac{\sum_{s\in\mathcal{S}} \sum_{g\in\mathcal{G}} g_v \gamma(g|s)\mu(s) + g^L_v \mu(s^L)}{m}. \tag{14}$$

## 4.7 Very Large Establishments

There is a fixed mass $\lambda^L$ of infinitely-lived, very large establishments. They have the same production function as other establishments, with constant productivity $z^L$. Each employs a continuum of workers. We assume that they always provide insurance. Net revenue is

$$R^L(g_h, g_u, g_v) \equiv z^L F(g_h + g_u) - c_f - p(g)(g_h + g_u) - c_v g_v.$$

Healthy workers are made indifferent between staying and leaving, which implies a wage $w^L(V^s_h, V^s_u)$. The Bellman equation is

$$J^L(g_h, g_u) = \max_{g_v} \left\{ R^L(g_h, g_u, g_v) - w^L(V^s_h, V^s_u)(g_h + g_u) + \beta J^L \left[ \mu^L(g|\theta, \nu) \right] \right\}$$

where

$$\begin{aligned} \mu^L_h(g|\theta,\nu) &\equiv \rho(1-\delta) \left[ q^H_{hh}\psi_h + q^H_{uh}\psi_u \right] + q(\theta)\nu g_v, \\ \mu^L_u(g|\theta,\nu) &\equiv \rho(1-\delta) \left[ q^H_{hu}\psi_h + q^H_{uu}\psi_u \right] + q(\theta)(1-\nu) g_v. \end{aligned}$$

Since they are not subject to shocks, these establishments are always in steady state. Let $g^L(\theta, \nu, V_h^s, V_u^s) \in \mathbb{R}_+^3$ denote optimal steady state levels of employment and vacancies. Then we obtain the equilibrium relationship

$$g^L = g^L(\theta, \nu, V_h^s, V_u^s). \tag{15}$$

## 4.8 Entry of New Establishments

Let $\mu^Z$ denote the invariant distribution induced by the transition probabilities $q_{zz'}^Z$. The entry condition is

$$c_e = \sum_{z \in \mathcal{Z}} J(z, 0, 0, N)\mu^Z(z). \tag{16}$$

The first term reflects that a new establishment has zero employment and no past insurance coverage.

## 4.9 Definition of Equilibrium

**Definition 1** *$J(\cdot)$, $\gamma(\cdot,\cdot)$, $g^L$, $w(\cdot,\cdot)$, $\mathcal{G}(\cdot)$, $V_h(\cdot)$, $V_u(\cdot)$, $C_h(\cdot,\cdot)$, $C_u(\cdot,\cdot)$, $V_h^s$, $V_u^s$, $\mu(\cdot)$, $m$, $\nu$, $\phi$, and $\theta$ constitute a stationary equilibrium if they satisfy equations* (2), (6), (7), (8), (9), (10), (11), (16), (13), (14), *and* (15).

An algorithm for computing an equilibrium is outlined in Appendix II.

# 5 Quantitative Analysis

## 5.1 Preliminary Calibration

### 5.1.1 Functional Forms

To conduct a quantitative analysis we must choose functional forms for the utility, production, and matching functions and assign parameter values.

**Utility Function.** We use a constant absolute risk aversion utility (CARA) function[10]

$$U(c_t) = -\frac{\exp(-\varsigma c_t) - 1}{\varsigma}. \tag{17}$$

**Production Function.** Output of an establishment with productivity $z$ is given by

$$zF(g_e) = zg_e^\eta$$

[10]We use CARA instead of the more standard CRRA utility function since it facilitates some steps in the computation.

where $g_e$ is the number of workers. We assume that $\eta \in (0,1)$ so that the production function exhibits decreasing returns to scale and satisfies the usual Inada conditions.[11] The parameter $z$ varies across establishments and across time generating cross-sectional and time-series variation in establishment productivity.

Following Hopenhayn and Rogerson (1993), we assume that productivity shocks evolve according to the process

$$\ln(z') = \zeta(1-\varphi) + \varphi \ln(z) + \epsilon', \tag{18}$$

where $0 \geq \varphi < 1$, $\zeta \geq 0$ and $\epsilon' \sim N(0, \sigma_\epsilon^2)$. We denote the transition function for $z$ as $Q(z, dz')$. This process has a parsimonious representation with the parameters corresponding to objects that are of intuitive interest given the nature of the employer provided health insurance system that we study. For example, the volatility and persistence of this process have important impact in the variability of insurance provision at the establishment level. More complicated process can easily be incorporated into the analysis, but this one appears a reasonable first pass and is a standard process in the establishment dynamics literature.

**Matching Function.** For comparability with much of the literature we choose the Cobb-Douglas functional form of the matching function between workers and employers:

$$M(m, v) = \chi m^\alpha v^{1-\alpha}. \tag{19}$$

Thus, the are two parameters, $\chi$, $\alpha$, that characterize the matching function.

### 5.1.2 Parameters

The model parameters to be calibrated are:

1. $\beta$ – the time discount rate,
2. $\rho$ – the probability of a worker leaving the labor market,
3. $\varsigma$ – coefficient of absolute risk aversion,
4. $\eta$ – curvature of the production function,
5. $\zeta$ – unconditional mean of idiosyncratic productivity,

[11] Our model is a single-good model in which a non-degenerate distribution of establishment sizes is sustained by decreasing returns at the establishment level. An alternative framework is to assume differentiated products and constant returns at the establishment level. In this alternative framework, the nondegenerate distribution of establishment sizes is sustained by curvature in preferences. As discussed in, e.g. Restuccia and Rogerson (2008), conceptually these frameworks are very similar.

6. $\varphi$ – persistence of idiosyncratic productivity,
7. $\sigma_\epsilon$ – st. dev. of innovations in idiosyncratic productivity,
8. $\lambda^L$ – mass of very large establishments,
9. $z^L$ – productivity of very large establishments,
10. $\chi$ – scale parameter of the matching function,
11. $\alpha$ – elasticity of the matching function,
12. $c_f$ – establishments' flow cost of operating,
13. $c$ – cost of maintaining a vacancy,
14. $b$ – value of non-market activity,
15. $\tau$ – proportional tax on labor income,
16. $q_I$ – probability of commitment lapsing,
17. $\kappa(g_e)$ – administrative load schedule for employer provided insurance,
18. $b_I(g_e)$ – cost of initiating coverage,
19. $q^H$ – the health status transition matrix,
20. $e_u$, $e_h$ – health expenditures.

We choose the model period to be one month in order to accommodate the high frequency of transitions in the labor market. The data used to compute some of the targets have monthly, quarterly or annual frequency, and we aggregate the model-generated data appropriately when matching those targets.

The empirical counterpart of very large establishment in the model are the establishments employing more than 250 workers. Such establishments represent only 0.7% of all establishments but account for 30% of employment. The average size of an establishment with 250 workers or less is 11.03. The average size of an establishment with more than 250 workers is 679.37. We use the latter statistic to determine the productivity level of the very large establishment in equilibrium. The number of very large establishments is chosen so that they represent 0.7% of all establishments.

We choose a relatively low coefficient of absolute risk aversion of $\varsigma = 0.5$ since the model does not include other channels of insurance beyond health insurance. We choose $\rho = 0.9979$ to generate an expected working lifetime of 40 years. We set $\beta = 1/(1 + r)$, where $r$ corresponds to an annual interest rate of 4%.

The extent of decreasing returns in the establishment-level production function is an important parameter in our analysis. As described in Restuccia and Rogerson (2008), direct estimates of establishment-level production functions and different calibration procedures point to a value for $\eta = 0.85$.

Hagedorn and Manovskii (2008) report the average monthly job finding rate of 0.45, and the average value for labor market tightness $\theta = 0.634$. They also estimate that the average flow cost of posting a vacancy equals 58% of the average labor productivity, i.e., $c = 0.58p$. Labor productivity $p$ is defined as output per worker. We target these values. Since the average flow cost of posting a vacancy is a function of the equilibrium output per worker, its exact value is pinned down only once the equilibrium size distribution of establishments is determined.

The matching function elasticity parameter $\alpha$ is selected to match the elasticity of the job-finding probability with respect to labor market tightness. Petrongolo and Pissarides (2001) survey the empirical evidence and conclude that the value of $\alpha = 0.5$ for the elasticity of the job-finding rate with respect to labor market tightness is appropriate. (See also Brügemann (2008).) The value of the matching function efficiency parameter, $\chi$, is chosen to match the data the average values for the job-finding rate.

There is only limited evidence on administrative costs (marketing, billing, employee enrollment and education, payments to benefit consultants and insurance sales agents, risk charges, underwriting, etc.) and how they vary with establishment size. A study by the Congressional Research Service (1988) reports that administrative costs represent 8% of premiums on average, but up to 40% for small establishments. This study estimates costs for several size categories. We set $\kappa(g_e = 1) = 0.4$ and approximate the loads for larger establishments in larger size classes using the continuous schedule: $\kappa(g_e) = 0.37 - 0.07(g_e - 2.0)^{0.3}$ for $g_e > 1$.

For the benchmark calibration we adopt a constant schedule for the cost of initiating coverage $b_I(g_e) = b_I$. Higher cost of initiating coverage lead to more stable coverage decisions, so we choose to identify $b_I$ by matching the observed rate of discontinuance. We choose the probability of commitment lapsing $q^I$ to match the fraction of establishments providing insurance.

