Reclassification Risk in the Small Group Health Insurance Market*

Sebastian Fleitas† Gautam Gowrisankaran‡ Anthony Lo Sasso§

August 8, 2016

Abstract

Health insurance without long-run contracts does not necessarily provide risk protection over time. Enrollees with a bad and persistent health shock in one year may be faced with higher premiums in the subsequent year. We consider the small group insurance market for a context where insurers could largely pass through expected risk in the form of higher premiums. Using a panel of claims, plan characteristics, and premium data from a large, national insurer, we find that the insurer passes on 39% of an expected increase in mean health risk for an employer in the form of higher per-person premiums. We find no significant impact of health shocks on plan selection. Community rating—as will occur over time under the ACA—would lower an individual's expected standard deviation of the following year's healthcare expenditures by $252, while full experience rating would increase it by $794. Using published risk aversion numbers, community rating would increase average welfare by a dollar equivalent of $200 annually, assuming no changes in plan prices or attributes.

`Preliminary and incomplete.`

JEL Codes:
Keywords: reclassification risk, small group insurance, private insurance markets, imperfect competition

*We thank Tim Dunne, Ben Handel, and seminar and conference participants for helpful comments.
†University of Arizona
‡University of Arizona, HEC Montreal, and NBER
§University of Illinois-Chicago

1 Introduction

How well do competitive markets for health insurance function? This is a central question for health policy research (Bundorf et al., 2011). Most recent healthcare reforms in the U.S. have emphasized decentralized solutions, with the idea that competition will ensure efficient outcomes. For instance, the 2010 Affordable Care Act (ACA) mandates that individuals purchase health insurance from private providers. It is well understand that competition can help ensure efficient outcomes in decentralized markets, including health insurance.

Yet, even *perfect* competition without long-run contracts may not be efficient. To illustrate, consider a market for health insurance that is perfectly competitive but without the possibility of long-term contracts. Such a market will arrive at an insurance premium for each risk pool that is exactly equal to its expected risk, calculated based on factors that are both observable and contractible. This outcome is not sufficient to maximize welfare for health insurance markets, because it does not provide risk protection over time. Specifically, in the absence of pricing regulations, competitive insurers will pass on a health shock at one pool in one year in the form of premium changes in the next year that match the expected future costs of treatment. The pass through of higher expected risk in the form of higher premiums is called *experience rating* while the risk of higher insurance premiums or worse future coverage from a health shock is called *reclassification risk*. The possibility of reclassification risk from competitive markets and the relation of this risk to the lack of long-term health insurance contracts have been long recognized in the health economics literature (Cutler, 1994).

This goal of this paper is to examine reclassification risk in the small group market. This market currently insures individuals at groups with 1 to 50 or 100 members, depending on the state.[1] For the time period and states in our sample, insurers could experience rate small employers with few regulatory restrictions. Our paper estimates the extent to which higher expected claims for a small group are passed through in the form of higher premiums. We

[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.

then quantify the extent to which this experience rating creates reclassification risk. The ACA is phasing in *community rating* for the small group market, whereby insurers will have to charge the same price to all small groups in an area, with price variation allowed only for age and smoking status. Accordingly, the ACA will eventually prohibit passing through an increase in expected risk from a small employer back to that employer. We examine how community rating and full experience rating would affect reclassification risk and welfare in the small group insurance market.

We examine the small group insurance market because experience rating in this market is potentially very important. The small risk pools here leave open the potential for large reclassification risk and hence a large efficiency loss to risk-averse enrollees. For example, consider an individual who works for an employer with 5 employees. Suppose that the individual is diagnosed with a serious disease, perhaps diabetes, with an expected cost of $25,000 per year going forward. A perfectly competitive insurer will increase the premiums to this employer by $25,000, which will then result in an extra $5,000 per employee per year in extra charges. Thus, a competitive market for small groups may provide limited insurance value, since the individual with the health shock may bear a substantial part of the extra cost of her illness in future years. This reclassification risk will be exacerbated if some of the other employees drop coverage in response to the premium increase.

Notably, an insurer with pricing power may have different incentives from a perfectly competitive market in its pricing and benefit decisions for the small group market. While insurers with pricing power may act differently for a number of reasons, we focus here on how pricing power may affect the provision of risk protection. While a competitive market will raise premiums a dollar for every dollar increase in expected risk, an oligopolistic insurer will pass through a potentially different amount, which depends on the change in the demand elasticity that occurs with the premium change resulting from the extra risk. Depending on the shape of the demand curve, this may generate reclassification risk protection.

Moreover, when dealing with enrollees who incur inertia in their choice of health plans,[2]

[2]It is well-documented that there is inertia for health plan purchasers (Handel, 2013) and job lock among individuals (Madrian, 1994). In our case, the purchaser is the employer. Purchaser inertia and job lock together generate inertia at the enrollee level.

a forward-looking, oligopolistic insurer may charge relatively high markups before health shocks are realized but then not raise premiums proportionally in response to changes in expected risk. With inertia, enrollees are effectively committing to restrict their switching of insurers. Given this restriction, the insurer may in turn find it optimal to provide a credible implicit commitment to provide risk protection, since this adds value to enrollees which the insurer can partly capture. Interestingly, total welfare may then be higher in an oligopoly market that charges markups above costs but provides reclassification risk protection than it would be in a perfectly competitive market without long-run contracts.[3]

Our paper builds on a substantial literature that analyzes reclassification risk (Bundorf et al., 2011; Handel et al., 2015; Kowalski, 2015; Cutler, 1994; Einav et al., 2010). Bundorf et al. (2011) seek to understand the welfare impact of consumer choice of plans under different risk pricing mechanisms, using a dataset of 11 employers. Handel et al. (2015) seek to evaluate the equilibrium adverse selection and reclassification risk from a competitive market of exchange firms, while Handel et al. (2016) seek to understand 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 observe the extent to which experience-rated health insurance creates reclassification risk in the real world. Specifically, we empirically recover the extent to which current claims and expected future claims are passed through into future premiums, in a context in which this is permitted. Combining the pass-through measures with the distribution of health shocks then allows us to empirically understand reclassification risk under the current and counterfactual environments and understand the dollar equivalent costs of the reclassification risk. Thus, we can understand the extent to which community rating and other protection against reclassification risk add consumer value relative to potentially higher prices. We are not aware of any other study that has attempted to empirically quantify the reclassification risk from experience-rated health insurance.

Second, our estimation of the small group market is novel. Small group health insur-

[3] A related point has been made by Mahoney and Weyl (2014), who note that standard welfare results do not apply to the insurance market if market power changes the selection of customers.

ance is important in its own right, covering about 18 million people in 2013 (Kaiser Family Foundation). Moreover, there is substantial potential reclassification risk in the small group market stemming from small risk pools. Finally, there is substantial evidence of market failure from adverse selection in this segment. For instance, fewer than 50 percent of small firms even offer health insurance. The ACA may add to adverse selection in this market, by removing the ability of insurers to contract on health status. Thus, understanding the tradeoff here between reclassification risk and adverse selection is important. Note also we study reclassification risk at the group level, while the literature has generally studied it at the individual level.[4]

We analyze data covering 18,260 employers and approximately 1.3 million enrollee-years over 2013-14 observed in 10 states in the small group market. Our data are for enrollees of plans offered by one large insurer, which we refer to as "United States Insurance Company (USIC)" from now on. Prior to 2014, most states—including all the states in our sample—allowed for health insurers to experience rate plans in the small group market, although many states imposed caps on the amount of experience rating via ratings bands.[5] Starting in 2014, the small group market technically started being subject to *community rating* regulations under the ACA, whereby each insurer must pool risk in this segment over all its enrollees regionally. However, the extent of community rating was very small in 2014 and will continue to be small for 3-4 years. With encouragement from the Obama administration, forty states (including 9 of the 10 in our sample) essentially allowed existing insurers to experience rate in 2014 with a gradual planned phase-in to community rating over the subsequent three years. All states and the District of Columbia allowed indefinite experience rating for existing customers who chose to keep their plans.[6] Overall, we believe that rating requirements in our sample are very similar between 2013 and 2014.

Our data are provided by USIC and include enrollment, plan characteristic, and claims

[4]Here, the fact that employers generally cannot base the premiums that they charge to their employees on risk factors limits the reclassification risk at the group level.

