

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

Björn Brügemann
Yale University
bjoern.bruegemann@yale.edu

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

Preliminary and Incomplete.¹

Abstract

The US health insurance system for those younger than 65 is and will remain largely employer-based after the implementation of the 2010 reform. We develop an equilibrium model of the existing system and study quantitatively the likely effects of the enacted legislation. We show that to match the key empirical regularities of the existing system it is essential to model workers as being discrete, match the firm size distribution and the extent of search frictions, and to model explicitly government regulations affecting the system, such as tax-advantaged treatment of employer purchased coverage and legal non-discrimination restrictions. The model predicts that the enacted reform will achieve universal coverage. The premium regulations of the reform alone would induce a collapse of coverage due to adverse selection. But the accompanying tax credits for small businesses and penalties for individuals are sufficient to prevent this collapse.

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 Sagiri Kitao, Aysegul Sahin, Marcelo Veracierto, and 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, 2009 Yale Health Economics Workshop, 2010 Philadelphia Workshop on Monetary and Macroeconomics, 2010 Society for Economic Dynamics annual meetings, and “Labor market search and policy applications” workshop at Konstanz 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.

¹First version: April 28, 2008. This version: September 12, 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

There has been much legislative activity over the recent years attempting to reform the US health insurance system. At the federal level, this work culminated in 2010 with the passage of the health insurance reform bill (Patient Protection and Affordable Care Act, PPACA) with legislated full implementation date of 2014. There is an ongoing debate about modifying the bill further or even repealing it altogether. Despite the continuing uncertainty about the shape of the health insurance system in the future, there is one feature that seems certain: the health insurance system for those younger than 65 was and will remain largely employer-based.

Despite the active policy debate, there is very little work attempting to model and to evaluate quantitatively the performance of the existing employer-based health care system and of the effects of the current and possible alternative reforms. In particular, employers' coverage decisions are poorly understood. The role of existing regulations and policies in sustaining and shaping the patterns of coverage observed in the data are also unclear. Consequently, the response of coverage to changes in the regulatory environment are difficult to predict. This paper is a step towards filling this gap in the literature.

We present a quantitative equilibrium model of the existing employer-based health insurance system that is consistent with a number of empirical regularities characterizing the existing system and study quantitatively the likely consequences of the enacted legislation. The fact that the system is employer-based implies that the labor market and the employment relationship play an important role in the sharing of medical expenditure risk. In the course of developing the model we aim to understand employers' coverage decisions and the mechanisms that sustain the observed extent of risk sharing in the existing system. Since the enacted reform does not change the employer-based nature of the system, our model will likely remain useful not only for understanding the performance of the new system but also of the effects of its possible further reforms.

Some of the key regularities of the existing employer-provided health insurance system are as follows. Almost half of employers do not provide insurance. Large employers are much more likely to offer health insurance than small employers. Only 46% of establishments with up to 10 employees offer health insurance, compared to 96% of establishments with more than 50 employees. There is substantial turnover in coverage provision: about 11% of employers that provide coverage discontinue over the subsequent 2-year period. Turnover also varies with employer size, with smaller employers more likely to discon-

tinue coverage. Finally, there is little sorting of workers by health status into employers of different sizes. These patterns motivate our modeling choices.

Our model contains four features that seem essential to enable quantitative analysis of the existing system and of possible reforms. They are as follows.

1. We model employer size as being discrete rather than continuous. This is necessary to capture how health-expenditure dynamics at the worker level aggregate to fluctuations in the health composition at the employer level – a key determinant of employers’ coverage decisions.

The fact that an individual worker has positive mass 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. We will show that this is important to understand the dynamics of coverage across employers of different sizes.

2. An employee’s valuation of health insurance offered by an employer depends on how difficult it is for the employee to find another employer providing coverage in the event of becoming unhealthy. This leads us to incorporate search frictions into the model following Mortensen and Pissarides (1994).
3. To the extent that small employers are less likely to provide coverage, this has equilibrium effects on coverage decisions of other employers. To capture such effects quantitatively, it is important for the model to capture the size distribution of employers, which we model following the establishment-size dynamics literature, specifically the model of Hopenhayn and Rogerson (1993).
4. Finally, it would not be possible to provide a reasonable model of the existing health insurance system without incorporating into the model the government regulations affecting it. Indeed, government policy is widely considered to be the primary reason for the predominance of employer-provided insurance. It is tax deductible, while insurance purchased individually by employees is not. In addition to being tax-subsidized, employer-provided insurance is subject to numerous state and federal regulations. 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. Such regulations affect the extent of risk sharing provided by the system by limiting changes in coverage and in the terms of employment following health shocks. We incorporate these features of the regulatory environment into the model.

There are two aspects of the existing system on which we have no readily available outside information but that are important for understanding the performance of the system and its outcomes. These are the extent to which risk is shared between employers and insurance carriers and the extent to which employers are able and willing to make credible promises about future coverage. We use the model to infer the answers to these questions.

The allocation of risk between employers and insurance carriers is a matter of active but so far inconclusive debate in the health economics literature (discussed below). Employers can purchase group insurance from insurance carriers. Long-term contracts are typically not available, with the standard duration of contracts being one year. Most larger employers choose to self-insure. The conventional wisdom is that contracts offered by health insurers are to a large extent risk-rated, so that premiums fully adjust to reflect expected medical expenditures of the group. This implies that employers and employees bear much of the risk associated with persistent shocks to their employees' medical expenditures. In the 1990s concerns about risk rating (also referred to as experience rating) instigated legislation at the state level, with most states imposing rating bands to limit risk rating. Typically wide bands were adopted due to concerns about adverse selection. The 2010 health reform bill bans risk rating starting in 2014 and provides subsidies for coverage provided by small employers to counteract the effects of adverse selection. While there has been substantial legislative activity, there is no direct evidence on the extent of risk rating. We measure the extent of risk rating by comparing the implications of the model with the data for varying degrees of risk rating. We find that the calibrated model with full risk rating implies strong sorting of unhealthy workers into large employers. Only if risk rating is far from complete can the model match the low level of sorting across firm sizes based on health observed in the data.

The extent of the employers' ability to make credible promises is relevant because, as insurers do in other contexts, employers have an incentive to renege on the promise to provide coverage when high medical expenditures are incurred. In the absence of external enforcement, employers' ability to commit to providing coverage is then an

important determinant of the extent of risk sharing that can be achieved. Thus, we assume that employer have access to an imperfect commitment technology. We calibrate this technology to match the patterns of provision in the data and find that in equilibrium large employers choose to commit to future coverage while smaller employers do not. Allowing for commitment is necessary as we find that existing government regulations alone are not sufficient to sustain the observed level and patterns of coverage.

Having developed a model that quantitatively captures key empirical regularities of the employer-provided health insurance system, we use it to understand the likely effects of the recently enacted health insurance reform, the Patient Protection and Affordable Care Act of 2010. The reform maintains the employer-based system, but modifies the incentives of employers to provide and of individuals to obtain coverage through several components. Risk rating is banned in the small group and individual markets. The reform contains tax credits for coverage provided by small employers, penalties for large employers that do not provide insurance, as well as penalties on individuals who do not carry insurance. Most provisions are scheduled to take effect in 2014. We study the joint effect of all components, as well their interaction. We find that imposing community rating in the small group market by itself induces a collapse of coverage in the small group market driven by severe adverse selection. This has a negative spillover effect to coverage in the large group market, as unhealthy workers select into large employers. The small employers tax credit is sufficient to prevent the collapse of coverage. As long as community rating of small employers is complemented by this tax credit, there is little additional impact of the penalty for large employers. In the absence of the tax credit, the large employer penalties contain the spillover effect from the collapse of the small group market. Giving individuals access to the small group community rate without imposing individual penalties once again induces a collapse of coverage, even in the presence of the small employer tax credit. But we find that the individual penalty is sufficiently large to prevent this collapse, and our model predicts universal coverage as the outcome of the full reform.

We also use the model to address several additional questions. In particular, we study the distortions associated with the system of employer-provided insurance. In particular we quantify the effects of the current design of the system on worker flows across establishments and establishment-size dynamics and the likely effects of the PPACA on these variables.

The paper is organized as follows. In the rest of the Introduction we review some related literature. In Section 2 we describe key empirical regularities of employer-provided health insurance. Section 3 presents the model. In Section 4 we define equilibrium. The model is calibrated in Section 5. Section 6 shows the importance of the key elements of the model discussed above for the model's ability to account for the empirical regularities of Section 2. Section 7 uses the calibrated model to study the likely effects of the PPACA. In Section 8 we conduct a series of experiments that illustrate labor market effects of the employer-provided health insurance system and its response to stylized policy changes. Section 9 concludes.