We use estimates of the average marginal tax on labor provided by Lucas (1990) and Prescott (2004) and set the proportional tax rate on labor $\tau = 0.4$.[12]

To calibrate the health expenditure process we use data from the Medical Expenditure Panel Survey (MEPS). The MEPS is based on a series of national surveys conducted by the U.S. Agency for Health Care Research and Quality (AHRQ). In the model idiosyncratic health shocks follow a two state Markov process: high $h$ or low $u$ with transition

[12]Mendoza, Razin, and Tesar (1994) report a lower value $\tau = 0.3$ for the average labor tax rate.

probabilities $q^H_{ii'}$ for $i, i' \in \{h, u\}$. In the data we identify health status with medical expenditures $e_u > e_h$. We chose the parameters of the monthly health shock process so that after it is aggregated to an annual frequency, it matches the mean, median, variance, skewness, and kurtosis of the distribution of health expenditures in the MEPS data. We find $e_u = .5555$, $e_h = .0158$, where health expenditures are expressed as a function of the mean wage in the economy. The implied transition matrix is given by $q^H_{hh} = 0.9892$, $q^H_{hu} = 0.0108$, $q^H_{uh} = 0.1636$, and $q^H_{uu} = 0.8364$. We assume that all the new labor market entrants are healthy so that $q^0_h = 1$ and $q^0_u = 0$.

The mass of operating establishments is pinned down by the free entry condition. All active employers must pay a flow cost $c_f$ in each period of operation. If an establishment does not pay the fixed cost in any period then it ceases to exist. The parameter $c_f$ is important for evaluating the effects of alternative health insurance policies. For example, if $c_f > 0$ potential health insurance policies that subsidize smaller and less efficient establishments may induce selection that lowers the aggregate productivity.

Remaining Parameters. Five parameters remain to be determined: the values of non-market activity, $b$, the entry cost of establishments $c_e$, the unconditional mean $\zeta$ and persistence $\varphi$ of idiosyncratic productivity shocks process, and the variance of its innovations $\sigma_\epsilon$. We choose the values for these parameters to match the data on the average value for labor market tightness described above, the mean and median sizes of establishments with 250 workers or less equal to 4 and 11.03, respectively,[13] the job creation rate of continuing establishments with 250 workers or less of 7% that we compute from the Business Employment Dynamics (BED) statistics provided by the BLS, and to generate the normalized value of wages equal to one. Thus, there are five targets to pin down five parameters. To find the values of these parameters we solve the model numerically according to the computational algorithm described in Appendix II.

The calibrated parameter values are summarized in Table 1. The performance of the calibrated model in matching the calibration targets is described in Table 2.

[13]These numbers are computed from the RWJ sample. All the establishment size distribution statistics from RWJ sample are nearly identical to those from the 1997 US Economic Census reported in Guner, Ventura, and Yi (2008).

Table 1: Calibrated Parameter Values.

| Parameter | Definition | Value |
|---|---|---|
| $\beta$ | the time discount rate | 0.996 |
| $\varsigma$ | coefficient of absolute risk aversion | 0.5 |
| $\rho$ | the probability of a worker staying in the labor market | 0.9979 |
| $\eta$ | curvature of the production function | 0.85 |
| $\zeta$ | unconditional mean of idiosyncratic productivity | 1.5385 |
| $\varphi$ | persistence of idiosyncratic productivity | 0.9963 |
| $\sigma_\epsilon$ | st. dev. of innovations in idiosyncratic productivity | 0.0207 |
| $\lambda^L$ | mass of very large establishments | 0.007 |
| $z^L$ | productivity of very large establishments | 3.34 |
| $\chi$ | scale parameter of the matching function | 0.6364 |
| $\alpha$ | elasticity of the matching function | 0.5 |
| $c_f$ | establishments' flow cost of operating | 0.0 |
| $c$ | cost of maintaining a vacancy | $0.58p$ |
| $b$ | value of non-market activity | 0.6247 |
| $\tau$ | proportional tax on labor income | 0.4 |
| $\kappa(g_e)$ | administrative load for employer provided insurance | $40\%^a$ |
| $b_I(g_e)$ | cost of initiating coverage | 1.5 |
| $q^I$ | probability of commitment lapsing | 0.0278 |
| $e_u$ | health expenditures of unhealthy workers | 0.5555 |
| $e_h$ | health expenditures of healthy workers | 0.0158 |
| $q^H_{hh}$ | the health status transition probability | 0.9892 |
| $q^H_{uu}$ | the health status transition probability | 0.8364 |
| $q^0_h$ | fraction of new entrants who are healthy | 1.00 |

Note - The table contains the calibrated parameter values in the benchmark calibration. [a] The value of 40% represents the administrative load for a one-worker establishment. The loads for larger establishments are given by the following schedule: $0.37 - 0.07(g_e - 2.0)^{0.3}$.

Table 2: Matching the Calibration Targets.

| | Target | Data | Model |
|---|---|---|---|
| 1. | Fraction of establishments that employ > 250 workers | 0.007 | 0.007 |
| 2. | Mean size of an establishment with > 250 workers | 679.37 | 680.6 |
| 3. | Mean size of an establishment with <= 250 workers | 11.03 | 11.06 |
| 4. | Median size of an establishment with <= 250 workers | 4 | 4 |
| 5. | Job creation rate of establishment with <= 250 workers | 0.07 | 0.07 |
| 6. | Labor market tightness | 0.50[a] | 0.50 |
| 7. | Job finding rate | 0.45 | 0.45 |
| 8. | Fraction of establishments providing health insurance | 0.56 | 0.46 |
| 9. | Fraction of establishments discontinuing insurance over 2 years | 0.11 | 0.14 |
| 10. | Average wage (normalization) | 1.00 | 1.01 |

Note - The table describes the performance of the model in matching the calibration targets.

[a] We targeted the value of labor market tightness of 0.5 by mistake. The correct value to target is 0.634 and we will target it in the next draft. We do not expect this to affect the findings in a substantively important way.

Table 3: Benchmark Results and Experiments

| | Benchmark | Experiments | | | | | |
|---|---|---|---|---|---|---|---|
| | (I) | (II.1) | (II.2) | (II.3) | (II.4) | (II.5) | (II.6) |
| *Fraction of Establishments Providing Coverage* | | | | | | | |
| 3-10 | 0.138 | 1.000 | 0.875 | 0.916 | 0.905 | 0.073 | 0.998 |
| 11-25 | 0.705 | 1.000 | 0.767 | 0.888 | 0.801 | 0.460 | 0.995 |
| 26-50 | 0.940 | 1.000 | 0.604 | 0.921 | 0.827 | 0.889 | 0.983 |
| 51- | 0.932 | 1.000 | 0.564 | 0.930 | 0.884 | 0.932 | 0.999 |
| all | 0.461 | 1.000 | 0.785 | 0.911 | 0.866 | 0.356 | 0.995 |
| *Fraction of Establishments Providing Coverage that Commit* | | | | | | | |
| 3-10 | 0.838 | 0.369 | 0.000 | 0.066 | 0.020 | 0.000 | 0.063 |
| 11-25 | 0.972 | 0.688 | 0.000 | 0.507 | 0.158 | 0.000 | 0.273 |
| 26-50 | 0.987 | 0.858 | 0.000 | 0.917 | 0.652 | 0.000 | 0.985 |
| 51- | 0.997 | 0.931 | 0.000 | 0.997 | 0.997 | 0.000 | 0.998 |
| all | 0.960 | 0.561 | 0.000 | 0.374 | 0.223 | 0.000 | 0.329 |
| *Fraction of Establishments Discontinuing Insurance over a 2 Year Period* | | | | | | | |
| 3-10 | 0.346 | 0.000 | 0.125 | 0.085 | 0.096 | 0.625 | 0.000 |
| 11-25 | 0.106 | 0.000 | 0.233 | 0.112 | 0.199 | 0.075 | 0.002 |
| 26-50 | 0.056 | 0.000 | 0.396 | 0.079 | 0.173 | 0.000 | 0.015 |
| 51- | 0.067 | 0.000 | 0.436 | 0.070 | 0.116 | 0.000 | 0.001 |
| all | 0.136 | 0.000 | 0.216 | 0.090 | 0.134 | 0.158 | 0.003 |
| *Establishment Size Distribution* | | | | | | | |
| 3-10 | 0.322 | 0.348 | 0.314 | 0.311 | 0.321 | 0.312 | 0.312 |
| 11-25 | 0.154 | 0.153 | 0.146 | 0.147 | 0.153 | 0.150 | 0.156 |
| 26-50 | 0.084 | 0.087 | 0.084 | 0.083 | 0.084 | 0.078 | 0.087 |
| 51- | 0.056 | 0.059 | 0.048 | 0.051 | 0.054 | 0.054 | 0.052 |
| mean | 16.28 | 16.99 | 15.29 | 15.39 | 15.96 | 15.66 | 16.14 |
| median | 4 | 5 | 4 | 4 | 4 | 4 | 4 |
| *Fraction of Unhealthy Workers among all Employees* | | | | | | | |
| 3-10 | 0.009 | 0.065 | 0.000 | 0.005 | 0.002 | 0.001 | 0.061 |
| 11-25 | 0.047 | 0.063 | 0.001 | 0.034 | 0.012 | 0.037 | 0.068 |
| 26-50 | 0.062 | 0.057 | 0.011 | 0.058 | 0.041 | 0.066 | 0.062 |
| 51- | 0.073 | 0.068 | 0.076 | 0.077 | 0.081 | 0.078 | 0.066 |
| Avg. Wage | 1.006 | 0.960 | 1.016 | 1.001 | 1.004 | 1.010 | 0.996 |
| Avg. Labor Cost | 1.048 | 1.012 | 1.035 | 1.034 | 1.030 | 1.050 | 1.032 |
| Output per Worker | 1.287 | 1.286 | 1.294 | 1.291 | 1.291 | 1.288 | 1.283 |
| $m$ | 0.074 | 0.031 | 0.085 | 0.067 | 0.076 | 0.077 | 0.065 |
| $\nu$ | 0.855 | 0.925 | 0.715 | 0.829 | 0.790 | 0.862 | 0.944 |
| $\theta$ | 0.500 | 2.800 | 0.580 | 0.650 | 0.581 | 0.450 | 0.545 |

Notes to Table 3 -

The table contains the results of quantitative experiments using the calibrated model.

