# Online Appendix for “Market Segmentation and Competition in Health Insurance”

Michael J. Dickstein, Kate Ho, Nathaniel Mark

September 28, 2022

Dickstein: New York University and NBER, michael.dickstein@nyu.edu; Ho: Princeton University, NBER, and CEPR, kate.ho@princeton.edu; Mark: U.S. Department of Justice, nathaniel.mark@usdoj.gov. The views expressed herein do not represent the views of the U.S. Department of Justice. Throughout this paper, we use the term "market" in ways that do not necessarily reflect the product and geographic boundaries of antitrust markets.

# A Institutional Details

## A.1 Individual market

**Pricing constraints.** In 2012, prior to the implementation of ACA regulations, only six states required insurers to issue individual insurance to any applicant and only seven states used a form of community rating to prohibit premium setting based on health status (The Kaiser Family Foundation, 2012). Oregon did not have a state-level 'guaranteed issue' rule prior to the ACA; only those households who met the requirements of a federal statute could be guaranteed coverage.[2] Carriers in Oregon could also price based on health status, though not as freely as in some states. Along with 10 other states and the District of Columbia, Oregon applied rate bands in the individual market prior to the ACA, prohibiting insurance companies from charging premiums beyond a specified share of the average premium in the market (The Kaiser Family Foundation, 2012).

The implementation of the ACA in 2014 harmonized formerly divergent state-based regulation of the individual market. Under current law, insurers in all states must guarantee issue all plans to all consumers.[3] In setting rates, carriers may only vary premiums with family size, state-defined geographic regions, tobacco use, and age. Further, premiums may only vary by age following a standard age curve, with a ratio of 3 to 1 from the oldest to youngest enrollee. We exploit the formulaic variation in premiums by age in demand analyses.

**Plan design.** Plans offered in the individual market after the ACA are also subject to regulation on both covered benefits and patient out-of-pocket costs. Under the ACA, all plans must cover a set of ten essential health benefit categories, including outpatient services, emergency room visits, pregnancy and maternity visits, mental health care, and prescription drugs. All plans offered fall into metal tiers classified by their actuarial value, the percentage of health costs the plan is expected to cover. The plan tiers are Bronze, Silver, Gold, and Platinum, with actuarial values of 60%, 70%, 80% and 90%, respectively.

Oregon added regulation in the marketplace beyond the federal requirements on two dimensions: required tier offerings and plan standardization (Blumberg et al., 2013). First, while the federal regulations require insurers to sell at least one gold and one silver plan in the marketplace in each geographic market they enter, Oregon requires that all insurers entering the marketplace must offer a bronze, silver, and gold plan. If an insurer chooses to offer individual plans outside the marketplace, it must offer at least one bronze and one silver plan. Second, Oregon requires insurers to offer a standardized plan in each of the required metal tiers. The standardized plan features a specific cost-sharing and benefit design; under the broader federal regulations, insurers have flexibility to design plan copayments, deductibles, and other benefits within a metal tier so long as it achieves the specified actuarial value for the tier.

Households eligible for cost-sharing subsidies must purchase a silver plan in order to receive them. These subsidies shift the standard silver plan design to a more generous actuarial value of between 73% and 94% for consumers with incomes on 100% to 250% of the FPL, with the lowest incomes receiving the higher actuarial values.

[2] For example, under federal law, individuals leaving group coverage of at least 18 months duration could not be turned down for individual coverage provided they enrolled within 63 days of losing group coverage.

[3] 'Grandfathered' plans—that is, plans that existed before the ACA and currently in effect—were not subject to these new regulations.

## A.2 Small group insurance market

Small employers, defined as firms with up to 50 full-time employees, have the option of offering health insurance coverage for their employees. In 2015, approximately 1 in 4 full-time employees worked for a small employer. The share of these employees covered by employer-provided insurance equaled 35% for workers at firms with 3-24 workers and 49% for firms between 25 and 50 workers. These rates combine both the likelihood of firms' offering coverage to classes of employees and employees taking up that coverage. 47% of firms with 3-9 workers and 68% of firms with 10-49 workers offered coverage to at least some of their workers. Within the set of small firms that offer insurance, approximately 76% of eligible workers take-up coverage (Claxton et al., 2015).

The small group market is subject to many of the same plan design restrictions as the individual market. Plans must cover the same essential health benefits, must be structured according to the same metal tiers, and must be "guaranteed issue". The small group and individual markets differ in the purchasing channel and the pricing rules.

**Purchasing channel.** As in the individual market, the ACA intended states to offer a marketplace for small group employers to shop for plans, known as the Small Business Health Options program or SHOP. During the span of our data, from 2014 through 2016 Oregon did not operate a SHOP marketplace; all small groups purchase coverage through agents.

Employees in a small group face two levels of agency: the employer and the insurance broker. A typical small group offers one to two broker-recommended plans to its employees, often from the same carrier (Agency for Healthcare Research and Quality, 2016). By contrast, consumers shopping in the individual market face a much larger choice set.

After choosing which plans to offer, employers contract with the relevant carrier(s) and pay premiums on behalf of the group. Employees pay their share of the group premium from their pre-tax earnings– that is, all premiums for insurance obtained through an employer are federal and state tax-exempt, regardless of whether the employer or the employee pays. This creates a tax wedge relative to the individual market, where households typically pay for insurance with post-tax dollars.[4] Employers also typically subsidize the cost of employee premiums, covering as much as two thirds of the premium cost (Claxton et al., 2015).[5]

**Premium setting.** Oregon requires insurers in the small group market to use a form of community rating known as 'tiered composite' rating. Employees at a firm do not face premiums proportional to their family size or the ages of members of their family. Instead, these ages and family sizes contribute to a group level premium cost which is then divided into four premium prices, based on the household size: (a) employee only, (b) employee and spouse, (c) employee and children, and (d) employee, spouse, and children.