1.1 Related Literature

Two other papers study related issues using equilibrium models. Jeske and Kitao (2008) study tax deductibility of employer-provided insurance in a heterogeneous agent model with incomplete markets. They find that removing this deduction induces healthy employees to switch to individual coverage, leading to a partial collapse of employer-provided coverage and lower welfare. An individual in their model can always purchase risk-rated individual coverage and save in a riskless asset, but does not always have an offer of employer-provided coverage. The latter is modeled as an exogenous stochastic process. Thus Jeske and Kitao abstract from two features which are the focus of our analysis. First, coverage decisions of employers are not modeled explicitly. Thus their model is not designed to address how coverage decisions vary with employer size. Moreover, their policy experiments must rely on empirical estimates of how employers respond to changes in premiums. Second, in their model individuals with high medical expenditures cannot sort into employers that offer coverage. Meanwhile, our model abstracts from mechanisms that are central to their analysis, specifically effects of health insurance operating through the aggregate level of precautionary saving.

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.

Our conclusion that risk rating of small employers is far from complete relates to research by Pauly and Herring (1999) and Herring and Pauly (2001, 2006) on private individual insurance. As for the small group market, conventional wisdom holds that premiums of individual coverage are based strongly on risk. Here perfect risk rating implies a unit elasticity of premiums with respect to expected medical expenses at the household level. Investigating this prediction for various household surveys, Herring and Pauly estimate elasticities far below unity, indicating substantial pooling. How to test for perfect risk rating of small groups using household survey data is less clear. First, even under complete risk rating, pooling within groups mutes the relationship between expected expenses and premiums at the individual level, in a way that depends on group size. Second, expected expenses of an employee may be correlated with expected expenses of coworkers, which affect premiums under risk rating but are not observed in household surveys.² Thus, we evaluate the extent of risk rating by studying the implications of various rating regimes for worker sorting in our equilibrium model.³

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 employee contributions), benefits, cost-sharing, and employer self-insurance status. The

²For instance, if premiums are indeed risk rated, then coverage may be sustainable for a small group only if at most one employee has an expensive chronic condition. Observing such an employee in an insured small employer implies that coworkers cannot also have an expensive chronic condition.

³Cutler (1994) documents large variation in premiums in the small group market that is not explained by actuarial values or demographic composition. He argues that the residual likely reflects risk factors, and concludes that risk rating is quantitatively important. However, the findings of Pauly and Herring (1999) for the individual market suggest that inferring the extent of risk rating from residual dispersion in premiums is problematic. They document large residual dispersion, yet find it to be unrelated to risk factors, and conclude that it must stem from other factors such as search frictions in shopping for plans.

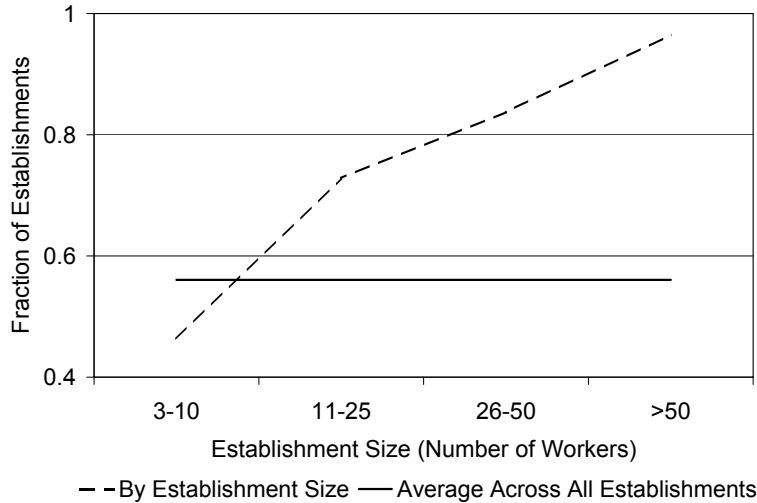


Figure 1: Fraction of Establishments Providing Insurance by Establishment Size.

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 establishment size is almost independent of the average payroll. Strikingly, the lowest

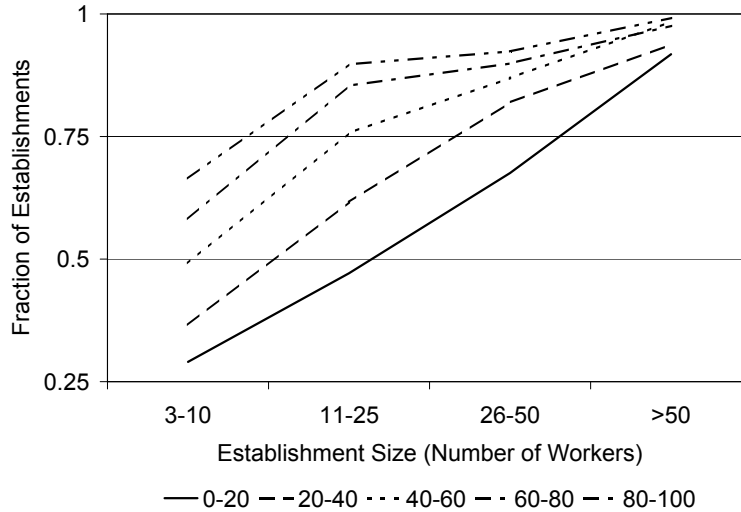


Figure 2: Fraction of Establishments Providing Insurance by Establishment Size and Average Payroll.

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.⁴

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.⁵

One drawback of the statistic above is that it might be affected by time-aggregation. The RWJ survey asks employers that do not provide insurance at the time of the interview

⁴The establishments were ranked by their percentile in the overall establishment distribution, not the distribution within a size class, making such a comparison meaningful.

⁵The 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.

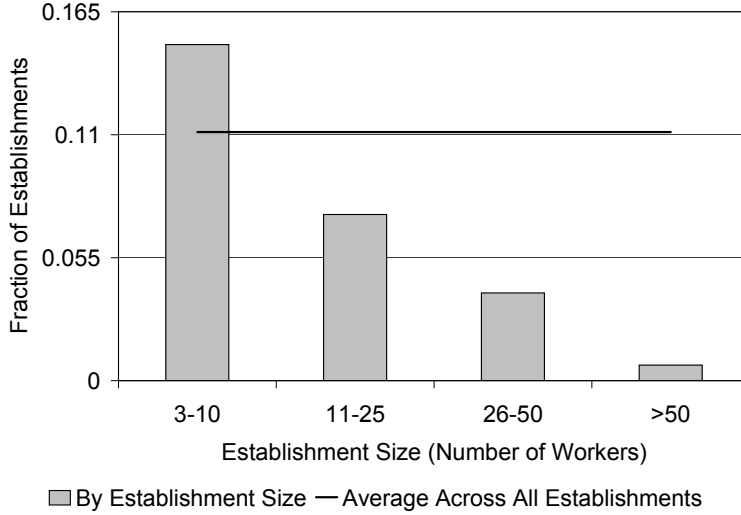


Figure 3: Fraction of Establishments Discontinuing Insurance over a 2 Year Period by Establishment Size.

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.⁶ Suppose there is a unit mass of employers that start coverage at rate α and stop coverage at rate σ . 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 instantaneous rate of discontinuing insurance equals 0.6 for establishments of all sizes and

⁶We 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.

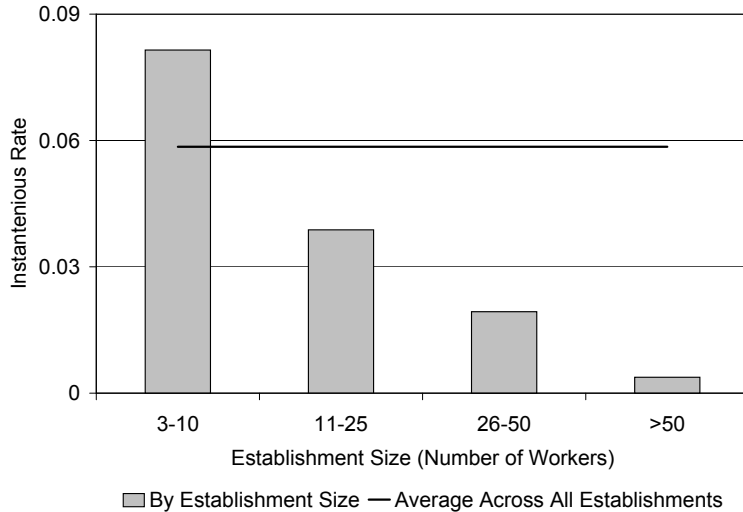


Figure 4: Instantaneous Rates of Discontinuing Insurance by Establishment Size.