Description of Columns:

Column (I) − Results from benchmark calibration of the model.

Column (II.1) − Medical expenditures are assumed to be the same for the healthy and unhealthy workers.

Column (II.2) − No commitment, uniformly low loads, no cost of starting.

Column (II.3) − Commitment, uniformly low loads, no cost of starting.

Column (II.4) − Commitment, benchmark load schedule, no cost of starting.

Column (II.5) − No commitment, benchmark load schedule, high cost of starting.

Column (II.6) − Community rated insurance for establishments of all sizes where insurers are not permitted to adjust rates based on the health composition of the employer.

Description of Rows:

$m$ − The mass of searchers.

$\nu$ − Fraction of healthy workers among searchers.

$\theta$ − Labor Market tightness.

## 5.2 Benchmark Results

The results of the benchmark calibration are presented in Column (1) of Table 3. The model qualitatively matches the pattern of coverage provision with establishment size although the slope is somewhat steeper than in the data. It also qualitatively captures the decline in he probability of discontinuing coverage which establishment size. In the next subsection we explore which ingredients of the model are important for the ability of the model to capture these features of the data.

## 5.3 Properties of the Model

Without non-discrimination provisions 100% of employed workers would be covered in the model and there would be no discontinuance of coverage. The reasons for this are clear. Employees have to pay for their medical expenditures. Without non-discrimination provisions, employed workers would effectively have individual insurance contracts the cost of which is equal to the expected medical expenditures of that employee. However, if they pay out of pocket (or through purchasing individual insurance), the payment would be made with after-tax income. Instead, if the employer pays for these expenses, the payment is made with before-tax dollars. Clearly, workers and employers will take advantage of the tax subsidy and employers will contract on behalf of each employee. Of course, such a system will result in no risk pooling across individuals.

With the non-discrimination provisions in place, employers face a trade-off between the tax subsidy for healthy workers vs. rents for the unhealthy. Effectively, there is a threshold level of the fraction of unhealthy workers in an establishment, $\frac{\psi_u}{\psi_h}$, such that when it is exceeded medical expenditures on the unhealthy outweigh the wage savings associated with providing coverage. In this event the establishment discontinues coverage unless it is currently committed to providing coverage. Thus fluctuations in the health composition induce transitions into and out of providing coverage, and we refer to this as the fragility channel.

To illustrate the role of various economic forces in the model we perform a sequence of experiments the results of which are described in Columns (II.1) through (II.6) of Table 3. When performing these experiments we hold the values of all parameters not directly changed in the experiment fixed at their levels in benchmark calibration.

### 5.3.1 Uniform Medical Expenditures

First, we assume that medical expenditures are the same for the healthy and unhealthy workers and report results in Column (II.1). This experiment effectively eliminates the role of the fragility channel and the distortions associated with it. Since non-discrimination provisions have no meaning in this set-up, just as we discussed above, 100% of employed workers would be covered by employer-provided insurance to take advantage of the tax subsidy and there would be no discontinuance of coverage. Since this experiments shuts down shocks to the health composition of an establishment's workforce, it allows us to assess the impact of the interaction of these shocks with non-discrimination provisions on the establishment size distribution. Comparing the distributions in the benchmark calibration to the one in this experiment we find that they are surprisingly similar. This is surprising because one might expect the benchmark distribution to be shifted in favor of larger establishments. The reason is that smaller establishments are less able to provide coverage and to benefit from the tax subsidies that come along with it. This makes them relatively less competitive. This effect is counteracted by the fact that because small establishments in the benchmark tend to not provide insurance, unhealthy workers tend to select into larger establishments, making labor more expensive for them. For example, the fraction of unhealthy workers in establishments with more than 50 workers is 7.3% in the benchmark calibration compared to 6.8% with uniform expenditures. Similarly, the fraction of healthy workers among searchers is 85.5% in the benchmark and 92.5% with uniform expenditures.

Another interesting effect is evident from the fact that unemployment is as low as 3.1% in the economy where medical expenditures are the same for the healthy and unhealthy workers and is 7.4% in benchmark economy. The reason for this effect is that average wages and average labor cost in the benchmark calibration are higher than in the experiment of equalizing medical expenditures. This is because non-discrimination provisions raise the level of compensation. In the absence of such provisions, or when they are immaterial as in the case of equal medical expenditures, the wage setting mechanism implies that workers of both types never receive more than their reservation wage. In search equilibrium this implies the Diamond Paradox, with workers receiving the same payoff as from not participating in the labor market at all. Legal non-discrimination provisions break this implication. The logic is similar to Albrecht and Axell (1984), where employers cannot discriminate between workers with different valuation of leisure, as the

latter is private information of workers. Workers with different reservation wages are paid the same, thus workers with low reservation wages receive rents. A slight difference of our model relative to Albrecht and Axell's is that here workers stochastically switch between types, hence non-discrimination provisions raise the reservation wage of both types. The experiment of equalizing medical expenditures shows that the heterogeneity of medical expenditures in the benchmark calibration is sufficiently large to generate a sizable effect on the overall wage level.

The next four experiments assesses the role of employers' ability to commit to providing insurance, of the extent to which administrative costs of maintaining insurance vary with establishment size, and of the costs of starting insurance in shaping aggregate outcomes.

### 5.3.2 No Commitment, Uniformly Low Loads, No Cost of Starting.

In this experiment we shut down the ability of establishments to commit to provide insurance coverage ($q^I = 1$). The only choice available to an establishment is whether to provide coverage today or not with no ability to make credible promises about provision in the future. In addition we assume that there are no costs associated with initiation of coverage ($b_I = 0$). Finally, we assume that the administrative load is independent of the size of the establishment and is set at a relatively low level faced by large establishments in benchmark calibration ($\kappa(g_e) = 0.09$). The results of this experiment are described in Column (II.2) of Table 3.

This experiment results in two striking outcomes. First, insurance coverage *declines* in establishment size. Second, the fraction of establishments discontinuing coverage over a two year period *increases* in establishment size. These patterns are opposite of those found in the data.

The reason for these outcomes is as follows. The data imply that the probability of a healthy worker to become unhealthy and face substantial medical expenditures is relatively low. This implies that small establishments with healthy workers only rarely have an unhealthy worker. Thus, most of the time they are able to provide coverage that provides tax advantaged coverage for routine medical expenditures of healthy workers. When even a single worker becomes unhealthy, small establishments discontinue coverage which induces the unhealthy worker to leave the employer. Once the unhealthy worker leaves, the coverage is reinstated (reinstating coverage is cheap in this experiment, we

explore its effects below). In equilibrium of this experiment 87.5% of establishments with 10 workers or less are providing coverage. At the same time, they employ no unhealthy workers. This shows that the popular coverage statistics are not informative on their own. Coverage rate might be high but there might be virtually no insurance provided against large medical expenditures.

In large establishments the situation is different. They are willing to provide insurance as long as the fraction of unhealthy workers does not exceed a certain threshold. Since employers cannot commit to provide insurance in the future this threshold is determined by the static trade-off between tax savings on the healthy and rents conceded to the unhealthy. It turns out that for the calibrated medical expenditure process this threshold lies below the fraction of unhealthy in the population. When below the threshold the establishment provides coverage and the fraction of unhealthy trends up. When above the threshold coverage is discontinued and some unhealthy workers leave, but only up to the point where it becomes sufficiently likely that coverage will restart. While small establishments can maintain a healthy workforce throughout, the fraction of unhealthy in large establishments stays close to the threshold. The result of these dynamics is that large establishments provide only about half the time.

### 5.3.3 Commitment, Uniformly Low Loads, No Cost of Starting.

Next we reintroduce the ability of employers to commit to provide coverage in the future into the environment studied in the previous subsection. We maintain the assumptions that there are no costs associated with initiation of coverage ($b_I = 0$) and that the administrative load is independent of the size of the establishment and is set at a relatively low level faced by large establishments in the benchmark calibration ($\kappa(g_e) = 0.09$). The results of this experiment are described in Column (II.3) of Table 3.

Performing this experiment we find that coverage is no longer declining in size. Instead, the coverage rate is high and virtually independent of the establishment size. Given the ability to commit we observe that large establishments choose to do so while small employers prefer to retain the flexibility of the provision decision. Indeed, 99.7% of establishments with more than 50 workers choose to commit when they provide coverage. Only 6.6% of establishments with 10 workers or less choose to do so.

The reason for this outcome is as follows. Large establishments that hover around the provision threshold chose to commit when given the opportunity to do so. Commitment

is profitable for them. Healthy workers have a chance of becoming unhealthy. Being risk averse, they are willing to take wage concessions if the establishment commits to covering their medical expenditures should they become unhealthy. The cost of commitment is that moderately sized establishments will tend to accumulate unhealthy workers. Consequently, they will discontinue coverage temporarily at the expiration of the commitment period to induce some of the unhealthy workers to leave the establishment.

For small establishments, on the other hand, commitment is not profitable. They obtain the same benefit from committing as large establishments, in form of a wage concession from their healthy workers. But they have an alternative which is not available to large establishments, namely maintaining a health workforce throughout by discontinuing coverage whenever there is an unhealthy worker. This alternative remains superior even if the option to commit to coverage is available.

#### 5.3.4 Commitment, Benchmark Load Schedule, No Cost of Starting.

Next we reintroduce the administrative load schedule from benchmark calibration into the environment studied in the previous subsection. We maintain the assumptions that there are no costs associated with initiation of coverage ($b_I = 0$). The results of this experiment are described in Column (II.4) of Table 3.