To determine these four premium levels, the employer creates a list of all households who would be covered in the plan, including the family size and the ages of each household member.[6] With this

[4] Premiums in the individual market are part of itemized deductions, but subject to limitations: only medical expenses exceeding 7.5% of adjusted gross income are deductible.

[5] Small businesses with fewer than 25 employees are also eligible for tax credits of up to 50% of premium costs if they satisfy a number of qualifications, including: they must buy a plan certified for SHOP, average employee pay must be less than $50,000, and the employer must cover at least 50% of the premium (Oregon Health Insurance Marketplace, 2020).

[6] The prices also vary geographically based on the location of the employer. If an employer has multiple locations that span more than one pricing region in the state, employees may face different premiums by location.

list, and for a chosen plan design (e.g. a silver managed care plan), the insurer applies its baseline plan-specific premium to each employee and dependent according to the individual's age. Summing all of these premiums generates the total amount the employer pays the insurer for coverage under the chosen plan.

To determine the four premiums that households will face (before any employer subsidy), the insurer needs two measures. The first is the total age-based premiums in the group. The second is a sum of 'rating factors'; state regulators fix the factors to a specific level by family type. The rating factor equals 1 for employee-only, 2 for employee and spouse, 1.85 for employee and children, and 2.85 for employee, spouse, and children. The rating factors do not vary with the number of children. With these measures, the insurer calculates the average employee premium by dividing the total age-based premiums by the total rating factors. The four actual employer premium levels equal this average employee premium multiplied by the respective rating factor for each household size category.

This tiered composite system creates a cross-subsidy within the employer pool between older and younger enrollees and between employees covering only themselves and those covering families. The extent of the subsidy depends on how well the fixed rating factors match the differential in premium costs the families contribute to the small group pool.

## A.3 Individual Coverage HRA

Finally, we describe the structure of individual coverage health reimbursement arrangements, or ICHRAs, which became available as an option for employers on January 1, 2020. Employers can choose to offer ICHRAs to all of their employees or to certain classes of employees, such as full-time vs. part-time workers.[7] Generally, if an employer offers a traditional health plan to one class and an ICHRA to another, such as salaried versus non-salaried employees, the classes offered ICHRAs must exceed a minimum size to prevent adverse selection in the indivudal insurance market (Centers for Medicare & Medicaid Services, 2019). For example, for employers with less than 100 full-time employees, the minimum class size for an employer to offer an ICHRA plan is 10 employees.

When offering an ICHRA, the reimbursements that employers provide to employees do not count toward an employee's taxable wages. Thus, similar to a traditional plan, the employees' health spending and premium spending are excluded from federal income and payroll taxes. The employee will have greater choices for coverage, as they can use the ICHRA funds to purchase any qualified health plan offered in the individual insurance market in their geographic region. Larger employers subject to the employer mandate (generally those with greater than 50 full-time employees) can also avoid employer mandate penalties by offering an ICHRA, as long as the funds provided in the ICHRA allow the employee to purchase an "affordable" plan in the marketplace.[8]

If an employee is offered an "affordable" ICHRA plan, the employee and her household are not eligible for federal premium subsidies for marketplace coverage. However, if the ICHRA does not meet the minimum standards for affordability, and an employee declines the coverage, she is eligible for premium subsidies on the individual marketplace according to the usual income schedule.

[7] The allowable classes under ICHRA regulations include categories such as part-time vs. full-time workers, employees distinguished by geographic area of their workplace, salaried vs. non-salaried, and a few other enumerated classes (Health Reimbursement Arrangements, 2019).

[8] An ICHRA plan is not considered affordable if the premium for the lowest-cost silver plan for single coverage in the employee's rating region, less the monthly amount in the ICHRA, exceeds 9.61% (in 2022) of the household's annual income divided by 12 (Health Reimbursement Arrangements, 2019).

Finally, if an employee purchases coverage that exceeds the monthly reimbursement provided by the employer in the ICHRA, the employee may be able to use pre-tax money to pay the remaining cost through use of a 'cafeteria' plan, where the employer shifts funds from an employee's pay into the cafeteria plan. In such a scenario, the funds in the cafeteria plan can only be used to purchase an 'off-exchange' plan, not a plan offered on an marketplace created under the Affordable Care Act.

# B Dataset Construction

## B.1 Household dataset creation

Our primary data come from Oregon Health Authority's All Payer All Claims (APAC) dataset for the years 2014-2016. These data include all medical claims, drug claims, and insurance plan enrollment records for commercially insured individuals in the state of Oregon. The medical and drug claim data include financial variables, diagnoses, and dates of claims. The insurance plan enrollment records are at the individual – month level and include details of both medical and drug insurance plans. Before 2016, reporting of all claims and enrollment records were mandatory by law for all commercial insurers. In 2016, the US Supreme Court, in *Gobeille v. Liberty Mutual Insurance Company*, created an exemption for reporting for self-insured plans. This exemption does not affect our research, as we focus on the fully-insured individual and small group insurance markets.

The sample we use in the analysis is a subset of the APAC dataset. The original insurance plan enrollment data include approximately 50 million observations per year (49,952,770, 48,957,463, and 38,364,318 in 2014, 2015, and 2016 respectively). We keep only unique medical plan observations (leaving 27,882,303, 27,037,082, and 20,034,911 observations). Next, we collapse the data to the individual – year observation level from the individual – month observation level (2,160,301, 2,035,972, 1,643,776).