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.⁷

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 statistic, consider the difference between the 90th and 10th percentile of the premium

⁷E.g., The U.S. Department of Health and Human Services Fact Sheet available at <http://www.hhs.gov/news/press/2000pres/20000110.html>.

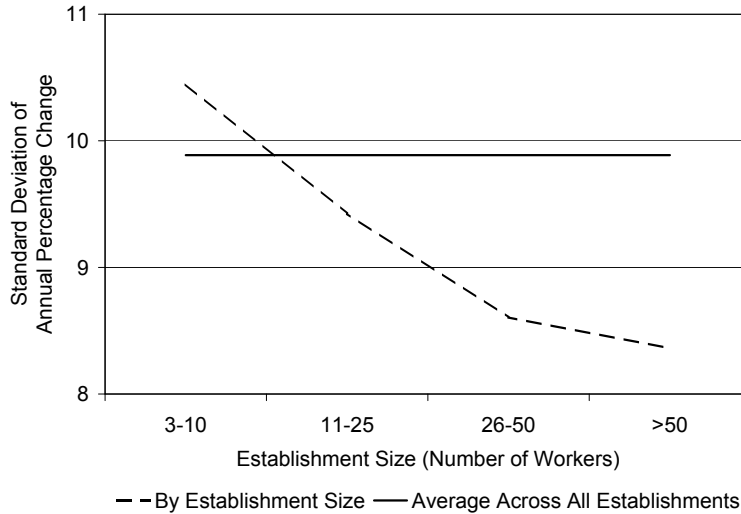


Figure 5: Variability of Insurance Premiums.

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.⁸

2.5 Sorting by Health Status across Employers

Despite the fact that large employers are more likely to provide coverage there is little sorting of workers across firms of different size based on their health status. For example, the 1996 wave of the Survey of Income and Program Participation (SIPP) provides information on workers' employer size (three categories: less than 25 employees, 25 to 99 employees and 100 or more employees) and self-reported worker health status (five

⁸Cutler (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 adjust for the different generosity of the plans. Finally, while our unit of analysis is an establishment, it is a firm in Cutler's study.

Table 1: Average Health Status of Employees by Establishment Size.

< 25	25 – 99	>= 100
2.00	1.97	1.98

categories: 1- Excellent, 2- Very Good, 3- Good, 4- Fair, 5- Poor). Table 1 contains the average health status of full-time workers depending on the employer size. The results suggest virtually no sorting across employer sizes based on health. Pauly and Herring (1999) who use the 1987 National Medical Expenditure Survey and Kapur, Escarce, Marquis, and Simon (2008) who use data from the Medical Expenditure Panel Survey from 1996 to 2001 also find very little evidence of sorting across employer sizes based on health.

Bundorf, Herring, and Pauly (2008) find a positive relationship between health risk and being covered by private health insurance. They find that this relationship is present for small, medium, and large groups.

2.6 Features of Observed Insurance Contracts

In practice group health insurance contracts have a typical length of one year. Explicit long-term group health insurance contracts are virtually non-existent. Legally, premiums can be adjusted almost freely upon renewal (subject to some restrictions imposed in several states). The conventional wisdom is that renewal premiums at least in part reflect changes in expected medical expenses of the insured establishment, a practice referred to as risk rating or experience rating.

To the extent that premiums are risk rated, establishments are insured against the risk that medical expenses within a year exceed expectations, but not against the risk of persistent shocks to medical expenses beyond the horizon of one year.⁹

There is no consensus in the literature as to what forces have led the system to rely on annual contracts with the possibility of risk rating. Some of the hypothesis offered in the literature include the presence of adverse selection and moral hazard, whereby had the insurers not risk 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

⁹For 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.

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 make no attempt to explain the features of observed contract. We take them as given as a feature of the environment and focus on the coverage decisions of employers in this environment. Since there is no direct evidence on the extent of risk rating, we ask for what level of risk rating the model comes closest to account for observed patterns of coverage and sorting.

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

2.7.1 No Mandate

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

2.7.2 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.¹⁰

¹⁰Under 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

2.7.3 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 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

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 from the tax-deductibility of contributions and earnings in 401(k) retirement plans (Table 19-1 in Office of Management and Budget (2008)).

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_{ii}^H for $i, i' \in \{h, u\}$. Health status in turn determines medical expenditures $e_u > e_h$. Let q_i^0 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 maximize the present discounted value of profits, with discount factor β . Establishments are subject to idiosyncratic productivity shocks: z indexes the productivity of an establishment, which follows a discrete Markov process with transition probabilities q_{zz}^Z . 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 δ every period.

An establishment can recruit new workers by posting vacancies. It posts $g_v \in \{0, 1, \dots\}$ 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

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

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. The health insurance premium may depend on the health composition of the employer. The extent of risk rating is parametrized

parsimoniously through a single parameter $\omega \in [0, 1]$. Specifically, we assume that a fraction ω of expected medical expenditures at the employer level is risk-rated, while the remainder is pooled across all employers. Health insurers charge an administrative load which may depend on the number of employees $\kappa(g_e)$ and applies to both the risk-rated and the pooled component of the premium.

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 $\mathcal{S}^L \equiv \mathcal{S} \cup \{s^L\}$ 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 0 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} . Let $\mathcal{G}^L \equiv \mathcal{G} \cup \mathbb{R}_+^3$ denote the set of all pure policies including those that can be chosen by very large establishments. 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 $J(\cdot) : \mathcal{S} \rightarrow \mathbb{R}$, with $J(s)$ giving the value of an establishment in state s .
2. An establishment policy function $\gamma(\cdot|\cdot)$; for s in \mathcal{S} , $\gamma(\cdot|s) \in \Gamma$ is the mixed policy for establishments in state s , and $\gamma(g|s)$ is the probability that an establishment in state s implements pure policy $g \in \mathcal{G}$; for $s = s^L$, $\gamma(g^L|s^L) = 1$ if $g^L = (g_h^L, g_u^L, g_v^L) \in \mathbb{R}_+^3$ is the policy vector of very large establishments, and $\gamma(g|s^L) = 0$ for all $g \neq g^L$.
3. A wage function $w(\cdot, \cdot) : \mathcal{S} \times \mathcal{G} \rightarrow \mathbb{R}$ with $w(s, g)$ giving the wage an establishment in state s would pay if it were to adopt policy g .
4. An insurance premium function $p(\cdot) : \mathcal{G}^L \rightarrow \mathbb{R}_+$.

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 $V_h(\cdot), V_u(\cdot) : \mathcal{S}^L \rightarrow \mathbb{R}$, 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 $C_h(\cdot, \cdot), C_u(\cdot, \cdot) : \mathcal{S} \times \mathcal{G} \rightarrow \mathbb{R}$, 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 $\mu(\cdot) : \mathcal{S}^L \rightarrow [0, 1]$ 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 ϕ .
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

$$u_i(w, g_I) = \iota(g_I)U((1 - \tau)w) + (1 - \iota(g_I)) \left[q_{ih}^H U((1 - \tau)w - e_h) + q_{iu}^H U((1 - \tau)w - e_u) \right],$$

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 $w(s, g)$:

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

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

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

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

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

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

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

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

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

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:

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

where $p(g)$ is the health insurance premium as a function of the pure policy g . The current state s , a pure policy g , tightness θ , and the fraction of healthy workers among searchers ν , 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

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

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

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

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

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

The current state s , a pure policy g , tightness θ , and the fraction of healthy workers among searchers ν , together induce a distribution over the establishment's future state. It is denoted $\mu_{ii'}(s' | s, g, \theta, \nu)$ 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

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

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

$$V_i^s = u_i^s + \beta\rho(1 - f(\theta)) [q_{ih}^H V_h^s + q_{iu}^H V_u^s] + \beta\rho f(\theta) \left\{ q_{ih}^H \sum_{s' \in \mathcal{S}^L} V_h(s') \mu_h^s [s' | \mu(\cdot), \gamma(\cdot|\cdot), \theta, \nu] + q_{iu}^H \sum_{s' \in \mathcal{S}^L} V_u(s') \mu_u^s [s' | \mu(\cdot), \gamma(\cdot|\cdot), \theta, \nu] \right\} \quad (11)$$

The second part of the continuation utility corresponds to contacting an establishment, while the first part applies in the absence of a contact. Here $\mu_i^s (s' | \mu(\cdot), \gamma(\cdot|\cdot), \theta, \nu)$ 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

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

4.5 Mass of Searchers

Let μ_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 μ_i the mass of workers which remain employed by an establishment at the time of matching:

