NBER WORKING PAPER SERIES

RECLASSIFICATION RISK IN THE SMALL GROUP HEALTH INSURANCE MARKET

Sebastián Fleitas
Gautam Gowrisankaran
Anthony Lo Sasso

Working Paper 24663
http://www.nber.org/papers/w24663

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2018, Revised March 2019

We thank Steve Berry, John Bertko, Allan Collard-Wexler, Jan De Loecker, Tim Dunne, Liran Einav, Antonio Galvao, Josh Gottlieb, Matt Grennan, Jon Gruber, Ben Handel, Igal Hendel, Dan Herbst, Hide Ichimura, Neale Mahoney, Tom McGuire, Mark Pauly, Daniel Prinz, Jim Rebitzer, Bob Town, Erin Trish, and numerous seminar and conference participants for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w24663.ack

Reclassification Risk in the Small Group Health Insurance Market
Sebastián Fleitas, Gautam Gowrisankaran, and Anthony Lo Sasso
NBER Working Paper No. 24663
May 2018, Revised March 2019
JEL No. D25,I13,L13

ABSTRACT

We evaluate reclassification risk in the small group health insurance market from a period before ACA community rating regulations. Using detailed individual-level data from a large insurer, we find a pass through of 16% from health risk to premiums with enrollee fixed effects, and 70% without fixed effects. The fixed effects estimates identify the extent to which the insurer passes through changes in risk to changes in premiums while the higher estimates without fixed effects may be due to more risk rating for new accounts. Our estimates control for selection into insurance take-up with a non-parametric selection model, using individual risk and industrial sector as exclusion restrictions. Our results are also robust to other possibilities, including potential measurement error of risk scores and slow pass through over time. We seek to explain why our fixed effects estimates are much closer to community rating than full experience rating. The limited pass through may be due to implicit “guaranteed renewability” contracts with one-sided pricing commitment on the part of the insurer, as our results are broadly consistent with the equilibrium pass through that would occur under these contracts. Our results cannot be explained by market power, search frictions, or slow pass through over time. We simulate the value that is generated by the insurer’s pricing policy relative to counterfactual pricing policies. The insurer’s policy generates 60% of the welfare gain from community rating relative to full experience rating. Even community rated plans generate substantial reclassification risk due to high out-of-pocket costs.

Sebastián Fleitas
Department of Economics
University of Leuven
3000 Leuven
Belgium
sebastian.fleitas@kuleuven.be

Gautam Gowrisankaran
Department of Economics
University of Arizona
P.O. Box 210108
Tucson, AZ 85721-0108
and HEC Montreal
and also NBER
gowrisankaran@eller.arizona.edu

Anthony Lo Sasso
Health Policy and Administration Division
School of Public Health
University of Illinois at Chicago
1603 W. Taylor
Chicago, IL 60612
losasso@uic.edu

1 Introduction

One of the most important concerns in designing health insurance markets is reclassification risk. Reclassification risk, which occurs when an adverse and persistent health shock leads to higher future premiums or worse coverage, has the potential to lead to market failure by limiting the long-run risk protection from insurance. The 2010 Affordable Care Act (ACA) sought to reduce reclassification risk—in the individual and small group markets—through community rating provisions. In isolation, restrictions on pricing based on risk may drive low risk people out of the market, leading to adverse selection.

This paper considers reclassification risk in the context of the small group health insurance market. This market provides insurance to individuals at employers with 2 to 50 employees.[1] In 2013, this market covered 18 million people in the U.S. (Kaiser Family Foundation, 2013), representing about $100 billion in revenues.[2] Reclassification risk is particularly salient for this market because of the small sizes of the employers. To illustrate, consider an individual who works for an employer with 5 employees with an annual health insurance contract.[3] Suppose that the individual or her co-worker is diagnosed with a serious illness, perhaps diabetes, with an expected cost of $25,000 per year going forward. A market that fully passes through risk to each employer will increase the premiums to this employer by $25,000, which will in turn raise per-employee costs by $5,000 per year.

A number of influential studies have documented substantial variation in premiums across employers in the small group market (Cutler, 1994; Cebul et al., 2011; Bundorf et al., 2012). Using an employer survey, Cutler finds that the 90th percentile of premiums is 2.74 times the 10th percentile for this market. Due to data limitations, Cutler did not explicitly tie the variation in premiums to health risk. Nonetheless, Cutler's findings are suggestive that the premium variation in this market is mostly due to reclassification risk from

[1]Prior to the ACA, the small group market included groups with 1 to 50 members. The ACA originally mandated a change in the market definition to include groups with up to 100 members. This change was eliminated in the 2015 Protecting Affordable Coverage for Employees (PACE) Act, so that the federal definition remains 1-50 members. However, four states use the 100 members maximum in their definition (Jost, 2015).

[2]Authors' calculation using premium information from Blavin et al. (2014).

[3]Annual insurance contracts are typical in this market.

experience rating, i.e., from employers with higher health risks facing higher premiums (Gruber, 2000). In contrast, Pauly and Herring (1999) find low experience rating in both the individual and small group insurance market. In particular, using survey data on reported premiums and individual but not group risk, they find premium elasticities with respect to claims costs that ranges from -0.06 to 0.44, compared to an elasticity of 0 for community rating and 1 for full experience rating. Given this mixed body of evidence, we believe that the question of how much reclassification risk exists in this market is still open. Data that link health risks to premiums seem crucial to obtaining more definitive evidence.

This paper has two main goals related to health insurance in the small group market. Our first goal is to examine the extent of reclassification risk in this market and quantify the resulting welfare loss. Our second goal is to understand the welfare consequences of alternative pricing policies, such as community rating and full experience rating, relative to the current environment.

Our study makes use of a unique dataset on the small group insurance market provided to us by a large health insurance company, which we refer to as the United States Insurance Company, or USIC. USIC provided us with a panel of claims and premiums for their small group market products in 10 states over the period 2012-15.[4] Our analysis data contain information on over 300,000 USIC enrollees at more than 12,000 employers. Our study is unique in having access to a large dataset on the small group market that includes both individual level claims and group level premiums. This dataset allows us to estimate how USIC responds to shocks across the small employers that it serves.

To motivate our empirical specifications, we first develop a simple two-period model of insurance in the small group market. Our model specifies that USIC offers health insurance to a small employer, charging premiums that are potentially based on the health risk (or equivalently, the expected claims cost) of the employer. Potential enrollees decide whether or not to enroll in insurance if offered, given their health risk and the premium charged. Our model shows that the welfare loss from reclassification risk is increasing in

[4]This time period was immediately before most of the ACA regulations for the small group market were effective. For the time period and states in our sample, insurers could experience rate small employers without significant regulatory restrictions.

the pass through from mean health risk at an employer to premiums. We highlight three cases of pass through: full experience rating—where claims risk at an employer is fully passed through in the form of higher premiums; community rating—where pass through is zero; and USIC's actual pass-through rate—which we estimate. Community rating is equivalent to binding two-period contracts as in both cases, the pass through from health risk to premiums is 0.

We show that the pass-through coefficient from mean health risk to premiums forms a sufficient statistic for understanding reclassification risk in this market. In the spirit of Chetty (2008, 2009); Chetty and Saez (2010) and Einav et al. (2010), we evaluate the extent of reclassification risk here by estimating a pass-through regression, instead of estimating a structural model of insurer, employer, and enrollee choice. We estimate USIC's pass through by regressing the extent to which changes in mean health risk at an employer in a year translates into changes in premiums for the employer. We compute health risk for each individual as the ACG score, using the previous year's claims data.[5]

Because movement in and out of small group insurance with USIC occurs frequently, our estimation controls for selection. Following Newey (2009), we estimate a selection equation for insurance take-up in a first stage and use it to create a non-parametric power series selection correction in our pass-through regressions. We estimate the selection equation using two different datasets. First, we use the USIC data. These data allow us to calculate individual and group health risk but can only control for individuals who were insured by USIC and then terminated coverage. Second, we use the Medical Expenditure Panel Survey (MEPS) data. The MEPS data allow us to control for take-up among everyone who was offered insurance in the small group market, but they do not allow us to calculate group health risk, as they do not provide group characteristics. Exclusion restrictions in the reclassification risk (treatment) regression are needed to credibly identify selection effects. Individual risk and industrial sector provide useful exclusion restrictions and our multiple data sources help with robustness. Our fixed effects estimates will be identified by the extent to which changes in health risk at an employer translate into changes in premiums, conditioning non-parametrically for the probability of being in the

[5]The ACG score, which was developed by Johns Hopkins University, is widely used in this context (see, for instance Gowrisankaran et al., 2013; Handel, 2013; Handel et al., 2019).

sample.

Overview of findings. We find that a unit increase in mean ACG score for an employer increases its mean annual claims cost by $4003.[6] Our base estimates—which control for selection using the USIC data—show that this unit increase causes premiums to rise by $624 with enrollee fixed effects or $2,811 with only market fixed effects.[7] Thus, our estimated pass through from expected claims to premiums is 16% without fixed effects and 70% with fixed effects. Idiosyncratic premium risk—i.e., premium variation not correlated with health risk—is also present in this market, with a standard deviation of $576. Though not correlated with health risk, idiosyncratic premium risk also lowers consumer welfare, all else equal.

Our specifications with employer or enrollee fixed effects uncover USIC's pass through for existing employers while our non-fixed-effects specifications uncover the equilibrium relationship between health risk and premiums, which includes the pass through on new accounts. Thus, USIC appears to pass through very little risk for existing customers, though new customers do receive premiums that are more risk-based. Low pass through from costs to prices is observed in a variety of contexts, including energy costs (Ganapati et al., 2016), tariffs (De Loecker et al., 2016), and beverage taxes (Cawley et al., 2018).

Given this evidence that the market is closer to community rating than to experience rating, we seek to understand what might explain this evidence. Pauly et al. (1995); Herring and Pauly (2006) show that "guaranteed renewability" contracts with one-sided price commitment on the part of the insurer can provide protection from reclassification risk that is similar to community rating. Handel et al. (2015) consider the optimal guaranteed renewability contract accounting for enrollee liquidity constraints, finding positive but limited pass through. We simulate our pass-through coefficient under Handel et al. (2015)'s data and optimum and find that it is $1,821 without enrollee switching costs. The results are likely lower with enrollee switching costs; inertia in health insurance choice has been widely documented (Crocker and Moran, 2003; Handel, 2013). Thus, these re-

[6]An ACG score of 1 is the population mean score, so a unit increase would occur from an employer having double the expected health cost of the population mean.

[7]We find no significant effect of extra risk on plan benefits. Our results are also robust to alternative measures of risk scores.

sults suggest that USIC may implicitly provide reclassification risk protection similar to guaranteed renewability contracts. Health insurance experts also believe that insurers typically offered guaranteed renewability contracts prior to the ACA regulations.[8]

We consider three other explanations for our results, but do not find support for any of them. First, we consider whether our results could be driven by USIC passing through a health risk shock to an employer slowly over time. In this case, we should see that the coefficient on the lag of the risk score should be positive, in a specification with employer or enrollee fixed effects. We estimate specifications where we allow the risk score and lagged risk score to both affect premiums but do not find a significant effect on the lagged risk score. Second, we consider whether our results could be driven by static pricing power. We interact our pass through across markets with different measures of market concentration and find that they do not significantly predict pass through. Since concentration should affect the curvature of the residual demand curve, this suggests that our results are not caused by static market power (though the power of this finding is limited by the fact that our data are from only 10 states). Third, we consider whether our results could be driven by search frictions (Pauly et al., 2006; Cebul et al., 2011). Applying the search model estimates from Cebul et al., we find that approximately 87% of expected claims costs would be passed through as higher premiums. Given that our fixed effects pass-through estimates are much smaller, our results are also unlikely to be explained by search frictions.

Using our estimated coefficients, we simulate counterfactuals to evaluate the extent to which the insurance provided by USIC provides value in the form of protection from reclassification risk in the small group market. We non-parametrically simulate the evolution of health risk over a ten-year horizon to evaluate how this would translate into financial risk for individuals. Our counterfactuals incorporate selection into and out of USIC insurance. When individuals remain in USIC insurance with the same employer, they receive a new premium based on the new employer risk and the estimated coefficient from the model with enrollee fixed effects and the selection correction. In other

[8] Pauly and Lieberthal (2008) note that individual insurance were generally offered with guaranteed renewability contracts. A private conversation with Prof. Mark Pauly, December 10, 2018, Philadelphia, PA, verifies this point for the small group market.

cases, we assume that they work at an employer of a similar size and risk but receive a premium based on the pass-through coefficient estimated without enrollee fixed effects. We calculate the certainty equivalent loss in income from financial risk, using a CARA functional form and estimated risk preferences from the literature (Handel, 2013).

We find that USIC's current policies result in a mean annual certainty equivalent loss from financial risk of $3,050 over the ten years after the initial period. Even though our estimates with enrollee fixed effects show low pass through similar to community rating, idiosyncratic premium risk, out-of-pocket costs, and employees in this market frequently switching insurance coverage limit the reclassification risk protection of these contracts. Without idiosyncratic premium risk, this annual certainty equivalent loss would decline to a mean of $2,750 annually. With community rating, the mean certainty equivalent loss is $1,950, all of which is due to the financial risk from out-of-pocket costs. The certainty equivalent income loss is $4,500 under full experience rating implying that USIC's current pricing policy generates about 60% of certainty equivalent gains in risk protection of community rating regulations relative to full experience rating.

Finally, we evaluate the extent to which pooling in the small group market adds value relative to an individual market with identical guaranteed renewability contracts and selection. We find that individuals in such a market would have a certainty equivalent loss of $3,650 over the ten year horizon, implying that pooling across beneficiaries within a small employer adds an average of $600 in risk protection annually.