This experiment has relatively limited effects. Since the cost of providing coverage has increased for all but the largest establishments, the share of establishments providing coverage declines from 91.1% to 86.6%. The most pronounced effect is on the decision of employers to commit to providing coverage. Higher administrative loads make it costlier to provide insurance, especially to unhealthy workers. This induces more establishments to retain coverage flexibility. The share of establishments providing insurance that commits drops from 50.7% to 15.8% among establishments with 11 to 25 workers and from 91.7% to 65.2% among establishments with 26 to 50 workers. A flip side of the decision to retain flexibility is that employers do exercise this option more frequently and push out unhealthy workers. Thus, the fraction of unhealthy workers employed in small establishments drop substantially.

The only difference between this experiment and the benchmark is the absence of cost of starting. Maintaining a healthy workforce throughout is associated with discontinuing coverage whenever an unhealthy worker is present. The main effect of a fixed cost of starting is to make this strategy costly for small establishments. They respond by either

not providing coverage at all, or by committing to coverage.

### 5.3.5 No Commitment, Benchmark Load Schedule, High Cost of Starting.

An alternative modeling choice is to not to allow the choice of whether to commit to provide coverage but instead to calibrate the cost of starting coverage to match the average fraction of establishments providing coverage. Specifically, we now assume that the cost of starting coverage is proportional to establishment size $b_I(g_e) \equiv b_I g_e$ and instead of $q^I$ (now set to one) we use the slope parameter $b_I$ to target the level of coverage. This yields $b_I = 1.1$, that is a cost of starting of slightly more than one month of wages per employee. The results of this experiment are reported in Column (II.5) of Table 3. With this high cost of starting the economy without the commitment choice is quite similar to the economy where the choice of whether to commit is allowed. This is because a high cost of starting acts very similarly to commitment, in that it makes the employer reluctant to discontinue coverage. Similar results are implied by the high cost of stopping coverage.

While one can choose the cost of starting to match a number of statistics we are interested in, it does lead to some counterfactual predictions. To generate the coverage rates that increase in establishment size and probability of discontinuance that is decreasing in establishment size, the cost of starting has to increase with establishment size. This is captured by the specification we adopted in which the cost of starting is directly proportional to establishment's employment level. This specification implies, however, that establishments start coverage only when they are small. Indeed, an establishment with, say, 10 healthy workers will find it optimal to fire all these workers, start coverage, and post a lot of vacancies. Larger establishments that do not provide coverage never start providing it. This is in contrast to the data where the probability of starting coverage conditional on not providing in a given period is increasing in establishment size and is quite high for large establishments. The model in which employers are allowed to choose whether to commit to coverage decisions is consistent with these dimensions of the data and thus appears to be a preferred modeling choice.

While calibrating the cost of starting in the model without commitment appears an inferior modeling choice for matching the observed patterns in the data, experimenting with the cost of starting provides interesting insights into how the model works and suggests potentially interesting policy tools. In Appendix Table A-1 we describe the

behavior of the model economy as we vary $b_I$ from 0.0 to 1.4 (maintaining the proportional specification $b_I(g_e) \equiv b_I g_e$) and do not allow establishments to commit to future coverage provision.

The case $b_I = 0.0$ is identical to the no-commitment experiment discussed in section 5.3.2, except that here administrative loads are as in the benchmark calibration. The results are very similar, and 78% of establishments provide coverage. This high coverage rate masks the fact that a large fraction of unhealthy workers is unable to obtain insurance. The reason is that small employers drop coverage whenever they have an unhealthy worker and restart coverage when this worker leaves. The mass of unhealthy workers in the economy is close to 0.06. The mass of unemployed unhealthy workers in this experiment is 0.027, implying that almost half of all unhealthy workers are unemployed and do not receive employer-provided health insurance. Excluding very large establishments that by assumption always provide coverage, only 52.9% of employed unhealthy workers are covered by employer-provided insurance. Moreover, only 28.2% of employed unhealthy workers are in establishments that also employ healthy workers. In these establishments the employer sets wages to leave healthy workers indifferent between staying and leaving. Here unhealthy workers receive rents, as non-discrimination provisions insure that an adverse health-expenditure shock does not result in a commensurate drop in the wage. The remaining covered unhealthy work in establishments that exclusively employ unhealthy workers. Here non-discrimination provisions have no bite, and unhealthy workers receive no rents as employers make them indifferent between staying and leaving.

As the cost of starting coverage increases to, say, $b_I = 0.4$ so that it equals 40% of the monthly wage bill of the establishment, the coverage rate declines to only 18.7% of establishments. However, the measure of unemployed unhealthy workers declines to 0.015 so that now about 75% of unhealthy workers are employed. Moreover, 99.9% of employed unhealthy workers receive employer-provided health insurance and 97.8% of employed unhealthy workers receive rents. Thus, despite the overall decline in the coverage rate both healthy and unhealthy workers are better off in the economy with a higher cost of coverage. Increasing the cost of starting coverage up to about 1.4 keeps increasing the effective amount of insurance provision and makes workers even better off. Increasing the cost of coverage beyond 1.4 has no effect.

### 5.3.6 Community Rating.

In Column (II.6) we introduce complete community rating into the benchmark economy. Complete community rating means that all employers are charged exactly the same premiums irrespective of their size and health composition. The community-rated premium in this experiment pools both health expenditures and administrative loads across employers. We solve for the lowest level of the premium that permits the system to break even.[14]

The result of this experiment is virtually complete coverage. To shed light on what features are important for this outcome, we conduct a sequence of additional experiments. In particular, we explore the quantitative implications of whether

1. very large establishments are required to be part of the community rating (Y/N),
2. administrative loads are pooled across employers of different sizes (Y/N),
3. there is a cost of starting insurance as in the benchmark calibration (Y/N).

We conduct a total of eight experiments representing all combinations of answers to the three questions above. The results are collected in Appendix Table A-2. The label of each experiment represents answers to these questions. For example, YNY denotes the experiment in which very large establishments participate, loads are not pooled, and there are costs of starting insurance.

The results of these experiments indicate that participation of very large establishments in the community-rating scheme is essential to sustaining coverage. Their low administrative loads contribute to a low community-rated premium. Without their participation there is no level of the premium that yields positive coverage while permitting the scheme to break even. What may be unexpected is that even establishments with only unhealthy worker do not participate. In the absence of the cost of starting such establishments would clearly gain from participating due to the tax subsidy, but the cost of starting is sufficiently high that this is not worthwhile.

The cost of starting is also critical for sustaining coverage. In its absence an intermediate level of coverage is sustained as long as very large establishments participate.

[14] Mas-Colell, Whinston and Green (1995, Proposition 13.B.1) show that this is the unique equilibrium in a model of Bertrand competition among sellers (insurance carriers in our model). Alternatively the community-rating scheme could be run by the government. Given the latter interpretation the premium we determine is the lowest premium that the government can implement without subsidizing the scheme (above and beyond the tax deductability built into the benchmark calibration).

But this is an artifact of our assumption that very large establishments always provide coverage. One can verify that the premium in this experiment is sufficiently high that these establishments would prefer to drop coverage. Without their presence only establishments employing exclusively unhealthy workers participate in the scheme.

Interestingly, partial coverage is sustained if the only departure from the full community-rating experiment is that administrative loads are not pooled. In this case high loads in conjunction with the cost of starting prevent very small establishments, including those with a large fraction of unhealthy workers, from participating in the scheme. This mitigates adverse selection. Establishments of intermediate size no longer cross-subsidize the administrative expenses of very small employers, making it attractive for many of them to participate. The resulting level of coverage is positive, but lower than in the benchmark calibration.

This last result suggests that a switch from experience rating to community rating may increase or decrease coverage, depending on the extent to which the scheme pools administrative expenses.

Table 4: Evaluating Current Reform Proposals

| | Benchmark | Experiments | | | |
|---|---|---|---|---|---|
| | (I) | (II.1) | (II.2) | (II.3) | (II.4) |
| *Fraction of Establishments Providing Coverage* | | | | | |
| 3-10 | 0.138 | 0.000 | 1.000 | 0.000 | 1.000 |
| 11-25 | 0.705 | 0.000 | 0.998 | 0.000 | 0.996 |
| 26-50 | 0.940 | 0.000 | 0.849 | 0.000 | 0.865 |
| 51- | 0.932 | 0.610 | 0.954 | 1.000 | 1.000 |
| all | 0.461 | 0.065 | 0.983 | 0.099 | 0.989 |
| *Fraction of Establishments Providing Coverage that Commit* | | | | | |
| 3-10 | 0.838 | 0.000 | 0.612 | 0.000 | 0.593 |
| 11-25 | 0.972 | 0.000 | 0.854 | 0.000 | 0.736 |
| 26-50 | 0.987 | 0.000 | 0.960 | 0.000 | 0.977 |
| 51- | 0.997 | 0.195 | 0.998 | 0.128 | 0.984 |
| all | 0.960 | 0.195 | 0.749 | 0.128 | 0.703 |
| *Fraction of Est. Discontinuing Insurance over a 2 Year Period* | | | | | |
| 3-10 | 0.346 | 1.000 | 0.000 | 1.000 | 0.000 |
| 11-25 | 0.106 | 1.000 | 0.002 | 1.000 | 0.004 |
| 26-50 | 0.056 | 1.000 | 0.146 | 1.000 | 0.131 |
| 51- | 0.067 | 0.357 | 0.045 | 0.000 | 0.000 |
| all | 0.136 | 0.609 | 0.016 | 0.424 | 0.011 |
| *Establishment Size Distribution* | | | | | |
| 3-10 | 0.322 | 0.327 | 0.308 | 0.326 | 0.309 |
| 11-25 | 0.154 | 0.144 | 0.187 | 0.145 | 0.187 |
| 26-50 | 0.084 | 0.077 | 0.045 | 0.083 | 0.045 |
| 51- | 0.056 | 0.065 | 0.061 | 0.060 | 0.061 |
| mean | 16.28 | 16.48 | 15.55 | 16.33 | 15.54 |
| median | 4 | 4 | 4 | 4 | 4 |
| *Fraction of Unhealthy Workers among all Employees* | | | | | |
| 3-10 | 0.009 | 0.000 | 0.067 | 0.000 | 0.065 |
| 11-25 | 0.047 | 0.000 | 0.070 | 0.000 | 0.068 |
| 26-50 | 0.062 | 0.000 | 0.061 | 0.010 | 0.062 |
| 51- | 0.073 | 0.062 | 0.064 | 0.087 | 0.064 |
| Avg. Wage | 1.006 | 1.035 | 0.998 | 1.027 | 0.998 |
| Avg. Labor Cost | 1.048 | 1.035 | 1.043 | 1.027 | 1.045 |
| Output per Worker | 1.287 | 1.287 | 1.286 | 1.288 | 1.286 |
| $m$ | 0.074 | 0.105 | 0.058 | 0.088 | 0.058 |
| $\nu$ | 0.855 | 0.712 | 0.925 | 0.775 | 0.937 |
| $\theta$ | 0.500 | 0.430 | 0.697 | 0.456 | 0.687 |