[5]See Kaiser Family Foundation, http://kff.org/other/state-indicator/small-group-health-insurance-market-rate-restrictions/.

[6]See http://www.commonwealthfund.org/publications/blog/2014/jun/adoption-of-the-presidents-extended-fix for details.

information. We observe claims information for 2012-13 and enrollment and plan characteristic information for 2013-14. The plan characteristic data provide information on coinsurance rates, copays, deductibles, and covered services for each plan. The enrollment data include the premiums charged by USIC to each employer in the small group market, the eligible number of subscribers at each employer, and the actual number of subscribers at each employer. For each subscriber, these data include the age and gender of the members covered, which include the subscriber and potentially dependents. Finally, we observe detailed information on the medical and pharmaceutical claims for each member, including charges and the amounts paid by USIC and the member for the claim.

Using these data, we first compute a risk score for each enrollee in each year using the ACG methodology developed by Johns Hopkins University. ACG scores have been widely used in the literature as a useful predictor of observable health claims risk (Carlin and Town, 2009; Gowrisankaran et al., 2013; Handel, 2013). We use the medical and pharmaceutical claims from the previous year to predict expected costs in a given year.

In a series of linear regressions, most of which are at the employer-year level and include fixed effects for employers and plans, we examine the extent to which an increase in the mean ACG score of an employer results in increases in premiums. We find that a one standard deviation increase in mean ACG score for an employer increases its mean annual premium by $1,649. We then examine whether factors other than the ACG score predict premium changes from USIC. Most other factors that we examine—including lagged claims and the prevalence of chronic diseases—do not significantly affect the per-enrollee premium. We find little change in the benefits chosen by employers in response to changes in the mean ACG score. Finally, we also find little impact of healthy enrollees leaving a health plan in response to a premium increase, suggesting that inertia is important here.

We then examine the extent to which the ACG score predicts claims. We find that a one-standard deviation increase in the ACG score increases annual claims by an average of $4,218. Dividing the the increase in premiums by the increase in claims, we find that USIC passes on only about 39% in its future expected risk in the form of higher premiums, and thus essentially provides protection from reclassification risk for the remaining 61%.

Using our sample and estimates, we investigate the extent to which the risk protection provided by USIC provides value in the form of protection from reclassification risk in the small group market. We find that under the current regime, out-of-pocket expenditures and reclassification risk lead individuals in 2013 to have an average standard deviation of \$663 in their 2014 expenditures for healthcare and health insurance. With community rating, this standard deviation would go down to \$411, while it would increase to \$1,204 with full experience rating. The mean healthcare and health spending in our sample in 2014 is \$9,432. Applying a CARA model and measures of risk aversion from Handel (2013), we find that the reclassification and out-of-pocket expenditure risk generates a certainty equivalent income loss is \$9,772 under the base model, \$9,572 with community rating, and \$10,548 with full experience rating. Thus, the gains in welfare from community rating are only \$200, or equivalent to a 2.1% decrease in welfare.

The remainder of our paper is organized as follows. Section 2 describes our model of firm pricing and enrollee risk. Section 3 describes our data. Section 4 describes our empirical approach. Section 5 describes our estimation and counterfactual results. Finally, Section 6 concludes.

2 Model

We develop a simple and stylized model of reclassification risk and pricing in the health insurance industry. The model has two time periods, periods 1 and 2. Period 2 payoffs are discounted at the rate β. Utility is additively separable across the time periods. We consider potential enrollees who work for a small-group employer and obtain health insurance through their employer. Denote the potential enrollee by i, her employer by j, and the time period by t. Let I_j denote the number of enrollees at employer j.

Each potential enrollee starts each period with an expected risk score r_{ijt}, which is based on her previous year's healthcare use. The risk score is proportional to her total expected costs of healthcare at time t, is normalized to one for the mean individual in the sample, and is observable to both the potential enrollee and the insurer. The employer is faced with a per-person premium amount, p_{jt}, which is based on the mean risk score of its employees, $R_{jt} \equiv \frac{1}{I_j}\sum_{i=1}^{I_j} r_{ijt}$, and its history with the insurer. Thus, we can write $p_{jt} = p(R_{jt}, j)$.

Each period, each potential enrollee is faced with a distribution of potential health shocks, which is a function of her current risk score. Let the random variable $H(r_{ijt})$ denote the period t health shock and let $c(H(r_{ijt}))$ denote the cost of treating an individual with health shock $H(r_{ijt})$. We can separate costs into the portion that is paid by the insurer, $c^{ins}(H(r_{ijt}))$ and the portion that the enrollee pays out of pocket, $c^{oop}(H_{ijt}(r_{ijt}))$.

Importantly, our model allows for health shocks to be serially correlated over time. A costly health shock in period 1 will likely increase the period 2 risk score which will correlate with costly health shocks in period 2. Since the potential enrollee's time 2 expected health risk is a function of her time 1 realized health shock, we can write $r_{ij2} = f(H_{ij1})$. We assume that the potential enrollee and insurer learn the realization of H_{ij1} during time 1 from the potential enrollee's health claims and determine p_{j2} in part using the mean realized values of r_{ij2} for employees of employer j. Since the expected costs are proportional to the risk score, we can write $E[c(H(r_{ijt}))] = \gamma_1 r_{ijt}$, where γ_1 is the constant of proportionality.

We now exposit the utility at each period prior to the realization of the period health shock. We start by defining the per-period utility from obtaining insurance at each time period. This is a function of the potential enrollee's income Y_{ijt}, her premium, her out-of-pocket health costs, and an idiosyncratic shock to the value of purchasing insurance:

$$U(r_{ijt}, p(R_{jt}, j)) = \int u\left[Y_{ijt} - p(R_{jt}, j) - c^{oop}(H(r_{ijt}))\right] dF_H(H(r_{ijt})) + \varepsilon_{ijt}, \tag{1}$$

where $dF_H(H(r_{ijt}))$ is the distribution of health shocks conditional on a risk score, $u(\cdot)$ is the utility conditional on a particular health shock realization, and ε_{ijt} is an idiosyncratic shock of the value of insurance, that is known to the potential enrollee at the time that she chooses whether or not to purchase insurance. We assume that $u(\cdot)$ follows a CARA functional form. We further assume that each potential enrollee pays the full cost of her health premium to her employer, in the form of higher actual premiums or lower wages.[7]

[7]We abstract from the differential tax treatment of wage earnings versus employer sponsored health insurance benefits.

We write the utility from not purchasing insurance as:

$$U^0(r_{ijt}) = \overline{c}(r_{ijt}) + \varepsilon_{ijt0}. \tag{2}$$

In equation (2), $\overline{c}(r_{ijt})$ is a constant term that provides the expected utility from obtaining health shocks and facing the full price for treating these shocks—or perhaps foregoing treatment for some shocks due to the lack of health insurance, while ε_{ijt0} is an idiosyncratic shock to the value of having no insurance, and is known to the potential enrollee at the time that she chooses whether or not to purchase insurance.

Using equations (1) and (2), consider period 2 first. At the beginning of this period, the ex ante utility from purchasing insurance is given by $U(r_{ij2}, p(R_{j2}, j))$. The potential enrollee will observe her risk status and premium and purchase insurance if $U(r_{ij2}, p(R_{j2}, j)) \geq U^0(r_{ij2})$. Thus, an enrollee will purchase insurance if the benefits from the risk protection provided by the insurance outweigh the premium that the enrollee has to pay for the insurance. The only potential risk that is faced by an enrollee at period 2 who purchases insurance is the relatively small risk of paying high out-of-pocket costs given a bad health shock.

Consider now period 1. At this point, even purchasers of insurance face an additional source of risk: the possibility of reclassification risk caused by a health shock for themselves or one of their co-workers. In particular, a bad and persistent health shock may yield a high realization of R_{jt} which may in turn raise premiums for the individual. Accounting for reclassification risk, we can write the value function for an individual at period 1 as:

$$\begin{aligned} V(r_{1j1}, \ldots, r_{I_j j1}, i, j) &= \max\{U(r_{ij1}, p(R_{j1}, j)), U^0(r_{ij1})\} \\ +\beta \int \max\{U(r_{ij2}, p(R_{j2}, j)), U^0(r_{ij2})\} & dF_r(r_{ij2}|r_{ij1}) dF_R(R_{j2}|r_{1j1}, \ldots r_{I_j j1}), \end{aligned} \tag{3}$$

where $dF_r(r_{ij2}|r_{ij1})$ is the conditional distribution of the risk score for the individual at period 2 and $dF_R(R_{j2}|r_{1j1}, \ldots r_{I_j j1})$ is the conditional distribution of the mean risk score at the employer at period 2.