Relation to literature. Our paper builds on a substantial empirical literature that analyzes reclassification risk (Cutler, 1994; Cutler and Reber, 1998; Pauly and Herring, 1999; Gruber, 2000; Buchmueller and DiNardo, 2002; Herring and Pauly, 2006; Einav et al., 2010; Cebul et al., 2011; Bundorf et al., 2012; Handel, 2013; Handel et al., 2015; Kowalski, 2015). Cutler and Reber (1998); Einav et al. (2010); Handel (2013) and Kowalski (2015) examine large employers, evaluating the premiums that they charge their employees for the different plans that they offer and the resulting adverse selection and reclassification risk. Buchmueller and DiNardo (2002) consider the impact of community rating on the small group and individual markets, using New York's implementation of community rating in these markets as the treatment. Bundorf et al. (2012) focus on the small group market, evaluating the welfare impact of employee choice of plans under different premium pass-through mechanisms from employers to enrollees. For the individual market, Handel et al. (2015) evaluate the equilibrium adverse selection and reclassification risk from a competitive market of exchange firms, while Handel et al. (2019) examine reclassification risk in a competitive market of long-term contracts with one-sided commitment.

We add to this literature in two ways. First, our data are unique and allow us to identify the extent to which experience-rated health insurance creates reclassification risk in the real world. Specifically, we recover how much expected future claims are passed through into future premiums, in a context in which pass through is permitted. We believe that our results here are salient in the ongoing policy debate regarding the role and value of private insurance and in understanding the extent of reclassification risk in private insurance markets, which is important since researchers have had different beliefs on the extent of reclassification risk in the individual and small group markets. Our results can also help understand how ACA regulations are affecting reclassification risk in these markets.

Second, we develop a simple theoretical framework that allows us to estimate risk, selection, and welfare in a straightforward way. We believe that this framework may be useful in evaluating these issues for other markets.

The remainder of our paper is organized as follows. Section 2 describes our model of enrollee choice, risk, and selection. Section 3 describes our data sources and estimation sample. Section 4 describes our empirical approach. Section 5 describes our estimation results, Section 6 presents our counterfactuals, and Section 7 concludes.

2 Model

2.1 Enrollee utility and choice

We develop a simple and stylized model of reclassification risk and selection in the health insurance industry. The model has two time periods, periods 1 and 2. Period 2 payoffs are discounted at the rate δ. A period is meant to represent a year, the typical length of a health insurance contract.[9] We consider potential enrollees who can obtain health insurance through a small group employer.[10] Denote the potential enrollee by i, the employer by j, the time period by t, and the number of potential enrollees at employer j by I_j.

Each potential enrollee starts each period with risk score r_{ijt}. The risk score is based on her previous year's healthcare claims, is proportional to her total expected costs of healthcare at period t, is normalized to 1 for the mean individual in the population, and is observable to both the potential enrollee and the insurer. Let $H \sim dF_H(r_{ijt})$ denote the period t health shock as a function of risk score and $c(H)$ denote the claims cost given health shock H. We separate costs into the portion that the insurer pays, $c^{ins}(H)$, and the portion that the enrollee pays out of pocket, $c^{oop}(H)$. Insurer-paid claims satisfy

$$E[c^{ins}(H)] = \gamma r_{ijt} \tag{1}$$

where γ is a constant of proportionality.[11] The individual risk scores imply that the employer mean risk score over its population of potential enrollees is $R_{jt} \equiv \frac{1}{I_j} \sum_{i=1}^{I_j} r_{ijt}$.

The timing in our model is as follows. Each period, the insurer first observes R_{jt} and then decides on the per-person premium, $p_{jt}(R_{jt})$, which is potentially based on both R_{jt} and the employer's history with the insurer. Following this, the employer decides whether to offer insurance. Potential enrollees then decide whether to take up the insurance if offered. Finally, the health shocks H are realized.

We now exposit the utility at each period. We assume that utility is additively separable across the time periods. Consider first the per-period utility from obtaining insurance when offered, which we denote $U^A_{ij}(r_{ijt}, p_{jt}(R_{jt}))$. This utility is a function of the potential enrollee's income Y_{ijt}, her employer's premium, and her out-of-pocket health costs:

$$U^A_{ij}(r_{ijt}, p_{jt}(R_{jt})) = \int u_{ij} \left[Y_{ijt} - p_{jt}(R_{jt}) - c^{oop}(H) \right] dF_H(r_{ijt}), \tag{2}$$

where $u_{ij}(\cdot)$ is her money utility function. We assume that, through adjustments to pre-

[9] We make this assumption for ease of notation. Our empirical work allows for more than two periods.

[10] Our theoretical analysis does not distinguish between potential enrollees who are employees and dependents.

[11] While risk scores typically concern overall costs, we assume here that the proportional relationship holds for the costs borne by the insurer.

mium or wage, the potential enrollee pays her employer the full cost of her health premium,[12] and that enrollees are risk averse.

Consider now the per-period utility from not having insurance, which we denote $U^N_{ij}(r_{ijt})$. Without insurance, the individual bears the full cost of her health expenditures:

$$U^N_{ij}(r_{ijt}) = \int u\left[Y_{ijt} - c(H)\right] dF_H(r_{ijt}). \tag{3}$$

Combining the utility from both choices, the potential enrollee's per-period utility is then:

$$U_{ij}(r_{ijt}, p_{jt}(R_{jt})) = \max\{U^A_{ij}(r_{ijt}, p_{jt}(R_{jt})), U^N_{ij}(r_{ijt})\}. \tag{4}$$

Finally, the discounted value of the potential enrollee over the two periods is:

$$V_{ij}(\vec{r}_{j1}) = U_{ij}(r_{ij1}, p_{jt}(R_{j1})) + \delta \int U_{ij}(r_{ij2}, p_{jt}(R_{j2}))dF_{R,r}(R_{j2}, r_{ij2}|\vec{r}_{j1}), \tag{5}$$

where $\vec{r}_{1j1}$ is the vector of period 1 risk scores and $dF_{R,r}(R_{j2}, r_{ij2}|\vec{r}_{j1})$ is the joint conditional risk score distributions at period 2, for the potential enrollee and her employer.

2.2 Risk rating and reclassification risk

Reclassification risk can enter in our model because a bad and persistent health shock at period 1 for the individual or her coworker will raise R_{j2}. With experience rating, this will in turn raise premiums for the individual. The extent of reclassification risk depends on the distribution of F_R and on $p_{j2}(\cdot)$. If the individual were in a large risk pool, then reclassification risk would not be a substantial issue because the distribution of F_R would be very concentrated and degenerate to a point in the limit. Even if the individual were in a small risk pool, if $p_{jt}(R_{jt})$ did not vary much in response to R_{jt}, then she would not be faced with much reclassification risk. Thus, individuals employed by large employers or in settings without much experience rating do not face much reclassification risk. In contrast, individuals in small risk pools without significant restrictions on experience

[12] The literature has shown positive but sometimes incomplete pass through from higher premiums to lower wages (Baicker and Chandra, 2006; Bhattacharya and Bundorf, 2009).

rating—i.e., individuals in our sample—may be faced with significant reclassification risk.

We now consider the impact of different risk rating policies, focusing on the case where potential enrollees take up insurance. For ease of notation, we assume that the insured have no out-of-pocket costs.[13] This implies that $E[c(H(r_{ijt}))] = E[c^{ins}(H(r_{ijt}))] = \gamma R_{jt}$.

First, we examine full experience rating. In this case, the insurer sets premiums exactly equal to expected equilibrium insured risk, so that $p_{jt}(R) = \gamma R$. Suppressing dependencies on variables that no longer enter, equation (5) specializes to:

$$V_{ij}(\vec{r}_{j1}) = U_{ij}^A(\gamma R_{j1}) + \delta \int U_{ij}^A(\gamma R_{j2}) dF_R(R_{j2}|\vec{r}_{j1}). \tag{6}$$

Individuals here are faced with reclassification risk: an increase in the expected equilibrium mean risk score among the insured in period 2, R_{j2}, is passed through into an increase in expected insurance costs at the employer in period 2. This occurs even though contracts are actuarially fair.

Next, we consider long-run contracts with a binding commitment to future premiums. Consider such a contract with a period 1 premium of $p_{j1} = \gamma R_{j1}$ and a period 2 premium of $p_{j2} = \gamma E[R_2|\vec{r}_{j1}]$. This contract would have premium equal to expected marginal cost and would eliminate reclassification risk. Given our assumption that consumers are risk averse,

$$\int U_{ij}^A(Y_{ij2} - \gamma R_{j2}) dF_R(R_{j2}|\vec{r}_{j1}) < U_{ij}^A(Y_{ij2} - \gamma E[R_{j2}|\vec{r}_{j1}),$$

implying that such a contract would improve enrollee welfare for individuals who take-up insurance over state-contingent one-period contracts of the same actuarial value. Consider further the case where income and mean risk are the same across periods, so that $Y_{ij1} = Y_{ij2},\ \forall i$, and $E[R_{j2}|\vec{r}_{j1}] = R_{j1}$. In this case, the above contract would maximize U^A among long-run break-even contracts. This implies that a competitive insurance industry with full take-up would result in employers signing these two-period contracts.[14] Note further that this two-period contract is equivalent to a community rating provision in that

[13]Our empirical work accounts for out-of-pocket costs.

[14]In the real world, it is difficult to enforce long-run contracts with commitment on both sides. Without such enforcement, a competitive insurance industry might provide partial protection against reclassification risk (Handel et al., 2019).

there is no pass through from health risk to premiums.

Finally, we consider the general case with different levels of passed through. A simple functional form for premiums here is:

$$p_{jt} = c_{jt} + \beta R_{jt}, \tag{7}$$

for some constant c_{jt}, which reflects baseline prices at period t, and might vary due to changes in healthcare provider prices or general expected increases in health risk over time.

If $\beta = \gamma$, then this is the full experience rating case. For $0 < \beta < \gamma$, there will be positive but incomplete pass through from risk to premiums. Under community rating or binding two-period contracts, we would have $\beta = 0$. Given that preferences are risk averse, for $\beta' < \tilde{\beta}$,

$$\begin{aligned} &\int U_{ij}^A(Y_{ij2} - p_{j1} - c - \tilde{\beta}(R_{j2} - E[R_{j2}|\vec{r}_{j1}])dF_R(R_{j2}|\vec{r}_{j1}) \\ < &\int U_{ij}^A(Y_{ij2} - p_{j1} - c - \beta'(R_{j2} - E[R_{j2}|\vec{r}_{j1}])dF_R(R_{j2}|\vec{r}_{j1}), \end{aligned} \tag{8}$$

for a constant c. This follows since the left side is a mean-preserving spread of the right side. Note that in comparing (8) to (7), $c_{jt} = c + \tilde{\beta}E[R_{j2}|\vec{r}_{\cdot j1}]$ with $\tilde{\beta}$ pass through, while $c_{jt} = c + \beta' E[R_{j2}|\vec{r}_{\cdot j1}]$ with β' pass through. This is consistent with c_{jt} decreasing as β increases for contracts with the same actuarial value. Equation (8) shows that since utility is decreasing in β, β is a sufficient statistic to evaluate the reclassification risk from a contract, conditional on preferences and the distribution of health shocks scaled in dollars using γ.

In sum, our main estimable parameter is β, which measures the degree of reclassification risk. In the tradition of Chetty (2009) and Einav et al. (2010), our empirical analysis will estimate β and use it to characterize reclassification risk in this market. This does not require us to specify or estimate all preferences parameters. To understand welfare under the observed and counterfactual environments, we then combine our estimates of β with risk preference parameters from the literature, the estimated distribution of health shocks scaled in dollars, and models of selection into insurance.

2.3 Selection of enrollees

We now discuss our model of selection of potential enrollees into insurance. While selection based on unobservables would not affect the main intuition of the model, selection may affect the estimation of the pass-through coefficient and the computation of the counterfactuals. We model the combined offer and take-up decision for the employer and enrollee as:

$$D_{ijt} = \mathbb{1}\{f(R_{jt}, r_{ijt}, x_{ijt}) + \varepsilon^s_{ijt} > 0\} \tag{9}$$

where $D_{ijt} = 1$ indicates offer and take-up of insurance, $f(\cdot)$ is a flexible mean utility function to be estimated, and ε^s_{ijt} is an unobservable. We then model premiums as:

$$p_{ijt} = c_{jt} + \beta R_{jt} + \varepsilon^p_{ijt}, \tag{10}$$

where we are now indexing p by 'i' and where ε^p_{ijt} is also an unobservable, capturing idiosyncratic premium risk unexplained by other factors and uncorrelated with risk, for instance due to variation in employer or insurance broker bargaining ability.

Our data contain premiums only for individuals who take up insurance and hence for whom $D_{ijt} = 1$. Hence, (9) is our selection equation while (10) is our treatment equation. As is typical in selection models, we allow for correlations between $\vec{\varepsilon}^s_{it}$ and $\vec{\varepsilon}^p_{it}$. We generally expect that there would be a negative correlation between the two unobservables since individuals who received a higher premium than expected given observables would also be less likely to select into insurance than expected given observables. Note also that ε^p_{ijt} will be highly correlated for individuals at the same employer and year since individuals at a given employer and year pay the same premium.

Our estimation seeks to recover consistent treatment effects for β in (10) in the presence of selection. Specifically, we non-parametrically control for selection in (10) using the methods of Newey (2009).

3 Data and Estimation Sample

3.1 Data

Our principal data are from employers who purchase health insurance for employee and dependent coverage from "United States Insurance Company" (USIC) in the small group market during the years 2012 to 2015. USIC provided us with data from 10 different states: AR, DE, IL, PA, OK, MO, TN, TX, WI, and WY. USIC further classified the data into 19 different geographic markets, e.g., Texas is divided into Central Texas, Dallas, Houston, North Texas, and South Texas. Employers in our sample purchased fully-insured insurance products from USIC, not third-party administrative services. Figure A.1 in on-line Appendix A provides a map of the states in our estimation sample.

While all states regulate small group insurance, they vary in the degree of their regulation. The states that we use were all lightly regulated prior to the ACA. For instance, none of the states had community rating regulations during this period. One measure of state regulation is the extent to which premiums are allowed to vary across groups for all reasons apart from plan generosity, which are known as ratings bands. Prior to the start of ACA regulations on this market, DE, PA, TX, IL, WI, and WY allowed premiums to range across groups by a ratio of 25-to-1 or greater (a total of 12 states had bands this large); MO and OK had rating bands between 19- and 25-to-1; and AR and TN had rating bands between 13- and 19-to-1.[15] During this period, all states had provisions that essentially implied that USIC would not be able to cancel a group's policy even if the mean health risk for the group rose substantially.

