Reclassification Risk in the Small Group Health Insurance Market*

Sebastian Fleitas† Gautam Gowrisankaran‡ Anthony Lo Sasso§

November 7, 2016

Abstract

Health insurance with annual contracts does not provide complete risk protection since enrollees with persistent adverse health shocks will be faced with higher future premiums. We consider the small group insurance market in 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 expected mean health risk in the form of higher premiums, compared to 100% pass through with perfect competition. Assuming CARA preferences with published risk aversion estimates, this risk protection adds an equivalent mean of $776 annually in consumer welfare. Community rating—as will occur over time under the ACA—would increase annual welfare $200 more.

`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 markets for health insurance function? Since most recent healthcare reforms in the U.S. have emphasized decentralized solutions, this has become a central question for health policy research. For a decentralized health insurance market to be efficient, it must offer insurance at premiums close to average costs and also provide risk protection. An important dimension of risk protection is protection from the risk that an adverse and persistent health shock will lead to higher future premiums or worse coverage, called *reclassification risk*.

Even perfectly competitive markets do not necessarily provide reclassification risk protection. To illustrate, consider a market for health insurance that is perfectly competitive but with annual contracts, as is typical for individual- and employer-based health insurance. In the absence of pricing regulations, 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. Thus, the market will *experience rate*, i.e., it will pass on an adverse health shock at a pool in one year in the form of premium increases in the subsequent year that match the increase in expected future costs of claims. This will lead to reclassification risk and ultimately market failure in the form of inadequate risk protection. 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).

The goal of our paper is to examine reclassification risk in the small group market, which represents employer 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. Using a unique data set provided by a large health insurance company, our paper estimates the extent to which higher expected claims for a small group

[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).

are passed through in the form of higher premiums. We then quantify the extent to which such experience rating creates reclassification risk and affects welfare in the small group market. The ACA is phasing in *community rating* for the small group market, whereby insurers will be prohibited from risk rating premiums based on the health status of an individual group in an area, with premium variation allowed only for age and smoking status.[2] 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 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 in turn cost each employee an extra $5,000 per year. 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. Focusing on how pricing power affects 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.

The ability of an insurer with pricing power to mitigate reclassification risk may be even

[2]Different states have introduced the reforms at different paces. In particular, CMS has allowed states to have their own premium method (CMS, 2016).

higher when there is inertia in health plan choice.[3] With inertia, enrollees are effectively committing to restricting their switching of insurers. The commitment implied by inertia may help forward-looking insurers with pricing power implicitly commit to charge relatively high markups but to not raise rates too much based on health risk, i.e., to provide reclassification risk protection. The presence of such a "hedge" or buffer against reclassification risk potentially adds value to enrollees that the insurer can partly capture. Interestingly, consumer welfare and industry profits may both be higher in an oligopoly market that charges markups above costs but provides reclassification risk protection than they would be in a perfectly competitive market without long-run contracts.[4]

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). For instance, 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) evaluate the equilibrium adverse selection and reclassification risk from a competitive market of exchange firms, while Handel et al. (2016) 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. We 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 understand how much reclassification risk protection the current market provides relative to the benchmark of perfect competition with annual contracts. We also evaluate how community rating regulations add consumer value relative to the pass through observed in our data. While a number of empirical studies (e.g., Bundorf et al., 2011) have examined individual-level

[3]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.

[4]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.

reclassification risk stemming from inefficient choices of health plan attributes,[5] we are not aware of any other study that has attempted to empirically quantify the reclassification risk from experience rating.

Second, our estimation of the small group market is novel. Small group health insurance is important, covering about 18 million people in 2013 (Kaiser, 2013a); has substantial potential reclassification risk stemming from small risk pools; and displays substantial evidence of market failure.[6] The ACA may affect this market both positively or negatively. On one hand, community rating will raise welfare by providing reclassification risk protection. On the other hand, removing the ability of insurers to contract on health status may increase adverse selection, potentially requiring firms to raise premiums for all to remain profitable in this market. Thus, it is important to understand the extent of reclassification risk in this segment, and how it will be affected by community rating.

We analyze data covering 9,281 employers and 371,752 enrollees from 2012-14 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 specify ratings bands that effectively cap the amount of experience rating (see Kaiser, 2013b). 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 (see Lucia et al., 2014, for details). Overall, we believe that rating

[5]Since employers generally cannot base the premiums that they charge to their employees on risk factors, individual-level reclassification risk in employer-sponsored insurance does not stem from individual experience rating.

[6]For instance, fewer than 50 percent of small firms even offer health insurance (MEPS, 2013).

requirements in our sample are very similar between 2013 and 2014.

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

We estimate the extent of reclassification risk by evaluating how much a change in the mean ACG score for an employer leads to a change in premiums. Specifically, we estimate linear specifications at the employer-year level which regress premiums on the ACG score and include fixed effects for employers and years. Since we include firm fixed effects, our identifying assumption is that changes in the mean ACG score for an employer are not correlated with any unobservable changes in the premiums that would have occurred in the absence of the health shock. We find that a one standard deviation increase in mean ACG score for an employer—calculated using its previous year's claims—increases its mean annual premium by \$103 with employer fixed effects or \$1,296 without employer fixed effects. The true long-run pass through likely lies between these two numbers. We also 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 most chronic diseases—do not significantly affect the per-enrollee premium. Benefits chosen by employers rise slightly in response to increases in mean ACG score, although the effect is quantitatively small.