Notes to Table 4 -

The table contains the results of quantitative experiments using the calibrated model.

Description of Columns:

Column (I) − Results from benchmark calibration of the model.

Column (II.1) − Establishments with less than 50 workers are subject to community rating.

Column (II.2) − Same as Column (II.1) but purchase of insurance by establishments with less than 25 workers is subsidized.

Column (II.3) − Same as Column (II.1) but a tax is imposed on establishments with more than 50 workers that do not provide health insurance.

Column (II.4) − Same as Column (II.2) but a tax is imposed on establishments with more than 50 workers that do not provide health insurance.

Description of Rows:

See Notes to Table 3 for the description of the rows.

## 5.4 Evaluating Current Reform Proposal

Achieving comprehensive health insurance reform has emerged as a leading priority of the current U.S. Congress. A number of reform proposals have been announced as the debate proceeds over how to overhaul the health insurance system. The House of Representative has passed on November 7, 2009 the Affordable Health Care for America Act. The Senate has passed on December 24, 2009 the Patient Protection and Affordable Care Act. Both Acts have many features in common. In particular,

1. The health insurance system will continue to be employer-based. Tax deductibility of employer contributions to health insurance remains in effect. Individual purchase of insurance remains not tax deductible. All non-discrimination provisions remain in effect.

2. Individuals will be required to have "acceptable health coverage." This is enforced through a tax penalty of 2.5% of income up to the cost of the average national premium (House) or a tax penalty of $750 per adult per year (Senate).

3. Employers will be required to offer coverage or pay a tax equal to 8% of payroll (House) or $400 per employee (Senate). In the Senate's version of the legislation, employers with 50 employees or less are exempt. In the House's version the exemption applies to small employers with annual payroll of $750,000.

4. Employers with less than 25 employees will be provided with a tax credit up to 50% of the employer's contribution toward the employee's health insurance premium under both proposals. The amount of the tax credit will vary with the average wage of the employees and the size of the employer, but the schedule has not yet been fully specified.

5. Insurance exchanges will be created through which individuals and small employers can purchase insurance. Guaranteed issue and renewability of coverage are required; limited rating variation based only on age, premium rating area, and family enrollment, is allowed. Purchases of those with incomes less than 400% of the Federal Poverty Line will be subsidized. Insurers will be prohibited from rescinding coverage.

There are numerous other proposed changes (the passed reform Acts are over 2,000 pages long!) that are less directly applicable for our analysis and we omit them from consideration.

We evaluate the consequences of a reform that is similar to the proposed one. In particular, we introduce the following changes to the model.

1. A community rated insurance market for establishments with less than 50 employees where premiums cannot depend on the health composition of workforce.

2. A tax subsidy of 30% of premiums paid by employers with less than 25 employees.

3. A tax equal to 8% of payroll for employers with more than 50 employees if they choose to not provide coverage.

The results of this experiment are presented in Column (II.4) of Table 4. In Columns (II.1) through (II.3) we introduce the components of the reform sequentially, starting with community rating for small groups in Column (II.1), and then introducing the premium subsidy in Column (II.2), or a tax on employers with more than 50 employees if they choose to not provide coverage in Column (II.3).

Column (II.1) shows that introducing community rating for establishments with less than 50 workers by itself results in an effective collapse of coverage by these establishments. This is because of severe adverse selection induced by such a reform. Since coverage among small establishments collapses, the health composition among workers in large establishments deteriorates; as a consequence they also become less likely to provide coverage and discontinue coverage more frequently.

Column (II.2) shows that introducing an additional tax subsidy of 30% of premiums paid by employers with less than 25 employees is sufficient to prevent unraveling of coverage by community rated establishments with less than 50 workers. Interestingly, although only the smallest establishments are given the subsidy, establishments of all sizes respond by providing essentially full coverage. Once again this is due to the equilibrium effect that when small establishments provide coverage, the health composition of larger establishments improves, making it easier for them to provide coverage as well. Finally, we note that the peculiar design of this policy, where establishments with less than 50 workers are community rated but only establishments with 25 workers get the subsidy, implies sizable effects on the establishment size distribution. Establishments with 26 to

50 workers before the reform tend to shrink in order to qualify for the subsidy. Smaller establishments, however, tend to expand since the subsidy reduces their labor costs. Overall, the fraction of establishments with 11 to 25 workers increases by 21.4% while the fraction of establishments with 26 to 50 workers declines by 46.4%.

Column (II.3) combines the community rating for establishments with less than 50 workers with the tax equal to 8% of payroll for employers with more than 50 employees if they choose to not provide coverage. The tax on large establishments is insufficient to overcome the collapse of insurance for establishments with less than 50 workers.

The full reform is simulated in Column (II.4). The results indicate that such a reform can result in a virtually complete coverage. A subsidy to smallest employers is essential to achieve coverage in the community rated segment of the market. The tax on largest establishments is sufficient to achieve full coverage among large establishments so that this tax is not actually paid in equilibrium.

# 6 Conclusion

In this paper we proposed a framework to quantitatively study how the design of the employer-based US health insurance system affects employers' decisions on whether to offer coverage to their employees as well as the effects of the system's design on the labor market and establishment dynamics. Our analysis suggests that three features of the system are very important in determining its effects. These are (1) tax deductibility of coverage, (2) legal non-discrimination provisions within establishments with respect to hiring, firing, access and price of health insurance, and (3) legal restrictions on the ability of health insurers to price discriminate across establishments based on the health composition of their workforce.

Using the calibrated model we evaluated the quantitative importance of the fact that small establishments have a more variable health composition and found that, given the current design of the system, it is very important in accounting for the low and more volatile health insurance provision by small establishments. It is also important for understanding the flows of individuals in and out of the pool of the uninsured. We also found that the current system is associated with sizable effects on worker mobility, unemployment, the size distribution of establishments, and other aggregate measures. We also found that equilibrium effects are very powerful in shaping the patterns of insurance provision. Changes in the economic environment that affect the behavior of a small subset

of establishments have very strong effects on other establishments. For example, starting from low coverage by establishments of all sizes, a subsidy to the smallest establishments only may lead to complete coverage by the largest establishments.

We also used the calibrated model to assess the likely outcomes of the reforms currently being debated in Congress and the role of various elements of the proposed reform. Our results suggest that to achieve sizable coverage large subsidies to small employers will be required. Introducing community rating without such subsidies will likely result in unraveling of health insurance provision by small employers.

Our analysis has abstracted from several potentially important features of the system that we would like to incorporate in future work. First, since at least Feldstein (1973), it has been recognized that since health insurance is subsidized workers may demand too much of it leading to excessive insurance coverage and costs. Second, the tax treatment of health insurance is argued to be regressive because workers' tax savings depend on their marginal tax rates. Since marginal tax rates generally increase with income, higher income individuals and families obtain greater tax savings. This might be important for our analysis in light of the well-documented fact that larger employers tend to pay higher wages. Finally, it is not uncommon to receive insurance through ones spouse's employer. Thus introducing couples into the model may be an important extension. An interesting question is whether large employers subsidize small employers through this channel.

This model and the algorithm we developed to compute it represent interesting contributions in themselves. The key innovation is to solve an industry dynamics model where each worker has a positive mass and behaves strategically. The model could be applied to study other issues such as the fragility of small groups induced by the decisions of key members to join or leave them.

# References


Albrecht, J. W., and B. Axell (1984): "An Equilibrium Model of Search Unemployment," *Journal of Political Economy*, 92(5), 824–840.

Brügemann, B. (2008): "What Elasticity of the Matching Function is consistent with U.S. Aggregate Labor Market Data?," Working paper, Yale University.

Cochrane, J. H. (1995): "Time-Consistent Health Insurance," *Journal of Political Economy*, 103(3), 445–473.

Congressional Research Service (1988): *Costs and Effects of Extending Health Insurance Coverage*. U.S. Government Printing Office, Washington, DC.

Cutler, D. M. (1994): "Market Failure in Small Group Health Insurance," Working Paper 4879, National Bureau of Economic Research.

Dey, M. S., and C. J. Flinn (2005): "An Equilibrium Model of Health Insurance Provision and Wage Determination," *Econometrica*, 73(2), 571–627.