We further restrict the individual - year observation level dataset to reflect the goals of our analysis. We restrict the data to large carriers to enable us to solve for equilibrium insurer behavior in our empirical framework. We define a large carrier as a carrier with more than 5% of individuals in any of the years in the individual market or small group market. Eight carriers qualify as large. This removes 3.3% and 2.1% of individuals in the individual and small group markets. We further restrict our sample to include only individuals in the individual and small group markets, leaving 506,356, 466,320, and 494,539 individuals in 2014, 2015, and 2016. Lastly we collapse the dataset to the household – year observation level from the individual – year observation level. After these changes, the sample includes 312,122 households in 2014, 294,745 households in 2015, and 320,273 households in 2016.

## B.2 Switcher dataset creation

To estimate the preferences of small group market households, we use a sample of small group households whose plans were closed in 2014 or 2015. We then track these households as they choose plans in the individual market, under the assumption that we can estimate small group consumers' preferences, conditional on observables, in this setting. The APAC data allow us to conduct this exercise, as they include an identifier that allows us to track individuals across plans, markets, and years. We exploit this feature of the data to determine which plans close, which households move to the individual market in the following year, and which plans attract market share in the individual market.

We identify households whose plans were closed by tracking the members of employer-health plans in the following year. We define an employer-plan as a closed plan if 80% or more of its members are either uninsured, in a plan in the individual market, or obtain coverage through a spouse or as a dependent in the following year. We identify 6,779 closed employer-plans. These plans had a total of 16,631 enrollees.

Table A1 describes the destinations of these enrollees in the year after their plan closed. 5,321 choose plans in the individual market. 1,148 obtain coverage through a spouse or as a dependent, 920 obtain coverage in other markets, including group coverage or Medicare coverage.[9] 9,242 enrollees do not obtain coverage in any of the above markets. These enrollees may be uninsured, may be covered under Medicaid, may have moved out of Oregon, or may be untracked for some other reason. We assume that these enrollees choose the outside option of uninsurance.

### B.3 Variable construction

We collect household characteristics, costs, and plan choices from the claims data and the insurance plan enrollment data. We also collect insurer administrative costs from public medical loss ratio data. Last, we collect plan data from the National Association of Insurance Commissioners' SERFF database.

**Rating area.** Rating areas are sets of counties. We assign rating areas to individuals using zipcodes from claims data and city names from the plan enrollment data.

First, we assume that all members of a household live in the same location. We assign the mode of zipcodes observed in a household to every member of that household. We also assign the mode of city names in a household to every member of that household. Second, we assign rating areas in the following order:

An individual's rating area is assumed to be the rating area where their city is located. We allow the individual's city name to approximately match with the city name in an external dataset according to the minimal optimal string alignment distance with a maximum distance of 1. The external database of city names and their rating areas was partially manually created and partially obtained from publicly available census datasets. Some individuals are not assigned a rating area based on their city name due to missing city names or incorrectly spelled city names. We assign these individuals to the rating area that is associated with their zipcode. If the zipcode falls within two rating areas, we choose the rating area with the largest share of zipcode population. We use the HUD-USPS zipcode-county crosswalk from 2010 census data for this matching.

**Employer-health plan identifier.** For the small group market, we construct an employer-health plan identifier. For most insurance carriers, this identifier corresponds to a variable in the data named contract number. For two carriers, there are multiple health plans associated with each contract number. For these carriers, we set the employer-health plan identifier equal to a contract number-health plan index.

[9] We assign Medicare coverage for an enrollee who turns 65 in the plan year.

**Income.** While we do not observe household income, we develop a predictive model to approximate income for enrollees in the individual and small group markets. Our goal is to use this measure of income to identify a household's eligibility for government premium and cost-sharing subsidies in the individual market.

For a subset of households in our APAC data, we observe the household's net premium– that is, premium net of subsidies.[10] Because government premium subsidies scale with income, we can compute the household's income as a function of demographics (age and family composition) and the observed net premiums. We recover a predicted measure of income for 73,440 households with observed net premiums.

Using the recovered incomes, we estimate a predictive Tobit model of income. In the model, incomes $y_{imt}$ follow a normal distribution, conditional on household characteristics $x_{imt}$:

$$\begin{gathered} y_{imt} = \beta' x_{imt} + u_{imt} \\ u_{imt} \sim N(0, \sigma^2) \end{gathered} \tag{B.1}$$

However, the econometrician observes a truncated version of income:

$$y^*_{imt} = min\{y_{imt}, y^{max}_{imt}\}$$

The upper bound on the predicted measure of income, $y^{max}_{imt}$, is the lowest income that would make premium subsidies equal to zero. In almost all cases, $y^{max}_{imt} = 4$. We estimate the model parameters $\beta, \sigma^2$ via maximum likelihood. We use household characteristics including family size, city of residence, year, age, household health status score (both mean and max within family), whether the household is insured for all 12 months of the year, and the 10th and 50th income percentile of the household's zipcode.[11] We also include non-linear transformations of these characteristics. If households have missing age or gender, we estimate income using a Tobit model with these variables omitted.

We use the estimated model to predict the income of households in both the individual and small group markets. Using the estimated income of each household, we determine the household's federal and state average tax rates using the NBER's Internet TAXSIM version 27. We account for both marital status and the number of dependents when estimating average tax rates. We assume that married households file taxes jointly and that households have zero tax deductions.

**Metal tier.** Metal tiers and health plans are not consistently reported by all insurance carriers in the enrollment data. When they are missing, we estimate the metal tier of a household's chosen plan using their contract number, premiums, zipcodes, and external data on the number of subscribers in each metal tier and year by insurer and zipcode from Oregon's Department of Consumer and Business Services, Division of Financial Regulation.