We then examine the the relation between ACG score and claims. This relation is meant to be causal conditioning on medical prices, and hence we estimate it with linear regressions that include market fixed effects—which allow for different medical prices across different markets—but not employer fixed effects. We find that a one-standard-deviation increase in the ACG score increases annual claims by an average of $3,121. Dividing 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 expected standard deviation of $660 or less in their 2014 expenditures for healthcare and health insurance. With full experience rating—as would occur with perfect competition and annual contracts—this standard deviation would rise to $1,201. Applying a CARA model and measures of risk aversion from Handel (2013), we find that this extra variation in expected income implies that full experience rating at the same mean premiums lowers the certainty equivalent income level by an average of $756 or more. Thus, the fact that USIC does not fully experience rate adds substantial value in this market, value that might be shared between the firm and and consumers.

We also investigate the impact of community rating regulations, as will occur under the ACA. With community rating, the expected standard deviation of health spending would fall from $660 or less to $411. The reason that the standard deviation is still positive is because community rating does not eliminate out-of-pocket expenditures. The relatively small decrease in the expected standard deviation is also reflected in a certainty equivalent utility gain of only $193 or less from community rating, relative to the pricing policy reflected in our data. Thus, community rating regulations are only likely to add value if they do not result in substantial increases in markups.

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. It serves two purposes. First, the model allows us to fix ideas regarding how to think about reclassification risk. Second, it provides a formal framework with which we consider the welfare and distributional effect of alternative risk rating policies.

2.1 Enrollee utility and choice

We start by discussing the enrollee side of the market. We consider utility and choice for 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, the time period by t, and the number of enrollees at employer j as I_j. 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.

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 population, 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 the function $c(H(r_{ijt}))$ denote the claims cost for an individual with health shock $H(r_{ijt})$. We 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

$$E[c(H(r_{ijt}))] = \gamma_1 r_{ijt}, \tag{1}$$

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 assume that utility is additively separable across the time periods. Per-period utility is a function of the potential enrollee's income Y_{ijt}, her premium, and her out-of-pocket health costs:

$$U(p(R_{jt}, j), r_{ijt}) = \int u\left[Y_{ijt} - p(R_{jt}, j) - c^{oop}(H(r_{ijt}))\right] dF_H(H(r_{ijt})), \tag{2}$$

where $dF_H(H(r_{ijt}))$ is the distribution of health shocks conditional on a risk score and $u(\cdot)$ is the utility conditional on a particular health shock realization. We assume that $u(\cdot)$ follows a CARA functional form, which is often used to model health expenditures (e.g., Handel, 2013). 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]

We now further consider the utility from health insurance, starting with period 2. At the beginning of this period, the ex ante utility for an enrollee is given by $U(p(R_{j2}, j), r_{ij2})$. The only potential risk that is faced by an insurance enrollee at period 2 is the (relatively small)

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

risk of paying high out-of-pocket costs given a bad health shock; the enrollee does not face the reclassification risk of higher future premiums.

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:

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

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 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. 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.

We now turn to the decision of employer j to purchase health insurance and offer it to its employees. We assume that employer j partially internalizes the utility of its enrollees in purchasing insurance but also has an idiosyncratic component to its utility from purchasing insurance, ε_j. We let each ε_j be an *i.i.d.* draw from some known distribution, whose realization is not known to the insurer when its sets p_j but is known to employer j when it makes its insurance purchase decision. Employer j's per-period utility from purchasing insurance and offering it to its employees is then:

$$U^j(p, r_{1j}, \ldots, r_{I_j j}) = \overline{U}^j(p, r_{1j}, \ldots, r_{I_j j}) + \varepsilon_j,$$

for some mean utility function $\overline{U}^j$.

We assume that the mean risk score R_j is a sufficient statistic for $r_{1j}, \ldots, r_{I_j j}$ in $\overline{U}^j$. Rewriting the argument of employer utility, employer j will purchase insurance when $U^j(p, R_j) > 0$. We also assume that the enrollee take-up decision is inframarginal so that every potential enrollee at employer j will take up insurance if the employer offers it. These assumptions simplify the insurer decision problem: since either everyone or no one at employer j takes up insurance, the mean risk score conditional on selling insurance is always R_j, which the insurer can then take as given.

Since ε_j is not known to the insurer, the insurer assesses some probability that the employer will purchase insurance that depends on R_j and the offered premium. This then results in an expected quantity function from the point of view of the market, $Q_j(p, R)$, where premiums p are a choice variable and the mean risk score R is given.

2.2 Competitive insurance market

Having discussed the enrollee side, we now turn to the insurer side, starting with a perfectly competitive insurance industry. We assume that insurers are risk neutral and maximize expected profits by setting a potentially different premium for each employer. We focus on the premiums charged by this market to employer j.

To simplify our theoretical analysis, we assume that out-of-pocket costs are zero and that the only costs to insurers are claims, implying that insurer costs are the same as total claims costs.[8] With this assumption, we can write $E[c^{ins}(H(r_{ijt}))] = E[c(H(r_{ijt}))] = \gamma_1 r_{ijt}$, so the insurer's expected costs from covering individual i are proportional to her risk score. This assumption also allows us to simplify notation by omitting r as an argument of U, which only entered because out-of-pocket expenses were a function of risk scores. We also assume

[8] In contrast, our empirical results do allow for out-of-pocket costs to be positive.

that every employer purchases insurance when premiums are equal to costs. We can thus write that $Q_j(\frac{1}{I_j}E[c(H(r_{1jt})) + \ldots + c(H(r_{I_jjt}))], R) = Q_j(\gamma_1 R, R) = I_j$.

We first consider the case where insurers can only offer annual contracts. In this case, there is no linkage between the two periods of the model. Since firms observe risk scores and the competitive market will set premiums equal to expected claims costs, we have that $p(R, j) = \gamma_1 R$. For this case then, equation (3) specializes to:

$$V(r_{1j1}, \ldots, r_{I_jj1}, i, j) = U(\gamma_1 R_{j1}) + \delta \int U(\gamma_1 R_{j2}) dF_R(R_{j2}|r_{1j1}, \ldots r_{I_jj1}). \qquad (4)$$

Even though consumers in this market pay premiums equal to their expected costs, they are still faced with reclassification risk: an increase in mean risk score at employer j would translate into an increase in expected insurance costs at the employer in period 2.