The ACA implemented community rating regulations for the small group market—specifically a ban on health status underwriting and a requirement that plans in the market have a common small group risk pool—that were originally supposed to start in January, 2014. However, almost all small group plans were exempt from the ACA market reforms during our sample period, for two reasons. First, some of these plans were "grandfathered," meaning that the ACA included a clause that allowed consumers to keep their

[15] See `http://www.naic.org/documents/topics_health_insurance_rate_regulation_brief.pdf`.

existing health plans, conditional on the plan not significantly changing its benefits.[16] Second, a transitional rule let states allow "grandmothered" plans in the small group market, meaning that they could permit insurers to continue offering non-ACA compliant plans to small employers. The great majority of states opted to allow the sale of grandmothered plans past our sample period, and indeed through 2018.[17] Importantly for our analysis, both grandmothered and grandfathered plans are exempt from the ACA's community rating regulations noted above.

Our data include information at both the enrollee-year (employee or dependent) and employer-year levels. At the employer-year level, for all the employers that contract with USIC, we observe the total number of employees that are eligible for health coverage, the number of health insurance plans available to their enrollees in each year, the characteristics of each plan, and the total premium paid by the employer to the insurer for each plan in each month of each year.

We observe data for each enrollee that takes up insurance in each year. Specifically, we observe age, gender, the health plan chosen, the relationship of the enrollee to the employee (e.g., self, spouse, child), and information to link enrollees to the employer and to the employee with employer-sponsored coverage. We also observe claim-level data—for both medical and pharmaceutical claims—for every healthcare encounter. These data provide diagnosis, procedure, date of service, and premium information and are linked to the enrollee identifier.

We calculate a per-enrollee premium by dividing the total premium paid by the employer to USIC in a year for a plan by the number of enrollees (employees and dependents) at that employer and plan during that year. We use the January premium and enrollee information for this calculation and multiply the monthly premium by twelve to annualize it.[18]

[16]The concept of grandfathering of health plans was popularized by President Obama's statement that "if you like your health plan, you can keep your health plan."

[17]See Jost (2017) and CMS (2017) for further details on this discussion.

[18]Because individuals typically make enrollment decisions annually with contracts starting in January, the total premiums paid by the employer to USIC in January is a good representation of annual per-person premiums charged by USIC. We also computed per-enrollee premiums using the mean and mode of the monthly premiums paid by the employer over different months, and obtained similar results with these alternative measures.

To measure the predicted health expenditure risk for each enrollee, we use the ACG risk prediction software developed at Johns Hopkins Medical School. The software outputs an "ACG score" for each enrollee in each year, which corresponds to r_{ijt} in our model. The ACG score indicates the predicted relative healthcare cost for the individual over the year, and has a mean of 1 in a reference group chosen by ACG. The ACG score is based on past diagnostic codes, expense, prescription drug consumption (code and length of consumption), age, and gender for each individual. In our case, we use the twelve months of data from the previous year to generate the ACG score for a given year. Similarly to the ACG score, USIC also uses a proprietary system to derive a risk score for each enrollee. While we do not have access to the USIC scores, we believe that the ACG and USIC scores are very similar. For new groups, information on enrollee health is generally, if imprecisely, obtained via questionnaire.

Since our risk score measures are calculated using the previous year claims data, we need to observe an employer or individual for two consecutive years in order to have a complete observation where we can observe the risk score and the premium. Thus, for instance if we observed an employer in 2012 and 2013, this would allow us to compute the 2013 premium and mean risk score for the employer, where the risk score was computed from 2012 data.

Most of our regressions use employer or enrollee fixed effects. Since we obtain the risk score calculation from the previous year, we need three continuous years of data (which generates two years with complete observations) to compute an employer/enrollee fixed effect. For comparability across estimates, we drop employers/enrollees for which we observe fewer than three continuous years of data for all our specifications, even those without employer/enrollee fixed effects.[19]

As an additional source to evaluate selection, we use data from the Medical Expenditure Panel Survey (MEPS), which is a nationally representative survey regarding individual insurance decisions and health care expenditures. The MEPS data allow us to understand selection into the small group market since it has information about individuals that did not take up insurance. We construct the sample by using panel 18 from the consoli-

[19]We also drop employers with missing information for premiums, plan characteristics, or enrollment.

dated dataset for years 2013 and 2014. To select individuals who could participate in the small group market, we select individuals that were (a) working but not self-employed at the beginning of the period, (b) who worked at an establishment size less than or equal to 50 individuals, and (c) were offered insurance via their employer. We use age, gender, and health conditions of individuals and establishment size and industrial sector as characteristics to predict the probability of taking up insurance.

3.2 Summary statistics on estimation sample

Table 1: Descriptive statistics on estimation sample at the enrollee-year level

	Full Sample	Stayers	Joiners	Quitters
Unique individuals	336,755	80,031	87,107	113,124
Observations	646,904	240,093	176,163	186,012
Relation (%)				
Employees	56.57	56.25	56.19	56.46
Spouses	15.50	16.18	15.28	15.49
Children	27.56	27.35	28.12	27.78
Others	0.37	0.22	0.41	0.28
Age	38 (18)	40 (18)	36 (18)	38 (18)
Female (%)	47	46	47	48
In dollars:				
Lagged paid total claims	3,388 (17,468)	3,778 (16,251)	3,287 (18,250)	3,272 (17,839)
Lagged out-of-pocket claims	902 (1,854)	1,009 (1,881)	894 (1,844)	845 (1,918)
Annual premiums	5,219 (1,955)	5,493 (2,028)	4,977 (1,698)	5,105 (2,106)
Health risk, r_{ijt}	1.00 (1.46)	1.01 (1.41)	0.92 (1.40)	1.05 (1.58)
$r_{ijt} - r_{ij,t-1}$	0.05 (1.07)	0.05 (1.03)	0.06 (1.04)	0.06 (1.19)
Conditions (%)				
Cancer	2.47	2.57	2.03	2.60
Acute myocardial infarction	0.16	0.16	0.16	0.17
Transplant	0.14	0.15	0.12	0.16
Diabetes	5.57	5.66	4.90	5.90
Hypertension	14.12	14.64	12.26	14.55
Heart disease	0.39	0.38	0.34	0.43
Chronic kidney disease	0.48	0.49	0.43	0.51
Asthma	3.38	3.27	3.35	3.59

Note: each observation in table is one enrollee during one year, 2013-15. Table reports mean values with standard deviations in parentheses. "Stayers" are enrollees always in sample; "joiners" are enrollees with one or more full observation but without a full observation in 2013; and "quitters" are enrollees with one or more full observation but without a full observation in 2015.

Table 1 provides summary statistics on the enrollees in our estimation sample. Our full sample includes over 330,000 unique individuals and over 650,000 observations.

We first analyze enrollee turnover. To do this, we characterize enrollees based on whether they have joined or quit available USIC coverage during our sample. A "joiner" is an enrollee for whom we did not have a complete observation in the first year but for whom we had a complete observation in a later year and whose employer was in the sample prior to her being there. A "quitter" is the opposite: an enrollee for whom we did not have a complete observation in the last year but for whom we had a complete observation in an earlier year and whose employer remained in the sample after she was no longer there. A "stayer" is an enrollee for whom we have three complete observations. Note that an individual can be both a joiner and a quitter, which would occur if she were in our data in the middle two years only. Also, note that enrollees for whom we do not observe for two consecutive years would not fit any of these three categories.

We find a lot of enrollee turnover. Only 37% of observations in our sample are for stayers. Approximately 27% of observations are for joiners while 29% of observations are for quitters.

Overall, there is a clear though moderate pattern of differences between the groups, where joiners have lower health risk than quitters who have lower health risk than stayers. Specifically, joiners have a mean ACG score—or expected claims cost—of 0.92, compared to 1.01 for stayers, and 1.05 for quitters. Consistent with this, joiners are on average two years younger than quitters, who are themselves two years younger than stayers.

On average, people paid $5,219 in annual premiums, had $3,388 in total claims, and $902 in out-of-pocket claims. We measure a number of chronic conditions from the claims data. The most prevalent is hypertension, occurring in 14% of observations. The next most common is diabetes, which occurs in 6% of enrollees.

Table 2 provides summary statistics on the employers in our estimation sample. Our sample includes 12,242 employers. Similarly to Table 1, we report the employers which are stayers, joiners, or quitters. We define an employer to be a stayer if it had at least one enrollee with complete data in each year; a joiner if it had no enrollee with complete data in 2013 but enrollee with complete data in 2014 and 2015 and a quitter if it had no enrollee with complete data in 2015 but enrollees with complete data in 2013 and 2014. The majority of employers in our sample, 54%, were stayers and hence present throughout the sample period, with complete observations from 2013-15. Similarly to at the individual level, more employers quit than joined coverage.

On average, employers in our sample have 21 subscribers. Eligible potential enrollees include employees, spouses, children, and other family members. Employees constitute 65% of covered lives. The mean take-up rate among eligible employees was 54%.

Table 2: Descriptive statistics at the employer-year level

	Full Sample	Stayers	Joiners	Quitters
Employers	12,242	6,560	2,281	3,401
Observations	31,044	19,680	4,562	6,802
Subscribers	21 (27)	21 (26)	23 (27)	20 (28)
Take up rate (%)	54 (22)	54 (22)	57 (21)	53 (23)
Relation (%)				
Employees	64.80	64.45	63.90	66.40
Spouses	12.82	13.01	13.08	12.12
Children	22.17	22.32	22.85	21.28
Others	0.21	0.21	0.18	0.21
Age	41 (9)	41 (9)	39 (8)	41 (10)
Female (%)	46	46	46	47
In dollars:				
Lagged paid total claims	4,076 (8,456)	4,003 (8,272)	3,775 (6,951)	4,490 (9,783)
Lagged out-of-pocket claims	1,092 (889)	1,051 (812)	1,061 (835)	1,232 (1,098)
Annual premiums	6,162 (2,837)	6,248 (2,689)	5,385 (2,067)	6,433 (3,529)
2013	5,954 (2,839)	5,881 (2,711)		6,095 (3,066)
2014	6,276 (3,103)	6,394 (2,808)	5,196 (2,157)	6,772 (3,908)
2015	6,238 (2,402)	6,469 (2,499)	5,574 (1,955)	
Health risk for enrolled, R_{jt}	1.07 (0.72)	1.05 (0.70)	0.97 (0.59)	1.17 (0.82)
$R_{jt} - R_{j,t-1}$	0.02 (0.51)	0.01 (0.49)	0.04 (0.45)	0.05 (0.62)
Conditions (%)				
Cancer	3.02	3.04	2.40	3.38
Acute myocardial infarction	0.18	0.17	0.17	0.21
Transplant	0.19	0.21	0.10	0.19
Diabetes	6.15	5.95	5.29	7.33
Hypertension	15.67	15.43	14.15	17.39
Heart disease	0.46	0.45	0.40	0.52
Chronic kidney disease	0.57	0.55	0.54	0.66
Asthma	3.34	3.28	3.18	3.61

Note: each observation in table is one small group employer during one year, 2013-15. Table reports mean values with standard deviations in parentheses. “Stayers” are employers always in sample; “joiners” are employers with one or more full observation but without a full observation in 2013; and “quitters” are employers with one or more full observation but without a full observation in 2015.

We observe a similar pattern of selection at the employer level to at the enrollee level. The mean of the employer mean risk scores, R, is 0.97 for joiners, 1.05 for stayers, and 1.17 for quitters. We also observe a substantial standard deviation in the *change* in R over time. This variation will provide us with power to identify the rating behavior that USIC uses, even with employer or enrollee fixed effects.

Table 2 also presents the same statistics on enrollees that we reported in Table 1, but at the employer-year level. We find similar values of the statistics regarding age, gender, premiums, claims, and out-of-pocket costs using this measure. Premiums in this market rose a moderate 5%

over our two-year sample period.

Finally, Table 2 presents the mean incidence of eight chronic conditions at an employer—cancer, transplants, acute myocardial infarctions (heart attacks), diabetes, hypertension, heart disease, chronic kidney disease, and asthma—defined as the percentage of enrollees with a diagnosis of the condition during the year. In Section 5, we use the presence of these chronic conditions at the employer as a robustness check. While the incidence of transplants and AMI is less than 1%, the mean incidence of cancer is 3% and diabetes is 6%.

Overall, the previous tables show lots of movement in and out of USIC insurance. The small businesses that are in this market frequently start and stop coverage with USIC. Potential enrollees at these businesses also frequently start and stop insurance take-up. This movement is driven by at least three different factors. First, businesses may opening or close for reasons orthogonal to health insurance premiums. Second, individuals can also be changing jobs for reasons that are orthogonal to premiums. Both of these factors are likely to be true given that small businesses enter and exit frequently and also change employees frequently. Third, there can be selection of health insurance based on premiums. Our results show moderate evidence of selection based on health risk, e.g., quitter employers have 9% higher expected costs than stayer employers, while quitter enrollees have 4% higher expected costs than stayer enrollees. This selection based on risk suggests that there may be selection based on premiums, which would bias our estimates of the pass-through coefficient β. In order to address this potential selection, our estimates control for the effect of selection and our counterfactuals use bounds for individuals who leave the sample.

We next present the patterns of persistence over time for the ACG risk, in Table 3. Panel A presents the results at the individual level for an AR(1) process in columns 1 and 2 and an AR(2) process in column 3. Column 1 reports the AR(1) process for the full sample while column 2 reports the AR(1) process for the same sample as in column 3. Mean health risk exhibits substantial persistence but at the same time a reversion to the mean. For instance, in the specification with only one lag, the autocorrelation coefficient is 0.733. In the specification with two lags, reported in column 3, the autocorrelation coefficients sum to 0.802. Thus, all the autocorrelated models are stationary. Moreover, the sum of the coefficients when we include two lags is similar to the results when we include only one lag, although these two processes imply different risk effects over time.