From the point of view of an enrollee, the extent to which health shocks lead to reclassification risk depends on the distributions of F_R and $p(\cdot)$. If the enrollee is in a large risk pool, then reclassification risk is not a substantial issue because the distribution of F_R is very concentrated. Thus, individuals employed by large firms or in settings with community rating do not face much reclassification risk. In contrast, individuals in a small risk pool without community rating—i.e., individuals in our sample—will be faced with the potential for reclassification risk.

Having discussed the enrollee side, we now turn to the insurer side. We assume that insurers are risk neutral and maximize expected profits. To simplify our analysis, we further assume that out-of-pocket costs are zero, implying that insured costs are the same as total costs.[8] With this assumption, expected insurer costs for an individual are only a function of the mean risk score: $E[c^{ins}(H(r_{ijt}))] = E[c(H(r_{ijt}))] = \gamma_1 r_{ijt}$. This assumption also allows us to simplify notation by omitting r as an argument of U.

We start by examining a perfectly competitive insurance industry and with the assumption that insurers cannot sign binding two-period contracts. In this case, there is no linkage between the two periods of the model. For the competitive industry, for simplicity we consider the case where U^0 is inframarginal and every potential enrollee purchases insurance. Since firms observe risk scores and the competitive market will set premiums equal to expected marginal costs, in this case we have that $p(R, j) = \gamma_1 R$. For this case then, equation (3) specializes to:

$$V(r_{1j1}, \ldots, r_{I_j j1}, i, j) = U(\gamma_1 R_{j1}) + \beta \int U(\gamma_1 R_{j2}) dF_R(R_{j2} | r_{1j1}, \ldots r_{I_j j1}). \tag{4}$$

Even though consumers in this hypothetical industry pay premiums equal to their expected costs, they are still faced with reclassification risk: an increase in risk score for their pool would translate into an increase in expected insurance costs in period 2.

Next consider the case where the perfectly competitive insurance industry can offer binding long-run contracts with commitments on both sides, maintaining the assumption that

[8] While out-of-pocket costs are not the focus of our model, our empirical results do allow for out-of-pocket costs to be positive.

every potential enrollee purchases insurance. Consider a two-period contract with a period 1 premium of $p_1 = \gamma_1 R_{j1}$ and a period 2 premium of $p_2 = \gamma_1 E[R_{j2}|r_{1j1}, \ldots r_{I_j j1}]$. Note that this contract would have premium equal to expected marginal cost and would eliminate the reclassification risk. Because of this, with the assumption of CARA utility,

$$\int U(\gamma_1 R_{j2}) dF_R(R_{j2}|r_{1j1}, \ldots r_{I_j j1}) < U(\gamma_1 E[R_{j2}|r_{1j1}, \ldots r_{I_j j1}]),$$

implying that such a contract would improve consumer welfare over the state-contingent one-period contracts examined above. Under the further assumptions that income is identical across periods and that mean risk is the same across periods so that $E[R_{j2}|r_{1j1}, \ldots r_{I_j j1}] = R_{j1}$, this contract would be the utility-maximizing contract among break-even contracts; otherwise some other contract that eliminated risk but provided savings would dominate. In either case, the perfectly competitive insurance industry would result in every employer signing a binding two-period contract that eliminated reclassification risk at period 2 by basing its period 2 premium on its period 1 risk.

In the real world, it is difficult to enforce long-run contracts with commitment on both sides. Without such enforcement, there are a variety of possibilities for what might occur with a perfectly competitive insurance industry. For instance, Handel et al. (2016) consider the case of a multi-period competitive insurance industry where only the insurer can commit to long-run contracts. In this case, it is not generally optimal for the insurer to completely eliminate reclassification risk because this would lower consumption in the first period too much. However, Handel et al. show that the competitive insurance industry provides partial protection against reclassification risk where, in equilibrium, individuals with high risk shocks end up paying more for their insurance, but not as much as in the complete absence of long-run contracts. Overall, our simple model shows that the ability to sign long-run contracts may improve welfare even in a competitive setting.

We next consider insurance provided by a single insurer with pricing power. In this case, to understand the insurer's optimal pricing behavior, we no longer assume that every potential enrollee purchases insurance. Instead, we assume that enrollee substitution to the outside option in response to a premium increase leads to a twice-differential demand curve for insurance. We further assume that the per-unit demand for insurance is the same across firms conditional on mean risk score and premium. Let $Q(p, R)$ denote the per-unit expected demand for the insurer for risk score R when the mean premiums are p.

We again first consider the insurer with pricing power which cannot offer binding two-period contracts. Similarly to the competitive case, the firm decision problem here is a repeated static pricing decision. Here, the insurer can set a separate premium for each pool based on its mean risk score and does not need to consider incentive compatibility constraints across mean risk scores since it can observe and contract on the mean risk score. Let $EC^{ins}(p, R)$ denote the expected cost to the insurer which sets a premium of p when faced with a pool with a mean risk score of R. This expression encapsulates the fact that higher premiums might change the mean risk scores of the potential enrollees who choose insurance but assumes that R is a sufficient statistic for determining this selected mean risk score.

Then, we can write the expected profits for the insurer from premium p for a group with risk score R as:

$$\pi(p, R) = [p - EC^{ins}(p, R)]Q(p, R). \tag{5}$$

Assume that ${\partial EC^{ins}(p,R)}/{\partial p} = 0$, which rules out adverse or advantageous selection. Then, we can write the first-order necessary condition for profit maximization as:

$$\frac{\partial \pi}{\partial p} = \left[p - EC^{ins}(p, R)\right] Q_p + Q(p, R) = 0, \tag{6}$$

where the subscript indicates a derivative; e.g., $Q_p \equiv \frac{\partial Q}{\partial p}$. In order to understand the impact of risk score on the insurer premium, we can implicitly differentiate (6) with respect to the risk score. Doing so and rearranging terms, we obtain:

$$\frac{dp}{dR} = \frac{\frac{\partial EC^{ins}}{\partial R} Q_p - Q_R - (p - EC^{ins}(R))Q_{pR}}{2Q_p + Q_{pp}(p - EC^{ins}(R))}, \tag{7}$$

where the double subscripts indicate second derivatives.

From equation (7), the insurer knows that an increase in R causes both an increase in cost, as given by ${}^{\partial EC^{ins}}/_{\partial R}$, and a potential change in demand, as given by Q_R. Together these imply that an insurer with pricing power will not necessarily pass through the extra expected costs in the form of extra premiums but may instead pass through more or less than the full amount.

Since the pass through expression is relatively involved, it is worth considering a very simple example, with a linear demand curve, so $Q_{pp} = 0$, and $Q_R = 0$, so that a change in expected risk changes only costs and not demand. In this case, $\frac{dp}{dR} = \frac{1}{2}\partial EC^{ins}/\partial R$.

This simple result is analogous to the well-understood result that such a firm would pass through one-half of an expected cost increase to consumers. It implies that an oligopolistic insurer who is faced with a linear demand curve and $Q_R = 0$ will pass through only one-half of the increased costs from an increase in expected risk in the form of higher premiums. In this case, since the pass through from expected costs to premiums is less than one, this implies that here, insurer pricing power lowers reclassification risk relative to the perfectly competitive case.

Now consider one-sided commitment in the context of firms with pricing power. As in the perfectly competitive case, firms with pricing power that can provide one-sided commitment may choose to lower enrollee reclassification risk since this adds value. Consider also enrollee inertia for health plans, which is generally believed to exist in the context of individual health plan choice (Handel, 2013). While other researchers have focused on the fact that inertia may increase markups, inertia may also help with avoiding reclassification risk because it provides implicit commitment to not switch plans on the part of enrollees. While we typically do not observe formal long-run contracts in the health insurance market, it is possible that large insurers use their reputation to provide implicit commitment to limit or eliminate experience rating with employers, in exchange for margins and an implicit commitment of inertia provided by the employers.