Now suppose that the perfectly competitive insurance industry can offer two-period contracts to employer j with binding commitments on both sides. Consider such a contract with a period 1 premium of $p_{j1} = \gamma_1 R_{j1}$ and a period 2 premium of $p_{j2} = \gamma_1 E[R_{j2}|r_{1j1}, \ldots r_{I_jj1}]$. Note that this contract would have premium equal to expected marginal cost and would eliminate reclassification risk. Because of this, with CARA utility,

$$\int U(Y_{ij2} - \gamma_1 R_{j2}) dF_R(R_{j2}|r_{1j1}, \ldots r_{I_jj1}) < U(Y_{ij2} - \gamma_1 E[R_{j2}|r_{1j1}, \ldots r_{I_jj1}]),$$

implying that such a contract would improve enrollee welfare over the state-contingent one-period contracts considered above. If income were identical across periods and mean risk were the same across periods so that $E[R_{j2}|r_{1j1}, \ldots r_{I_jj1}] = R_{j1}$, this contract would be the utility-maximizing contract among break-even contracts and so the perfectly competitive insurance industry would result in employer j always signing this two-period contract.[9]

[9] 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., 2016)

2.3 Insurance market with pricing power

We next consider insurance provided by a single insurer with pricing power. We maintain the assumptions that out-of-pocket costs are zero and that insurers are risk-neutral and maximize expected profits. We also again focus on the premium that the insurer charges to employer j and assume that the insurer can charge different premiums to different employers. However, in contrast to the competitive case, since premiums might be higher than expected claims costs, we no longer assume that employer j always purchases insurance. Instead, we simply require that $Q_j(p, R)$ be twice differential.

We again first consider an insurer with pricing power which can only offer annual contracts. The insurer decision problem here is a repeated static pricing decision. The insurer will set a premium for group j based on the group's realized mean risk score and its expected demand. Thus, we can write the expected profits for the insurer from premium p for a group with risk score R as:

$$\pi(p, R) = [p - \gamma_1 R]Q_j(p, R), \tag{5}$$

where again $\gamma_1 R$ reflects the insurer per-patient costs.

We can then write the first-order necessary condition for profit maximization as:

$$\frac{\partial \pi}{\partial p} = Q_j(p, R) + [p - \gamma_1 R]\,\frac{\partial Q_j}{\partial p} = 0. \tag{6}$$

In order to understand the impact of risk score on the insurer premium, we implicitly differentiate (6) with respect to employer j's mean risk score R. Doing so and rearranging terms, we obtain:

$$\frac{dp}{dR} = \frac{\gamma_1 \frac{\partial Q_j}{\partial p} - \frac{\partial Q_j}{\partial R} - (p - \gamma_1 R)\frac{\partial^2 Q_j}{\partial p \partial R}}{2\frac{\partial Q_j}{\partial p} + \frac{\partial^2 Q_j}{\partial p^2}(p - \gamma_1 R)}. \tag{7}$$

Since equation (7) is relatively involved, it is worth considering a simple special case as a benchmark. Assume that that demand is linear, so $\frac{\partial^2 Q_j}{\partial p^2} = 0$, and that firm risk R does not affect its demand for insurance, so $\frac{\partial Q_j}{\partial R} = 0$. Under these assumptions, (7) specializes to $\frac{dp}{dR} = \frac{1}{2}\gamma_1$. This simple result is analogous to the well-understood result that a monopolist with linear demand would pass through one-half of an expected cost increase to consumers.

Recall that experience rating—or equivalently full pass-through from expected risk to premiums—implies that $\frac{dp}{dR} = \gamma_1$. Since the above example implies that pass-through is only one-half of the expected risk, here the insurer with pricing power provides partial reclassification risk protection, unlike in the competitive case. However, it is possible to construct other examples where the insurer with pricing power passes through more than the full amount. These can stem from either a different curvature of demand with respect to premium increases (Weyl and Fabinger, 2013) or from a non-zero response of expected quantity to a change in employer mean risk score.

Now consider commitment in the context of firms with pricing power. As in the perfectly competitive case, a binding two-period contract at a fixed premium adds value to risk-averse employees by shielding them from reclassification risk. Hence, an insurer with pricing power and the ability to enforce two-period contracts may choose to offer such a contract, as it will be able to both earn more profits from such a contract and add consumer value.

Relatedly, consider the role of inertia in affecting pass through. Inertia is generally believed to exist in the context of individual health plan choice (Handel, 2013) and may also exist for small employers' choices of health plans for their employees. While other researchers have focused on the fact that inertia may increase markups, employer inertia here may also help with providing equilibrium reclassification risk protection. This is because inertia provides implicit commitment on the part of the employer to not switch insurers. It may then be optimal for a large insurer to use its reputation to provide implicit commitment to limit or eliminate experience rating with employers, in exchange for the implicit commitment of employers without adverse health shocks to not switch health plans.

Combining these factors, suppose there is potentially incomplete pass from period 2 risk to premiums due to static pricing power and/or implicit commitments. A simple functional form for the pass through from period 2 risk to the premium offered to employer j is:

$$p_{j2} - p_{j1} = c_2 + \beta_1(R_{j2} - E[R_{j2}|r_{1j1}, \ldots, r_{I_j j1}]), \tag{8}$$

for some constant c_2 that is the same across employers. Note that the time 1 expectation of the period 2 premium increase is simply c_2, and thus this term would include any extra costs or markups in period 2.