Panel B presents the autocorrelation results at the employer level. The results show that the AR(1) and AR(2) processes are stable but also relatively persistent. The fact that persistence at the employer level is smaller than at the individual level implies that the shocks for different enrollees

Table 3: Persistence in health risk over time

	(1)	(2)	(3)
Panel A: dependent variable individual risk (r_{ijt})			
Enrollee ACG score, $r_{ij,t-1}$	0.733***	0.718***	0.561***
	(0.005)	(0.007)	(0.010)
Lagged enrollee ACG score, $r_{ij,t-2}$			0.241***
			(0.010)
Sample	2013-15	2014-15	2014-15
Market FE	Yes	Yes	Yes
Observations	523,679	264,153	264,153
Panel B: dependent variable employer risk (R_{jt})			
Health risk for enrolled, $R_{j,t-1}$	0.667***	0.630***	0.506***
	(0.003)	(0.004)	(0.006)
Lagged health risk for enrolled, $R_{j,t-2}$			0.193***
			(0.007)
Sample	2013-15	2014-15	2014-15
Employer FE	Yes	Yes	Yes
Observations	31,044	18,802	18,802

Note: for panel A (B), each observation is one enrollee (employer) during one year. Standard errors are clustered at the employer level. Markets are defined by USIC and roughly represent an MSA or state. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

are not completely correlated, so they partially cancel each other out over time.

Table 4: Descriptive statistics for MEPS sample

	Choose insurance	Age	Female	Employer size
Mean	0.72	41.77	0.52	21.47
Standard deviation	0.45	12.91	0.50	14.66
Observations	1,355	1,355	1,355	1,355

Note: each observation in table is one individual in the consolidated Panel 18 for years 2013 and 2014 in MEPS. We selected working individuals not self-employed, who worked in establishments with less than 50 individuals and were offered health insurance via the employer.

Finally, we turn to descriptive statistics on the MEPS sample are reported in Table 4. From the full sample of people offered insurance, 72% of individuals choose insurance through the employer. The average age of these individuals is 42 years old and 52% of them are females. The percentages for the incidence of the conditions are similar to in the USIC sample. The average employer size is 21 which is similar but somewhat larger than the average employer size in our USIC sample.

4 Empirical Approach

The primary goal of our estimation is to recover γ, the impact of risk score on expected insurer costs ($dE[c^{ins}]/dR$), and β, the impact of employer mean insured risk on premiums (dp/dR). We use these parameters together to understand the pass through from insurer costs to premiums, $dp/dE[c^{ins}]$:

$$\frac{dp}{dE[c^{ins}]} = \frac{\frac{dp}{dR}}{\frac{dE[c^{ins}]}{dR}} = \frac{\frac{dp}{dR}}{\frac{dE[c^{ins}]}{dr}} = \frac{\beta}{\gamma}. \tag{11}$$

We also use these parameters separately in our counterfactual analysis. Note that these parameters regard insurer behavior; we do not estimate any enrollee utility parameters and our estimation algorithm does not impose potential enrollee utility maximization.

We now discuss our estimation of γ, which is the technology parameter that scales risk scores into dollar costs. Following (1), we estimate regressions that take the form:

$$c_{ijt}^{ins} = \gamma r_{ijt} + \gamma_2 x_{jt} + \varepsilon_{ijt}^{r}, \tag{12}$$

where c_{ijt}^{ins} measures the total dollar value of claims for the individual over the year. Equation (12) considers the impact of the individual's current risk score—estimated using the previous year's claims—on current claims to the insurer. Comparing equations (12) and (1), the empirical specification uses the actual insurer costs while the theoretical model is based on the expectation of costs. Thus, ε_{ijt}^{r} in equation (13) will capture the difference between actual claims and expected claims for an individual in a year.[20]

The empirical specification also includes controls x_{jt}. We include market fixed effects here (using USIC's market definition) to control for different provider prices across markets. Unless individuals at different employers systematically use different-cost providers for the same conditions in a way that correlates with the risk at that employer, we do not need to include employer or enrollee fixed effects. We estimate γ using only data from 2014, to not have to worry about changes in provider prices over time.[21]

We now discuss our estimation of β, the impact of employer mean risk score on premiums. Following Newey (2009), we estimate a two-step semi-parametric selection model. Specifically,

[20] We also estimate empirical specifications with $c^{oop}(H(r_{ijt}))$ as the dependent variable.

[21] We also investigated estimating γ using other years in our sample and obtained similar results.

we first estimate the following selection equation:

$$D_{ijt} = \mathbb{1}\{f(R_{jt}, r_{ijt}, x_{ijt}) + \varepsilon^s_{ijt} > 0\} \tag{13}$$

where R_{jt} and r_{ijt} are the risk scores for the employer and the individual respectively, x_{ijt} are time-varying individual characteristics, and ε^s_{ijt} is an unobservable component to the utility from selection into the sample. We estimate (13) using a probit specification and a flexible functional form for $f(\cdot)$.

We estimate two different specifications for the selection equation $f(\cdot)$. The first specification uses USIC data. Since our USIC data only tracks individuals insured by USIC, observations at period t include both individuals who left USIC after period $t-1$ and those who were with USIC at period $t-1$ and remained at period t. These observations cannot include people who joined USIC at period t and hence this estimation controls for dropped coverage but not new take-up of insurance. In this specification, we include industry fixed effects, employer size, age, and gender in x_{ijt}. The main advantage of this specification is that the inclusion of R_{jt} increases the accuracy of the selection equation. The main disadvantage is that it only controls for individuals who left USIC.

Our second specification for our selection equation uses the MEPS data. In this specification, we include in $f(\cdot)$ proxies to approximate r_{ijt}, industry fixed effects, employer size, age, and gender. The proxies for risk are indicators for hypertension, heart disease, AMI, ischemic stroke, respiratory failure, cancer, diabetes, and asthma, and are observable in both the MEPS data and in the USIC data. In this specification we include fewer regressors but we are able to control for everyone who is offered insurance. Using the estimated parameters of (13), we define

$$S_{ijt} \equiv Pr(f(R_{jt}, r_{ijt}, x_{ijt}) + \varepsilon^s_{ijt} > 0). \tag{14}$$

We compute S_{ijt} using the predictions from our estimates of (13).

We use S_{ijt} as a control in the second stage treatment effects equation. The treatment effects equation is:

$$p_{ijt} = \beta R_{jt} + \alpha x_{jt} + \overline{FE}_i + FE_i + g(S_{ijt}) + \varepsilon_{ijt} \tag{15}$$

where p_{ijt} is the premium charged to enrollee i working at employer j at period t and R_{jt} is the employer mean ACG risk score among enrollees who take up insurance at period t. In equation (15),

$\overline{FE}_i$ are enrollee fixed effects, FE_t are year dummies, and x_{jt} are time-varying enrollee attributes, and ε_{ijt} is the econometric unobservable. The non-parametric selection correction $g(S_{ijt})$ (using power series approximation) approximates the inverse Mills ratio from Heckman (1979). Comparing (10) to (15), $\varepsilon^p_{ijt} = g(S_{ijt}) + \varepsilon_{ijt}$. While equation (15) specifies premium as the dependent variable, we also report specifications where plan characteristics are the dependent variables.

We adjust our standard errors of β in (15) for the fact that we estimated β with a two step estimator, by modifying the formula proposed by Newey (2009). Newey accounts for the standard errors of the selection parameters on the second stage variance formula; we modify his formula to two-way cluster the standard errors, by employer and year.

Finally, we discuss identification. To obtain consistent estimates for the parameters of interest, identification of both the selection equation (13) and the treatment equation (15) is required. It is well understand that selection models are most credibly identified with exclusion restrictions in the treatment equation. In our case, enrollee risk and industry fixed effects provide useful exclusion restrictions. In particular, we assume that individual risk and variation across industries in employment turnover rates affect the probability of leaving the sample, but do not affect the pass through given the set of controls in the second stage. We also use multiple data sources to evaluate the robustness of our selection equation to specification.

Note further that our treatment equations include employer or enrollee fixed effects in most specifications. In this case, our identification of β is based on changes in p_{jt} following changes in R_{jt}. The selection correction further non-parametrically corrects β to estimate its value if everyone selected into USIC insurance. Because we control for the baseline health status with fixed effects and also control for selection, we believe that it is reasonable to consider changes in the risk score—which reflect changes in expected health expenditure for the population of potential enrollees conditional on the base level—to be exogenous.

5 Estimation Results

We now present our estimation results. We start with the impact of health risk on claims costs, then exposit our results on sample selection, and finally discuss the impact of health risk on premiums.

Table 5: Pass through from expected risk to claims

	Dependent variable:		
	Paid amount ($)	Allowed amount ($)	OOP amount ($)
Regressor:	(1)	(2)	(3)
Enrollee ACG score, r_{ijt}	4,003***	4,483***	480***
	(129)	(131)	(9)
Market FE	Yes	Yes	Yes
Observations	204,913	204,913	204,913

Note: each observation is one enrollee during one year. The dependent variables indicate three measures of the total claims amount for that enrollee. The sample is covered individuals with an ACG score in 2014 only. Markets are defined by USIC and roughly represent an MSA or state. Standard errors are clustered at the employer level. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

5.1 Impact of health risk on claims costs

Table 5 presents the estimated relationship between health risk and claims, which is γ. From column 1, we find that an increase in ACG score of 1—which would mean doubling the score relative to the population mean—would lead to an expected increase in USIC-paid claims of γ =\$4,003. From column 2, an increase in ACG score increases the allowed amount of the claims by \$4,483. This latter figure includes the portion for which payment is the responsibility of the enrollee as well as the amount that USIC expects to pay for the claim. Not surprisingly, the coefficient on the out-of-pocket amount—which is reported in column 3—at \$480, is the difference between these two coefficients.

Robustness. Table A.1 in on-line Appendix A provides robustness on the evidence presented in Table 5 by using splines. Columns 1 and 2 use splines with cut points of 1, 2.5, and 5, chosen as round numbers that differentiate enrollees with serious chronic diseases from others. These results are very similar to our base results. Columns 3 and 4 use splines defined by quartiles of our in-sample ACG score distribution. The results here show more non-linearity, which might be due to a small number of outliers with high medical costs. Overall, our takeaway is that our base coefficient of \$4,003 is a reasonable approximation of the impact of risk score on expected claims.

5.2 Selection equation estimates

Table 6 presents the results of the sample selection equations that we then use to estimate β. Columns 1 and 2 present the average marginal effects and the standard errors, respectively, of the selection equation estimated using the USIC sample. Consistent with the findings that quitters had a higher R in Table 2, R here also positively predicts leaving the USIC sample. Enrollee age, employer size, and the industry fixed effects are also statistically significant.

Columns 3 and 4 of Table 6 present the average marginal effects and the standard errors, respectively, of the selection equation with the MEPS sample. Enrollee age and the industry fixed effects are the only variables that are statistically significant here. The lack of significance may be due to the MEPS sample size being much smaller.

Table 6: Selection equation estimates using USIC and MEPS samples

	Sample USIC		Sample MEPS	
Dependent variable:	*Drop coverage*$_{ijt}$		*Decline insurance*$_{ijt}$	
	(1)	(2)	(3)	(4)
	Average marginal effect	Standard error	Average marginal effect	Standard error
Health risk for enrolled, R_{jt}	0.067***	(0.009)		
Individuals ACG score, $r_{ij,t}$	0.008	(0.008)		
Age$_{ijt}$	-0.001***	(0.0001)	0.005***	(0.001)
Female$_{ijt}$	0.003	(0.003)	-0.039	(0.261)
Employer size$_{jt}$	0.001***	(0.0002)	0.001	(0.001)
Hypertension$_{j,t-1}$			-0.001	(0.030)
Heart disease$_{j,t-1}$			0.089	(0.092)
AMI$_{j,t-1}$			-0.177	(0.121)
Ischemic stroke$_{j,t-1}$			-0.116	(0.124)
Respiratory failure$_{j,t-1}$			0.064	(0.063)
Cancer$_{j,t-1}$			-0.054	(0.061)
Diabetes$_{j,t-1}$			0.019	(0.051)
Asthma$_{j,t-1}$			0.027	(0.041)
Industry fixed effects	Yes	Yes	Yes	Yes
Observations	204,913		1,355	

Note: in the USIC sample each observation is one enrollee during one year. "Drop coverage$_{ijt}$" indicates that individual was in sample in period t but not $t+1$. In the MEPS sample each observation is one individual in the consolidated panel 18 for 2013-14. "Decline insurance$_{ijt}$" indicates that the individual was offered insurance through the employer and declined this insurance coverage. R_{jt} is calculated based on individuals that worked in the employer last year and had a ACG score. Standard errors are clustered at the employer level. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

5.3 Main results: impact of health risk on premiums

We now investigate our main parameter of interest, the pass through from health risk to premiums, which is β. Table 7, panel A, column 1 regresses the employer mean ACG score among the insured, R, on the mean per-enrollee premium for the employer. This regression is at the employer/year level and it includes employer fixed effects but does not control for selection. Recall that our sample is employers with at least two complete observations. We find that a unit increase in employer mean ACG risk score for an employer results in a $188 increase in premiums. Column

2 presents similar specifications but at the enrollee/year level and it includes enrollee fixed effects. With this specification we find similar results: a unit increase in enrollee mean ACG risk score for an employer results in a $195 increase in premiums.

Panel A, columns 3 and 4 present the results where we control for sample selection by specifying first- and sixth-order polynomials for $g(S_{ijt})$, respectively. Both specifications include enrollee fixed effects. The estimates of β are $663 for the first-order polynomial and $624 for the sixth-order polynomial. The increase in the coefficient is consistent with enrollees disproportionately leaving the sample if they receive a high pass through from claims to premiums. While the sample selection controls increase the estimate of β, the coefficient remains small relative to γ.