Overall, our takeaway is that insurers with pricing power may provide a certain amount of reclassification risk protection just as might a competitive market of insurers. In the case of insurers with pricing power, this risk protection might come at the cost of markups over cost. Understanding the nature of these tradeoffs is thus an empirical question.

3 Data

Our data are from employers who purchase health insurance for employee coverage from "United States Insurance Company" (USIC) in the small group market during the years 2012 to 2014. We obtain data from 10 different states: AR, DE, IL, PA, OK, MO, TN, TX, WI, and WY. They are further classified by USIC into 19 different markets, e.g., Texas is divided into Central Texas, Dallas, Houston, North Texas, and South Texas. Our study is based on proprietary data provided to us by USIC. The states that we use are all lightly regulated states prior to the ACA, for instance, without community rating regulations. Employers in this market are purchasing fully-insured insurance products from USIC, not third-party administrative services.

Our data include information at both the enrollee-year (employee or dependent) and firm-year levels. At the firm-year level, for all the employers that contract with USIC, we observe the number of health insurance plans available to their employees in each year, the characteristics of each plan, and the total premium paid by the employer to the insurer for each plan. At the enrollee-year level, we observe age, gender, the health plan chosen, 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 price 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 premium by twelve to annualize it.

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. 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.

From the data provided to us from USIC, some employers are missing information about premiums, plan characteristics, or enrollment. We keep employers without missing values in these fields. In addition, because one of our central variables, the ACG score, is calculated using the previous year claims data for an individual, we need to observe an individual for two consecutive years to have a complete observation on the individual. Further, much of our estimation is based on within-employer variation, controlling for employer fixed effects. As such, we limit our estimation sample to employers for whom we observe at least one individual in both 2012 and 2013, and at least one individual in both 2013 and 2014. Our estimation sample of enrollees then consists of enrollees covered by these employers and with coverage in either 2013 or 2013, or both. Overall, we start with xxxx employer-year observations and xxxx employee-year observations, of which our estimation sample keeps xxxx and xxxx respectively.

Table 1 provides summary statistics on the enrollees in our estimation sample, and explains our calculation of the different firm-level variables. We partition enrollees in the estimation sample into one of five groups, based on the years in which the enrollee is in our sample. The first two groups are enrollees who are not in our sample for two consecutive years. We cannot calculate ACG scores for these enrollees, and hence they do not enter into the employer mean risk score calculation. Nonetheless, they enter into the employer per-enrollee premium calculation because this calculation is based on the total premiums and the total enrollees in any year.

The third group is what we call "joiners"—individuals who start coverage in 2013 and keep it through 2014. These individuals' risk scores enter into the 2014 employer mean risk

Table 1: Enrollees by years in sample

	2014 but not 2013	2013 only	2013 & 2014	2012 & 2013	2012–14
Number of 2013-14 observations	1	1	2	1	2
Number of missing risk scores	1	1	1	0	0
Number of complete observations	0	0	1	1	2
Percentage of observations	9%	3%	11%	7%	70%
Individuals	33,175	11,126	20,024	27,821	129,791
Percentage of individuals	15%	5%	9%	13%	58%
In 2013 employer mean risk score?	No	No	No	Yes	Yes
In 2014 employer mean risk score?	No	No	Yes	No	Yes
In 2013 employer lagged claims?	No	No	No	Yes	Yes
In 2014 employer lagged claims?	No	No	Yes	No	Yes
In 2013 premium calculation?	No	Yes	Yes	Yes	Yes
In 2014 premium calculation?	Yes	No	Yes	No	Yes
Descriptive name for group	No ACG score available	No ACG score available	Joiners	Quitters	Stayers

Note: statistics are calculated based on individuals in estimation sample, as defined in text.

score but not the 2013 employer mean risk score. Similarly, “quitters” factor into the 2013 but not the 2014 employer mean risk score. “Stayers” enter into all data elements. The bulk of our observations, 70%, consistent of stayers.

Table 2 provides summary statistics on our estimation sample at the enrollee-year level. The first column provides information on the full sample. Overall, our sample consists of about 370,000 enrollee-year observations, each corresponding to one of the five groups of individuals in Table 1. The majority of the individuals in the sample are covered employees (57%), while the other main categories are spouses (15%) and children (27%). The mean age for these individuals is 38 years old and 47% of them are women. Sample mean total paid claims are $2,834, with medical claims accounting for 95% and prescription drugs expenditures accounting for the remaining 5%. We also report the out-of-pocket claims and the allowed claims. The latter figure indicates the total claims amount that the provider should expect to receive, and should be roughly equal to the sum of paid and out-of-pocket claims. Out-of-pocket claims have a mean of $891 and allowed claims have a mean of $3,725, which empirically verifies this proposition. Finally, the sample mean ACG score is 0.96 with

Table 2: Descriptive statistics on estimation sample at the enrollee-year level

	Full sample	Joiners	Quitters	Stayers
Relation (%):				
Employees	57	56	54	56
Spouses	15	15	16	16
Children	27	28	29	27
Others	1	1	1	1
Age	38 [18]	33 [18]	38 [18]	40 [18]
Female	47	48	49	47
Lagged paid total claims ($)	3,320	2,793	2,704	3,608
	[16,116]	[11,872]	[17,176]	[16,941]
Lagged paid medical claims ($)	2,697	2,377	2,301	2,880
	[15,114]	[11,264]	[16,657]	[15,776]
Lagged paid pharmaceutical claims ($)	623	416	403	729
	[4,070]	[2502]	[2,310]	[4,575]
Lagged out-of-pocket claims ($)	906	834	654	982
	[1,809]	[1,632]	[1,980]	[1,859]
Lagged allowed claims ($)	4,226	3,627	3,358	4,590
	[16,906]	[12,706]	[17,885]	[17,743]
ACG score, r_{jt}	1.03	0.86	1.08	1.04
	[1.48]	[1.26]	[1.63]	[1.47]
Observations	371,752	40,048	27,821	259,582

Note: each observation is one enrollee during one year, either 2013 or 2014 for individuals in estimation sample, as defined in text. Standard deviations of variables in parentheses.

a standard deviation of 1.37, which implies that enrollees in our sample are slightly healthier on average than in the ACG reference group.

People enter and leave employment and employer-sponsored health insurance for many reasons, including potentially selection based on their risk scores. To analyze selection further, the last three columns of Table 2 present data on the subsamples of joiners, quitters, and stayers. It is useful to compare these three groups to understand the differences across them. In general, the three samples are very similar in their mix between employees and dependents and in gender. In terms of their health expenditures, the stayers are similar to the full sample, but with slightly higher average claims, while joiners have lower risk scores and are younger. The breakdown of the paid claims among medical services and pharmacy claims is also similar across the samples. Joiners will look different from the other samples in the ways that we observe—younger and lower risk—in part because babies are "joiners." Our takeaway is that there is little evidence that quitters are different than stayers in observable ways.

Table 3: Descriptive statistics at the employer-year level

Risk pool characteristics	Mean	Std. dev.
Subscribers	20.03	25.97
% Employees	64.98	22.08
Mean age	40.55	8.83
% Female	46.45	20.49
Average premium and risk score		
Employer mean annual premium	11,636	15,810
Employer mean ACG score, R_{jt}	1.15	0.81
Δ employer mean ACG score, $R_{j,2014} - R_{j,2013}$	0.03	0.62
Lagged presence of chronic conditions at employer level (%)		
Cancer (% of employees)	6.79	11.33
Transplant (% of employees)	0.20	2.15
Acute myocardial infarction (% of employees)	0.18	1.91
Diabetes (% of employees)	6.32	11.39
Number of unique employers	9,281	
Number of employer-year observations	18,562	
Number of employer-plan-year observations	21,079	

Note: statistics calculated based on employers in our estimation sample, as defined in text.

Table 3 provides summary statistics at the employer-year level. We observe 18,564 employer-year observations and 21,081 employer-plan-year observations. This provides substantial variation in the employer mean risk score that allows us to identify the pass through from employer mean risk scores to premiums. This richness of variation is not found in most other studies.