If $\beta_1 = \gamma_1$, then the insurer fully experience rates the health risk R_{j2}. For $0 < \beta_1 < \gamma_1$, there will be positive but incomplete pass-through from expected risk to premiums. Under community rating or binding two-period contracts, we would have $\beta_1 = 0$. It is also easy to see that, for $\beta' < \beta_1$, given CARA utility,

$$\int U(Y_{ij2} - p_{j1} - c_2 - \beta_1(R_{j2} - E[R_{j2}|r_{1j1}, \ldots, r_{I_j j1}]))dF_R(R_{j2}|r_{1j1}, \ldots r_{I_j j1})$$
$$< \int U(Y_{ij2} - p_{j1} - c_2 - \beta'(R_{j2} - E[R_{j2}|r_{1j1}, \ldots, r_{I_j j1}]))dF_R(R_{j2}|r_{1j1}, \ldots r_{I_j j1}),$$

since the left side is a mean-preserving spread of the right side. Thus, the utility for individuals at employers which offer insurance will be higher the lower is β_1. Because consumer utility will be higher with a lower β_1, an insurer with pricing power faced with potential enrollees with inertia would try to have a low β_1 to maximize consumer welfare and a high c_2 to capture some of this welfare.

Overall, our takeaway is that insurers with pricing power may provide a certain amount of reclassification risk protection even in the absence of formal multi-period insurance contracts. 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. The goals of our empirical analysis are to estimate γ_1 and β_1, ascertain the extent of pass through, and understand the implications of this pass through for consumer risk and welfare.

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 40,341 employer-year observations and 891,953 employee-year observations, of which our estimation sample keeps 18,562 and 371,752 respectively.

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	32,483	10,837	39,212	26,934	127,301
Percentage of individuals	14%	4%	17%	11%	54%
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.

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 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: 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 ($)	3,314	2,792	2,678	3,605
	[16,045]	[11,884]	[16,110]	[16,962]
Lagged paid medical claims ($)	2,692	2,376	2,278	2,877
	[15,040]	[11,270]	[15,550]	[15,799]
Lagged paid pharmaceutical claims ($)	622	416	400	728
	[4,065]	[2,512]	[2,278]	[4,574]
Lagged out-of-pocket claims ($)	907	832	656	982
	[1,814]	[1,632]	[2,000]	[1,864]
Lagged allowed claims ($)	4,222	3,624	3,334	4,587
	[16,839]	[12,716]	[16,864]	[17,763]
ACG score, r_{ijt}	1.03	0.86	1.08	1.04
	[1.48]	[1.27]	[1.63]	[1.47]
Observations	364,068	39,212	26,934	254,602

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.

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 \$3,314, with medical claims accounting for 81% and prescription drugs expenditures accounting for the remaining 19%. We also report the out-of-pocket claims and the allowed claims. The latter figure indicates the total claims amount that the provider expects 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 \$907 and allowed claims have a mean of \$4,222, which empirically verifies this proposition. Finally, the sample mean ACG score is 1.03 with a standard deviation of 1.48, 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 provides summary statistics at the employer-year level. We observe 18,206 employer-year observations and 20,663 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 enrollees per employer is approximately 20

Table 3: Descriptive statistics at the employer-year level

Risk pool characteristics	Mean	Std. dev.
Subscribers	19.99	25.94
% Employees	64.99	22.06
Mean age	40.57	8.83
% Female	46.45	20.47
Average premium and risk score		
Employer mean annual premium ($)		
Both years	6,308	2,930
2013	6,024	2,850
2014	6,592	2,981
Employer mean ACG score, R_{jt}	1.15	0.80
Δ employer mean ACG score, $R_{j,2014} - R_{j,2013}$	0.03	0.60
Lagged presence of chronic conditions at employer level (%)		
Cancer (% of employees)	6.80	11.34
Transplant (% of employees)	0.21	2.16
Acute myocardial infarction (% of employees)	0.19	1.93
Diabetes (% of employees)	6.34	11.38
Number of unique employers	9,103	
Number of employer-year observations	18,206	
Number of employer-plan-year observations	20,663	

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

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 $6,308, with a large standard deviation of $2,930.

In addition, Table 3 shows that the mean ACG score across employers is 1.15, slightly higher than the mean ACG score at the individual level, implying that small groups tend to have higher risk scores than average. The standard deviation of the employer mean risk score is 0.80, 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 (0.03) but the standard deviation is quite large, 0.60, 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 7% and diabetes is 6%.

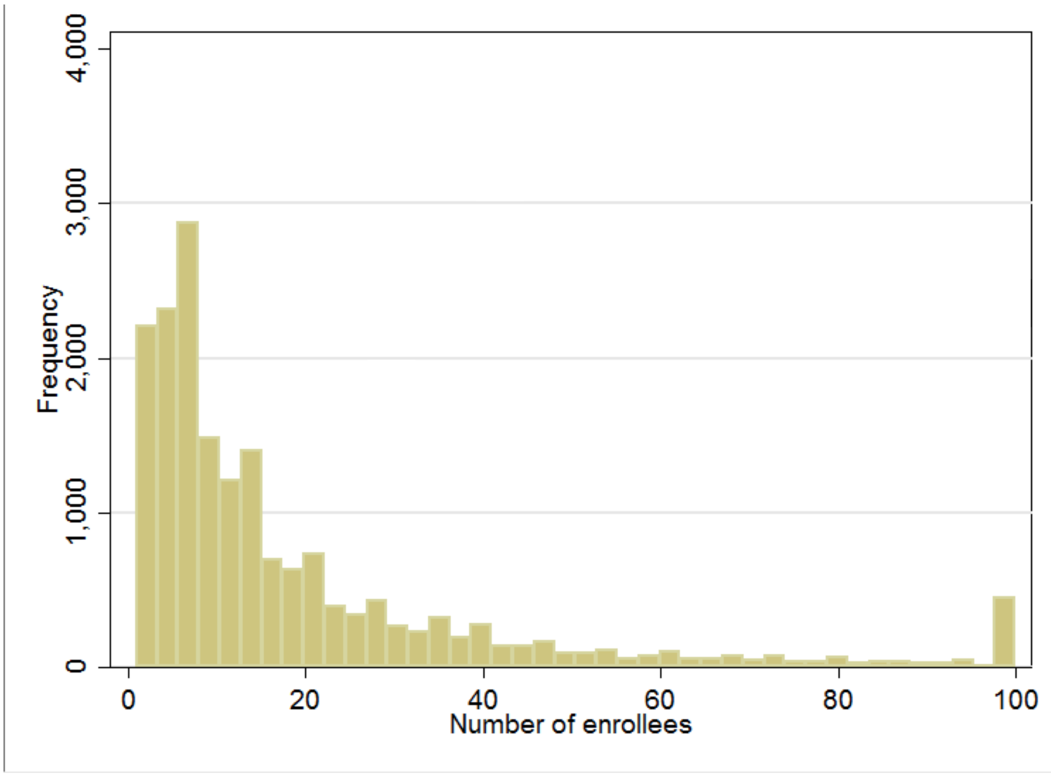
Finally, Figure 1 provides more evidence on the number of enrollees in our sample, by employer/year. The median number of enrollees per employer/year is 11.