Table 7: Pass through from risk to premiums with USIC sample correction

	Observation level:			
	Employer/year	Enrollee/year	Enrollee/year	Enrollee/year
	No selection correction		With selection correction	
	(1)	(2)	(3)	(4)
Panel A: specifications with employer/enrollee fixed effects				
Health risk for enrolled, R_{jt}	188**	195***	663***	624***
	(87)	(82)	(132)	(121)
Panel B: specifications without employer/enrollee fixed effects				
Health risk for enrolled, R_{jt}	1,749***	2,263***	2,594***	2,811***
	(120)	(88)	(174)	(116)
FE Year	Yes	Yes	Yes	Yes
Polynomial Order	No	No	1^{st}	6^{th}
Observations	31,044	448,259	448,259	448,259

Note: each observation is either one employer or enrollee during one year. The dependent variable is the premium charged the employer by USIC divided by the number of covered lives. R_{jt} is calculated based on individuals that worked in the employer last year and had a ACG score. Column (1) in Panel A includes employer fixed effects. Columns (2) to (4) in Panel A include enrollee fixed effects. Panel B includes market fixed effects. Columns (3) and (4) control non-parametrically for selection with polynomial of the selection probability. Markets are defined by USIC and roughly represent an MSA or state. Standard errors are two-way clustered at the enrollee and year levels. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Panel B of Table 7 presents analogous results to panel A but without employer/enrollee fixed effects (though with market fixed effects, using USIC's market definitions). Again, column 1 is at the employer/year level and all the other columns at the enrollee/year level. The estimates show substantially higher pass through than our estimates with employer/enrollee FE estimates. Estimates in columns 1 and 2 do not include sample selections controls, while columns 3 and 4 include the first- and sixth-order polynomials, respectively. The enrollee/year level coefficient in Column 2 is $2,263 and increases to $2,811 when we control by sample selection. Thus, the selection correction changes the estimate of β in the same direction as in the enrollee fixed effects specification.

Takeaways. We offer four main takeaways from these estimates. First, they imply a much smaller pass through than under our benchmark model with full experience rating. In particular, among our four coefficients, the pass through from expected costs to premiums ranges from 16 to 70%, versus 100% under the benchmark model.

Second, we interpret the differences between our estimates with enrollee fixed effects and those without the fixed effects as reflecting the difference in pass through to existing enrollees and to new enrollees. In particular, our specifications with enrollee fixed effects estimate USIC's pass through for enrollees already insured by USIC while our non-fixed-effects specifications estimate the equilibrium relationship between health risk and premiums, which includes the pass through on new accounts. Thus, USIC appears to pass through very little risk for existing customers, while new customers receive premiums that are more risk-based.[22]

Third, our selection-corrected results are somewhat larger than uncorrected results. This implies that people with a high pass through from risk to premiums disproportionately quit USIC insurance. It is consistent with our finding that higher risk people disproportionately quit USIC insurance and with our prior that individuals with unobservably higher premiums would disproportionately quit USIC insurance.

Finally, idiosyncratic premium risk—i.e., premium variation not correlated with health risk—is also present in this market, with a standard deviation of $576 in our selection controlled estimates in panel A, column 4. Though not correlated with health risk, idiosyncratic premium risk also lowers consumer welfare for risk averse consumers, and hence our counterfactuals need to take into account.

Robustness. We consider a number of robustness checks of our main results, with tables and figures in on-line Appendix A. First, we examine robustness to our selection correction. Table A.2 presents the results of our main estimates when we control for selection using the MEPS sample instead of the USIC sample, with analogous specifications to Table 7. The estimates of the pass through are smaller than in Table 7 for both the fixed-effects and the non-fixed-effects specifications. Overall, the estimates using the MEPS sample correction show a very similar pattern to the main estimates.

Second, we consider the bias from measurement error from the fact that we use the ACG risk

[22]Our counterfactuals incorporate this view, specifying that enrollees obtain the fixed-effects pass through when they stay with the existing policy and the non-fixed-effects pass through when they switch policies.

score instead of USIC's risk score. For a subset of 3,759 enrollees in 2013, we have USIC's own risk score.[23] We found a high correlation between the USIC score and the ACG score for this subsample. In particular, the linear (Pearson) correlation between the ACG risk score and the USIC risk score is 0.835 and the rank Spearman correlation is 0.881.[24] We then computed the size of the true pass-through coefficient assuming that the measurement error is uncorrelated with the regressors,[25] using $\beta^{estimate} = \beta^{true} \frac{\sigma_{RR}}{\sigma_{RR}+\sigma_{\omega\omega}}$ (Wooldridge, 2010, p. 75), where $\sigma_{\omega\omega}$ is the variance of the measurement error and σ_{RR} is the variance of the (USIC) true risk score. We calculated $\sigma_{RR} = 1.92$ for our estimation sample and estimated $\sigma_{\omega\omega}$ as the standard deviation of the residual from a regression of the ACG score on the USIC score, obtaining 0.74, which yields $\beta^{true} = \$863$. Thus, the measurement error here can only explain a very small part of the pricing relative to full experience rating. We also constructed our own risk score (ORS) measure using a random forest algorithm, and instrumented for the USIC risk score with ORS. We find a pass through that is smaller than the estimated using the ACG score that is not statistically different from zero.

Third, we show robustness of β to a non-linear specification. Table A.3, columns 1 and 2 present the analogous results to our main results in Table 7, panel A, column 4, when we consider a linear spline relationship between the risk score and claims. The results generally show a roughly linear relationship between risk scores and premiums, which is somewhat lower for employers with higher risk scores. Table A.3, columns 3 and 4 present the same relationship, but stratified across smaller and larger employers. The results here are very similar to the results with all employers.

Fourth, Table A.4 presents similar specifications to our main results in Table 7, panel A, column 4 but with the addition of the percent of enrollees with specific chronic diseases. We chose cancer, transplants, AMIs (heart attacks) and diabetes (in Panel A), and hypertension, heart failure, kidney chronic disease and asthma (in Panel B), as these diseases result in persistent increases in the costs of healthcare, and they may serve as markers that insurers use to price risk. The pass through from the risk score to premiums is very similar, ranging from \$625 to \$648, which are not very different from our main estimate of \$624. While increases in the percent of enrollees with cancer and heart failure increase premiums, we do not find significantly significant and positive effects for the other conditions. Our takeaway from this is that the pass through from expected claims to premiums is

[23] USIC develops this score in-house for use in its risk rating decisions. USIC was unable to recover their risk score for the rest of the sample.

[24] We also estimated β for this subsample using USIC's risk score instead of the ACG score and we found a result that is not statistically different from our main result.

[25] The measurement error here is likely uncorrelated, since the point estimate of β does not change in different specifications when we add different controls.

very stable to the inclusion of these chronic diseases.

Fifth, we perform a robustness check where we explore the size of the pass-through coefficient when we include the presence of a large claim in the previous period. The idea is that USIC may only adjust premiums following a very large claims due to some perceived fixed cost of adjustment. In estimates not reported in the paper we add variables that indicate claims larger than $50,000, $75,000 and $100,000 to our base specification. In all these specifications, the pass-through coefficient is very similar to in our preferred specifications. As another unreported robustness check, we include the variance of risk across enrollees, to check if there is any conditional response of premiums to the variance. Here again we find that the pass-through coefficient from premiums to risk does not change in a statistically significant way and that the coefficient on the variance term is not statistically significant.

Sixth, in Table A.5 we examine whether changes in health risk lead to changes in the plan benefits that the employer chooses, again using similar specifications to our main results. We consider three measures of plan benefits: the out-of-pocket maximum, the coinsurance rate, and the in-network deductible. For each of our benefit measures, the estimated coefficient on health risk is small and not statistically significant, suggesting that employers do not systematically substitute to plans with different benefit structures following health risk shocks.

Seventh, we consider whether the low pass through may be due to the planned roll-out of community rating regulations, that may have affected insurers' pass through. We checked whether there was a change in the pass-through coefficient between 2013-2014 and 2014-2015, because of the gradual phase-in of ACA regulations. We present the results from our base specification but with the two different samples in Table A.6. The estimated coefficient of pass through for the first period is somewhat smaller than in our main specification but not statistically significant from our main estimate. The point estimate for the 2014-2015 sample is lower, and not statistically different from the estimate for the first period. Overall, the evidence does not suggest that our estimated passthrough is affected by changes in regulations after 2014.

Finally, our analysis suggest that there are not large differences in the pass-through results by employer size. To analyze this, we split the sample based on deciles of the distribution of employers size. Figure A.3 presents the results for our base specification but where we split the pass-through coefficient by deciles. The pass through remains stable across the size distribution.

Interpretation. Our effects paint a very different picture from the benchmark model of full experience rating. We find much less reclassification risk in the small group market than this model, particularly among groups continuously enrolled with USIC. This is also different from what many observers thought was likely occurring in this market (Gruber, 2000). However, it is consistent with the low pass through from costs to prices that is observed in a variety of contexts, including energy costs (Ganapati et al., 2016), tariffs (De Loecker et al., 2016), and beverage taxes (Cawley et al., 2018). We now consider potential explanations for this finding.

First, our results may be driven by an implicit commitment by USIC to offer "guaranteed renewability" dynamic contracts with one-sided price commitment (Pauly et al., 1995). Such contracts can partially or fully mimic long-run contracts (Handel et al., 2015). We consider pass through under Handel et al.'s equilibrium dynamic one-sided commitment contracts, by using the Handel et al. data and estimating a weighted regression of the change in health risk score on the change in equilibrium premiums.[26] We find that $\beta = 1,821$ with a standard error of $410, an estimate that is in the range of our estimates.[27] The optimal β is likely smaller with consumer switching costs, as found by (Handel, 2013) in similar settings, switching costs imply that healthy enrollees have a lower ability to renegotiate. Overall, our results suggest that USIC offering implicit guaranteed renewability contracts may explain much of the limited pass through.

We also consider three other different explanations for our findings, finding little support for any of them. First, one possibility is that when an employer has an increase in R, its average risk score, USIC passes through the expected costs to premiums slowly over time, rather than immediately. In order to test this proposition, Table 8 reports the pass through using the current and lagged ACG scores, with specifications analogous to our main specifications with and without fixed effects (which are repeated in columns 1 and 3, respectively). Our fixed-effects specification with a lagged risk score in column 2 shows no evidence that employers raise their premiums based on the lagged risk score. In contrast, our specification in column 4 without fixed effects shows a positive and significant estimate on lagged risk score, with the sum of the coefficients adding up roughly to the unique risk score coefficient in column 3. Our interpretation is again that employers that are new to USIC face a premium that is relatively risk based while employers with existing USIC accounts do not see much variation in premiums as their risk changes.

Second, our results may be driven by USIC having pricing power and choosing pass through to maximize revenues based on this pricing power. Suppose that the insurer set premiums as a

[26]We use Handel et al.'s health status transitions for ages 30-35 and consider the first two years of the equilibrium long-term contract for individuals aged 25 with a flat net income.

[27]Using the health statuses from people from 40-45 years old the beta estimate is similar ($\beta = 1,720$ with a standard error of $390).

Table 8: Pass through from risk to premiums with lagged risk score

Dependent variable: annual employer mean premium, p_{jt}				
Regressor:	(1)	(2)	(3)	(4)
Health risk for enrolled, R_{jt}	624**	450	2,811***	1,822**
	(121)	(92)	(127)	(112)
Lagged health risk for enrolled, $R_{j,t-1}$		218		1,311**
		(54)		(98)
Year FE	Yes	Yes	Yes	Yes
Enrollee FE	Yes	Yes	No	No
Market FE	No	No	Yes	Yes
Polynomial Order	6^{th}	6^{th}	6^{th}	6^{th}
Observations	448,259	160,062	448,259	264,145

Note: each observation is one enrollee during one year. The dependent variable is the premium charged the employer by USIC divided by the number of covered lives. R_{jt} is calculated based on individuals that worked in the employer last year and had a ACG score. Markets are defined by USIC and roughly represent an MSA or state. All estimates control non-parametrically for selection with a 6th order polynomial of the selection probability. Standard errors are two-way clustered at the enrollee and year levels. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

single firm maximizing static profits. In this case, we would expect that its price in each market would vary with its residual demand in that market. While it would be difficult to estimate USIC's residual demand across markets, we believe that residual demand would likely vary based on the market concentration in a particular area. Thus, we test whether our results may be generated by static pricing power by evaluating whether USIC's pricing policies vary across markets based on measures of concentration. Table 9 interacts the pass-through coefficient with three different measures of market concentration, specifically the Herfindahl Index (HHI), the market share of the leader insurer, and the number of insurers with more than 5% of market share, in panels A, B and C, respectively.[28] In all cases, the interaction measures are not statically significant. Thus, we do not find evidence that the low levels of pass through that we estimate are driven by insurer pricing power. Additionally, we present the variation across different states in Figure A.2 in on-line Appendix A. We find no significant pattern in the variation of pass through across states.

Third, our results may be driven by consumer search. To understand the size of the effect of consumer search on pass through, we calculate a back-of-the-envelope estimate of the potential pass through from search frictions using Cebul et al. (2011)'s estimated model of search frictions for health insurance. In equation (13), Cebul et al. define average premiums as $\bar{p} = c + \frac{\gamma}{1+\gamma}(p^R - c)$, where c is marginal cost, p^R is the maximum willingness to pay for insurance, and γ is the "market friction parameter." Using this equation and their estimated $\gamma = 0.153$ from Panel A of Table

[28]We define these indices at the state level using Kaiser Family Foundation data for 2013-15.

Table 9: ACG score and claims pass-through to premiums, by market concentration

Dependent variable: annual employer mean premium, p_{jt}			
	(1)	(2)	(3)
Panel A: market HHI			
Health risk for enrolled, R_{jt}	624** (121)	617** (146)	700** (116)
HHI		0.035*** (0.008)	0.057 (0.032)
R_{jt} × HHI			−0.023 (0.057)
Employer FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Observations	448,259	448,259	448,259
Panel B: share of largest insurer			
Health risk for enrolled, R_{jt}	624** (121)	617** (87)	711** (101)
Share of leader insurer		280** (58)	447** (245)
R_{jt} × Share of leader insurer			−174 (228)
Employer FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Polynomial Order	6^{th}	6^{th}	6^{th}
Observations	448,259	448,259	448,259
Panel C: number of insurers with 5% or larger market share			
Health risk for enrolled, R_{jt}	624** (121)	615** (116)	605 (231)
Number of insurers with 5%+		−58 (19)	− 60 (73)
R^p_{jt} × Number of insurers with 5%+			2.5 (61)
Employer FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Polynomial Order	6^{th}	6^{th}	6^{th}
Observations	448,259	448,259	448,259

Note: each observation is one enrollee during one year. Markets are defined by USIC and roughly represent an MSA or state. HHI indexes, share of leader insurer and number of insurer with 5%+ market share are taken from Kaiser Family Foundation State Health Facts database. All estimates control non-parametrically for selection with a 6th order polynomial of the selection probability. Standard errors are two-way clustered at the enrollee and year levels. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

12, the pass through can be expressed as: $\frac{\Delta \bar{p}}{\Delta c} = \frac{1}{1+\gamma}$=86.7%. Therefore, their estimated search model implies that about 87% of a cost increase would be passed through to the mean small group employer. Given our pass-through estimate of 16% in the fixed effects specification, our results are also unlikely to be mostly explained by search frictions.