Our determination of metal tier follows three steps. First, we set the household's metal tier to equal the metal tier of their health plan when available. Second, we use the metal tier reported by the insurer when it is available. Third, in the individual market, we set all remaining household's metal tiers to equal the mode of assigned metal tiers in their contract, if the household is in a contract that is likely to have consistent metal tiers for all of its members. We define these

[10] We confirm that the premiums we observe for these households are inclusive of subsidies by comparing the reported premiums to public data from SERFF that contain gross premiums by plan and household composition.

[11] When zipcode is unknown, we match percentiles of income using the household's city of residence.

contracts as those in carrier-markets where more than 95% of households have the metal tier that is most common in their contract. For data from most insurance carriers, these three steps generate numbers that match with external counts of metal tiers from Oregon's Division of Financial Regulation. For these carriers, we assume all other households to be in grandfathered plans without ACA-regulated actuarial values. For other carriers with less well populated metal tier identifiers, we use claims data, observed premiums, and data from Oregon's Division of Financial Regulation (DFR) to deduce metal tiers by plan.

**Medical spending.** To handle outliers, we winsorize medical spending above the 99th percentile of individual market medical spending ($8,937 per month). Here, we define medical spending as all medical costs covered under the insurance plan, omitting patient out-of-pocket expenses. Additionally, when we report statistics relating to medical costs, including expected non-discretionary spending ($E[\lambda]$), we winsorize using the same upper bound.

**Administrative costs.** We define administrative cost measures using publicly available data collected by the Centers for Medicare and Medicaid Services (CMS). CMS requires insurers to report administrative costs annually at the state-market level as an input into regulation of insurers' medical loss ratios. We collect this insurer-state level data for the individual market for the years 2014-2016.

From the CMS data, we define total per-enrollee monthly administrative costs as the sum of the following constituent elements of administrative costs: (1) taxes and fees, from Section 3 of the CMS form, which include federal and state taxes deductible from premiums, community benefit expenditures, contributions to the Federal Transition Reinsurance Program, and various other regulatory fees; (2) wellness activities, from Section 4 of the CMS form, which include activities to improve and promote health outcomes and prevent hospital readmission as well as expenses related to health information technology; and, (3) general administrative expenses, from Section 5 of the CMS form, which include direct sales salaries and benefits, agents and broker fees, and all other non-medical general and administrative expenses. We employ the total administrative cost measure in our empirical analyses reported in Table 6.

We report summary statistics of these cost variables in Table A2. In addition, in the table we report other measures available in the CMS data. We compute statistics on the count of annual subscribers in Oregon that each insurer covers in each possible business line. Finally, we present statistics on the number of states, including Oregon, where payers that operate in Oregon are active. We classify an insurer as active in a state when it files a report with CMS in that state and year.

**Constructed Plans.** To simplify the model of household plan choice, we bin similar plans together to create constructed plans. Specifically, we combine all plans with the same insurer, rating area, metal tier, on/off exchange status, and managed care type into a constructed plan.[12] We define the constructed plan's premium as an average of the constituent plan premiums.

Table A7 shows that there is little variation in plan characteristics within a constructed plan. In the individual market, the average constructed plan consists of 2.9 plans and more than 75% of constructed plans consist of 3 or less plans. Constructed plans in the small group market are

[12]Managed care type indicates whether a plan has a managed care design. We define a managed care as EPO or HMO plan types.

somewhat larger. There are 8.6 plans in the average small group constructed plan. In more than 50% of constructed plans, all plans have the same network and service area. In more than 75% of constructed plans, all plans are the same plan type and either all have a national network or do not.

### B.4 Uninsured population

Uninsured households are not observed in our data. We use the American Community Survey (ACS) to infer the number of uninsured households by county and age and we use the California Health Interview Survey (CHIS) to determine the likely characteristics of uninsured households. The ACS includes estimates of the uninsured populations by age, sex, and geography groups. The CHIS provides measures of the joint distribution of characteristics of the California uninsured population.

**Data.** To infer the size of the uninsured population, we use the 2014-2016 American Community Survey 1-year Estimates Health Insurance Coverage Status by Sex by Age datasets from the U.S. Census Bureau. These data include estimates of the number of civilian noninstitutionalized persons in each age and sex group that report "No health insurance" and "With health insurance." The estimates are available at the county level if county populations are above 65,000. Estimates are also available at the Public Use Microdata Area (PUMA) level. Counties with populations below 65,000 are proper subsets of PUMAs.

We chose to use the mixture of 1-year county level and 1-year PUMA level data over using only 5-year county data. Using 5-year county data offers improved precision of estimates, especially for small counties. However, the 5-year ACS relies on a moving average of 5 years of survey results. Therefore, its estimates are biased upwards by higher uninsurance rates in pre-2014 years.

To infer the characteristics of the uninsured population, we use the 2014-2016 California Health Interview Survey. The survey provides data on the joint distribution of family size and type, income, urban share, and health conditions for uninsured households. Unlike the ACS, the CHIS observations are at the household level. In using CHIS, we assume that the empirical distribution of characteristics observed in the CHIS is similar to the distribution of characteristics in Oregon. In 2016, California and Oregon had similar uninsurance rates (7% and 6%, respectively). This similarity remains in cross sections of the income distribution. Within the subset of incomes below 200% of the Federal Poverty Line, California had a non-elderly uninsurance rate of 13%. Oregon had a non-elderly uninsurance rate of 11% (The Kaiser Family Foundation, 2017).

**Procedure.** Using the ACS, we estimate the uninsured populations in each age group and rating area. Since rating areas are a collection of counties, we conduct these estimates at the county level and aggregate to the rating area. If a county has a small population, the ACS does not include county-level data. For these counties, we infer county age group uninsured populations using PUMA-level ACS data. Specifically, a small county's uninsured population for an age group is calculated as the percent of the PUMA population that is in that county multiplied by the uninsured population in the PUMA for the age group.[13] Out of 36 counties, 21 had populations below 65,000 and had to be estimated.