Feldstein, M. S. (1973): "The Welfare Loss of Excess Health Insurance," *Journal of Political Economy*, 81(2), 251–280.

Guner, N., G. Ventura, and X. Yi (2008): "Macroeconomic Implications of Size-Dependent Policies," *Review of Economic Dynamics*, 11(4), 721–744.

Hagedorn, M., and I. Manovskii (2008): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited," *American Economic Review*, 98(4), 1692–1706.

Henry J. Kaiser Family Foundation (1998): "The Uninsured: A Primer," Report No. 7451-04, The Kaiser Commission on Medicaid and the Uninsured.

Hopenhayn, H., and R. Rogerson (1993): "Job Turnover and Policy Evaluation: A General Equilibrium Analysis," *Journal of Political Economy*, 101(5), 915–938.

Jeske, K., and S. Kitao (2008): "U.S. Tax Policy and Health Insurance Demand: Can a Regressive Policy Improve Welfare?," *Journal of Monetary Economics*, Forthcoming.

Joint Committee on Taxation (2008): "Tax Expenditures for Health Care," JCX-66-08.

Long, S. H., and M. S. Marquis (1997): *ROBERT WOOD JOHNSON FOUNDATION EMPLOYER HEALTH INSURANCE SURVEY [COMMUNITY TRACKING STUDY AND STATE INITIATIVES IN HEALTH CARE REFORM PROGRAM], 1997 [Computer file].* ICPSR02935-v2. Washington, DC: RAND Corporation [producer], 2000. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2005-07-06. doi:10.3886/ICPSR02935.

——— (1998): "Trends in Offering Employer-Sponsored Coverage," Data Bulletin No. 15, Center for Studying Health Care Change.

Lucas, R. J. (1990): "Supply Side Economics: An Analytical Review," *Oxford Economic Papers*, 42(2), 293–316.

Lyke, R. (2008): "The Tax Exclusion for Employer-Provided Health Insurance: Policy Issues Regarding the Repeal Debate," *Congressional Research Service Report for Congress*, RL34767.

Mendoza, E., A. Razin, and L. Tesar (1994): "Effective Tax Rates in Macroeconomics: Cross-country Estimates of Tax Rates on Factor Incomes and Consumption," *Journal of Monetary Economics*, 34, 297–323.

Mortensen, D., and C. Pissarides (1994): "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, 61(3), 397–415.

Office of Management and Budget (2008): *Analytical Perspectives, Budget of the United States Government, Fiscal Year 2009.* U.S. Government Printing Office, Washington, DC.

Petrongolo, B., and C. Pissarides (2001): "Looking into the Black Box: A Survey of the Matching Function," *Journal of Economic Literature*, 39(2), 390–431.

Prescott, E. (2004): "Why Do Americans Work So Much More Than Europeans," *Federal Reserve Bank of Minneapolis Quarterly Review*, 28(1), 2–13.

Restuccia, D., and R. Rogerson (2008): "Policy Distortions and Aggregate Productivity with Heterogeneous Plants," *Review of Economic Dynamics*, 11(4), 707–720.

Selden, T. M., and B. M. Gray (2006): "Tax Subsidies for Employer-Related Health Insurance: Estimates for 2006," *Health Affairs*, 25(2), 1568–1579.

APPENDICES

# I Transition Probabilities

## I.1 Establishment Transition $\mu(s'|s, g, \theta, \nu)$

Consider an establishment in state $s$ with policy $g$. After workers have been induced to leave $(g_h, g_u)$ workers remain, so at this stage the workforce of the establisment is still deterministic. Next workers are dismissed at random, leaving $g_e$ workers in total. The probability of arriving at the workforce $(\psi_h, \psi_u)$ after random dismissal is given by

$$q^{dis}(\psi_h, \psi_u|g) = \begin{cases} \frac{\binom{g_h}{\psi_h}\binom{g_u}{\psi_u}}{\binom{g_h+g_u}{g_e}}, & \text{if } \psi_h + \psi_u = g_e, \\ 0, & \text{otherwise.} \end{cases} \tag{A1}$$

The logic behind this formula is as follows. The total number of workers before random dismissal is $g_h + g_u$, and there are $g_e$ slots. There are $\binom{g_h+g_u}{g_e}$ different ways of allocating these slots, all equally likely. The number of different ways of allocating these slots which have $\psi_h$ healthy and $\psi_u$ unhealthy workers are $\binom{g_h}{\psi_h}\binom{g_u}{\psi_u}$. Vandermonde's identity insures that these probabilities add up to one. Let $Q^{dis}(g)$ collect these probabilities in a vector, ordering workforces in the natural way: $(0, 0)$, $(1, 0)$, $(0, 1)$, $(2, 0)$, $(1, 1)$, $(0, 2)$ and so on.

Next workers draw their new health status, and we need to compute the probability of transiting from $(\psi_h, \psi_u)$ to $(\psi'_h, \psi'_u)$ in this step. The number of workers must remain the same $\psi'_h + \psi'_u = \psi_h + \psi_u$, otherwise the probability of this transition is zero. For a transition to $(\psi'_h, \psi'_u)$ it must be that the number of workers remaining in status $h$ is at least $\max\{\psi'_h - \psi_u, 0\}$, because no more than $\psi_u$ can join from status $u$. The probability of $j$ workers remaining in status $h$ is given by $B\left(j; \psi_h, q^H_{hh}\right)$. Here $B$ is the binomial distribution: the first argument is the number of successes, the second argument the number of trials, and the third argument the probability of success. If $j$ workers remain in status $h$, the transition to $\psi'_h$ requires that exactly $\psi'_h - j$ switch from status $u$ to status $h$. The latter happens with probability $B\left(\psi'_h - j; \psi_u, q^H_{uh}\right)$. Thus

$$q^{health}(\psi_h, \psi_u; \psi'_h, \psi'_u) = \sum_{j=\max\{\psi'_h-\psi_u,0\}}^{\min\{\psi'_h,\psi_h\}} B\left(j; \psi_h, q^H_{hh}\right) B\left(\psi'_h - j; \psi_u, q^H_{uh}\right).$$

Notice that this probability does not depend on the employer's policy $g$. Let $Q^{health}$ denote the transition matrix associated with this step.

In the next step workers exit the labor market with probability $1-\rho$, or quit exogenously with probability $\delta$. Thus a worker stays with the employer with probability $\rho(1-\delta)$. The probability of a transition from $(\psi_h, \psi_u)$ to $(\psi'_h, \psi'_u)$ in this step is

$$q^{exit}\left(\psi_h, \psi_u; \psi'_h, \psi'_u\right) = B(\psi'_h; \psi_h, \rho(1-\delta))B(\psi'_u; \psi_u, \rho(1-\delta)).$$

Let $Q^{exit}$ denote the transition matrix.

The final step is that searching workers are allocated to the employer. The employer has $g_v$ vacancies, each of which is filled with probability $q(\theta)$. The probability to transit from $(\psi_h, \psi_u)$ to $(\psi'_h, \psi'_u)$ is

$$\begin{aligned} & q^{vac}\left(\psi_h, \psi_u; \psi'_h, \psi'_u\right) \\ & \equiv B\left(\psi'_h + \psi'_u - \psi_h - \psi_u; g_v, q(\theta)\right) B\left(\psi'_h - \psi_h; \psi'_h + \psi'_u - \psi_h - \psi_u, \nu\right) \end{aligned}$$

The first term captures that out of $g_v$ vacancies it must be that $\psi'_h + \psi'_u - \psi_h - \psi_u$ make contact with a worker, with a probability of success $q(\theta)$. The second term captures that out of these contacts, $\psi'_h - \psi_h$ must be with a healthy worker, with a probability of success $\nu$. Let $Q^{vac}(g, \theta, \nu)$ denote the transition matrix.

Combining these transitions, the distribution $\mu(s'|s, g)$ is given by

$$\mu(s'|s, g) = q^Z_{z(s)z(s')} q^I_{I(s)I(s')}(g) \left[Q^{vac}(g, \theta, \nu) \cdot Q^{exit} \cdot Q^{health} \cdot Q^{dis}(g)\right]_{\psi_h(s')\psi_u(s')}$$

where $[Q]_{\psi_h\psi_u}$ extracts the element of vector $Q$ corresponding to the workforce $(\psi_h, \psi_u)$. The transition probabilities for insurance provision status are $q^I_{CC}(g) = 1 - q^I$, $q^I_{CE}(g) = q^I$, $q^I_{EC}(g) = g^I$, $q^I_{EN}(g) = 1 - g^I$, $q^I_{NC}(g) = g_I$, $q^I_{NN}(g) = 1 - g_I$, and zero for the remaining transitions.

## I.2 Worker Transition $\mu_{ii'}\left[s'|s, g, \theta, \nu\right]$

We derive $\mu_{hh'}\left[s'|s, g, \theta, \nu\right]$, the remaining cases are analogous. The calculations parallel the derivation of the establishment transition, with the twist that we need to condition on the worker being healthy both in this period and in the next period, and that the worker remains in the labor market and stays with the establishment.