6 Counterfactuals and Welfare

6.1 Simulation of counterfactuals

Using our estimates, we now examine the extent of reclassification risk and the resulting welfare loss under the current pricing environment and counterfactual environments, over a 10-year horizon after the initial insurance enrollment.

We calculate our counterfactuals with three steps. First, we iteratively construct the future distribution of enrollee health risk and mean employer health risk to which an enrollee is exposed, over a 10-year renewal period following the initial insurance enrollment. Following Table 3, which shows that two lags of risk scores are predictive of the current score, we predict the health risk using two lags of the score. Rather than using the coefficients from Table 3, we simulate future risk scores and out-of-pocket expenditures non-parametrically for each individual, using enrollees with similar ACG scores for the two previous periods, for each enrollee.[29]

Second, we evaluate how changes in risk translate into changes in future premiums and out-of-pocket costs. This reclassification risk occurs through two mechanisms. First, health shocks (for the enrollee or others in her group) result in enrollees facing higher premiums; second, enrollees may drop health coverage due to the higher premiums caused by these health shocks. Since both of these sources of risk are potentially important, our counterfactuals capture both sources. Specifically, to evaluate reclassification risk under the observed environment, after simulating the new risk scores R_{ijt} and r_{ijt} each period, we simulate a joint draw from the estimated ε^s and ε^p distributions. We then use ε^s to simulate whether the individual selects into insurance. If so, she receives new premiums based on the estimated β from our main specification with enrollee fixed effects in Table 7, panel A, column 4 (and using the draws of ε^s and ε^p). If she selects out of her current insurance, we assume that she pools with other similar individuals (by selecting insurance at a new but similar employer). To provide conservative bounds on her reclassification risk, we assume that in this case, she receives new premiums based on β estimated without employer/enrollee fixed effects or selection corrections (Table 7, panel B, column 2) and a draw from the unconditional estimated distribution of ε^p.

Third, we examine how this distribution of premiums and out-of-pocket costs translates into a certainty equivalent income level. We use a CARA functional form for our money utility function $u_{ij}(\cdot)$. We do not estimate the CARA risk aversion parameter, but instead use 0.000428, the value

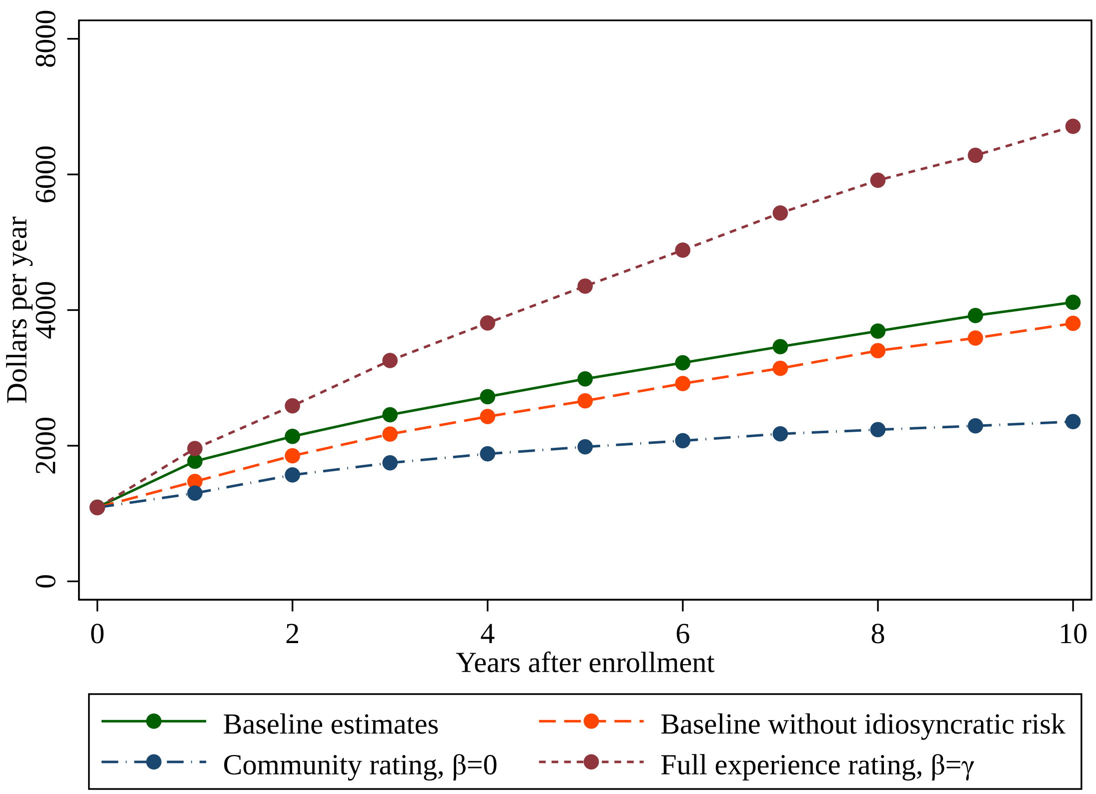
[29]We use a uniform kernel and choose the bandwidth based on Silverman's rule (Hansen, 2018).

from Handel (2013), who estimates risk in a similar context of health insurance choice.[30]

6.2 Counterfactual results

We consider four different pricing environments. First, as the baseline, we examine the observed pricing environment, using our simulation methods described above. Second, we examine the baseline but without idiosyncratic risk. In this case, we follow the same procedure as in the baseline, but we set the residual on premiums to 0. Third, we examine community rating, where $\beta = 0$. Finally, we examine full experience rating, under which $\beta = \gamma$.

Figure 1: Simulated mean certainty equivalent loss from risk across pricing policies

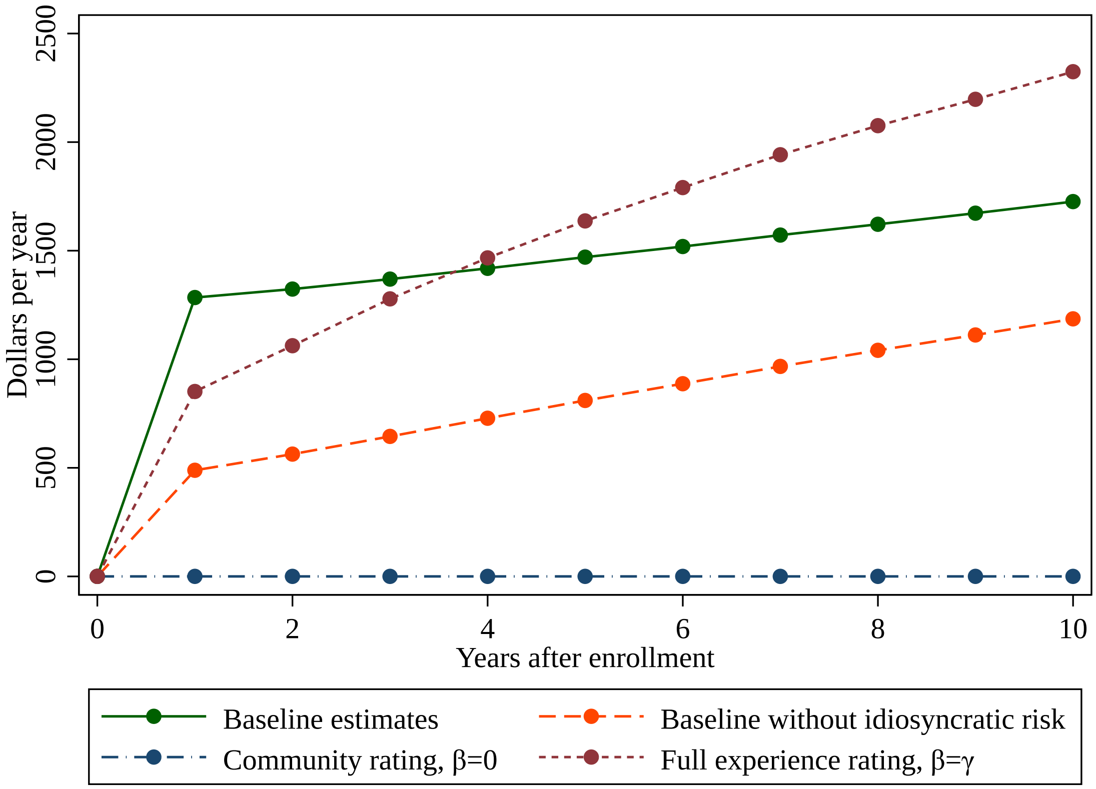


Note: Figure based on authors' calculations as described in paper.

Figure 1 reports the mean across individuals in the certainty equivalent loss from risk across these pricing environments. The baseline estimates show a mean annual certainty equivalent loss

[30]With the CARA utility function, the certainty equivalent income loss of a lottery does not depend on the base income level, and hence we do not specify the income for each enrollee. We also examine robustness using the CARA parameter of 0.00008, which is also used by Handel et al. (2019).

Figure 2: Simulated mean standard deviation in premiums across pricing policies

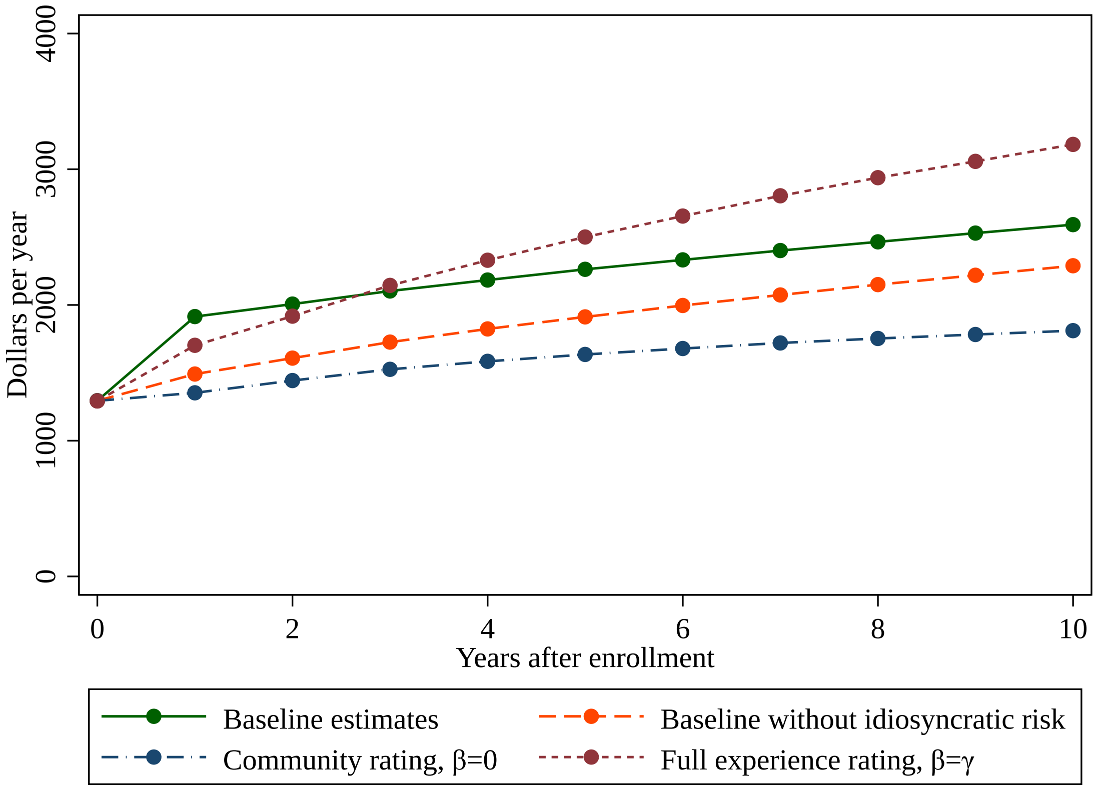


Note: Figure based on authors' calculations as described in paper.

from risk of $3,050 in the ten years after the initial enrollment. Some of this is caused by idiosyncratic premium risk. Without idiosyncratic risk, the mean certainty equivalent loss from risk drops to $2,750. Under community rating, the mean annual certainty equivalent income loss is $1,950, while it is $4,500 under full experience rating. Thus, USIC's observed pricing policy provides approximately 60% of the consumer welfare protection from reclassification risk as does community rating, relative to full experience rating. Even though USIC's pricing policy for existing customers exhibits very little experience rating, turnover in this market is large, which limits this protection. In addition, even under community rating, the possibility of large out-of-pocket expenditures generates a substantial certainty equivalent loss from risk, particularly later in the sample.