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

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

$$m = m_h[\gamma(\cdot|\cdot), g^L, \mu(\cdot)] + m_u[\gamma(\cdot|\cdot), g^L, \mu(\cdot)], \quad (13)$$

$$\nu = \frac{m_h[\gamma(\cdot|\cdot), g^L, \mu(\cdot)]}{m}.$$

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

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

4.7 Very Large Establishments

There is a fixed mass λ^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 $w^L(V_h^s, V_u^s)$. The Bellman equation is

$$J^L(g_h, g_u) = \max_{g_v} \{ R^L(g_h, g_u, g_v) - w^L(V_h^s, V_u^s)(g_h + g_u) + \beta J^L[\mu^L(g|\theta, \nu)] \}$$

where

$$\begin{aligned} \mu_h^L(g|\theta, \nu) &\equiv \rho(1 - \delta) [q_{hh}^H \psi_h + q_{uh}^H \psi_u] + q(\theta) \nu g_v, \\ \mu_u^L(g|\theta, \nu) &\equiv \rho(1 - \delta) [q_{hu}^H \psi_h + q_{uu}^H \psi_u] + q(\theta)(1 - \nu) g_v. \end{aligned}$$

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

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

4.8 Entry of New Establishments

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

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

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

4.9 Insurance Premiums

The parameter $\omega \in [0, 1]$ governs the extent of risk rating. Specifically, a fraction $1 - \omega$ of expected medical expenditures is pooled across all employers that purchase coverage. Thus to determine the premium at the employer level we first need to determine average expected medical expenditures across these employers. They are given by

$$e[\gamma(\cdot|\cdot), \mu(\cdot)] \\ \equiv \sum_{s \in S^L} \sum_{g \in \mathcal{G}^L} \frac{g_h(q_{hh}^H e_h + q_{hu}^H e_u) + g_u(q_{uh}^H e_h + q_{uu}^H e_u)}{g_h + g_u} \frac{g_e \ell(g_I) \gamma(g|s) \mu(s)}{\sum_{s \in S} \sum_{g \in \mathcal{G}^L} g_e \ell(g_I) \gamma(g|s) \mu(s)}$$

The premium of an employer following policy g is then given by the weighted average of average expected medical expenditures for this employer and $e[\gamma(\cdot|\cdot), \mu(\cdot)]$, to which the administrative load is applied $\kappa(g_e)$ is applied:¹¹

$$p(g) = (1 + \kappa(g_e)) \left[\omega \frac{g_h(q_{hh}^H e_h + q_{hu}^H e_u) + g_u(q_{uh}^H e_h + q_{uu}^H e_u)}{g_h + g_u} + (1 - \omega) e[\gamma(\cdot|\cdot), \mu(\cdot)] \right] \quad (17)$$

4.10 Definition of Equilibrium

Definition 1 $J(\cdot)$, $\gamma(\cdot, \cdot)$, $w(\cdot, \cdot)$, $g(\cdot)$, $\mathcal{G}(\cdot)$, $V_h(\cdot)$, $V_u(\cdot)$, $C_h(\cdot, \cdot)$, $C_u(\cdot, \cdot)$, V_h^s , V_u^s , $\mu(\cdot)$, m , ν , ϕ , and θ constitute a stationary equilibrium if they satisfy equations (2), (6), (7), (8), (9), (10), (11), (16), (17), (13), (14), and (15).

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

5 Calibration

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

¹¹The 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.

5.1 Functional Forms

Utility Function. We use a constant absolute risk aversion utility (CARA) function¹²

$$U(c_t) = -\frac{\exp(-\varsigma c_t) - 1}{\varsigma}. \quad (18)$$

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

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

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.¹³ 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

$$\ln(z') = \zeta(1 - \varphi) + \varphi \ln(z) + \epsilon', \quad (19)$$

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:

$$M(m, v) = \chi m^\alpha v^{1-\alpha}. \quad (20)$$

Thus, there are two parameters, χ , α , that characterize the matching function.

¹²We use CARA instead of the more standard CRRA utility function since it facilitates some steps in the computation.

¹³Our 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.

5.2 Parameters

The model parameters to be calibrated and their definition are shown in the first two columns of Table 2. We begin by discussing parameters which are calibrated independently, and then discuss the moments we match to jointly calibrate the remaining parameters.

We choose the model period to be one month in order to accommodate the high frequency of transitions in the labor market. We set $\beta = 1/(1+r)$, where r corresponds to an annual interest rate of 4%. 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.

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 $\eta = 0.85$.

The empirical counterpart of very large establishment in the model are establishments employing more than 250 workers. While accounting for 30% of employment, such establishments represent only 0.7% of all establishments, hence we set $\lambda_L = 0.007$.

The matching function elasticity parameter α 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).)

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 adopt a piecewise linear schedule such that the average administrative load within each size category matches these estimates.

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 $\tau = 0.4$.¹⁴

¹⁴Mendoza, Razin, and Tesar (1994) report a lower value $\tau = 0.3$ for the average labor tax rate.

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 probabilities $q_{ii'}^H$ 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 $q_{hh}^H = 0.9892$, $q_{hu}^H = 0.0108$, $q_{uh}^H = 0.1636$, and $q_{uu}^H = 0.8364$. We assume that all the new labor market entrants are healthy so that $q_h^0 = 1$ and $q_u^0 = 0$. Notice that this specification is such that even healthy workers obtain some value from health insurance. First, even if they remain healthy they have some medical expenditures e_h , which could be interpreted as routine medical expenditures for healthy individuals. Second, there is a chance that they may become unhealthy within the period.

The remaining parameters are determined jointly by matching a set of moments. For each parameter we selected a moment that we consider informative about the value of this parameter. Therefore, while the parameters are determined jointly, the following discussion associates each parameter with a specific moment.

The median and average size of establishments with 250 workers or less are 4 and 11.03, respectively. The mean ζ and standard deviation σ_ϵ of idiosyncratic productivity are chosen to match these values. We choose its persistence to match the quarterly job creation rate of continuing establishments with 250 workers or less, which is 7% in the Business Employment Dynamics (BED) statistics provided by the BLS. 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 very large establishment in equilibrium z^L .

Hagedorn and Manovskii (2008) report an average monthly job finding rate of 0.45, and an 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 p , where labor productivity p is defined as output per worker. These three targets are used to calibrate the matching function scale parameter χ , the entry cost c_e , and the vacancy cost c .

The fixed cost of initiating coverage b_I and the probability of commitment lapsing q_I are chosen to match evidence on the level and dynamics of coverage from Section 2. Specifically, we target average coverage of 56%, and an average two-year discontinuation probability of 11%, respectively.

Finally, the extent of risk rating ω is chosen to match a moment that captures the extent of sorting of unhealthy workers into large employers. Specifically, using MEPS data we estimate a model to predict medical expenditures based on age, gender, and chronic conditions.¹⁵ We then compute average predicted medical expenditures for employees working for employers with at least 26 employees, and the corresponding measure for employers in the 3-25 range. Their ratio is 1.15, indicating that employees of large employers are relatively less healthy. This ratio is used as the calibration target for ω .

Given the calibration approach of the other parameters described above, the choice of the value of non-market activity is simply a normalization, and we choose it to normalize the average wage in the economy to one.

Calibrated parameter values are shown in the third column of Table 2.

¹⁵¹⁶ We then compute average predicted medical expenditures for employees working for employers with at least 26 employees, and the corresponding measure for employers in the 3-25 range.

Table 2: Calibration.

Parameter	Definition	Value
β	the time discount rate	0.996
ς	coefficient of absolute risk aversion	0.5
ρ	the probability of a worker staying in the labor market	0.9979
η	curvature of the production function	0.85
λ^L	mass of very large establishments	0.007
α	elasticity of the matching function	0.5
$\kappa(g_e)$	administrative load for employer provided insurance	see text
τ	proportional tax on labor income	0.4
e_u	health expenditures of unhealthy workers	0.5555
e_h	health expenditures of healthy workers	0.0158
q_{hh}^H	the health status transition probability	0.9892
q_{uu}^H	the health status transition probability	0.8364
q_h^0	fraction of new entrants who are healthy	1.00
ζ	unconditional mean of idiosyncratic productivity	1.5385
σ_ϵ	st. dev. of innovations in idiosyncratic productivity	0.0207
φ	persistence of idiosyncratic productivity	0.9963
z^L	productivity of very large establishments	3.34
χ	scale parameter of the matching function	0.6364
c_e	establishment entry cost	478.44
c	cost of maintaining a vacancy	0.58 <i>p</i>
b_I	fixed cost of initiating coverage	0.65
q^I	probability of commitment lapsing	0.0278
ω	extent of risk rating	0.15
b	value of non-market activity	0.6247

Note - The table contains the parameters that are invariant across alternative calibrations.

Table 3: Matching the Calibration Targets.