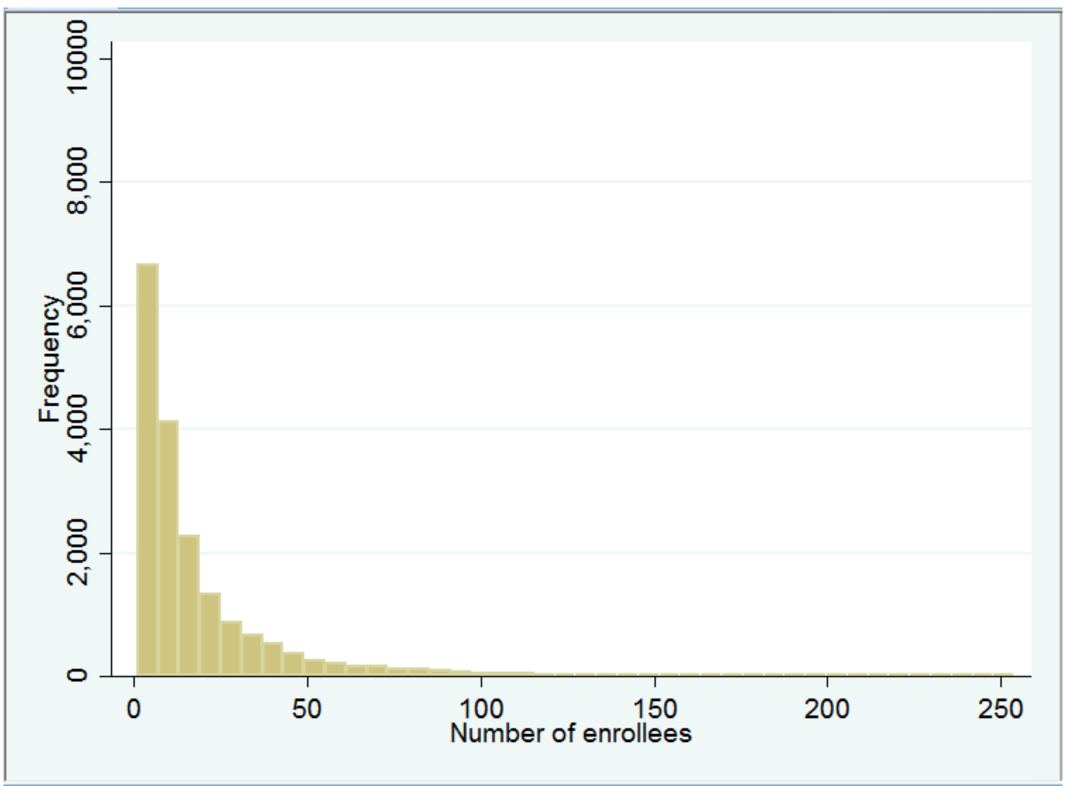
Table 3 further shows that the mean number of subscribers per employer is approximately 20 with a mean ratio of 65% employees out of total subscribers. The average age of subscribers at these employers is 41 years and the average percentage of females at these employers is 46%. The mean annual premium per subscriber at these employers is $7,336, with a large standard deviation of $8,408.

In addition, Table 3 shows that the mean ACG score across employers is 1.06, slightly higher than the mean ACG score at the individual level, implying that smaller employers tend to have higher risk scores. The standard deviation of the employer mean risk score is 0.74, which is approximately one half of the standard deviation of the ACG score in the sample of stayers. Thus, risk pooling at the small group level reduces risk substantially relative to risk pooling at the individual level, but still leaves a large amount of risk. The change in employer mean risk score is very close to 0 (-0.09) but the standard deviation is quite large, 0.56, implying that reclassification risk within an employer is a large part of the overall risk from pooling at a small employer.

Table 3 also presents the mean incidence of four chronic conditions at an employer—cancer, transplants, acute myocardial infarctions (heart attacks), and diabetes—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 10.1% and diabetes is 6%.

Finally, Figure 1 provides more evidence on the distribution of enrollees by employer. The figure shows that the vast majority of employers have 10 enrollees or fewer.

Figure 1: Histogram of number of enrollees by employer



Note: histogram excludes two observations which are from an employer with more than 300 enrollees in 2014 and 2015.

4 Empirical Approach

4.1 Estimation Approach

The goal of our estimation section is to recover the pass through for USIC from expected costs to premiums. We can express this as:

$$\frac{\partial p}{\partial EC^{ins}} = \frac{\partial p/\partial R}{\partial EC^{ins}/\partial R}. \tag{8}$$

In words, equation (8) states that the pass through from expected insurer costs to premiums can be expressed as the pass through from risk score to premiums divided by the pass through from risk score to expected insurer costs. We first discuss our estimation of $\partial p/\partial R$ and then turn to our estimation of $\partial EC^{ins}/\partial R$.

To estimate the pass through from risk at the employer level to premiums, we estimate a two-year panel data regression based on small group claims data for 2012-13 and premium data, for 2013-14. Specifically, we estimate regressions of the form:

$$p_{jt} = \beta_1 R_{jt} + \beta_2 x_{jt} + FE_j + FE_t + \varepsilon^A_{jt}. \tag{9}$$

In equation (9), p_{jt} is the premium charged to employer j at time t and R_{jt} is the ACG risk score at time t. Note that R_{jt} is calculated using claims data from year $t-1$. FE_j are employer level fixed effects, FE_t are annual fixed effects, while ε^A_{jt} is the residual in this regression.

Our key variable of interest is β_1, the pass through from ACG risk score to premiums. We would like to understand the impact of changes in risk scores on changes in premiums, controlling for other firm level attributes. Thus, our base specification includes fixed effects at the level of the firm. We also include year fixed effects to account for the fact that health plan premiums have been increasing over time. Given the inclusion of fixed effects, we need a minimum of two years of data to estimate equation (9), which implies using three years of claims data, since risk scores are calculated using lagged claims.

The residual here will capture changes in premiums unexplained by other factors, for instance due to variation in firm or insurance broker bargaining ability. Our regressions based on equation (9) cluster standard errors at the employer level.

Although our model specifies that insurers should base their changes in premiums solely on changes in expected risk scores, we would also like to test whether other health factors result in changes in the premiums that are charged to an employer. One possibility is that the insurer directly considers chronic diseases in its pricing decision in addition to the risk score. Hence, in some specifications, we allow the mean percentage of enrollees with chronic diseases, in addition to the ACG risk score, to affect future premiums. In other specifications, we allow for the current year claims to directly affect future claims, rather than being mediated through the ACG risk score.

One empirical limitation is that, since our data are from USIC, we only know the risk scores and claims for people who were enrolled in the previous year while our model considers the mean risk score of the overall pool. These two terms are not identical since we do not use the risk scores for new enrollees or for potential enrollees at an employer who decline coverage. Moreover, adverse selection in health insurance markets may be very important (Einav et al., 2010; Rothschild et al., 1976).

It is possible that USIC obtains risk data on new enrollees at an employer in updating their premiums for the employer. It is possible, though less likely, that USIC updates premiums for an employer based on employees who terminate coverage. Our base specifications of equation (9) determine the risk score based on all enrollees with a risk score in a given year; thus we use stayers and quitters for the 2013 risk score and stayers and joiners for the 2014 risk score, as defined in Section 3. While we cannot verify that the people who stopped obtaining insurance became different after they stopped their insurance, the fact that their observable characteristics are similar to the people who remain in the insurance pool suggests that selection into or out of insurance based on changes in risk score is not very important here. We further check this point by estimating robustness specifications that calculates R_{jt} using the risk score only for stayers.

We now turn to estimation of the pass through from risk score to insurer expected costs. Here, we estimate regressions of the form:

$$claims_{ijt} = \gamma_1 r_{ijt} + \gamma_2 x_{jt} + \varepsilon^B_{ijt}, \tag{10}$$

where $claims_{ijt}$ is the mean dollar value of claims for an individual in an year, or alternately put, the insurer's realized costs for that individual. Our specification in equation (10) considers the impact of the current risk score—estimated using the previous year's claims—on current claims to the insurer. The residual in this regression, ε^B_{ijt}, will capture the difference between actual claims and expected claims for an individual in a year. Our regressions based on equation (10) also cluster standard errors at the employer level.