Figures 2 and 3 report the mean standard deviation in premiums and total health spending, respectively. The numbers here follow Figure 1 pretty close here, with the baseline policy without idiosyncratic risk having about half of the standard deviation in premiums of the full experience rating case relative to the community rating case. Consistent with the welfare loss from out-of-

Figure 3: Simulated mean standard deviation in healthcare expenditures across pricing policies



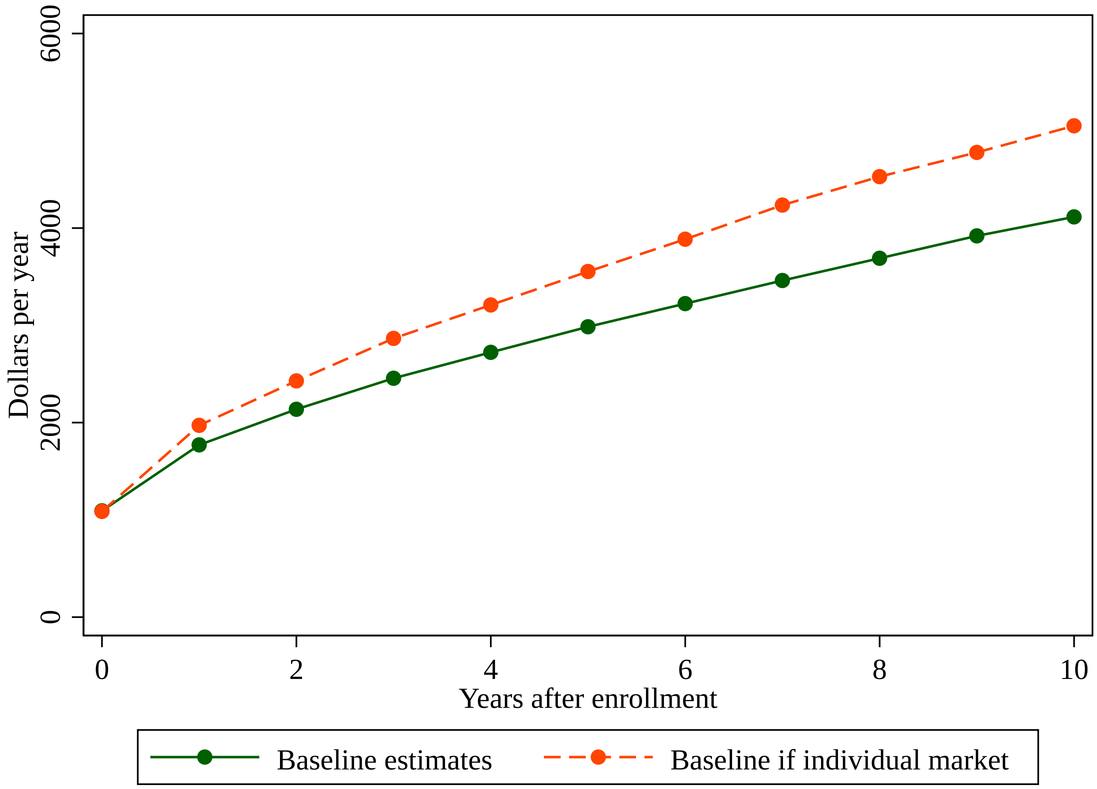
Note: Figure based on authors' calculations as described in paper.

pocket costs in Figure 1, the standard deviation on total healthcare expenditures is $1,300 in the initial year, stemming exclusively from out-of-pocket expenditures. This rises gradually over time under all of the four scenarios.

One difference here is that the baseline policy has a greater standard deviation in the early years than other policies. But, because this risk is $i.i.d.$, this standard deviation does not generate as big a welfare impact over time as the premium risk from experience rating.

Finally, Figure 4 examines the extent to which pooling in the small group market adds value. Specifically, it performs a simulation similar to our baseline estimate but where each individual's premiums is based on her risk score, r_{ijt}, rather than on her employer's mean risk score, R_{jt}. We find that the value generated by pooling in the small group market is moderate. Specifically, we find that the mean difference in the certainty equivalent income loss from being exposed to USIC's current pricing policies and being in a small group versus being exposed to the same pricing policies but purchasing individual insurance is $600 per year over the 10-year horizon.

Figure 4: Simulated mean certainty equivalent loss with no pooling within group



Note: Figure based on authors' calculations as described in paper.

Robustness. Figure A.4 replicates Figure 1, considering the certainty equivalent income loss with the lower CARA risk aversion coefficient of 0.00008 used by Handel et al. (2019). It finds that the same relative pattern of certain equivalent losses hold as in Figure 1. However, the dollar values of the certainty equivalent income losses are much smaller. For instance, there is a $200 mean annual loss under the baseline instead of $3,050 with the base risk aversion parameter.

7 Conclusion

In this paper, we seek to understand the extent of reclassification risk in the small group insurance market from a period before ACA community rating regulations were effective. We develop a simple two-period model of group insurance. Our model considers employers and enrollees who, respectively, choose whether or not to offer and take-up employer-sponsored health insurance. We allow for the insurer's pricing to each employer to be based on the expected claims risk of the employer and for enrollees to select out of insurance based on their premiums and risk scores. The model highlights how the pass through from expected claims cost to premiums forms a sufficient statistic for understanding reclassification risk in this market and how to control for the impact of enrollee selection on pass through.

To estimate the pass-through coefficient, we use a unique dataset from a large U.S. health insurer, "United States Insurance Company" (USIC), with premium information on over 12,000 employers, and claims data from more than 300,000 enrollees at these employers. Because employers and individuals frequently start and stop health insurance coverage in this market and because there is some evidence that high-risk people leave coverage, our estimates control non-parametrically for selection. When controlling for selection, we find that the pass through from mean health risk to premiums is 16% with enrollee fixed effects and 70% without employer fixed effects; the pass-through estimates without controls for selection are moderately smaller. The fixed effects results are much closer to community rating than experience rating. This limited reclassification risk may be due to "guaranteed renewability" one-sided commitment contracts where USIC imposes little risk rating on existing customers. It cannot be explained by slow pass through over time, market power, or search frictions.

Finally, we simulate counterfactuals to evaluate the extent to which USIC insurance provided value in the form of protection from reclassification risk in the small group market. To compute this, we non-parametrically simulate the evolution of health risk for an employer over a ten-year horizon and evaluate how this would translate into selection into and out of insurance and a welfare loss from financial risk. We use our estimated pass-through parameters, our estimated evolution of selection and health risk, and a CARA risk aversion parameter taken from the literature. We find that the observed USIC policy adds about 60% of the difference in consumer welfare between full experience rating and community rating. The high turnover of enrollees limits the value from the guaranteed renewability feature of USIC's contracts. Substantial out-of-pocket costs in this market also generate substantial welfare loss. The value from pooling within a small group is relatively small, at $600 per year under the baseline policy. Hence, the individual insurance market would have only moderately lower welfare if the same pricing policies were used.

References

Baicker, K. and Chandra, A. (2006). The labor market effects of rising health insurance premiums. *Journal of Labor Economics*, 24(3):609–634.

Bhattacharya, J. and Bundorf, M. K. (2009). The incidence of the healthcare costs of obesity. *Journal of health economics*, 28(3):649–658.

Blavin, F., Garrett, B., Blumberg, L., and Buettgens, M. (2014). Monitoring the impact of the affordable care act on employers, 2014; accessed: 05.19.2018. `https://www.urban.org/sites/default/files/publication/33696/413273-Monitoring-the-Impact-of-the-Affordable-Care-Act-on-Employers.PDF`.

Buchmueller, T. and DiNardo, J. (2002). Did community rating induce an adverse selection death spiral? evidence from New York, Pennsylvania, and Connecticut. *American Economic Review*, 92(1):280–294.

Bundorf, M. K., Levin, J., and Mahoney, N. (2012). Pricing and welfare in health plan choice. *American Economic Review*, 102(7):3214–48.

Cawley, J., Willage, B., and Frisvold, D. (2018). Pass-through of a tax on sugar-sweetened beverages at the philadelphia international airport. *Jama*, 319(3):305–306.

Cebul, R. D., Rebitzer, J. B., Taylor, L. J., and Votruba, M. E. (2011). Unhealthy insurance markets: Search frictions and the cost and quality of health insurance. *American Economic Review*, 101(5):1842–71.

Chetty, R. (2008). Moral hazard versus liquidity and optimal unemployment insurance. *Journal of Political Economy*, 116(2):173–234.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics*, 1(1):451–488.

Chetty, R. and Saez, E. (2010). Optimal taxation and social insurance with endogenous private insurance. *American Economic Journal: Economic Policy*, 2(2):85–114.

CMS (2017). Insurance standards bulletin series. extension of transitional policy through calendar year 2018; accessed: 08.15.2016. `https://www.cms.gov/CCIIO/Resources/Regulations-and-Guidance/Downloads/Extension-Transitional-Policy-CY2018.pdf`.

Crocker, K. J. and Moran, J. R. (2003). Contracting with limited commitment: Evidence from employment-based health insurance contracts. *RAND Journal of Economics*, pages 694–718.

Cutler, D. (1994). Market failure in small group health insurance. Technical report. NBER Working Paper 4879.

Cutler, D. M. and Reber, S. J. (1998). Paying for health insurance: the trade-off between competition and adverse selection. *The Quarterly Journal of Economics*, 113(2):433–466.

De Loecker, J., Goldberg, P. K., Khandelwal, A. K., and Pavcnik, N. (2016). Prices, markups, and trade reform. *Econometrica*, 84(2):445–510.

Einav, L., Finkelstein, A., and Cullen, M. R. (2010). Estimating welfare in insurance markets using variation in prices. *The Quarterly Journal of Economics*, 125(3):877–921.

Ganapati, S., Shapiro, J. S., and Walker, R. (2016). The incidence of carbon taxes in us manufacturing: Lessons from energy cost pass-through. Technical report, National Bureau of Economic Research.

Gowrisankaran, G., Norberg, K., Kymes, S., Chernew, M. E., Stwalley, D., Kemper, L., and Peck, W. (2013). A hospital system's wellness program linked to health plan enrollment cut hospitalizations but not overall costs. *Health Affairs*, 32(3):477–485.

Gruber, J. (2000). Health insurance and the labor market. In *Handbook of health economics*, volume 1, pages 645–706. Elsevier.

Handel, B., Hendel, I., and Whinston, M. D. (2015). Equilibria in health exchanges: Adverse selection versus reclassification risk. *Econometrica*, 83(4):1261–1313.

Handel, B., Hendel, I., and Whinston, M. D. (2019). Welfare effects of dynamic health insurance contracts. Working Paper.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *The American Economic Review*, 103(7):2643–2682.

Hansen, B. E. (2018). Econometrics. https://www.ssc.wisc.edu/~bhansen/econometrics/.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, 47(1):153–161.

Herring, B. and Pauly, M. V. (2006). Incentive-compatible guaranteed renewable health insurance premiums. *Journal of Health Economics*, 25(3):395–417.

Jost, T. (2015). PACE and EACH Acts pass House; CMS addresses consumers enrolled in multiple plans (updated); accessed: 08.15.2016. http://goo.gl/JFYUFT.

Jost, T. (2017). Administration allows states to extend transitional policies again; accessed: 05.21.2018. https://www.healthaffairs.org/do/10.1377/hblog20170223.058912/full/.

Kaiser Family Foundation (2013). Market share and enrollment of largest three insurers- small group market, 2013; accessed: 08.15.2016. http://goo.gl/qk8mKg.

Kowalski, A. E. (2015). Estimating the tradeoff between risk protection and moral hazard with a nonlinear budget set model of health insurance. *International Journal of Industrial Organization*, 43:122–135.

Newey, W. K. (2009). Two-step series estimation of sample selection models. *The Econometrics Journal*, 12:S217–S229.

Pauly, M. V. and Herring, B. (1999). *Pooling health insurance risks*. American Enterprise Institute.

Pauly, M. V., Herring, B., and Song, D. (2006). Information technology and consumer search for health insurance. *International Journal of the Economics of Business*, 13(1):45–63.

Pauly, M. V., Kunreuther, H., and Hirth, R. (1995). Guaranteed renewability in insurance. *Journal of Risk and Uncertainty*, 10(2):143–156.

Pauly, M. V. and Lieberthal, R. D. (2008). How risky is individual health insurance? *Health Affairs*, 27(3):w242–w249.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

A Appendix

A.1 On-line Appendix A

Figure A.1: States in our estimation sample

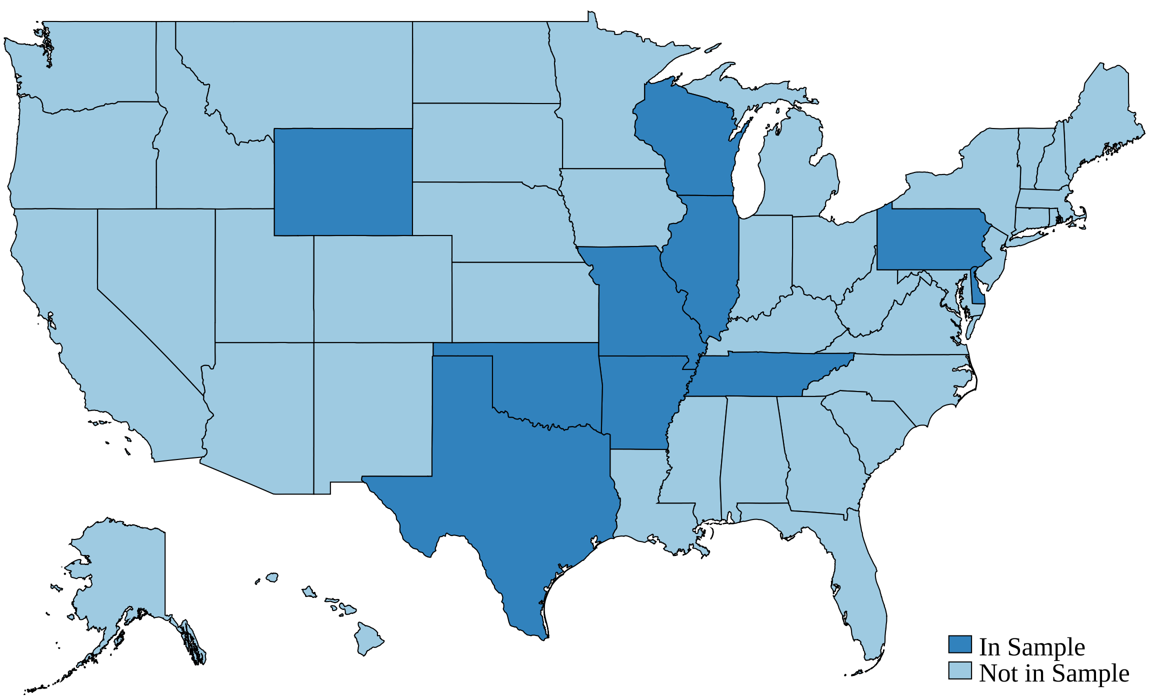
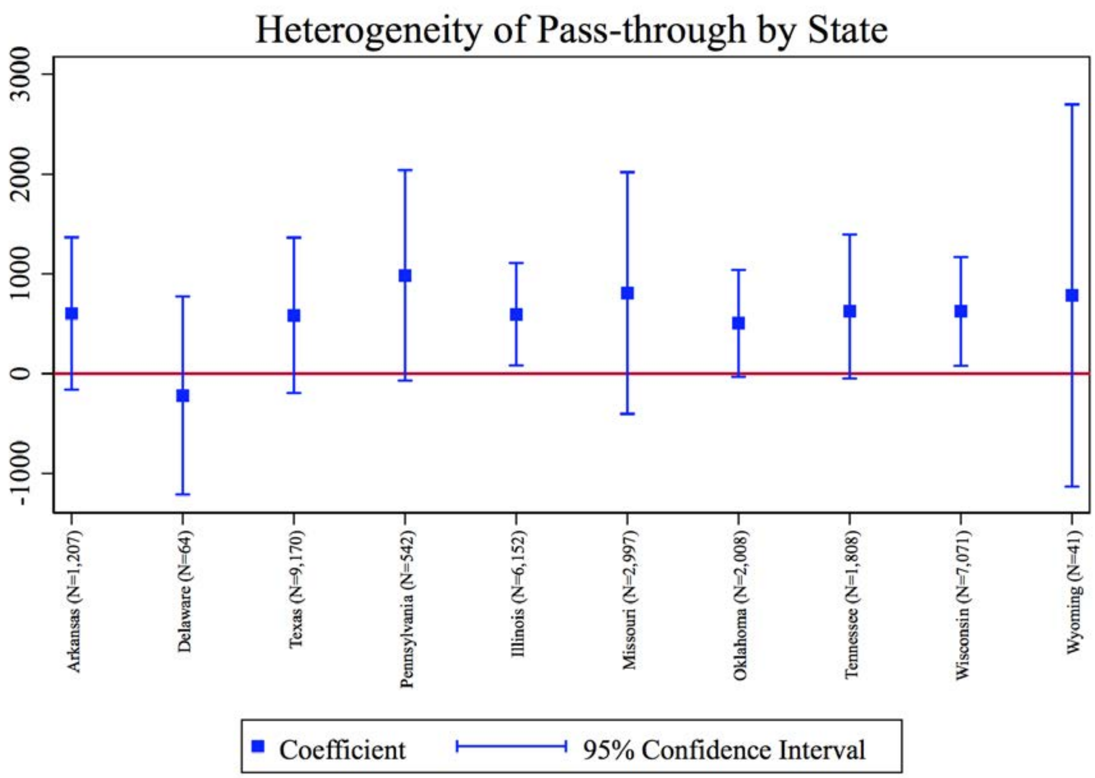