4 Empirical Approach

4.1 Estimation Approach

The goal of our estimation is to recover the impact of risk score on expected costs ($^{\partial p}/_{\partial R}$), which is γ_1, and the impact of employer mean risk score on premiums ($^{\partial E[c^{ins}]}/_{\partial R}$), which is β_1. We are interested in these parameters for two principal reasons. First, we use them together

Figure 1: Histogram of number of enrollees by employer/year



Note: histogram top codes employer/years with more than 100 enrollees.

to recover and report the pass through from insurer costs to premiums, $\frac{\partial p}{\partial E[c^{ins}]}$, since

$$\frac{\partial p}{\partial E[c^{ins}]} = \frac{\partial p/\partial R}{\partial E[c^{ins}]/\partial R} = \frac{\beta_1}{\gamma_1}. \tag{9}$$

Second, we use these parameters separately in our counterfactual analysis, in order to understand the reclassification risk and welfare from different rating mechanisms. Note that these parameters regard insurer behavior; we do not estimate any enrollee utility parameters and our estimation algorithm does not impose consumer utility maximization.

We first discuss our estimation of the impact of risk score on premiums and then turn to our estimation of the impact of risk score on insurer costs. To estimate the impact of employer mean risk score on premiums, we estimate an empirical analog to the pricing equation for firms with pricing power given in equation (8). Specifically, we estimate regressions of the form:

$$p_{jt} = \beta_1 R_{jt} + \beta_2 x_{jt} + FE_j + FE_t + \varepsilon^A_{jt}, \tag{10}$$

where p_{jt} is the premium charged to employer j at time t and R_{jt} is the employer mean ACG risk score at time t. Note that R_{jt} is calculated using claims data from year $t-1$. In equation (10), FE_j are employer level fixed effects, FE_t are annual fixed effects, x_{jt} are time-varying firm attributes, and ε^A_{jt} is the unobservable. The unobservable 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 (10) cluster standard errors at the employer level.

Our key parameter of interest in equation (10) is β_1, the pass through from ACG risk score to premiums. We estimate equation (10) with two-year panel data of premiums from 2013 and 2014, although R_{jt} is based on the lagged claims, from 2012 and 2013. Since we use firm fixed effects and a two-year estimation sample, we can rewrite equation (10) as:

$$p_{j2} - p_{j1} = \beta_1(R_{j2} - R_{j1}) + \beta_2(x_{j2} - x_{j1}) + (FE_2 - FE_1) + (\varepsilon^A_{j2} - \varepsilon^A_{j1}). \tag{11}$$

Comparing equation (11) to equation (8), our empirical specification is equivalent to the theoretical model treating $-\beta_1 R_{j1} + \beta_2(x_{2t} - x_{1t}) + (FE_2 - FE_1) + (\varepsilon^A_{2t} - \varepsilon^A_{1t})$ as the empirical analog of $-\beta_1(E[R_{j2}|r_{1j1}, \ldots, r_{I_j j1}]) + c_2$.

Our main identifying assumption in equation (10) is that R_{jt} is exogenous conditional on firm and time fixed effects and other characteristics, or equivalently, using (11), that the change in employer j's mean risk score between 2013 and 2014 is mean independent from changes in unobservable factors that affect the premiums that employer j pays for insurance from USIC. Because we control for the baseline health status with the risk score, we believe that it is reasonable to consider the change in the risk score—which reflect changes in expected health expenditure conditional on the base expectation—to be exogenous.

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 premium changes. One possibility is that the insurer directly considers chronic diseases, in addition to the risk score, in its pricing decision. Hence, in some specifications based on (10), we include the mean percentage of enrollees with chronic diseases as an additional regressor. 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 significant empirical limitation is that, since our data are from USIC, we have no information on enrollees prior to them starting health coverage or following them leaving health coverage. Moreover, adverse selection in health insurance markets may be very important (Einav et al., 2010; Rothschild et al., 1976). However, the evidence from Table 1 that "quitters" are very similar to "stayers" suggests that the adverse selection at the enrollee level is relatively minor in our setting.

Finally, note that our base specifications of equation (10) 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. It is possible that USIC removes risk score data on "joiners" or "quitters" in updating their premiums. Thus, we also estimate a robustness specification 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 that follow from (1), and take the form:

$$c^{ins}(H(r_{ijt})) = \gamma_1 r_{ijt} + \gamma_2 x_{jt} + \varepsilon^B_{ijt}, \tag{12}$$

where $c^{ins}(H(r_{ijt}))$ is measured as 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. Our regressions based on equation (12) also cluster standard errors at the employer level.