Unlike our estimation of the impact of risk score on premiums from equation (9), our specifications here conceptually can be estimated from a cross-section and do not need to include employer fixed effects. The reason for this is that the risk score is meant to be a causal and proportional predictor of health expenditures. Thus, we should expect a linear relationship between r_{ijt}, which is calculated using time $t-1$ data, and claims at time t. This relationship is exactly what we would like to recover, to understand ${}^{\partial EC^{ins}}\!/_{\partial R}$.

Our specifications based on (10) do include market level fixed effects in x_{jt}. The reason that we include these variables is because provider prices may vary across markets. By including fixed effects here, we are able to control in part for market-level provider price variation.

4.2 Calculation of welfare and counterfactuals

Our estimation results recover the pass through from changes in expected claims to changes in premiums. We then seek to uncover the extent to which this reclassification risk translates into a consumer welfare loss. We compare the estimated level of reclassification risk to other environments, notably full community rating and full risk rating. We also examine how the relative welfare levels of alternate risk sharing policies compare to the costs to consumers of insurance in the small group market.

Our counterfactual approach has three parts. First, we construct the future distribution of enrollee health risk and mean employer health risk to which an enrollee is exposed. This distribution depends on the claims exposure for the enrollee and other covered lives at the same employer. Second, we evaluate how the distribution of future risk translates into a distribution of future premiums and out-of-pocket costs. Third, we examine how this distribution of premiums and out-of-pocket costs translates into a utility level and an equivalent income that, with certainty, that would generate the same utility. We now discuss these three parts of our analysis in turn.

First, to construct the distribution of enrollee health risk and mean employer health risk for an enrollee, our main assumptions are that (1) the ACG score is a sufficient statistic for predicting the distribution of future ACG scores for an enrollee; and (2) every enrollee receives an *i.i.d.* draw from this distribution. Using these assumptions, we non-parametrically estimate the empirical distribution of period 2 ACG scores for enrollees (at any employer) with similar ACG scores to the enrollee in period 1. In order to construct the distribution of employer mean risk faced by an individual, we then simulate vectors of period 2 ACG score draws based on the observed period 1 ACG scores for all employees at the employer. Each simulation draw vector provides one possibility for the claims exposure at period 2. For each simulation draw vector, we calculate the enrollee ACG score and the employer mean ACG score in period 2. We take multiple simulation draw vectors to approximate the empirical distribution of employer mean ACG scores.

Second, to evaluate how changes in the period 2 mean employer ACG score translate into changes in premiums, we use our estimate of the pass through from employer health risk to premium, based on β_1. To evaluate how changes in the period 2 employee ACG score translate into changes in out-of-pocket costs, we use our estimates of the relation between ACG score and out-of-pocket costs. In both cases, we use the simulation draw vectors that provide distributions of employee and mean employer ACG risk scores for each enrollee. We then sum the financial risk imposed by the distribution of higher premiums with the financial risk imposed by the out-of-pocket costs to derive the period 2 distribution of healthcare/insurance expenditures. For this step, we also consider counterfactual exposures to reclassification risk.

Specifically, we examine complete pass through of reclassification risk to premiums, by setting $\beta_1 = \gamma_1$. In addition, we examine community rating, as will occur under the ACA, by setting $\beta_1 = 0$.

Third, we consider the utility loss and certainty equivalent utility from the healthcare and health insurance expenditure risk borne by individuals. Following Handel (2013), we assume that utility follows CARA preferences so that mean utility is:

$$u(Y_{ijt} - p(R_{jt}, j) - c^{oop}(H(r_{ijt}))) = -\frac{1}{\gamma} \exp\left(-\gamma \left[Y_{ijt} - p(R_{jt}, j) - c^{oop}(H(r_{ijt}))\right]\right). \quad (11)$$

We do not estimate γ, but instead take estimates from the literature that evaluates risk in similar contexts. Specifically, we use a value of $\gamma = 0.000428$ from Handel (2013). Step 2 provides us with simulation draws from the distributions of $p(R_{jt}, j)$ and r_{jt} (which we use to calculate out-of-pocket costs) for every individual under the observed and counterfactual risk rating environments. We use these simulated distributions to calculate the distribution of healthcare expenditures of each environment. We then calculate the standard deviation of income for each environment using these simulated distributions.

Finally, using (11), we calculate the certainty equivalent income for each environment, which is the income level that, when earned with certainty, would give a utility level equal to the utility level observed in the experiment. Because we assume that consumers are risk averse, the certainty equivalent income levels will be less than the actual income levels. One of the advantages of the CARA utility function is that the certainty equivalent utility of a position does not depend on the base income level. Hence, we do not need to approximate the base income levels of individuals in our sample to obtain the certainty equivalent utility levels.

5 Results

5.1 Estimation results

We first investigate results for the pass through from expected risk to ACG score. These regressions are at the employer/year level. They are based on equation (9). The regressions are based on individuals observed in 2013 and 2014, although the ACG score is calculated based on the individual's claims from the previous year.

Table 4: Pass through from expected risk to premiums

	Dependent variable: Annual employer mean premium, p_{jt} (\$)			
Regressor:	I	II	III	IV
Employer mean ACG score, R_{jt}	1,649***	3,212***		
	(544)	(118)		
Employer mean lagged total claims			0.038	0.128***
			(0.030)	(0.019)
Firm FE	Yes	No	Yes	No
Market FE	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	18,562	18,562	18,562	18,562

Note: each observation is one employer during one year. The main dependent variable is the premium charged the employer by USIC divided by the number of covered lives. R_{jt} is calculated based on covered individuals with an ACG score in both 2013 and 2014. 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 and ** indicates significance at the 5% level.

The base results are given in Table 4, column I. This specification regresses the employer mean ACG score on the mean premium for the employer. It includes employer and year fixed effects. The base results show that the mean ACG score in a year, which is calculated based on the previous year's claims, is significantly and positively related to the claims for an employer. An increase in ACG score of 1 predicts an increase in annual premium of \$1,649. The result is statistically significant at the 1% level. This compares to the employer mean ACG score of 1.06 and a standard deviation of 0.74. Thus, a one standard deviation increase in expected risk for an employer would raise annual per-person premiums by \$1,220 annually.

Column II of Table 4 displays results from a specification that is the same as column I except that it excludes employer fixed effects. This column shows an effect of ACG scores on premiums that is larger, \$3,212 instead of \$1,649. Thus, it appears that employers with high mean risk scores tend to pay higher premiums than do employers faced with an increase in their risk score during our sample. There are at least three possible reasons for this difference. First, employers with high mean risk scores may also have other attributes that lead to higher premiums; e.g. lower ability to bargain. Second, these employers may choose plans with more benefits, which will then be more expensive. Finally, risk may be transmitted to premiums over a longer horizon than one year.

Table 4 also shows the impact of lagged claims on premiums. Since the risk score is based on lagged claims, one may think that employers with higher claims in one year will generally face a significant increase in the premium the next year. In fact, this is not the case. Column III shows that with employer fixed effects, the pass through from lagged claims to premiums is not statistically significant and has a very small point estimate. Column IV, which does not include employer fixed effects, shows a significantly positive relationship, but the effect is small, with a \$100 increase in claims only raising premiums by \$12.80 in the subsequent year. We believe that these columns show that USIC is relatively sophisticated in basing its experience rating on future health risk rather than just current claims exposure.

We next turn to analyzing alternative specifications based on equation (9), with results in Table 5. For convenience, column I repeats the base specification from Table 4. Column II shows results from a specification that is the same as the base one, but with the omission of year fixed effects. Without year fixed effects, the pass through from ACG score to premiums is larger than the base results, showing that controlling for year fixed effects is important in understanding this relationship, likely because premiums are generally rising over this time period. Column III starts with the base specification and adds the mean age of beneficiaries as a control. This variable is significantly positive but results in a similar coefficient on mean ACG score. Column IV estimates a log-log version of Column I. The results show that a significant relationship between ACG score and premiums. However, a linear version specification is more interpretable, since it fits more closely with the fact that mean ACG score should linearly predict mean insurer costs and hence premiums.