Figure A.2: Effect of risk on premium by state

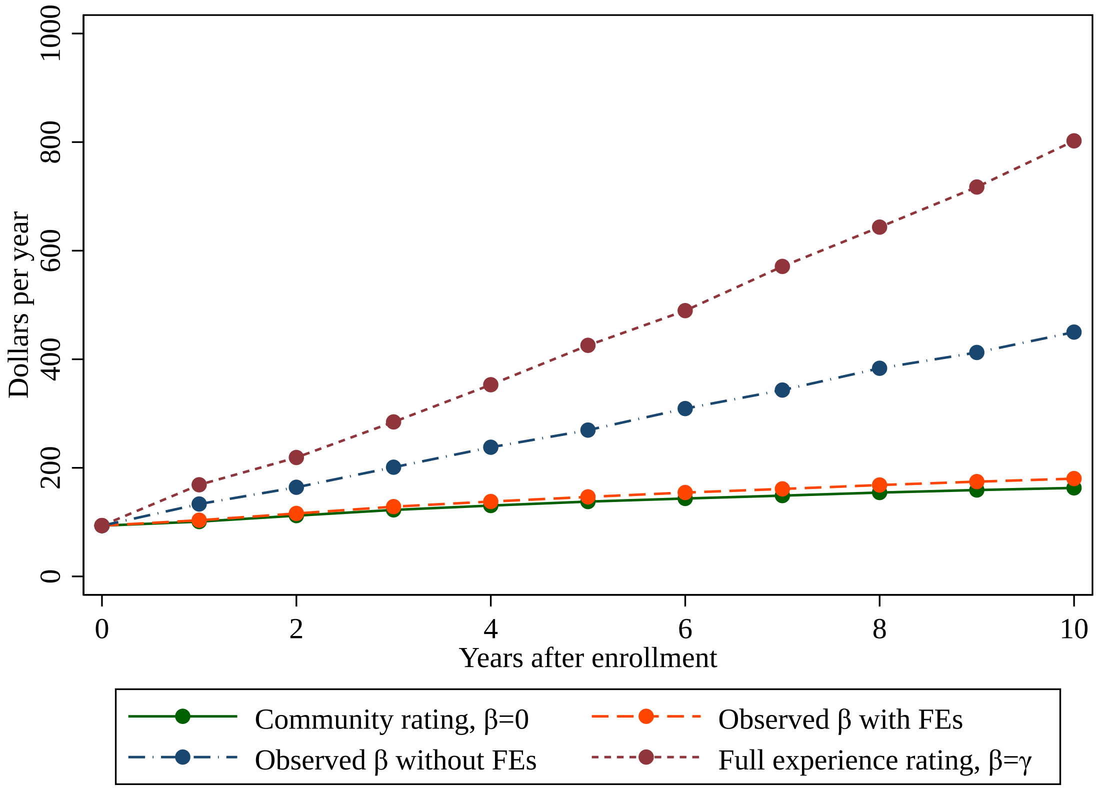


Note: Coefficients from our preferred specifications allowing different effects by state.

Figure A.3: Effect of risk on premium by employer size

Note: Coefficients from our preferred specifications allowing different effects by size.

Figure A.4: Simulated mean certainty equivalent loss from risk, with lower risk aversion



Note: Figure based on authors' calculations as described in paper. The CARA risk aversion parameter is 0.00008.

Table A.1: Pass-through from expected risk to claims using splines

	Dependent Variable: Paid amount ($)			
Regressor:	(1)	(2)	(3)	(4)
Spline enrollee ACG score, $r_{ijt} \in [0, 1)$	2,746***	2,836***		
	(94)	(96)		
Spline enrollee ACG score, $r_{ijt} \in [1, 2.5)$	3,174***	3,190***		
	(151)	(151)		
Spline enrollee ACG score, $r_{ijt} \in [2.5, 5)$	4,284***	4,282***		
	(361)	(361)		
Spline enrollee ACG score, $r_{ijt} \in [5, \infty)$	4,692***	4,689***		
	(398)	(398)		
Spline enrollee ACG score, $r_{ijt} \in [0, .32)$			2,503***	2,633***
			(559)	(563)
Spline enrollee ACG score, $r_{ijt} \in [.32, .57)$			3,756***	3,814***
			(411)	(411)
Spline enrollee ACG score, $r_{ijt} \in [.57, 1.13)$			1,189***	1,289***
			(421)	(420)
Spline enrollee ACG score, $r_{ijt} \in [1.13, \infty)$			4,345***	4,344***
			(185)	(185)
Market FE	No	Yes	No	Yes
Splines	Fixed cut points	Fixed cut points	Quartiles	Quartiles
Observations	204,913	204,913	204,913	204,913

Note: each observation is one enrollee during one year. The dependent variables indicate the total claims amount paid by USIC for that enrollee. The sample is covered individuals with an ACG score in 2014 only. Markets are defined by USIC and roughly represent an MSA or state. Standard errors are clustered at the employer level. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table A.2: Pass through from risk to premiums using MEPS sample correction

	Observation level:			
	Employer/year	Enrollee/year	Enrollee/year	Enrollee/year
	No selection correction		With selection correction	
	(1)	(2)	(3)	(4)
Panel A: specifications with employer/enrollee fixed effects				
Health risk for enrolled, R_{jt}	188**	195***	195	196
	(87)	(82)	(84)	(102)
Panel B: Estimations with market fixed effects				
Health risk for enrolled, R_{jt}	1,749***	2,263***	2,210**	2,175***
	(120)	(88)	(94)	(272)
FE Year	Yes	Yes	Yes	Yes
Polynomial Order	No	No	1^{st}	6^{th}
Observations	31,044	448,259	448,259	448,259

Note: each observation is either one employer or enrollee during one year. The dependent variable is the premium charged the employer by USIC divided by the number of covered lives. R_{jt} is calculated based on individuals that worked in the employer last year and had a ACG score. Column (1) in Panel A includes employer fixed effects. Columns (2) to (4) in Panel A include enrollee fixed effects. Panel B includes market fixed effects. Markets are defined by USIC and roughly represent an MSA or state. Columns (3) and (4) control non-parametrically for selection with polynomial of the selection probability. Standard errors are two-way clustered at the enrollee and year levels. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table A.3: Pass through from risk to premiums using splines

Dependent variable: annual employer mean premium, p_{jt}				
	(1)	(2)	(3)	(4)
Health risk for enrolled, R_{jt} (.,.72)	943**	2,641***	493	1,088**
	(173)	(241)	(374)	(159)
Health risk for enrolled, R_{jt} [.72,.91)	952**	2,157***	509**	1,063**
	(191)	(183)	(287)	(188)
Health risk for enrolled, R_{jt} [.91,1.12)	941*	4,627***	471	1,010*
	(216)	(201)	(364)	(306)
Health risk for enrolled, R_{jt} [1.12,.)	523**	2,143***	346	586 **
	(105)	(189)	(186)	(114)
Sample	All	All	Smaller employers	Larger employers
Enrollee FE	Yes	No	Yes	Yes
Market FE	No	Yes	No	No
Year FE	Yes	Yes	Yes	Yes
Polynomial Order	6^{th}	6^{th}	6^{th}	6^{th}
Observations	448,259	448,259	81,614	366,645

Note: each observation is one enrollee during one year. The dependent variable is the premium charged the employer by USIC divided by the number of covered lives. R_{jt} is calculated based on individuals that worked in the employer last year and had a ACG score. Smaller employers are those with 13 or fewer covered lives in all sample years; larger employers are all others. Markets are defined by USIC and roughly represent an MSA or state. All estimates control non-parametrically for selection with a 6th order polynomial of the selection probability. Standard errors are two-way clustered at the enrollee and year levels. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table A.4: Pass through from expected risk to premiums, with chronic conditions

Dependent Variable: Annual employer mean premium, p_{jt}					
Panel A: Effect controlling for chronic conditions					
Regressor:	(1)	(2)	(3)	(4)	(5)
Health risk for enrolled, R_{jt}	624**	648**	626**	625**	628**
	(116)	(117)	(135)	(124)	(116)
Lag % cancer at employer		2**			
		(4)			
Lag % transplant at employer			2		
			(2)		
Lag % AMI at employer				1	
				(0.5)	
Lag % diabetes at employer					1
					(0.3)
Year FE	Yes	Yes	Yes	Yes	Yes
Employer FE	Yes	Yes	Yes	Yes	Yes
Polynomial Order	6^{th}	6^{th}	6^{th}	6^{th}	6^{th}
Observations	448,259	448,259	448,259	448,259	448,259
Panel B: Effect controlling for chronic conditions					
Health risk for enrolled, R_{jt}	624**	627**	633**	627**	625**
	(116)	(120)	(119)	(116)	(119)
Lag % hypertension at employer		0.2			
		(0.1)			
Lag % heart failure at employer			2**		
			(0.4)		
Lag % kidney disease at employer				0.7	
				(0.3)	
Lag % asthma at employer					0.2
					(0.1)
Year FE	Yes	Yes	Yes	Yes	Yes
Employer FE	Yes	Yes	Yes	Yes	Yes
Polynomial Order	6^{th}	6^{th}	6^{th}	6^{th}	6^{th}
Observations	448,259	448,259	448,259	448,259	448,259

Note: each observation is one enrollee during one year. The dependent variable is the premium charged the employer by USIC divided by the number of covered lives. R_{jt} is calculated based on individuals that worked in the employer last year and had a ACG score. Chronic disease regressors indicate the mean percent of enrollees with a claim for the disease in the previous year. Markets are defined by USIC and roughly represent an MSA or state. All estimates control non-parametrically for selection with a 6th order polynomial of the selection probability. Standard errors are two-way clustered at the enrollee and year levels. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table A.5: Effects of expected risk on benefits

	Dependent variable		
	In-network maximum OOP ($)	Coinsurance rate (%)	In-network deductible ($)
Regressor:	(1)	(2)	(3)
Health risk for enrolled, R_{jt}	303	-0.43	159
	(113)	(0.57)	(58)
Year FE	Yes	Yes	Yes
Employer-plan FE	Yes	Yes	Yes
Polynomial Order	6^{th}	6^{th}	6^{th}
Observations	448,259	448,259	448,259

Note: each observation is one enrollee during one year. Each dependent variable is a measure of plan benefits. R_{jt} is calculated based on individuals that worked in the employer last year and had a ACG score. Markets are defined by USIC and roughly represent an MSA or state. All estimates control non-parametrically for selection with a 6th order polynomial of the selection probability. Standard errors are two-way clustered at the enrollee and year levels. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table A.6: Pass through from risk to premiums with heterogeneity by different periods

Dependent variable: annual employer mean premium, p_{jt}				
	(1)	(2)	(3)	(4)
Health risk for enrolled, R_{jt}	568*	2,903**	440*	2,766***
	(85)	(74)	(68)	(189)
Sample Years	2013-14	2013-14	2014-15	2014-15
Employer FE	Yes	No	Yes	No
Market FE	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
Polynomial Order	6^{th}	6^{th}	6^{th}	6^{th}
Observations	281,932	325,080	246,358	307,293

Note: each observation is one enrollee during one year. The dependent variable is the premium charged the employer by USIC divided by the number of covered lives. R_{jt} is calculated based on individuals that worked in the employer last year and had a ACG score. Markets are defined by USIC and roughly represent an MSA or state. All estimates control non-parametrically for selection with a 6th order polynomial of the selection probability. Standard errors are two-way clustered at the enrollee and year levels. *** indicates significance at the 1% level, ** indicates significance at the 5% level and * indicates significance at the 10% level.