[13] For greater accuracy, we use population data from the American Community Survey 5-year estimates.

We use the CHIS to map the number of individuals who are uninsured to the number of households who are uninsured. For each year, age group, and rating area, we estimate the expected number of adults in a household using the CHIS data. We estimate the number of uninsured households by dividing the number of uninsured individuals by the expected number of adults in a household.

Next, we distribute characteristics to the uninsured households. First, uninsured households are distributed to bins according to the empirical joint distribution of characteristics from the CHIS, conditional on metropolitan area status of the rating area, year, and age group. Bins are defined by rating area, year, age group ( $< 17$, 17-23, 24-33, 34-43, 44-53, 55-64, $\geq 65$), income ( $< 2.5$, $\geq 2.5$), marital status, and whether the household survey respondent has health conditions.[14]

Lastly, we draw specific values of characteristics for each uninsured household randomly from the insured population of households in the same bin. Here, we assume that the distributions of characteristics are similar between the insured and uninsured populations once we condition on metropolitan area status, age, income, marital status, and health conditions.

# C Estimation Details

## C.1 Further details on deriving equations for estimation

As noted in Section 6.2.1, the assumption that $\lambda \sim exp(\alpha)$ implies

$$u^*_{j,t} = \frac{1}{2}x^2\omega\lambda - (1-x)\lambda + y_t - p_{t,j} + g(X_{j,t}, \epsilon)$$

Now the expected utility over the distribution of $\lambda$ is:

$$v_{j,t}(F_{\lambda,t}, \omega, \psi) = -\int exp(-\psi u^*_{j,t})dF_{\lambda,t}(\lambda).$$

If $\lambda \sim exp(\alpha)$ so that $E(\lambda) = 1/\alpha$, we can apply the properties of the exponential distribution along with the monotonic transformation $-\frac{1}{\psi}ln(-v_{j,t})$ to find the order preserving utility function[15]:

$$U_{j,t} = y_t - p_{t,j} + \frac{1}{\psi}ln[\frac{\alpha - \psi(1-x) + \psi\frac{1}{2}x^2\omega}{\alpha}] + g(X_{j,t}, \epsilon)$$

.

Comparing this utility to the utility from the outside option

$$U_{0,t} = y_t + \frac{1}{\psi}ln[\frac{\alpha-\psi}{\alpha}] + g_0(\epsilon_{0,t}),$$

[14]Health conditions used are Asthma, Diabetes, High Blood Pressure, and Heart Disease. A rating area is defined as metropolitan if most of its counties are designated metropolitan counties by the U.S. Census Bureau. These are rating areas 1, 2, 3, and 7. The ACS reports 0 uninsurance in some metropolitan area-year-age-income-marital status-family status-health status groups. We found this to be unreasonable and replaced those uninsurance estimates with: first, the average of uninsurance rates across other years in the same bin, holding everything else stable (139); second, if the first method is unavailable, the average of other uninsurance rates across other metropolitan status, holding everything else stable (114).

[15]Marone (2020) describes this step as “estimating demand in certainty equivalent units”. In our setting, when $X \sim exp(\alpha)$, $kX \sim exp(\alpha/k)$. Further, $exp(kX) \sim Pareto(1, \alpha/k)$ so that $\int exp(kX)df_X = E(exp(kX)) = \frac{\alpha}{\alpha-k}$ provided $\alpha > k$.

we obtain the utility of the inside goods relative to the (non-stochastic component of the) outside option as

$$U_{j,t} = -p_{t,j} + \frac{1}{\psi} ln[1 + \frac{\psi x}{\alpha - \psi} + \frac{\psi}{2}(\frac{x^2 \omega}{\alpha - \psi})] + g(X_{j,t}, \epsilon)$$

and

$$U_{0,t} = g_0(\epsilon_{0,t}).$$

Recognizing that when $Ax + Bx^2$ is close to zero, we can approximate $ln(1 + Ax + Bx^2) \approx Ax + Bx^2$., we write our utility expression as:

$$U_{j,t} \approx -p_{t,j} + \frac{x}{\alpha - \psi} + \frac{x^2 \omega}{2(\alpha - \psi)} + g(X_{j,t}, \epsilon).$$

We specify $g(X_{j,t}, \epsilon) = (\beta_0 X_{j,t} + \epsilon_{j,t})/(\alpha - \psi)$ so that sicker or more risk aversion consumers put more weight on plan characteristics like carrier identity, in the same way that they put more weight on coverage. Making an analogous assumption for the outside option, we find:

$$U_{j,t} \approx -p_{t,j} + \frac{x}{\alpha - \psi} + \frac{x^2 \omega}{2(\alpha - \psi)} + \frac{\beta_0 X_{j,t} + \epsilon_{j,t}}{\alpha - \psi} \tag{C.1}$$

$$U_{0,t} = \frac{\epsilon_{0,t}}{\alpha - \psi}$$

## C.2 Likelihood derivation

Section 6.2.2 provides a joint likelihood for the household's plan choice and its health spending that can be rewritten as

$$\mathcal{L}(\theta) = \prod_{i=1}^{N} \Bigg\{ \left( \frac{1}{\sum_{k=1}^{J} exp(V_{i,k})} \right)^{d_{i,1}} \prod_{j=2}^{J} \Bigg[ \left( \frac{exp(V_{i,j})}{\sum_{k=1}^{J} exp(V_{i,k})} \right) \left( 1 - exp\left( \frac{-\alpha_i \underline{c}}{x_j + \omega_i x_j^2} \right) \right)^{\mathbf{1}\{c_{ij} \le \underline{c}\}}$$
$$* \left[ \left( \frac{\alpha_i}{x_j + \omega_i x_j^2} \right) exp\left( -c_{i,j} * \frac{\alpha_i}{x_j + \omega_i x_j^2} \right) \right]^{\mathbf{1}\{c_{ij} > \underline{c}\}} \Bigg]^{d_{i,j}} \Bigg\} \tag{C.2}$$