Table 6 presents further robustness specifications. Column II presents similar specifica-

Table 5: Pass through from expected risk to premiums, robustness checks

	Dependent variable:			
Regressor:	Annual employer mean premium	Annual employer mean premium		Annual employer log mean premium
	Base specification	Robustness specifications		
	I	II	III	IV
Employer mean ACG score	1,649***	2,645***	1,060**	
	(544)	(656)	(539)	
Mean age of beneficiaries			697***	
			(101)	
% Female			5,505	
			(6,157)	
Employer mean log ACG score				0.084***
				(0.029)
Year FE	Yes	No	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	18,562	18,562	18,562	18,562

Note: each observation is one employer 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 covered individuals with an ACG score in both 2013 and 2014. Standard errors are clustered at the employer level. *** indicates significance at the 1% level and ** indicates significance at the 5% level.

Table 6: Pass through from expected risk to premiums, robustness to ACG computation and split by employer size

	Dependent variable: Annual employer mean premium, p_{jt} (\$)			
	Base specification	Mean ACG for stayers	Smaller employers	Larger employers
Regressor:	I	II	III	IV
Employer mean ACG score, R_{jt}	1,649*** (544)		1,316** (563)	662 (507)
Employer mean ACG score for stayers		1,256** (572)		
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	18,562	18,504	9,686	8,876

Note: each observation is one employer 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 covered individuals with an ACG score in both 2013 and 2014. Smaller employers are those with 12 or fewer covered lives in both 2013 and 2014; larger employers are all others. 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.

tions to our base specification but including only the sample of stayers in the calculation of employer mean ACG score. We find similar results to the base specification. This suggests that, even if employers adjust their premiums based on people who leave coverage, the effect of this on the pass through from mean employer ACG scores to premiums is small. This result also corroborates the evidence from Table 2 that stayers and quitters are similar in observable characteristics. Columns III and IV present specifications similar to our base specification from Table 4, but splitting the sample based on smaller and larger firms, all within the small group market. The smaller employers within the small group market have a similar pass through coefficient to the base regression while the larger employers have a somewhat smaller pass through coefficient that is not significant. We believe that the larger employers here may not provide much identifying variation since the mean ACG risk score for their enrollees will not vary much due to their larger size.

Table 7: Pass through from expected risk to premiums, with chronic conditions

	Dependent Variable: Annual employer mean premium, p_{jt} (\$)				
Regressor:	I	II	III	IV	V
Employer mean ACG score, R_{jt}	1,649***	1,672***	1,527***	1,624***	1,820***
	(544)	(569)	(558)	(544)	(558)
Lag % cancer at employer		-512			
		(2,975)			
Lag % transplant at employer			24,540 *		
			(13,392)		
Lag % AMI at employer				10,050	
				(13,595)	
Lag % diabetes at employer					-5,982
					(5,523)
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Observations	18,562	18,562	18,562	18,562	18,562

Note: each observation is one employer 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 covered individuals with an ACG score in both 2013 and 2014. Chronic disease regressors indicate the mean percent of enrollees with a claim for the disease in the previous year. Standard errors are clustered at the employer level. *** indicates significance at the 1% level.

Table 7 also presents similar specifications to our base specification in Column I of Table 4 but with the addition of the percent of enrollees with particular chronic diseases. We chose cancer, transplants, AMIs (heart attacks), and diabetes, as these diseases result in persistent increases in the costs of healthcare, and they may serve as markers that insurers use to price risk. None of these diseases result in increases in premiums, when controlling for employer fixed effects, year fixed effects, and risk scores. Our takeaway from this is that the ACG risk score captures much of the variation in claims experience that USIC is using to adjust premiums for an employer from one year to the next.

Table 8: Effects of expected risk on benefits

	Dependent Variable		
	In-network maximum OOP ($)	Coinsurance rate (%)	In-network deductible ($)
Regressor:	I	II	III
Employer mean ACG score, R_{jt}	-15	-0.52**	-30**
	(19)	(0.22)	(10)
Year FE	Yes	Yes	Yes
Firm-Plan FE	Yes	Yes	Yes
Observations	21,079	21,079	21,079

Note: each observation is one employer/plan during one year. Each dependent variable is a measure of plan benefits. Mean risk score is calculated based on covered individuals with an ACG score in both 2013 and 2014. Standard errors are clustered at the employer level. ** indicates significance at the 5% level and * indicates significance at the 10% level.

Table 8 evaluates whether changes in expected risk lead to changes in the plan benefits that the employer chooses. Here, our unit of observation is one employer/plan during one year, rather than one employer during one year. Most employers in our sample choose one plan but some choose more than one plan, resulting in us having 21,081 observations here instead of 18,564 in the base sample. We consider three measures of plan benefits. The out-of-pocket dollar maximum for in-network services does not respond significantly to changes in the employer mean risk score. The coinsurance rate decreases by 0.52 percentage points, and the in-network deductible drops by $30; both drops are significant at the 5% level. Overall, it appears that the increases in premiums from higher employer mean ACG scores are somewhat mitigated by better plans benefits, although the impact is very small.

Table 9 presents the estimated relationship between expected risk and claims, based on equation (10). Together with our regressions above based on equation (9), this allows us to recover $\partial p/\partial EC^{ins}$, as in equation (8). From Column I, we find that an increase in ACG score

Table 9: Pass through from expected risk to claims

	Dependent variable:		
	Paid amount ($)	Allowed amount ($)	OOP amount ($)
Regressor:	I	II	III
Enrollee ACG score, r_{jt}	4,218***	4,736***	518***
	(181)	(182)	(11)
Market FE	Yes	Yes	Yes
Observations	157,612	157,612	157,612

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 2013 only. Standard errors are clustered at the employer level. *** indicates significance at the 1% level.

of one leads to an increase in USIC-paid claims for the enrollee by $4,218. From column II, an increase in ACG score increases the allowed amount of the claims by $4,736. 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. Given that the coefficient on allowed amount is larger than the coefficient on paid amount, it is not surprising that the coefficient on out-of-pocket amount—which is reported in column III—is positive, at $518. These three coefficients are all significant at the 1% level. From column I, a one-standard deviation increase in ACG score, or an increase in 0.74, would result in $3,121 more insurer-paid claims.

Table 10 provides robustness on the evidence presented in Table 9, column I, by considering a linear spline relationship between the risk score and claims. Columns I and II use splines with cut points of 1, 2.5, and 5, chosen as round numbers that differentiate enrollees with serious chronic diseases from others. The numbers here are generally show a roughly linear relationship between risk scores and claims, which is higher for enrollees with very high risk scores. Columns III and IV use splines defined by quartiles of our in-sample ACG score distribution. The results here show more non-linearity—with a third-quartile coefficient close to 0—which might be due to a small number of outliers with high medical costs near the top of the second quartile. Overall, our takeaway is that our base coefficient of $4,218 is a reasonable approximation of the pass through from risk score to expected claims.

Combining our base results from columns I of Tables 4 and 9, we find that a unit increase

Table 10: Pass through from expected risk to premiums using splines

	Dependent Variable: Paid amount ($)			
Regressor:	I	II	III	IV
Spline employee ACG score, $r_{jt} \in [0, 1)$	2,966***	3,053***		
	(102)	(104)		
Spline employee ACG score, $r_{jt} \in [1, 2.5)$	3,188***	3,206***		
	(186)	(185)		
Spline employee ACG score, $r_{jt} \in [2.5, 5)$	3,615***	3,611***		
	(506)	(506)		
Spline employee ACG score, $r_{jt} \in [5, \infty)$	5,804***	5,798***		
	(688)	(688)		
Spline employee ACG score, $r_{jt} \in [0, .32)$			2,547***	2,723***
			(360)	(363)
Spline employee ACG score, $r_{jt} \in [.32, .57)$			4,907***	4,964***
			(496)	(498)
Spline employee ACG score, $r_{jt} \in [.57, 1.13)$			214	308
			(630)	(631)
Spline employee ACG score, $r_{jt} \in [1.13, \infty)$			4,690***	4,688***
			(272)	(272)
Market FE	No	Yes	No	Yes
Splines	Fixed cut points	Fixed cut points	Quartiles	Quartiles
Observations	157,612	157,612	157,612	157,612

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 2013 only. Standard errors are clustered at the employer level. *** indicates significance at the 1% level.

in mean risk score at an employer increases premiums by $1,649 and expected costs by $4,218. Together, these coefficients imply that the pass through from expected costs to premiums, $\partial p / \partial EC^{ins}$, is 39%. This number, while significantly positive, is much less than the predictions of our perfectly competitive model without long-run contracts.