Comparing equation (12) to equation (1), the empirical specification uses the actual insurer costs while the theoretical model is based on the expectation of costs. Thus, in the empirical specification, ε^B_{ijt}, will capture the difference between actual claims and expected claims for an individual in a year. Note also that our empirical specifications use $c^{ins}(H(r_{ijt}))$ and $c^{oop}(H(r_{ijt}))$ as the dependent variables while our theoretical model concerns $c(H(r_{ijt})) \equiv c^{ins}(H(r_{ijt})) + c^{oop}(H(r_{ijt}))$. Because the ACG risk score is only meant to be a linear predictor of total health expenditures—and not of its insured or out-of-pocket components—we estimate splines for our empirical specifications, in addition to a linear model.

Note that unlike our estimation of the impact of risk score on premiums from equation (10), our specifications here conceptually can be estimated from a cross-section in a local market 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 healthcare usage. Thus, we should expect a linear relationship between r_{ijt}, which is calculated using time $t-1$ data, and $c(H(r_{ijt}))$, provided that provider prices are held constant.[10] This relationship is exactly what we would like to recover, to understand ${\partial E[c^{ins}]}/{\partial R}$. Finally, note that because our source of variation and unit of observation are both different between our estimation of β_1 and γ_1, we do not estimate IV specifications to jointly recover their ratio.

[10]We include market-level fixed effects in these regressions to control for variation in provider prices across markets.

4.2 Calculation of welfare and counterfactuals

Our estimation results recover both β_1 and γ_1. Using these estimates and the empirical transitions of risk scores, we examine the extent of reclassification risk in our sample and the utility loss from this risk. We then compare the estimated level of reclassification risk and welfare to counterfactual environments, notably community rating and full experience rating.

Consumer reclassification risk in our framework occurs through two mechanisms. First, health shocks (for the enrollee or others in her group) result in enrollees facing higher premiums; second, enrollees may lose health coverage due to the higher premiums caused by these health shocks. Our counterfactuals focus on the first mechanism and do not account for any change in the coverage decision by the employer following a health shock. We make this simplification because we do not have any information on the health insurance coverage that employers receive upon terminating coverage with USIC.

We calculate our counterfactuals with three steps. First, we construct the future distribution of enrollee health risk and mean employer health risk to which an enrollee is exposed. 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 certainty equivalent income level. We now discuss these three parts of our analysis in turn.

First, to construct the distribution of enrollee health risk, we non-parametrically estimate the empirical distribution of period 2 ACG scores for each enrollee, using enrollees with similar period 1 ACG scores to each enrollee.[11] For each vector of simulation draws, we then construct a draw of the distribution of period 2 employer mean risk scores faced by the enrollee, by using the period 2 simulated enrollee risk scores for other enrollees at the same employer as the enrollee.

Second, to evaluate how changes in the period 2 mean employer ACG score translate into

[11] Although our model assumes that an enrollee's future health shocks are a random function of only her risk score, it is straightforward to loosen this assumption here and construct the future distribution of enrollee health risk as a function of both the risk score and other enrollee attributes such as age.

changes in premiums, we use our estimate of 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 draws calculated in step 1 that provide distributions of employee and mean employer ACG risk scores for each enrollee. We then sum the financial risks imposed by the distribution of higher premiums and by the out-of-pocket costs to derive the period 2 distribution of healthcare/insurance expenditures. We also replicate this step with counterfactual exposures to reclassification risk. Specifically, we examine full experience rating, by setting $\beta_1 = \gamma_1$, and community rating, by setting $\beta_1 = 0$. In these counterfactual simulations, we assume that the insurance plan characteristics other than price, and hence out-of-pocket costs, would remain the same as in the baseline.

Third, we consider the variance in net income and certainty equivalent income from the healthcare and health insurance expenditure risk borne by individuals in period 2. Since utility follows a CARA functional form, period 2 income can be expressed as:

$$u(Y_{ij2} - p(R_{j2}, j) - c^{oop}(H(r_{ij2}))) = -\frac{1}{\gamma} \exp\left(-\gamma \left[Y_{ij2} - p(R_{j2}, j) - c^{oop}(H(r_{ij2}))\right]\right). \quad (13)$$

We do not estimate γ, but instead use $\gamma = 0.000428$ from Handel (2013), which estimates risk in a similar context of health expenditure risk. Step 2 provides us with simulation draws from the distributions of $p(R_{j2}, j)$ and r_{j2} (which we use to calculate out-of-pocket costs) for every individual under the observed and counterfactual risk rating environments. We assume a common Y_{ij2} and then calculate certainty equivalent income levels—which are the income levels that, when earned with certainty, would give utility equal to the observed lottery—using these three components.[12] Because we assume that consumers are risk averse, the certainty equivalent income levels will be less than the actual income levels.

[12] With the CARA utility function, the certainty equivalent utility of a lottery does not depend on the base income level. Hence, our assumption of a common Y_{ij2} is without loss of generality. Note that our data do not include information on income.

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 (10). 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}	127 **	1,600***		
	(51)	(65)		
Employer mean lagged total claims			0.004	0.072***
			(0.030)	(0.008)
Firm FE	Yes	No	Yes	No
Market FE	No	Yes	No	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	18,206	18,206	18,206	18,206

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, though small. An increase in ACG score of 1 predicts an increase in annual premium of β_1 =$127 with employer fixed effects or $1,600 without fixed effects. The results are statistically significant at the 5% and 1% level respectively. This compares to the employer mean ACG score of 1.15 and a standard deviation of 0.80. Thus, a one standard deviation increase in expected risk for an employer would raise annual per-person premiums by $102 with fixed effects or $1,280 without fixed effects annually.

The difference between the results with and without fixed effects implies 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 two 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 willingness to bargain. Alternately, risk may be transmitted to premiums over a longer horizon than one year.