With this notation, we write the log-likelihood:

$$L(\theta) = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{i,j} ln\left( \frac{exp(V_{i,j})}{\sum_{k=1}^{J} exp(V_{i,k})} \right) + d_{i,j}\mathbf{1}\{j \neq 1\}\mathbf{1}\{c_{i,j} \le \underline{c}\} ln\left( 1 - exp\left( \frac{-\alpha_i \underline{c}}{x_j + \omega_i x_j^2} \right) \right)$$
$$+ d_{i,j}\mathbf{1}\{j \neq 1\}\mathbf{1}\{c_{i,j} > \underline{c}\} \left[ ln\left( \frac{\alpha_i}{x_j + \omega_i x_j^2} \right) - c_{i,j} * \left( \frac{\alpha_i}{x_j + \omega_i x_j^2} \right) \right] \tag{C.3}$$

Given the exponential distribution, we need to constrain $\frac{\alpha_i}{x_j+\omega_i x_j^2} > 0$ as well. The score function of our log likelihood is:

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_{i=1}^{N}\sum_{j=1}^{J} d_{i,j}\left[\frac{\partial V_{i,j}}{\partial \theta} - \frac{1}{\sum_{k=1}^{J} exp(V_{i,k})}\sum_{k=1}^{J} exp(V_{i,k})\frac{\partial V_{i,k}}{\partial \theta}\right] +$$
$$d_{i,j}\mathbf{1}\{j \neq 1\}\mathbf{1}\{c_{i,j} \leq \underline{c}\}\left[\left(\frac{-exp\left(\frac{-\alpha_i \underline{c}}{x_j+\omega_i x_j^2}\right)}{1-exp\left(\frac{-\alpha_i \underline{c}}{x_j+\omega_i x_j^2}\right)}\right)\left(\left(\frac{-\underline{c}}{x_j+\omega_i x_j^2}\right)\frac{\partial \alpha_i}{\partial \theta} + \left(\frac{\alpha_i \underline{c} x_j^2}{(x_j+\omega_i x_j^2)^2}\right)\frac{\partial \omega_i}{\partial \theta}\right)\right] +$$
$$d_{i,j}\mathbf{1}\{j \neq 1\}\mathbf{1}\{c_{i,j} > \underline{c}\}\left[\frac{\partial \alpha_i}{\partial \theta}\frac{1}{\alpha_i} - \frac{\partial \omega_i}{\partial \theta}\frac{x_j^2}{x_j+\omega_i x_j^2} - \frac{c_{i,j}}{x_j+\omega_i x_j^2}\left(\frac{\partial \alpha_i}{\partial \theta} - \frac{\alpha_i x_j^2}{x_j+\omega_i x_j^2}\frac{\partial \omega_i}{\partial \theta}\right)\right] \tag{C.4}$$

### C.3 Specification

We further parameterize $\alpha_i, \omega_i, \psi_i$ as a function of household and/or plan level observables:[16]

$$\begin{aligned} ln(\alpha_i) &= W_{1,i}\beta_1 \\ ln(\omega_i) &= W_{2,i}\beta_2 \\ ln(\psi_i) &= W_{3,i}\beta_3 \end{aligned} \tag{C.5}$$

We simplify notation by defining the following vector:

$$\theta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

Under this specification, the score function has the form

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_{i=1}^{N}\sum_{j=1}^{J} d_{i,j}\left[\frac{\partial V_{i,j}}{\partial \theta} - \frac{1}{\sum_{k=1}^{J} exp(V_{i,k})}\sum_{k=1}^{J} exp(V_{i,k})\frac{\partial V_{i,k}}{\partial \theta}\right] +$$
$$d_{i,j}\mathbf{1}\{j \neq 1\}\mathbf{1}\{c_{i,j} \leq \underline{c}\}\left[\left(\frac{-exp\left(\frac{-\alpha_i \underline{c}}{x_j+\omega_i x_j^2}\right)}{1-exp\left(\frac{-\alpha_i \underline{c}}{x_j+\omega_i x_j^2}\right)}\right)\left(\left(\frac{-\alpha_i \underline{c}}{x_j+\omega_i x_j^2}\right)\begin{pmatrix} 0 \\ W_{1,i} \\ 0 \\ 0 \end{pmatrix} + \left(\frac{\omega_i \alpha_i \underline{c} x_j^2}{(x_j+\omega_i x_j^2)^2}\right)\begin{pmatrix} 0 \\ 0 \\ W_{2,i} \\ 0 \end{pmatrix}\right)\right] +$$
$$d_{i,j}\mathbf{1}\{j \neq 1\}\mathbf{1}\{c_{i,j} > \underline{c}\}\left[\begin{pmatrix} 0 \\ W_{1,i} \\ 0 \\ 0 \end{pmatrix} - \frac{x_j^2}{x_j+\omega_i x_j^2}\begin{pmatrix} 0 \\ 0 \\ W_{2,i} \\ 0 \end{pmatrix} - \frac{\alpha_i c_{i,j}}{x_j+\omega_i x_j^2}\left(\begin{pmatrix} 0 \\ W_{1,i} \\ 0 \\ 0 \end{pmatrix} - \frac{\omega_i x_j^2}{x_j+\omega_i x_j^2}\begin{pmatrix} 0 \\ 0 \\ W_{2,i} \\ 0 \end{pmatrix}\right)\right] \tag{C.6}$$

where

[16] Here, $\beta_1$ is $K_1$ x 1, $\beta_2$ is $K_2$ x 1, $\beta_3$ is $K_3$ x 1. $x_j$ is a scalar.