Target	Data	Model
Mean size of an establishment with ≤ 250 workers	11.03	11.06
Job creation rate of establishment with ≤ 250 workers	0.07	0.07
Median size of an establishment with ≤ 250 workers	4	4
Mean size of an establishment with > 250 workers	679.37	680.6
Job finding rate	0.45	0.45
Labor market tightness	0.634	0.634
Ratio Vacancy Cost and Labor Productivity	0.58	0.58
Fraction of establishments providing health insurance	0.56	0.56
Fraction of establishments discontinuing insurance over 2 years	0.11	0.14
Ratio Med. Exp. Large (≥ 26) vs. Small (3 – 25)	1.15	1.15
Average wage (normalization)	1.00	1.00

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

Table 4: Results from Benchmark and Alternative Calibrations

	Benchmark	Alternative Calibrations			
	(I)	(II.1)	(II.2)	(II.3)	(II.4)
<i>Fraction of Establishments Providing Coverage</i>					
3-10	0.264	1.000	0.888	0.903	0.236
11-25	0.894	1.000	0.733	0.825	0.916
26-50	0.966	1.000	0.632	0.827	0.942
51-	0.991	1.000	0.580	0.932	0.961
all	0.562	1.000	0.792	0.877	0.547
<i>Frac. of Est. Providing Coverage that Commit</i>					
3-10	0.683	0.000	0.000	0.041	0.609
11-25	0.870	0.000	0.000	0.198	0.656
26-50	0.999	0.000	0.000	0.599	0.962
51-	0.999	0.000	0.000	0.997	0.999
all	0.872	0.000	0.000	0.238	0.771
<i>Frac. of Est. Discontinuing Insurance over 2 Years</i>					
3-10	0.361	0.000	0.111	0.092	0.305
11-25	0.064	0.000	0.263	0.162	0.055
26-50	0.024	0.000	0.365	0.148	0.044
51-	0.008	0.000	0.505	0.060	0.034
all	0.141	0.000	0.188	0.113	0.127
<i>Establishment Size Distribution</i>					
3-10	0.346	0.348	0.357	0.341	0.347
11-25	0.136	0.153	0.123	0.145	0.137
26-50	0.090	0.087	0.097	0.088	0.090
51-	0.051	0.059	0.048	0.055	0.057
<i>Fraction of Unhealthy Workers among all Employees</i>					
3-10	0.026	0.065	0.000	0.003	0.009
11-25	0.065	0.063	0.000	0.015	0.010
26-50	0.061	0.057	0.000	0.036	0.045
51-	0.065	0.068	0.075	0.075	0.070
ν	0.919	0.925	0.715	0.805	0.860

Note - The table contains the results for the benchmark and alternative calibrations. Column (I) – Benchmark Calibration (Non-discrimination, commitment, partial risk rating, cost of starting).

Column (II.1) – No non-discrimination, no commitment, full risk rating.

Column (II.2) – Non-discrimination, no commitment, full risk rating.

Column (II.3) – Non-discrimination, commitment, full risk rating.

Column (II.4) – Non-discrimination, commitment, full risk rating, cost of starting.

ν Denotes the fraction of healthy workers among searchers.

Table 5: Parameter Values of Benchmark Calibration.

Parameter	Definition	Value
ζ	unconditional mean of idiosyncratic productivity	1.505
φ	persistence of idiosyncratic productivity	0.9967
σ_ϵ	st. dev. of innovations in idiosyncratic productivity	0.0207
z^L	productivity of very large establishments	3.34
c^e	establishment entry cost	478.12
b	value of non-market activity	0.6247
b_I	fixed cost of initiating coverage	0.65
q^I	probability of commitment lapsing	0.0278

Note - The table contains the parameters that together with Table 2 constitute the benchmark calibration.

6 Performance of the Calibrated Model

Table 3 shows that the calibrated model is successful in matching the calibration targets. Column (I) of Table 4 shows that the model also captures empirical regularities discussed in Section 2 which were not targeted in the calibration. Specifically, coverage is increasing in establishment size, while the probability of discontinuance is decreasing in size. Quantitatively the slope of both relationships is somewhat steeper than in the data, however.

Our goal in this section is to demonstrate the importance of several key elements of the model for its ability to account for these features of the data. Specifically, we study non-discrimination rules, commitment, discreteness of employer size, labor market frictions, costs of starting coverage, and the extent of risk rating. We calibrate versions of the model without these elements, and show how these versions of the model fail to capture the empirical regularities described in Section 2. In Section 7.2 we will assess the importance of these model elements from another angle, asking whether omitting them would change our conclusions concerning the effects of the health insurance reform of 2010.

6.1 Non-discrimination

We start by calibrating a very basic version of the model without non-discrimination rules. The model has no commitment ($q_I = 1$), no cost of starting coverage ($b_I = 0$), and full risk rating ($\omega = 1$). Since the three parameters q_I , b_I and ω are now constrained, we drop their associated targets in the calibration: the discontinuance probability, the fraction of establishments providing coverage, and the measure of sorting.

Results are shown in Column (II.1) of Table 4. All employers provide coverage continuously. Without non-discrimination provisions employers make both healthy and unhealthy employees indifferent between staying and leaving. It is then optimal for employers to provide coverage as long as the tax subsidy and the gain employees obtain from insurance against medical expenditure risk within the period are not outweighed by administrative loads. It turns out that this condition is met for all employer sizes. In equilibrium both types of workers obtain their autarky utility, hence there is no risk pooling across individuals.

Non-discrimination provisions are introduced into this basic version of the model in Column (II.2). Coverage is no longer complete in this calibration, but *decreasing* in employer size. Nonetheless, unhealthy workers sort into large employers.

To understand the source of these failures of the model without commitment, it is instructive to analyze the trade-off faced by employers in the presence of non-discrimination rules. With these rules in place, employers face a trade-off between the value of the tax subsidy and within-period insurance to healthy workers on one hand and rents left to unhealthy workers on the other hand. Consider an employer making retainment and recruitment decisions (g_h, g_u, g_e, g_v) , and who makes the healthy indifferent. Conditional on these choices, providing coverage is optimal if

$$\left[\frac{\tau}{1 - \tau} - \kappa(g_e) \right] e_h^E + \Delta w_h^{\text{risk}} \geq \frac{g_u}{g_u + g_h} (1 + \kappa(g_e)) (e_u^E - e_h^E), \quad (21)$$

where e_h^E and e_u^R denote the medical expenditures of healthy and unhealthy workers expected at the beginning of the period, respectively. The left-hand side is the gain to the employer of providing coverage to a healthy worker. The first term is the gain from the tax subsidy after deducting administrative loads. The second term Δw_h^{risk} is the amount by which the wage of the healthy can be reduced due to the fact that they value within-period insurance.¹⁷ If the employer has only healthy employees, then coverage

¹⁷It is a function of only e_h^E , e_u^E , τ , and the coefficient of absolute risk aversion ς .

is optimal if the left-hand side is non-negative. This is true for all employer sizes. The right-hand side is the additional cost associated with coverage due to the fact that a given worker may be unhealthy. This is given by the chance that a given worker is unhealthy times the additional expected medical expenditures associated with an unhealthy worker. Importantly, given the calibrated parameter values, condition (21) fails if the fraction of unhealthy $\frac{g_u}{g_u+g_h}$ in the establishment equals the fraction of unhealthy in the population.

6.2 Discreteness

We can now explain why the level of coverage is decreasing in employer size. According to condition (21) an employer provides coverage if the health composition of its workforce is sufficiently favorable, that is, if average medical expenditures are below some threshold. Now consider an employer for which this requirement is violated in the current period, thus no coverage is provided. Employees with high expected medical expenditures will tend to leave this employer. But as they leave, the odds of future coverage improve, and the exodus of the unhealthy will stop eventually. The key difference between small and large employers is where this process stops, and the source of this difference is that for small employers the last employee that leaves has a larger mass relative to the employer's size. The mechanism is most easily explained using a hypothetical example. For simplicity assume that employees only take into account the probability of coverage next period (of course, in the model they take into account the entire future path of coverage). Consider a worker who would be willing to stay if the probability of coverage in the next period is 50%. Now consider the following question: is it possible that in equilibrium this employee leaves although once he has left the probability of coverage is 75%. The answer is negative for sufficiently large employers: the impact of a single employee staying on the probability of future coverage is small. Hence the probability of coverage would be close to 75% if this employee stays, thus staying is optimal. The answer could be affirmative for a small employer: if this single employee stays, the adverse effect on the employer's health composition can be sufficiently large for the probability of coverage to drop below 50%. Thus, on average, not providing coverage in the current period induces a larger improvement of the health composition in smaller employers, thus a higher probability that coverage will resume in the future. This higher probability of resuming coverage for small employers generates the pattern that coverage rates are declining in employer size.