Table 4 also shows the impact of lagged claims on premiums. The results on pass through are somewhat similar to our base results, though smaller. 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 smaller, with a $100 increase in claims only raising premiums by $7.20 in the subsequent year.

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	127**	177***	2	
	(51)	(53)	(44)	
Mean age of beneficiaries			146***	
			(5)	
% Female			1,192***	
			(244)	
Employer mean log ACG score				0.054***
				(0.007)
Year FE	Yes	No	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	18,206	18,206	18,206	18,206

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.

We next turn to analyzing alternative specifications based on equation (10), 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 but similar to the base results. Column III starts with the base specification and adds an x_{jt} control variable, the mean age of beneficiaries. The coefficient on mean ACG score is no longer significant in this specification. Column IV estimates a log-log version of Column I. The results show a significant relationship between ACG score and premiums. However, a linear 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: 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}	127** (51)		101* (60)	241*** (49)
Employer mean ACG score for stayers		5 (54)		
Year FE	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Observations	18,206	18,154	9,530	8,676

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.

Table 6 presents further robustness specifications. Column II presents similar specifications 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 the effect of sample selection on the pass through from expected risk to premiums is small. This result also corroborates the evidence from Table 2 that stayers and quitters are similar in observable characteristics.

Table 6 columns III and IV present specifications similar to our base specification from Table 4, but splitting the sample based on smaller employers, with 12 or fewer covered lives in both 2013 and 2014, and larger firms, which is the remainder of the sample.[13] 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 larger pass through coefficient. We believe that the larger employers here may be more willing to accept higher pass through because their 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}	127**	108**	127**	124**	95*
	(51)	(50)	(51)	(51)	(52)
Lag % cancer at employer		437*			
		(230)			
Lag % transplant at employer			-121		
			(994)		
Lag % AMI at employer				1,268*	
				(696)	
Lag % diabetes at employer					1,120***
					(292)
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 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. After controlling for employer and year fixed effects and risk scores, all the chronic

[13] We chose the cutoff of 12 covered lives, because it is the sample median.

diseases except transplants result in a statistically significant increase in premiums. However, the level of pass-through is similar. Our takeaway from this is that the level of pass through from diseases in one year to premiums the next year at USIC is quite low.

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}	-14	-0.51**	-31***
	(19)	(0.22)	(11)
Year FE	Yes	Yes	Yes
Firm-Plan FE	Yes	Yes	Yes
Observations	20,663	20,663	20,663

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 20,663 observations here instead of 18,206 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 0.51 percentage points, and the in-network deductible drops by $31; the drops are significant at the 5% and 1% levels, respectively. 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 (12). Together with our regressions above based on equation (10), this allows us to recover ${}^{\partial p}/_{\partial E[c^{ins}]}$, as in equation (9). From Column I, we find that an increase in ACG score of one leads to an increase in USIC-paid claims for the enrollee of γ_1 =$4,232. From column II, an increase in ACG score increases the allowed amount of the claims by $4,750.

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}	4,232***	4,750***	519***
	(184)	(185)	(11)
Market FE	Yes	Yes	Yes
Observations	154,235	154,235	154,235

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.

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 $519. 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 0.80, would result in $3,386 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 $4,232 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 in mean risk score at an employer increases premiums by no more than $1,600 and expected costs by $4,232. Together, these coefficients imply that the pass through from expected costs

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

	Dependent Variable: Paid amount ($)			
Regressor:	I	II	III	IV
Spline employee ACG score, $r_{jt} \in [0, 1)$	2,926***	3,016***		
	(101)	(103)		
Spline employee ACG score, $r_{jt} \in [1, 2.5)$	3,178***	3,194***		
	(188)	(187)		
Spline employee ACG score, $r_{jt} \in [2.5, 5)$	3,662***	3,657***		
	(513)	(513)		
Spline employee ACG score, $r_{jt} \in [5, \infty)$	5,823***	5,818***		
	(697)	(697)		
Spline employee ACG score, $r_{jt} \in [0, .32)$			2,504***	2,680***
			(366)	(370)
Spline employee ACG score, $r_{jt} \in [.32, .57)$			4,982***	5,043***
			(504)	(506)
Spline employee ACG score, $r_{jt} \in [.57, 1.13)$			91	186
			(638)	(639)
Spline employee ACG score, $r_{jt} \in [1.13, \infty)$			4,717***	4,715***
			(276)	(276)
Market FE	No	Yes	No	Yes
Splines	Fixed cut points	Fixed cut points	Quartiles	Quartiles
Observations	154,235	154,235	154,235	154,235

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.

to premiums, $\partial p / \partial E[c^{ins}]$, is no more than 38%. 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. Although the relatively low pass through may be driven by pure static pricing power, the pass through is lower than in our stylized example in Section 2.3. Thus, the result is likely driven in part 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

Table 11: 2014 outcomes and welfare for base environment and full experience rating

	Full sample		Smaller employers	
	Base environment	Full experience rating	Base environment	Full experience rating
Premium paid	5,607	5,607	7,160	7,160
Premium paid within std. dev.	385	988	586	1,883
Out-of-pocket expenses	923	923	973	973
Out-of-pocket exp. within std. dev.	411	411	451	451
Health spending	6,529	6,529	8,133	8,133
Health spending within std. dev.	660	1,201	905	2,153
Certainty equivalent income loss	6,854	7,610	8,705	11,461

Note: results based on estimates and estimation samples. 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.

Using our estimates, we now examine the extent of reclassification risk and the resulting welfare loss for the base and counterfactual environments. We first compare the results of the base environment reflected in our data to full experience rating, with results in Table 11. Table 11 reports two sets of numbers. The first two columns of numbers report the results for the full sample for both environments, using the coefficient estimate of $\beta_1 = \$1,600$, which is the highest pass-through consistent with our estimation. The final two columns then report the results for the sample of smaller employers, using the coefficient estimate of $\beta_1 = \$1,316$, that was estimated from this sample.