After workers have been induced to leave $(g_h, g_u)$ workers remain. Next workers are dismissed at random, leaving $g_e$ workers in total. Conditioning on the worker staying and being healthy, the probability of arriving at the workforce $\left(\hat{\psi}_h, \hat{\psi}_u\right)$ after random

dismissal is given by

$$q_{hh}^{dis}\left(\hat{\psi}_h, \hat{\psi}_u | g\right) = \begin{cases} \frac{\binom{g_h-1}{\hat{\psi}_h-1}\binom{g_u}{\hat{\psi}_u}}{\binom{g_h+g_u-1}{g_e-1}}, & \text{if } \hat{\psi}_h + \hat{\psi}_u = g_e, \\ 0, & \text{otherwise.} \end{cases} \tag{A2}$$

The logic behind this formula is as follows. The worker under consideration is healthy and is not dismissed. The remaining number of workers at risk of random dismissal is $g_h + g_u - 1$, and there are $g_e - 1$ remaining slots. There are $\binom{g_h+g_u-1}{g_e-1}$ different ways of allocating these slots, all equally likely. To end up at $(\hat{\psi}_h, \hat{\psi}_u)$ it must be that $\hat{\psi}_h - 1$ of these slots go to healthy workers. The number of different ways of allocating these slots which have $\hat{\psi}_h - 1$ healthy and $\hat{\psi}_u$ unhealthy workers are $\binom{g_h-1}{\hat{\psi}_h-1}\binom{g_u}{\hat{\psi}_u}$. Let $Q_{hh}^{dis}(g)$ denote the vector of these probabilities.

Next workers draw their new health status. We compute the probability of a transition from $(\psi_h, \psi_u)$ to $(\psi'_h, \psi'_u)$ The number of workers must remain the same $\psi_h + \psi_u = \psi'_h + \psi'_u$, otherwise the probability of this transition is zero. Here we need to condition on the event that the worker under consideration stays healthy. For a transition to $(\psi'_h, \psi'_u)$ it must be that the number of workers remaining in status $h$ is at least $\max\{\psi'_h - \psi_u, 1\}$, because no more than $\psi_u$ can join from status $u$, and we already condition on one healthy worker staying healthy. The probability of $j$ workers remaining in status $h$ is given by $B\left(j-1; \psi_h - 1, q_{hh}^H\right)$. If $j$ workers remain in status $h$, the transition to $\psi'_h$ requires that exactly $\psi'_h - j$ switch from status $u$ to status $h$. The latter happens with probability $B\left(\psi'_h - j; \psi_u, q_{uh}^H\right)$. Thus

$$q_{hh}^{health}\left(\psi_h, \psi_u; \psi'_h, \psi'_u\right) = \sum_{j=\max\{\tilde{\psi}_h - \hat{\psi}_u, 1\}}^{\min\{\tilde{\psi}_h, \hat{\psi}_h\}} B\left(j-1; \hat{\psi}_h - 1, q_{hh}^H\right) B\left(\tilde{\psi}_h - j; \hat{\psi}_u, q_{uh}^H\right).$$

Let $Q_{hh}^{health}$ denote the associated transition matrix.

Next workers exit the labor force with probability $(1-\rho)$, or separate exogenously with probability $\delta$. We condition on the worker under consideration remaining with the establishment. The probability of a transition from $(\psi_h, \psi_u)$ to $(\psi'_h, \psi'_u)$ in this step is

$$q_{hh}^{exit}\left(\psi_h, \psi_u; \psi'_h, \psi'_u\right) = B(\psi'_h - 1; \psi_h - 1, \rho) B(\psi'_u; \psi_u, \rho).$$

Let $Q_{hh}^{exit}$ denote the transition matrix.

Again, the final step is that searching workers are allocated to the establishment. This step is not affected by conditioning.

Combining these transitions, the distribution $\mu(s'|s,g,\theta,\nu)$ is given by

$$\mu(s'|s,g,\theta,\nu) = q^{Z}_{z(s)z(s')} q^{I}_{I(s)I(s')}(g) \left[Q^{vac}(g,\theta,\nu) \cdot Q^{exit}_{hh} \cdot Q^{health}_{hh} \cdot Q^{dis}_{hh}(g)\right]_{\psi_h(s'),\psi_u(s')}$$

## I.3 Searching Worker Transition $\mu^s_{i'}\left[s'|\mu(\cdot),\gamma(\cdot|\cdot),\theta,\nu\right]$

A searching worker who makes contact is randomly allocated to a vacancy. Let $s$ denote the state of the worker's new employer at the beginning of the period when it posted the vacancy. This employer implements policy $g$ with probability $\gamma(g|s)$, in which case it has $g_v$ vacancies. Thus the searcher is matched with an employer in state $s$ implementing policy $g$ with probability

$$q^{match}\left[s,g|\mu(\cdot),\gamma(\cdot|\cdot)\right] = \frac{g_v \gamma(g|s)\mu(s)}{\sum_{\tilde{g}\in G}\sum_{s\in\mathcal{S}} \tilde{g}_v \gamma(\tilde{g}|s)\mu(s)}.$$

If matched with an employer implementing policy $g$, the distribution of that employer's workforce after random dismissal and exogenous separations is $Q^{dis}(g)$, and after health status changes and exit from the labor market it is $Q^{exit} \cdot Q^{health} \cdot Q^{dis}(g)$. Next the employer is allocated searchers. In this step we need to condition on the worker under consideration having new health status $i'$. Here we consider the case $i' = h$, the case $i' = u$ is analogous. If an establishment has workforce $(\psi_h, \psi_u)$ before vacancies are filled and follows policy $g$, then the probability to arrive at $(\psi'_h, \psi'_u)$ is

$$\begin{aligned} q^{vac}_h\left(\psi_h,\psi_u;\psi'_h,\psi'_u|g\right) \equiv\ & B\left(\psi'_h + \psi'_u - \tilde{\psi}_h - \tilde{\psi}_u - 1; g_v - 1, q(\theta)\right) \\ & \cdot B\left(\psi'_h - \tilde{\psi}_h - 1; \psi'_h + \psi'_u - \tilde{\psi}_h - \tilde{\psi}_u - 1, \nu\right). \end{aligned}$$

The worker under consideration has already filled one vacancy and is healthy. The first term captures that out of $g_v - 1$ remaining vacancies it must be that $\psi'_h + \psi'_u - \tilde{\psi}_h - \tilde{\psi}_u - 1$ make contact with a worker, with a probability of success $q(\theta)$. The second term captures that out of these contacts, $\psi'_h - \tilde{\psi}_h - 1$ must be with a healthy worker, with a probability of success $\nu$. Let $Q^{vac}_h(g,\theta,\nu)$ denote the associated transition matrix. Then the distribution after this step is $Q^{vac}_h(g,\theta,\nu) \cdot Q^{exit} \cdot Q^{health} Q^{dis}(g)$ Combining these three steps

$$\begin{aligned} \mu^s_h\left[s'|\mu(\cdot),\gamma(\cdot|\cdot),\theta,\nu\right] = \sum_{s\in\mathcal{S}}\sum_{g\in\mathcal{G}} \Big\{ & \left[Q^{vac}_h(g,\theta,\nu) \cdot Q^{exit} \cdot Q^{health} \cdot Q^{dis}(g)\right]_{\psi_h(s')\psi_u(s')} \\ & \cdot q^{Z}_{z(s)z(s')} q^{I}_{I(s)I(s')}(g) q^{match}\left[s,g|\mu(\cdot),\gamma(\cdot|\cdot)\right] \Big\} \end{aligned}$$

# II Computational Algorithm

The algorithm iterates on the policy function $\gamma(\cdot|\cdot)$ until convergence.

First, notice that for any policy function it is straightforward to compute $\mathcal{J}(\cdot)$, $V_h(\cdot)$, $V_u(\cdot)$, $V_h^s$, $V_u^s$, $\mu(\cdot)$, $m$, $\nu$, and $\theta$ consistent with that policy. Second, for any $\{\mathcal{J}(\cdot), V_h(\cdot), V_u(\cdot), V_h^s, V_u^s, \mu(\cdot), m_h, m_u, \theta\}$ we can use equation (8) to compute a set of policies which are optimal. Combining these two mappings, we get a correspondence $\Omega$ mapping policy functions into sets of policy functions. Stationary equilibrium policy functions are the fixed points of this correspondence, so we're looking for $\gamma(\cdot|\cdot)$ such that

$$\gamma(\cdot|\cdot) \in \Omega[\gamma(\cdot|\cdot)].$$

For a policy function $\gamma^k(\cdot|\cdot)$, let $\gamma^k(\cdot|\cdot;\tilde{\gamma}(\cdot|s))$ denote the policy given by

$$\gamma^k(\cdot|z',\psi_h',\psi_u';\tilde{\gamma}(\cdot|s)) = \begin{cases} \gamma^k(\cdot|z',\psi_h',\psi_u') & \text{for all} \quad (z',\psi_h',\psi_u') \neq (s), \\ \tilde{\gamma}(\cdot|s) & \text{for} \quad (z',\psi_h',\psi_u') = (s). \end{cases}$$

In words, $\gamma^k(\cdot|\cdot;\tilde{\gamma}(\cdot|s))$ is obtained from $\gamma^k(\cdot|\cdot)$ by switching out the policy at one point in the state space, replacing $\gamma^k(\cdot|s)$ with $\tilde{\gamma}(\cdot|s)$.

Given $\Omega[\gamma(\cdot|\cdot)]$, define the projection

$$\Omega_{(s)}[\gamma(\cdot|\cdot)] = \{\tilde{\gamma}(\cdot|s) \in \Gamma | \tilde{\gamma}(\cdot,\cdot) \in \Omega[\gamma(\cdot|\cdot)]\}.$$

This is the sets of mixed policies that are optimal for the point in the state space $(s)$ given that all other equilibrium objects are induced by the policy function $\gamma(\cdot|\cdot)$.

The algorithm starts with a guess $\gamma^0(\cdot|\cdot)$. The approach is to find a fixed point for just one point in the state space at each iteration, and to move randomly through the state space until convergence. Iteration $k$ comprises the following steps:

1. Pick a point in the state space $(z^k, \psi^k) \in \mathcal{S}$ at random.

2. Given the policy function $\gamma^k(\cdot|\cdot)$, find a mixed policy $\tilde{\gamma}(\cdot|z^k,\psi^k)$ such that

$$\gamma^k(\cdot|\cdot;\tilde{\gamma}(\cdot|z^k,\psi^k)) \in \Omega_{(z^k,\psi^k)}[\gamma^k(\cdot|\cdot;\tilde{\gamma}(\cdot|z^k,\psi^k))].$$

   This step is implemented using a heuristic algorithm.