$$\frac{\partial V_{i,k}}{\partial \theta} = \begin{pmatrix} X_k \\ -exp(W_{1,i}\beta_1)p_{i,k} \cdot W_{1,i} \\ \frac{1}{2}exp(W_{2,i}\beta_2)x_k^2 \cdot W_{2,i} \\ exp(W_{3,i}\beta_3)p_{i,k} \cdot W_{3,i} \end{pmatrix} \tag{C.7}$$

# D Consumer Surplus

## D.1 Derivation

In Section 6.4 in the main text, we define the certainty equivalent utility, $e_{ijt}$, for household $i$, plan choice $j$, and year $t$. Here we provide more details of its derivation. To simplify notation, we suppress market $m$ subscripts in our description.

As in (Einav et al., 2013), we define $e_{ijt}$ such that $-exp(-\psi_i e_{ijt}) = v_{ijt}$. Thus, $e_{ijt} = -\frac{1}{\psi_i}log(-v_{ijt})$. Here, $v_{ijt}$ is the expected utility of choice $j$. In Section 5.1, we derived a form for $v_{ijt}$ by combining our chosen functional form for utility with the assumption that the household's underlying health care need, $\lambda_i$, follows an exponential distribution:

$$v_{ijt}(F_{\lambda,t}, \omega_i, \psi_i) = -\int exp(-\psi_i \times u^*_{ijt})dF_{\lambda,t}(\lambda) \tag{D.1}$$

$$= -exp\left[-\psi_i(y_t - p_{jt} + g(X_{jt}, \epsilon))\right] \left( \frac{\alpha_i}{\alpha_i + \psi_i(\frac{1}{2}x_{jt}^2\omega_i - (1 - x_{jt}))} \right) \tag{D.2}$$

where $\psi_i$ is the household's coefficient of absolute risk aversion. Substituting $v_{ijt}$ into our definition of $e_{ijt}$, we find:

$$e_{ijt} = -\frac{1}{\psi_i}log\left[exp\left(-\psi_i\left[y_t - p_{jt} + g(X_{jt}, \epsilon)\right]\right)\left(\frac{\alpha_i}{\alpha_i + \frac{\psi_i}{2}x_{jt}^2\omega_i - \psi_i(1 - x_{jt})}\right)\right] \tag{D.3}$$

$$= y_t - p_{jt} + g(X_{jt}, \epsilon) + \frac{1}{\psi_i}log\left(\frac{\alpha_i + \frac{\psi_i}{2}x_{jt}^2\omega_i - \psi_i(1 - x_{jt})}{\alpha_i}\right) \tag{D.4}$$

Similarly, the certainty equivalent utility for the outside option under the same approach equals:

$$e_{i0t} = y_t + g(X_{0t}, \epsilon) + \frac{1}{\psi_i}log\left(\frac{\alpha_i - \psi_i}{\alpha_i}\right) \tag{D.5}$$

We then normalize $e_{ijt}$ by the deterministic component of the certainty equivalent utility for the outside option. Again, recognizing that when $Ax + Bx^2$ is close to zero, we can approximate $ln(1 + Ax + Bx^2) \approx Ax + Bx^2$, we find:

$$e_{ijt} \approx -p_{jt} + \frac{\omega_i}{2(\alpha_i - \psi_i)}x_{jt}^2 + \frac{1}{\alpha_i - \psi_i}x_{jt} + \frac{\beta_0 X_{jt}}{\alpha_i - \psi_i} + \frac{1}{\alpha_i - \psi_i}\epsilon_{jt} \tag{D.6}$$

$$e_{i0t} = \frac{\epsilon_{0t}}{\alpha_i - \psi_i} \tag{D.7}$$

We want to compute consumer surplus for each household $i$. In notation:

$$CS_{it} = E_\epsilon \left[\max_j e_{ijt}\right] \tag{D.8}$$

We substitute our expression for the certainty equivalent utility into $CS_{it}$. Multiplying and dividing by $\alpha_i - \psi_i$, we find:

$$\begin{aligned}
CS_{it} &= E_\epsilon \left[\max_j e_{ijt}\right] && \text{(D.9)}\\
&= E_\epsilon \left[\frac{1}{\alpha_i - \psi_i} \max_j (\alpha_i - \psi_i) * e_{ijt}\right] && \text{(D.10)}\\
&= \frac{1}{\alpha_i - \psi_i} E_\epsilon \left[\max_j (\alpha_i - \psi_i) * e_{ijt}\right] && \text{(D.11)}\\
&= \frac{1}{\alpha_i - \psi_i} log \left(\sum_{j=0}^{J} \left[exp\left(-(\alpha_i - \psi_i) p_{jt} + \frac{\omega_i}{2} x_{jt}^2 + x_{jt} + \beta_0 x_{jt}\right)\right]\right) && \text{(D.12)}
\end{aligned}$$

## D.2 Consumer surplus in the small group market

To address the main caveat of our small group market consumer surplus measure – observed choices in the small group market do not fit the plan choices of small group employees forced to switch to the individual market – we assume a simple model of employer plan choice. In the model, the employer chooses a default silver-tier plan for its employees, but employees may opt for an alternative plan by paying a fee, $f_{i,j}$.[17]

If we define $V_{i,j}$ as the mean latent utility that household $i$ obtains when $f_{i,j} = 0$, then the mean latent utility function for household $i$ choosing plan $j$ including the fee is:

$$V'_{i,j} = V_{i,j} - (\alpha_i - \psi_i) f_{i,j}$$

We recover $\alpha_i$, $\psi_i$, and the remaining parameters of $V_{i,j}$ using our estimation of small-group preferences based on the forced switcher population. We then compare the predicted choices using these parameters to the observed choices in the small group market overall using a logit model. Specifically, we specify the fee as:

$$f_{i,j} = W_{i,j}^4 \beta_6$$

We recover the parameters of $f_{i,j}$ by fixing all other choice parameters at the switcher-sample estimates and maximizing the following likelihood:

$$L(\beta_6) = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{i,j} ln \left(\frac{exp(V'_{i,j})}{1 + \sum_{k \in \mathcal{J}_i} exp(V'_{i,k})}\right) \tag{D.13}$$

The variables $W_{i,j}^4$ include interactions between plan metal tier and household characteristics such as age, health status score, household size, and employer-group size.