We normalize all four environments to have the same mean premium of $8,151 (our overall sample mean for 2014) in order to focus on the variation in premiums rather than their levels. While the environments are constructed to have the same mean premiums, they differ in their actual premiums. In particular, for the full sample under the base environment, individuals face a mean standard deviation in their 2014 premiums of $387. With full experience rating, the mean standard deviation in premiums increases to $990. By construction, the ratio between these two is the same as the 39% pass through for the base environment.

For the sample of smaller employers, the standard deviations in their 2014 premiums are substantially higher, at $607 for the base environment and $1,944 with full experience rating. The substantially higher standard deviations are due to the fact that the risk sharing is over a much smaller number of individuals with these smaller employers. The fact that the pass-through coefficient is estimated to be slightly smaller for this sample implies that full experience rating is even more different than the base environment for the smaller employers than for our full sample.

By construction, out-of-pocket expenses are the same across the two environments. Thus, in both environments, the mean out-of-pocket spending for the full sample is $922 and the standard deviation is $411, with slightly higher numbers for the sample of smaller employers. The health spending variable combines the premium paid and out-of-pocket expenses. Thus, the mean health spending variable for the full sample, at $9,432, is the sum of the mean premium paid and out-of-pocket expenses, a number that is slightly higher for the smaller employer sample due to out-of-pocket costs.

Note that the distributions of 2014 premiums and out-of-pocket expenses are not independent, in both the real world and our counterfactuals. In the real world, we would expect a positive correlation between them since a negative health shock, all else equal, may lead an enrollee to have both a higher future premium and higher future out-of-pocket expenses.

Our counterfactuals allow for correlations by allowing each draw on risk scores to generate both 2014 premiums and out-of-pocket expenses. Our counterfactuals also show a positive correlation: e.g., the variance of health spending, at $\$663^2$, is higher than the sum of the variances of the two components of spending, $\$387^2 + \$411^2$.

The mean standard deviation of 2014 health spending under the base environment for the full sample is \$663, a number that rises to \$1,204 for full experience rating. For the sample of smaller employers, the base environment has a \$934 standard deviation in health spending which rises to \$2,221 under full experience rating.

Applying the CARA functional form and calibrated risk aversion parameter, we find that, even though the health spending in 2014 costs a mean of \$9,073, under the base environment individuals would on average be willing to pay a fixed \$9,413 to avoid all healthcare risk, implying that the risk reduces welfare by an average equivalent of \$340. If individuals were faced with full experience rating but with the same mean premiums, their certainty equivalent income would drop by \$776 ($\$10,189 - \$9,413$). Thus, the fact that USIC does not fully experience rate its plans benefits enrollees in the small group market by a mean of \$776. This number is much larger for the smaller employers at \$3,257 ($\$13,102 - \$9,845$). The certainty equivalent income loss ratio from the two gambles is much larger than the standard deviation ratio between these two gambles because there is a convex loss from larger gambles.

Thus, 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 a relatively small fraction of what it would be with full experience rating. This effect is particularly pronounced for the smaller employers in our sample. It adds value to risk averse enrollees.

Table 12 compares the base environment to community rating. As with Table 11, we report overall results and results that pertain only to the smaller employers within our sample. Columns 1 and 3 of Table 12 are identical to the analogous columns of Table 11, and are reproduced here just for completeness. Columns 2 and 4 are new and pertain to the community rating case.

By construction, community rating lowers the standard deviation of premiums from the

Table 12: 2014 outcomes and welfare for base environment and community rating

	Full sample		Smaller employers	
	Base environment	Community rating	Base environment	Community rating
Premium paid	5,607	5,607	7,160	7,160
Premium paid within std. dev.	385	0	586	0
Out-of-pocket expenses	923	923	973	973
Out-of-pocket exp. within std. dev.	411	411	451	451
Health spending	6,529	6,529	8,133	8,133
Health spending within std. dev.	660	411	905	451
Certainty equivalent income loss	6,854	6,661	8,705	8,265

Note: results based on estimates and estimation samples. 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.

baseline of \$387 (or \$607 for the smaller employer sample) to \$0. It does not eliminate reclassification risk, since individuals with high health shocks still face higher out-of-pocket costs on average. Overall, the effect is to reduce the mean standard deviation for health spending by \$252 (\$663 − \$411) for the full sample or \$472 (\$934 − \$462) for the smaller employer sample.

Because of the non-linearity of the welfare function, the certainty equivalent income gains from community rating are relatively small: \$200 (\$9,413 − \$9,213) for the full sample and \$535 (\$9,845 − \$9,310) for the smaller employer sample. Thus, were community rating regulations, as mandated by the ACA, to increase margins in this market by more than \$200, they would lower consumer welfare despite lowering reclassification risk.

Finally, note that we have only considered reclassification risk and not the risk from changes in premiums that are orthogonal to health status changes, i.e., ε_{jt}^{A} from (10). 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 environment reflected in our data. 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*.

CMS (2016). Market rating reforms, state specific rating variations; accessed: 08.15.2016. `https://goo.gl/J9tteI`.

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.

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`.

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

Kaiser (2013b). Small group health insurance market rate restrictions, 2013; accessed: 08.15.2016. `https://goo.gl/Wi6biJ`.

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.

Lucia, K., Corlette, S., and Williams, A. (2014). The extended ”fix” for canceled health insurance policies: Latest state action, 2014; accessed: 08.15.2016. `https://goo.gl/Wi6biJ`.

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.

MEPS (2013). Meps insurance component tables; accessed: 08.15.2016. https://goo.gl/Wi6biJ.

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.

Weyl, E. G. and Fabinger, M. (2013). Pass-through as an economic tool: Principles of incidence under imperfect competition. *Journal of Political Economy*, 121(3):528–583.