Our model provides several potential explanations for this finding. While one hypothetical possibility for the limited pass through is a perfectly competitive market with long-run contracts with one-sided commitment (Handel et al., 2016), we do not observe such contracts. Alternately, it is useful to consider our result within the context of an insurer with pricing power. Within this context, the result is likely driven by choice inertia, which adds commitment on the part of the employers purchasing the insurance, and by the fact that USIC is providing an implicit commitment to not completely experience rate in the small group market.

5.2 Counterfactual outcomes and welfare

We now consider the uncertainty in future health care expenditures, reclassification risk, and welfare loss present in the current environment and in counterfactual policies on risk rating. Our base results here are presented in Table 11. We consider the current environment, with its 39% pass through from expected claims to premiums, as well as community rating, where the pass through is 0%, and full experience rating, where the pass through is 100%. In all counterfactuals, we set the average premiums to USIC to be the same. We also assume that consumers are fully informed of the portion of their future premiums that do not relate to risk, which includes everything in our premium determination regression (9) except $\beta_1 R_{jt}$.

We first examine the premiums paid to USIC. The mean premiums paid by individuals in our sample in 2014 are $8,510. By construction, the mean premiums are the same across the three policy environments that we consider. While the three environments that we consider have the same mean premiums, they do not have the same premiums for each individual. In particular, we find that, under the current environment, individuals face a mean standard deviation in their 2014 premiums of $387. With community rating, by definition, individuals

Table 11: 2014 outcomes and welfare under different environments

	Current environment	Community rating	Full experience rating
Premium paid	8,510	8,510	8,510
Premium paid within std. dev.	387	0	990
Out-of-pocket expenses	922	922	922
Out-of-pocket expenses within std. dev.	411	411	411
Health spending	9,432	9,432	9,432
Health spending within std. dev.	663	411	1,204
Certainty equivalent income loss	9,772	9,572	10,548

Note: results based on estimates and estimation sample. All numbers are measured in dollars per year and report the means of the variables noted. "Within" standard deviations are for the distribution faced by an individual given her 2013 health score.

face no standard deviation in their 2014 premiums. With full experience rating, the mean standard deviation in premiums is higher than in the current environment, at $990.

The relation between the standard deviation in premiums under the current environment and under full experience rating is driven by the fact that the pass through from expected costs to premiums is 39%. Moreover, the scale of the effects is a function of the estimated parameters, but also of the distribution of 2014 mean risk scores for the employer. The employer's mean risk score is in turn a function of the 2014 risk scores for all the employees at the employer.

By construction, out-of-pocket expenses are the same across the three environments. Thus, in each environment, the mean out-of-pocket spending is $922 and the standard deviation is $411. The health spending variable combines the premium paid and out-of-pocket expenses. Thus, the mean health spending variable, at $9,432, is the sum of the mean premium paid and out-of-pocket expenses.

The mean standard deviation of 2014 health spending is $663. Note that the distributions of 2014 premiums and out-of-pocket expenses are not independent. In general, we would expect that they are positively correlated. For instance, a negative health shock may lead an individual to have both a higher future premium and higher future out-of-pocket expenses. Indeed, we can see that since the variance of health spending, at $\$663^2$, is higher than the sum of the variances of the two components of spending, $\$387^2 + \$411^2$, there is a positive correlation between 2014 out-of-pocket expenses and premiums paid.

Applying the CARA functional form and calibrated risk aversion parameter, we find that, even though the health spending in 2014 costs a mean of $9,432, the uncertainty raises the utility cost, so that individuals would on average be willing to pay a fixed $9,772 in the current environment to avoid the reclassification risk that they face with uncertainty regarding premiums and out-of-pocket expenses. Thus, under the current environment, the reclassification risk in 2014 reduces welfare by an average of $340. Even with pure community rating, the out-of-pocket portion of costs generates some reclassification risk, with a reduction in welfare equal to $140 ($9,572-$9,432). However, the reclassification risk is most important for the full experience rating case, where it reduces welfare by an average amount equivalent to $1,116 ($10,548-$9,432).

Thus, overall it appears that, because USIC does not fully experience rate policies in this market, the reclassification risk to the small groups that purchase insurance in this market is only a small fraction of what it would be with full experience rating. It is also worth noting that the potential reduction in reclassification risk of $200 ($9,772 - $9,572) from community rating with the same plan prices and characteristics is small. Thus, were the ACA to increase margins in this market by more than $200, it would lower consumer welfare despite lowering reclassification risk.

Table 12: 2014 outcomes and welfare under different environments: smaller employers

	Current environment	Community rating	Full experience rating
Premium paid	8,510	8,510	8,510
Premium paid std. dev.	607	0	1,944
Out-of-pocket expenses	972	972	972
Out-of-pocket expenses std. dev.	462	462	462
Health spending	9,482	9,482	9,482
Health spending std. dev.	934	462	2,221
Certainty equivalent income loss	10,204	9,669	13,461

Note: results based on estimates and estimation sample. All numbers are measured in dollars per year and report the means of the variables noted. “Within” standard deviations are for the distribution faced by an individual given her 2013 health score. Smaller employers are those with 12 or fewer covered lives in both 2013 and 2014.

Table 12 considers similar results to Table 11, but considers only smaller employers within the small group market, those with twelve or fewer enrollees. This table uses the coefficient estimates and sample of employers from Table 6, column II. We would expect individuals at smaller employers to bear more reclassification risk since their risk pools are smaller. This will be mitigated somewhat by the fact that the estimated coefficient from expected risk to premiums is lower in this segment than for our sample overall.

We find that individuals in this subsegment do indeed bear much more risk than for the small group market overall. For instance, the standard deviation of premiums with full experience rating here is \$1,944, compared to \$990 for the base model, while the standard deviation of overall health spending with full experience rating is \$2,221 instead of \$1,204 in the base model. An even larger difference occurs for welfare: the welfare loss from reclassification risk here is \$3,792, compared to \$1,116 in the base model, with the larger effect coming from the non-linearity of welfare losses in the size of the gambles. However, even for this sample, the reclassification risk under the current environment reduces welfare by an average equivalent of \$722, a figure that is relatively small compared to the average cost of healthcare for this sample.

Finally, note that we have only considered reclassification risk and not the risk from changes in price that are orthogonal to health status changes, i.e., ε_{jt}^{A} from (9). If community rating were able to eliminate this portion of risk as well as reclassification risk, then it could increase welfare by much more than the amounts that we find here.

6 Conclusion

In this paper, we seek to understand the extent of reclassification risk in the small group health insurance market. We make use of a dataset from a large U.S. health insurer, with premium information on over 9,000 employers and claims data from all the enrollees at these employers.

We seek to understand the extent to which mean health risk at an employer is passed through to the employer in the form of higher premiums. We find that the pass through from mean health risk to premiums is 39%. This compares with 100% pass through for a perfectly competitive market. There is little evidence that factors other than health risk affect changes in premiums for an employer. Using our pass through measures, we examine the value of community rating regulations as will occur in this market due to the ACA. We find that community rating would have reduced the mean standard deviation of 2014 health spending by an average of \$252, resulting in a welfare gain equal to an average of \$200. Full experience rating would result in much more reclassification risk, with a welfare loss equivalent to \$772 relative to the current environment. The welfare effects from reclassification risk are much larger for the subsegment of employers in the small group market with twelve or fewer employees.

References

Bundorf, M. K., Levin, J. D., and Mahoney, N. (2011). Pricing and welfare in health plan choice. *American Economic Review*.

Carlin, C. and Town, R. (2009). Adverse selection, welfare and optimal pricing of employer-sponsored health plans. *University of Minnesota Working Paper*.

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

Einav, L., Finkelstein, A., and Levin, J. (2010). Beyond testing: Empirical models of insurance markets. *Annual Review of Economics*, 2:311.

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

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. (2016). The welfare impact of long-term 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.

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.

Madrian, B. C. (1994). Employment-based health insurance and job mobility: Is there evidence ofjob-lock? *Quarterly Journal of Economics*, 109(1):27–54.

Mahoney, N. and Weyl, E. G. (2014). Imperfect competition in selection markets. Technical report. NBER Working Paper 20411.

Rothschild, M., Stiglitz, J. E., et al. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, 90(4):630–49.