6.3 Commitment

Next we add commitment to the basic model just analyzed, while maintaining the restrictions of zero cost of starting ($b_I = 0$) and full risk-rating ($\omega = 1$). The calibration target associated with the probability of commitment lapsing is the average rate at which employers discontinue coverage over two years. Notice that this rate is 18.8% in the preceding calibration, while in the data it is 11%. This rate turns out to be decreasing in the extent of commitment, that is, it is increasing in the probability with which commitment lapses q_I , and the latter can be chosen to match the observed discontinuation rate of 11%. Results for this calibration are shown in Column (II.3) of Table 4. It exhibits a substantially higher coverage rate among large employers. The source of this difference is that commitment relaxes the static trade-off represented by condition (21) by introducing a dynamic component: committing to coverage makes it more likely that healthy employees will have coverage if they become unhealthy in the future, enabling the employer to reduce their wage more today. The magnitude of this dynamic component is endogenous. In particular, the more other employers commit to coverage, the less any given employer benefits from doing so, as it becomes easier for its employees to find a different employer with coverage when becoming unhealthy. This force limits the fraction of establishments that commit in the calibrated model: as we vary q_I while recalibrating to match previous targets, we find that for no value of q_I does the model predict that all employers commit. Notice that it is large employers who commit, while small employers do not. Even in the calibration without commitment, small employers are able to provide coverage most of the time, thus exploit the tax subsidy most of the time, and this remains true here. It is large employers that have a comparative advantage in committing, as without commitment they discontinue frequently, thereby foregoing the tax subsidy.

Level and dynamics of coverage among small employers remain similar to the calibration without commitment. They discontinue whenever an employee is unhealthy, but this does not lead to a low level of coverage as unhealthy workers leave and coverage is resumed in the following period. Thus this calibration fails to account for the low level of coverage observed among small employers.

6.4 Cost of Starting

Next we relax the restriction of zero cost of starting and recalibrate the model with b_I as an additional parameter. Recall that the target associated with this parameter is the

average fraction of establishments providing coverage. Results are shown in Column (II.4) of Table 4. The level of coverage is now increasing in size, and the rate of discontinuation is decreasing in size, although the slope of both relationships is somewhat steeper than in the data.

6.5 Role of Frictions under Full Risk Rating

With full risk rating and given the calibrated medical expenditure process, labor market frictions are essential for any risk sharing to occur between unhealthy and healthy workers. With a frictionless labor market healthy workers do not value any commitment their current employer may offer to provide coverage in the future. Hence the trade-off faced by employers is given by condition (21). Since this condition is violated for the fraction of unhealthy in the population, it follows that in the absence of labor market frictions there does not exist an equilibrium in which healthy and unhealthy employees are working for the same employer. In other words, any equilibrium must be separating.

It is then clear that the severity of labor market frictions will matter quantitatively for the extent of risk sharing. To illustrate, in Columns (I.1)-(I.3) of Table 6 we show calibration results obtained by targeting a counterfactual job-finding probability which is double the observed probability. As a reference point, Column (I.1) shows the calibration with full risk rating, commitment, and cost of starting, taken from Column (II.4) of Table 4. As an intermediate step, Column (I.2) shows the effect of doubling the job-finding probability (through the matching function scale parameter χ) without recalibrating the model. Everything else equal, the higher job-finding probability enables unhealthy workers to more quickly find an employer that is offering coverage. Thus currently healthy workers value commitment to coverage less, which induces a substantial drop in coverage to only 36%. When attempting to recalibrate the model, we encounter the difficulty that the model can no longer simultaneously match the level of coverage and the discontinuance probability. The reason for this failure is as follows. The only way to restore the level of coverage is to reduce the cost of starting for small employers. But this drives up the average discontinuance probability, as this probability is relatively high for small employers. Column (I.3) shows the results from recalibrating the model to match the level of coverage while minimizing the discontinuance probability. Notice that the reduction in frictions leads to much stronger sorting of unhealthy workers into large employers.

6.6 Extent of Risk Rating

One feature this calibration shares with the previous two calibrations with full risk rating is that there is very strong sorting of unhealthy workers into larger employers. Our measure of sorting, the ratio of average predicted medical expenditures for large (at least 26 employees) and small (3-25 employees) employers is 1.4, while it is 1.15 in the data based on our estimates from the MEPS. Small employers never offer coverage when they have both healthy and unhealthy employees when given the choice.¹⁸ Thus unhealthy workers tend to sort into large employers.

Sorting is weakened if risk rating is less than complete. If risk rating is sufficiently weak, then forces of adverse selection become dominant and employers offer coverage when the health composition of their workforce is sufficiently poor. It is instructive to first consider this extreme case before turning to the case in which the extent of risk rating is calibrated. With very weak risk rating, small employers are still less likely to offer coverage than large employers: since their health composition is more dispersed, they are more likely to have a very healthy composition, in which case they do not offer coverage. But here small employers are likely to offer coverage precisely when they have many unhealthy workers, hence unhealthy workers are less likely to leave and sort into larger employers.

Thus by varying the extent of risk rating through the parameter ω the model is able to match the observed value of our measure of sorting in our benchmark calibration. The value of $\omega = 0.15$ in the benchmark calibration is quite close to community rating, indicating relatively weak risk rating. As discussed in the introduction, this contrasts with the conventional wisdom that risk rating has been quite severe.

6.7 Role of Frictions in Benchmark Calibration

The limited extent of risk rating in the benchmark calibration implies that employers tend to provide coverage when the health composition of their workforce is relatively poor. Thus forces of adverse selection are at work, but they are muted by labor market frictions. In the absence of labor market frictions employers that offer coverage would have an incentive to hire only unhealthy employees. In the present model replacing healthy with unhealthy workers requires costly creation of vacancies. The role of frictions in limiting adverse selection is illustrated in Columns (II.1)-(II.3) of Table 6 through

¹⁸Some small employers are constrained to offer coverage because they committed to do so when large.

a calibration exercise that targets a level of vacancy costs which is half the value in the benchmark calibration. For comparison, Column (II.1) shows the benchmark calibration. As an intermediate step, Column (II.2) shows the effect of halving vacancy costs without recalibrating the model. Stronger adverse selection induces a large drop in average coverage to 23%. When attempting to recalibrate the model, we again face the difficulty that the model cannot simultaneously match the level of coverage and the discontinuance probability. Here the tension arises as restoring coverage requires a stronger commitment technology, but stronger commitment reduces the discontinuance probability. Column (I.3) shows the results from recalibrating the model to match the level of coverage while maximizing the discontinuance probability. All but the smallest employers never discontinue, and again the reduction in frictions induces stronger sorting of unhealthy workers into large employers.

6.8 Size Dependent Cost of Initiating Coverage

In this section we show that an alternative modeling choice to the commitment technology with similar implications is to make the cost of initiating coverage size dependent. Specifically, we now assume that the cost of starting coverage is proportional to establishment size $b_I(g_e) \equiv b_I g_e$. Instead of the probability of commitment lapsing q^I (now set to one) we use the slope parameter b_I to target the rate at which coverage is discontinued over two years. Appendix Table A-1 shows how results vary with the level of b_I for the case of full risk rating. The observed discontinuation rate is matched for $b_I = 1.4$, that is a cost of initiating coverage of more than one month of wages per employee. 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 a high cost of stopping coverage.

7 Effects of the Health Insurance Reform of 2010

7.1 Effects in the Benchmark Calibration

The Patient Protection and Affordable Care Act (PPACA) was signed into law on March 23, 2010. Its key provisions related to health insurance are

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. Rating variations for small group and individual coverage will be limited to geographic area, age, and tobacco use.
3. Individuals will be required to have “acceptable health coverage.” This is enforced through a tax penalty of the greater of \$695 per year up to a maximum of three times that amount per family or 2.5% of household income.
4. Employers with more than 50 employees will be required to offer coverage or pay a fee of \$2000 per full-time employee, excluding the first 30 employees from the assessment.
5. Employers with less than 25 employees will be provided with a tax credit of up to 50% of the employer’s contribution toward the employee’s health insurance premium. This credit is available throughout the years 2010-2013. Starting in 2014 the credit becomes more generous, but each employer can claim the credit only for two consecutive years.¹⁹
6. 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, 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.

In this section we evaluate the consequences of stylized versions of these changes. In particular, we introduce the following changes to the benchmark calibration:

1. A community rated insurance market for establishments with less than 50 employees where premiums cannot depend on the health composition of the workforce.
2. A tax subsidy of 30% of premiums paid by employers with less than 25 employees. Notice that this experiment does not yet capture the two-year limit on the tax credit employers will face starting in 2014, a point we return to when interpreting the results.

¹⁹See Peterson and Chaikind (2010).