3. Set $\gamma^{k+1}(\cdot,\cdot) = \gamma^k(\cdot|\cdot;\tilde{\gamma}(\cdot|z^k,\psi^k))$.

Table A-1: Experiments Varying the Cost of Starting Insurance

| | Cost of Starting | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 0.0 | 0.2 | 0.4 | 0.6 | 0.8 | 1.0 | 1.2 | 1.4 |
| *Fraction of Establishments Providing Coverage* | | | | | | | | |
| 3-10 | 0.874 | 0.047 | 0.027 | 0.028 | 0.058 | 0.078 | 0.171 | 0.186 |
| 11-25 | 0.766 | 0.002 | 0.052 | 0.095 | 0.340 | 0.506 | 0.601 | 0.593 |
| 26-50 | 0.599 | 0.102 | 0.531 | 0.695 | 0.941 | 0.985 | 0.794 | 0.667 |
| 51- | 0.546 | 0.823 | 0.989 | 0.999 | 0.999 | 0.994 | 0.862 | 0.773 |
| all | 0.782 | 0.105 | 0.187 | 0.220 | 0.332 | 0.390 | 0.423 | 0.404 |
| *Fraction of Establishments Discontinuing Insurance over a 2 Year Period* | | | | | | | | |
| 3-10 | 0.127 | 0.737 | 0.739 | 0.744 | 0.658 | 0.624 | 0.381 | 0.343 |
| 11-25 | 0.234 | 0.964 | 0.653 | 0.540 | 0.200 | 0.095 | 0.022 | 0.011 |
| 26-50 | 0.401 | 0.752 | 0.168 | 0.067 | 0.007 | 0.001 | 0.001 | 0.001 |
| 51- | 0.454 | 0.123 | 0.007 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| all | 0.219 | 0.537 | 0.277 | 0.237 | 0.181 | 0.160 | 0.118 | 0.112 |
| *Fraction of Unhealthy Workers among all Employees* | | | | | | | | |
| 3-10 | 0.000 | 0.000 | 0.000 | 0.000 | 0.001 | 0.001 | 0.008 | 0.011 |
| 11-25 | 0.000 | 0.000 | 0.003 | 0.007 | 0.026 | 0.039 | 0.049 | 0.051 |
| 26-50 | 0.008 | 0.003 | 0.040 | 0.057 | 0.071 | 0.072 | 0.060 | 0.053 |
| 51- | 0.072 | 0.080 | 0.088 | 0.085 | 0.080 | 0.078 | 0.075 | 0.074 |
| *Fraction of Employed Workers Covered, Excluding Very Large Establishments* | | | | | | | | |
| All | 0.653 | 0.284 | 0.492 | 0.547 | 0.673 | 0.724 | 0.575 | 0.592 |
| Healthy | 0.655 | 0.273 | 0.473 | 0.529 | 0.657 | 0.710 | 0.556 | 0.574 |
| Unhealthy | 0.529 | 0.944 | 0.999 | 0.999 | 0.999 | 0.999 | 1.000 | 1.000 |
| Unhealthy w. Rent | 0.282 | 0.791 | 0.978 | 0.998 | 0.999 | 0.999 | 0.999 | 0.999 |
| *Fraction of Employed Workers Covered, Including Very Large Establishments* | | | | | | | | |
| All | 0.770 | 0.524 | 0.659 | 0.696 | 0.780 | 0.814 | 0.714 | 0.725 |
| Healthy | 0.765 | 0.504 | 0.641 | 0.680 | 0.767 | 0.803 | 0.698 | 0.710 |
| Unhealthy | 0.915 | 0.985 | 0.999 | 0.999 | 0.999 | 0.999 | 1.000 | 1.000 |
| Unhealthy w. Rent | 0.870 | 0.945 | 0.989 | 0.999 | 0.999 | 0.999 | 1.000 | 1.000 |
| $m_1$ | 0.073 | 0.065 | 0.063 | 0.063 | 0.062 | 0.062 | 0.063 | 0.064 |
| $m_2$ | 0.027 | 0.024 | 0.015 | 0.013 | 0.010 | 0.009 | 0.010 | 0.012 |
| $V_1$ | 82.89 | 83.09 | 83.80 | 83.92 | 84.12 | 84.20 | 84.11 | 84.01 |
| $V_2$ | 81.35 | 81.64 | 82.69 | 82.86 | 83.16 | 83.27 | 83.15 | 83.00 |

Note - The table contains the results of quantitative experiments of varying the cost of starting insurance provision. The parameter $q^I$ is set equal to one implying that employers have no ability to commit to insurance provision. All other parameter values are fixed at their values in benchmark calibration. $m_1$ and $m_2$ denote the mass of healthy and unhealthy searchers, respectively, $V_1$ and $V_2$ denote the value of search for healthy and unhealthy searchers, respectively.

Table A-2: Community Rating Experiments

| | Specification | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | YYY | YYN | YNY | NYY | YNN | NYN | NNY | NNN |
| *Fraction of Establishments Providing Coverage* | | | | | | | | |
| 3-10 | 0.998 | 0.144 | 0.036 | 0.000 | 0.113 | 0.000 | 0.000 | 0.000 |
| 11-25 | 0.995 | 0.192 | 0.170 | 0.000 | 0.098 | 0.000 | 0.000 | 0.000 |
| 26-50 | 0.983 | 0.324 | 0.609 | 0.000 | 0.191 | 0.000 | 0.000 | 0.000 |
| 51- | 0.999 | 0.648 | 0.950 | 0.167 | 0.691 | 0.169 | 0.169 | 0.169 |
| all | 0.995 | 0.223 | 0.229 | 0.013 | 0.168 | 0.013 | 0.013 | 0.013 |
| *Fraction of Establishments Discontinuing Insurance over a 2 Year Period* | | | | | | | | |
| 3-10 | 0.000 | 0.806 | 0.754 | 0.793 | 0.836 | 0.976 | 0.837 | 0.974 |
| 11-25 | 0.002 | 0.739 | 0.459 | 1.000 | 0.818 | 1.000 | 1.000 | 1.000 |
| 26-50 | 0.015 | 0.557 | 0.211 | 1.000 | 0.644 | 1.000 | 1.000 | 1.000 |
| 51- | 0.001 | 0.255 | 0.032 | 0.000 | 0.238 | 0.000 | 0.000 | 0.000 |
| all | 0.003 | 0.700 | 0.338 | 0.000 | 0.737 | 0.339 | 0.002 | 0.342 |
| *Fraction of Unhealthy Workers among all Employees* | | | | | | | | |
| 3-10 | 0.009 | 0.055 | 0.007 | 0.000 | 0.048 | 0.000 | 0.000 | 0.000 |
| 11-25 | 0.047 | 0.030 | 0.019 | 0.000 | 0.020 | 0.000 | 0.000 | 0.000 |
| 26-50 | 0.062 | 0.035 | 0.054 | 0.000 | 0.021 | 0.000 | 0.000 | 0.000 |
| 51- | 0.073 | 0.070 | 0.080 | 0.069 | 0.075 | 0.071 | 0.071 | 0.071 |
| *Fraction of Employed Workers Covered, Excluding Very Large Establishments* | | | | | | | | |
| All | 0.994 | 0.332 | 0.515 | 0.000 | 0.286 | 0.003 | 0.000 | 0.003 |
| Healthy | 0.994 | 0.310 | 0.493 | 0.000 | 0.264 | 0.000 | 0.000 | 0.000 |
| Unhealthy | 0.999 | 0.808 | 0.974 | 0.000 | 0.803 | 0.988 | 0.000 | 0.988 |
| Unhealthy w. Rent | 0.940 | 0.664 | 0.879 | 0.000 | 0.627 | 0.000 | 0.000 | 0.000 |
| *Fraction of Employed Workers Covered, Including Very Large Establishments* | | | | | | | | |
| All | 0.996 | 0.544 | 0.670 | 0.334 | 0.515 | 0.344 | 0.343 | 0.344 |
| Healthy | 0.996 | 0.525 | 0.652 | 0.312 | 0.494 | 0.320 | 0.321 | 0.320 |
| Unhealthy | 0.999 | 0.891 | 0.985 | 0.963 | 0.895 | 0.999 | 0.959 | 0.999 |
| Unhealthy w. Rent | 0.959 | 0.809 | 0.930 | 0.963 | 0.801 | 0.946 | 0.959 | 0.946 |
| $m_1$ | 0.064 | 0.064 | 0.064 | 0.069 | 0.066 | 0.066 | 0.066 | 0.066 |
| $m_2$ | 0.004 | 0.011 | 0.011 | 0.031 | 0.013 | 0.030 | 0.030 | 0.030 |
| $V_1$ | 84.52 | 83.81 | 84.02 | 82.80 | 83.65 | 82.86 | 82.86 | 82.86 |
| $V_2$ | 83.75 | 82.64 | 82.99 | 81.22 | 82.40 | 81.31 | 81.31 | 81.31 |

Note - The table contains the results of quantitative experiments with different version of community rating. The name of each experiment represents answers to the following three questions: (1) participation of very large establishments (Y/N), (2) pooling of administrative loads (Y/N), (3) existence of the cost of starting (Y/N). For example, YNY denotes the experiment in which very large establishments participate, loads are not pooled, with cost of starting. All parameter values are fixed at their values in benchmark calibration. $m_1$ ($V_1$) and $m_2$ ($V_2$) denote the mass (the value of search) of healthy and unhealthy searchers, respectively.