3. A tax equal to 8% of payroll for employers with more than 50 employees if they choose to not provide coverage.
4. Individuals can participate in the community-rated small group market.
5. Individuals must pay a penalty of 2.5% of their income from wages if they do not have coverage.

The results for introducing all the elements of this stylized reform simultaneously are presented in Column (II.6) of Table 7. In Columns (II.1) through (II.5) we introduce the components of the reform sequentially, starting with community rating for small groups in Column (II.1). Column (II.2) introduces the premium subsidy for small employers. Column (II.3) considers community rating for small groups without a premium subsidy, but combined with the penalty for large employers. The joint effect of these three components of the PPACA is shown in Column (II.4). The next two columns combine this with the provisions for individual coverage. Column (II.5) gives individuals access to the community-rated small group market, while Column (II.6) contains the full reform including the individual mandate.

Column (II.1) shows that introducing community rating for employers with less than 50 workers by itself results in an effective collapse of coverage by these employers. This is because of severe adverse selection induced by such a reform. Since coverage among small employers collapses, the health composition among workers in large employers 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 the collapse of coverage by community rated employers that are eligible for the subsidy. This also prevents the negative equilibrium spillover effect on coverage of larger employers. Coverage remains low in the 26-50 range, that is, among employers subject to community rating but ineligible for the subsidy. The size-dependent design of this policy, where employers with less than 50 workers are community rated but only employers with 25 workers get the subsidy, implies sizable effects on the establishment size distribution. Employers with 26 to 50 workers before the reform tend to shrink in order to qualify for the subsidy. Smaller employers, however, tend to expand since the subsidy reduces their labor costs. Overall, the fraction of employers with 11 to 25 workers increases by 40.4%

while the fraction of employers with 26 to 50 workers declines by 47.8%. Notice that this experiment does not take into account the two-year limit on claiming this credit which employers will face starting in 2014. This limit implies that at a point in time only a fraction of small employers is eligible for the credit. This weakens the stabilizing effect of the credit, thus it may not suffice to stop the collapse in coverage, a possibility we will examine in the next draft.

Column (II.3) combines the community rating for employers 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 employers does not prevent the collapse of insurance for employers with less than 50 workers, but it counteracts the negative equilibrium effect on coverage of larger employers observed in column (II.1).

Column (II.4) shows that the penalty for large employers has little effect if the tax subsidy for small employers is in place, so results are very similar to Column (II.2). While the penalty counteracts the negative equilibrium spillover effect on coverage of larger employers if community rating is not accompanied by a subsidy, the subsidy also largely eliminates this spillover effect, leaving only a minor role for the penalty.

Column (II.5) shows that giving individuals access to the community-rated small group market once again leads to unraveling due to severe adverse selection. While even small employers with a poor health composition often bring some healthy workers into the risk pool, only the unhealthy enter the pool as individuals.

Simulation of the full reform in Column (II.6) shows that the individual penalty is sufficient to prevent this unraveling, and induces all employers to offer coverage. In particular, the increased demand of workers for coverage induced by the penalty induces a large increase in coverage for the 26-50 size group compared to Column (II.4), that is, the group that is too large for the tax credit, but too small for employer penalties.

Table 6: Role of Frictions

	Full Risk Rating, High f			Partial Risk Rating, Low c		
	(I.1)	(I.2)	(I.3)	(II.1)	(II.2)	(II.3)
<i>Fraction of Establishments Providing Coverage</i>						
3-10	0.24	0.10	0.36	0.26	0.03	0.31
11-25	0.92	0.64	0.77	0.89	0.13	0.66
26-50	0.94	0.75	0.74	0.97	0.61	0.95
51-	0.96	0.91	0.86	0.99	0.90	1.00
all	0.55	0.36	0.55	0.56	0.23	0.56
<i>Frac. of Est. Discontinuing Insurance over 2 Years</i>						
3-10	0.30	0.66	0.42	0.36	0.65	0.05
11-25	0.06	0.28	0.20	0.06	0.39	0.00
26-50	0.04	0.20	0.22	0.02	0.20	0.00
51-	0.03	0.06	0.09	0.01	0.08	0.00
all	0.13	0.32	0.27	0.14	0.27	0.02
<i>Fraction of Unhealthy Workers among all Employees</i>						
3-10	0.010	0.000	0.001	0.030	0.000	0.021
11-25	0.046	0.011	0.018	0.065	0.011	0.047
26-50	0.056	0.026	0.025	0.061	0.046	0.061
51-	0.065	0.084	0.079	0.064	0.073	0.067

Note - The table contains the results for the benchmark and alternative calibrations.
Column (I.1) – Full Risk Rating Calibration (Non-discrimination, commitment, cost of starting), also see Column (II.4) of Table 4.
Column (I.2) – As Column (I.1), but job-finding rate f doubled without recalibrating.
Column (I.3) – As Column (I.2), but recalibrated.
Column (II.1) – Benchmark Calibration.
Column (II.2) – As Column (II.1), but vacancy cost c halved without recalibration.
Column (II.3) – As Column (II.2), but recalibrated.

Table 7: Benchmark Calibration and the Effects of the Health Insurance Reform of 2010

	Benchmark	Reform Experiments					
	(I)	(II.1)	(II.2)	(II.3)	(II.4)	(II.5)	(II.6)
<i>Fraction of Establishments Providing Coverage</i>							
3-10	0.264	0.001	0.839	0.000	0.894	0.000	1.000
11-25	0.894	0.000	0.942	0.000	0.947	0.000	1.000
26-50	0.966	0.000	0.151	0.000	0.181	0.000	1.000
51-	0.991	0.621	0.946	1.000	1.000	1.000	1.000
all	0.562	0.056	0.829	0.103	0.870	0.102	1.000
<i>Fraction of Establishments Providing Coverage that Commit</i>							
3-10	0.683	0.972	0.012	0.000	0.524	NA	NA
11-25	0.870	NA	0.043	NA	0.114	NA	NA
26-50	0.999	1.000	0.726	NA	0.921	NA	NA
51-	0.999	0.742	0.946	0.124	0.940	0.124	NA
all	0.872	0.745	0.138	0.124	0.443	0.124	NA
<i>Fraction of Est. Discontinuing Insurance over a 2 Year Period</i>							
3-10	0.361	0.809	0.026	1.000	0.005	1.00	0.000
11-25	0.064	1.000	0.028	1.000	0.027	1.00	0.000
26-50	0.024	0.997	0.791	1.000	0.755	1.00	0.000
51-	0.008	0.198	0.046	0.000	0.000	0.00	0.000
all	0.141	0.390	0.077	0.366	0.056	0.376	0.000
<i>Establishment Size Distribution</i>							
3-10	0.346	0.366	0.330	0.369	0.330	0.334	0.353
11-25	0.136	0.115	0.191	0.117	0.191	0.148	0.181
26-50	0.090	0.084	0.047	0.076	0.044	0.078	0.061
51-	0.051	0.055	0.060	0.064	0.064	0.064	0.045
mean	16.00	16.22	15.96	16.40	16.02	16.77	15.66
median	4	4	5	4	5	4	5
<i>Fraction of Unhealthy Workers among all Employees</i>							
3-10	0.065	0.001	0.061	0.000	0.061	0.000	0.60
11-25	0.063	0.000	0.065	0.000	0.064	0.000	0.63
26-50	0.057	0.001	0.014	0.023	0.018	0.023	0.60
51-	0.068	0.077	0.065	0.081	0.064	0.079	0.61
Avg. Wage	1.006	1.016	1.000	1.011	1.000	1.010	0.962
Avg. Labor Cost	1.048	1.040	1.029	1.030	1.032	1.043	1.012
Output per Worker	1.287	1.262	1.266	1.262	1.265	1.256	1.268
m	0.068	0.093	0.065	0.083	0.064	0.092	0.029
ν	0.919	0.776	0.913	0.813	0.920	0.83	0.943
θ	0.634	0.535	0.720	0.578	0.728	0.45	3.300

Notes to Table 7 -

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.

Column (II.5) – Same as Column (II.4) but individuals have access to the community-rated small group market.

Column (II.6) – Same as Column (II.5) but a penalty is imposed on individuals without coverage.

Description of Rows:

See Notes to Table 4 for the description of the rows.

7.2 Effects in Alternative Calibrations

In Section 6 we assessed the importance of several elements of the model by demonstrating that without these elements the model cannot account for certain features of the data. This does not necessarily mean, however, than models without these elements would yield different conclusions concerning the effects of policies. In this section we will examine the latter question for the case of the 2010 health insurance reform. Results from these experiments will be provided in future drafts of the paper.

8 Further Analysis

8.1 Labor Market Effects of the Current System

In this section we will study the labor market effects of the employer-provided health insurance system, by contrasting it with the hypothetical scenario in which all workers are perfectly insured against medical expenditure risk, irrespective of their employment status. As workers are already insured, employers no longer play a role in providing health insurance in this scenario. The fact that larger employers are more likely to provide coverage than small employers may play a role in shaping the establishment size distribution of the benchmark economy, and the experiment of this section allows us to identify such effects. It also allows us to identify the effect on non-discrimination provisions on the level of wages, as well as effects on labor market transitions at the individual level. Results for this experiment will be added in the next draft.

8.2 Community Rating.

A key element of the 2010 health insurance reform is community rating for small employers. When analyzing the PPACA in Section 7, we found that by itself community rating among small employers induces a collapse of coverage, but that other provisions in the PPACA effectively prevent this collapse. Community rating was already implemented in several states in the 1990s. In this section we use our calibrated model to investigate community rating in more detail, as it can be implemented in different ways. In particular, we examine under what circumstances community rating does and does not induce a collapse of coverage when it is not accompanied by subsidies and penalties of the kind contained in the PPACA. One variable is the set of employers over which medical expenditures are pooled. Another choice is whether only medical expenditures

are pooled, or whether administrative costs are spread across employers as well.

In this section we explore community rating under varying assumptions. 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 together with pooling of administrative loads is essential to sustaining coverage. The low administrative loads of very large establishments 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. 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 risk rating to community rating may increase or decrease coverage, depending on the extent to which the scheme pools administrative expenses.

9 Conclusion

In this paper we studied the system of employer-provided health insurance in the US using a quantitative equilibrium model of the labor market. Employers vary in size and decide whether to provide health insurance coverage. Workers sort into employers subject to search frictions. The model incorporates government regulation of the system, specifically tax subsidies to employer-provided coverage and non-discrimination rules. The model is consistent with key empirical regularities describing the performance of the existing system. We showed that several elements of the model are important for the model's ability to account for these regularities. If we eliminate employers' ability to commit to providing health insurance, then the model counterfactually predicts that coverage is declining in employer size. This led us to conclude that government regulation alone is insufficient to account for observed patterns of coverage, and that employer commitment must play an important role. If we impose that coverage sold by health insurers is fully risk-rated, then the model counterfactually generates strong sorting of unhealthy workers into large employers. This led us to conclude that health insurers provide more pooling of health expenditure risk across employers than suggested by conventional wisdom.

We used the calibrated model to analyze stylized versions of changes to the system of employer-provided insurance enacted as part of the Patient Protection and Affordable Care Act (PPACA) of 2010. We found that these changes lead to large gains in aggregate coverage, and that both penalties imposed on individuals without coverage as well as tax credits for small employers included in the law play a key role in generating this outcome. We also found that the small employer tax credits induce sizable changes in the establishment size distribution.

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 one's 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

- BRÜGEMANN, B. (2008): “What Elasticity of the Matching Function is consistent with U.S. Aggregate Labor Market Data?,” Working paper, Yale University.
- BUNDORF, M. K., B. HERRING, AND M. V. PAULY (2008): “Health Risk, Income, and Employment-Based Health Insurance,” Working Paper 11677, National Bureau of Economic Research.
- 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.
- HAGEDORN, M., AND I. MANOVSKII (2008): “The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited,” *American Economic Review*, 98(4), 1692–1706.
- HERRING, B., AND M. V. PAULY (2001): “Premium Variation in the Individual Health Insurance Market,” *International Journal of Health Care Finance and Economics*, 1, 43–58.
- (2006): “The Effect of State Community Rating Regulations on Premiums and Coverage in the Individual Health Insurance Market,” Working Paper 12504, National Bureau of Economic Research.
- 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.
- KAPUR, K., J. J. ESCARCE, M. S. MARQUIS, AND K. I. SIMON (2008): “Where Do the Sick Go? Health Insurance and Employment in Small and Large Firms,” *Southern Economic Journal*, 74(3), 644–664.
- 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.

- PAULY, M., AND B. HERRING (1999): *Pooling Health Insurance Risks*. Washington, D.C.: The AEI Press.
- PETERSON, C. L., AND H. CHAIKIND (2010): “Summary of Small Business Health Insurance Tax Credit Under PPACA (P.L. 111-148),” Report 7-5700, Congressional Research Service.
- 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.

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 establishment is still deterministic. Next workers are dismissed at random, leaving g_e workers in total. The probability of arriving at the workforce (ψ_h, ψ_u) after random dismissal is given by

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

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 ψ_h healthy and ψ_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 (ψ_h, ψ_u) to (ψ'_h, ψ'_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 (ψ'_h, ψ'_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 ψ_u can join from status u . The probability of j workers remaining in status h is given by $B(j; \psi_h, q_{hh}^H)$. 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 ψ'_h requires that exactly $\psi'_h - j$ switch from status u to status h . The latter happens with probability $B(\psi'_h - j; \psi_u, q_{uh}^H)$. Thus

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

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 δ . Thus a worker stays with the employer with probability $\rho(1 - \delta)$. The probability of a transition from (ψ_h, ψ_u) to (ψ'_h, ψ'_u) in this step is

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

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 (ψ_h, ψ_u) to (ψ'_h, ψ'_u) is

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

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 ν . Let $Q^{vac}(g, \theta, \nu)$ denote the transition matrix.

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

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

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

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

We derive $\mu_{hh'}[s'|s, g, \theta, \nu]$, 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 $(\hat{\psi}_h, \hat{\psi}_u)$ after random

dismissal is given by

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

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 (ψ_h, ψ_u) to (ψ'_h, ψ'_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 (ψ'_h, ψ'_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 ψ_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 $B(j-1; \psi_h-1, q_{hh}^H)$. If j workers remain in status h , the transition to ψ'_h requires that exactly $\psi'_h - j$ switch from status u to status h . The latter happens with probability $B(\psi'_h - j; \psi_u, q_{uh}^H)$. Thus

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

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

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

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

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

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

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

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

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

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 (ψ_h, ψ_u) before vacancies are filled and follows policy g , then the probability to arrive at (ψ'_h, ψ'_u) is

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

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 ν . Let $Q_h^{vac}(g, \theta, \nu)$ denote the associated transition matrix. Then the distribution after this step is $Q_h^{vac}(g, \theta, \nu) \cdot Q^{exit} \cdot Q^{health} Q^{dis}(g)$ Combining these three steps

$$\mu_h^s [s'|\mu(\cdot), \gamma(\cdot|\cdot), \theta, \nu] = \sum_{s \in S} \sum_{g \in G} \left\{ [Q_h^{vac}(g, \theta, \nu) \cdot Q^{exit} \cdot Q^{health} \cdot Q^{dis}(g)]_{\psi_h(s'), \psi_u(s')} \cdot q_{z(s)z(s')}^Z q_{I(s)I(s')}^I(g) q^{match} [s, g|\mu(\cdot), \gamma(\cdot|\cdot)] \right\}$$

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 , ν , and θ 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 Ω 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

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

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
<i>Fraction of Establishments Providing Coverage</i>								
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
<i>Fraction of Establishments Discontinuing Insurance over a 2 Year Period</i>								
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
<i>Fraction of Unhealthy Workers among all Employees</i>								
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
<i>Fraction of Employed Workers Covered, Excluding Very Large Establishments</i>								
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
<i>Fraction of Employed Workers Covered, Including Very Large Establishments</i>								
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_h	0.073	0.065	0.063	0.063	0.062	0.062	0.063	0.064
m_u	0.027	0.024	0.015	0.013	0.010	0.009	0.010	0.012
V_h	82.89	83.09	83.80	83.92	84.12	84.20	84.11	84.01
V_u	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_h and m_u denote the mass of healthy and unhealthy searchers, respectively, V_h and V_u denote the value of search for healthy and unhealthy searchers, respectively.

Table A-2: Community Rating Experiments

	Specification							
	YYY	YYN	YNY	NYN	YNN	NYN	NNY	NNN
<i>Fraction of Establishments Providing Coverage</i>								
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
<i>Fraction of Establishments Discontinuing Insurance over a 2 Year Period</i>								
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
<i>Fraction of Unhealthy Workers among all Employees</i>								
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
<i>Fraction of Employed Workers Covered, Excluding Very Large Establishments</i>								
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
<i>Fraction of Employed Workers Covered, Including Very Large Establishments</i>								
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_h	0.064	0.064	0.064	0.069	0.066	0.066	0.066	0.066
m_u	0.004	0.011	0.011	0.031	0.013	0.030	0.030	0.030
V_h	84.52	83.81	84.02	82.80	83.65	82.86	82.86	82.86
V_u	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_h (V_u) and m_h (V_u) denote the mass (the value of search) of healthy and unhealthy searchers, respectively.