We report estimates of $\beta_6$ in Table A5. We estimate that fees are positive and large for non-silver plans, suggesting that employers induce their employees into silver plans at greater frequency

[17] Normalizing silver tier fees to be zero has consequences for our consumer surplus estimates. This assumption is informed by the empirical observation that 50% of small group employees are enrolled in silver plans.

than the estimated preferences would suggest. For example, using results from specification (1) in Table A5, households would have to pay \$145 per month to receive bronze plan coverage. Our estimates also suggest that healthier households pay higher fees to choose an alternative plan. Older employees are estimated to pay lower fees for bronze plans, but not for higher coverage plans.

When accounting for the estimated fees using specification (5), our measure of small group consumer surplus drops significantly from our main estimate of -\$86 to -\$179.[18] In both of these measures, we assume that households do not have the option to opt out of insurance coverage. This generates negative estimates, as the revealed preferences of the small group switchers show that many households would prefer to be uninsured. Indeed, if we assume that uninsurance is an option for households, our consumer surplus measures increase markedly. Under a zero-fee assumption, consumer surplus increases from -\$86 to \$164; with fees, the surplus increases from -\$179 to \$134.

# E Cost Censor

In the main specification, the spending cutoff is fixed. As a robustness exercise, we allow the data to recover the cutoff $\underline{c}$. Insurers may not submit a claim when a cost draw falls below the cost cutoff $\underline{c}$. The probability that an insurer does not submit a claim when cost is below the cutoff is $G(\underline{c}|x_j,\omega_i) = P(c_{i,j} = 0|x_j,\omega_i,c_{i,j} \leq \underline{c})$. Thus, the density of a cost observation is:

$$f(c_{i,j}|x_j,\omega_i,\alpha_i) = \begin{cases} 1 & x_j = 0, c_{i,j} = 0 \\ 0 & x_j = 0, c_{i,j} \neq 0 \\ \left[1 - exp(-\frac{\alpha_i \underline{c}}{x_j+\omega_i x_j^2})\right] G(\underline{c}|x_j,\omega_i) & x_j \neq 0, c_{i,j} = 0 \\ \frac{\alpha_i}{x_j+\omega_i x_j^2} exp\left(-c_{i,j}\frac{\alpha_i}{x_j+\omega_i x_j^2}\right) & x_j \neq 0, c_{i,j} \neq 0 \end{cases}$$

The probability that costs are not submitted, given a cost below the cutoff, $G(\underline{c}|x_j,\omega_i)$, is an unknown object. We estimate this object.

We assume that the probability of not submitting a claim is independent of moral hazard. The object is estimated non-parametrically, conditioning on only on actuarial value: $\hat{G}(\underline{c}|x)$.

For each potential actuarial value, we estimate $\hat{G}(\underline{c}|x)$ non-parametrically using Gaussian kernel methods. First, we define the number of cost observations associated with the actuarial value and the number of cost observations associated with the actuarial value for which insurers did not submit a claim.

$$N_x = \sum_{i=1}^{N}\sum_{j=1}^{J} d_{ij}\mathbf{1}\{x_j = x\}$$

$$N_x^0 = \sum_{i=1}^{N}\sum_{j=1}^{J} d_{ij}\mathbf{1}\{c_i = 0, x_j = x\}$$

[18] As we do when computing the main estimate, here we omit both outlier households (in terms of expected health spending) and households for whom consumer surplus is undefined when computing average surplus. Additionally, small group consumer surplus estimates are calculated using the premiums households face in the tiered pricing environment. Removing tiered pricing increases our estimate of consumer surplus from -\$86 to -\$78.

Second, the empirical probability and cumulative distributions of non-zero cost observations below the cutoff are estimated.

$$\hat{F}(z|x_j = x) = \frac{1}{N_x - N_x^0} \sum_{m=1}^{M} K(\frac{c_m - z}{h})$$

$$\hat{f}(z|x_j = x) = \frac{1}{(N_x - N_x^0)h} \sum_{m=1}^{M} k(\frac{c_m - z}{h})$$

The probability that claims are not submitted can be recovered from these objects.

$$\hat{G}(\underline{c}|x_j) = \frac{N_x^0}{N_x^0 + (N_x - N_x^0)\hat{F}(\underline{c}|x_j)}$$

$$\hat{g}(\underline{c}|x_j) = \frac{\partial \hat{G}(\underline{c}|x_j)}{\partial \underline{c}} = \frac{-N_x^0(N_x - N_x^0)\hat{f}(\underline{c}|x_j)}{[N_x^0 + (N_x - N_x^0)\hat{F}(\underline{c}|x_j)]^2}$$

The log likelihood has the form:

$$\begin{aligned} L(\theta, \underline{c}) = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{i,j} ln \left( \frac{exp(V_{i,j})}{\sum_{k=1}^{J} exp(V_{i,k})} \right) \\ + d_{i,j} \mathbf{1}\{j \neq 1\} \mathbf{1}\{c_{i,j} = 0\} \left[ ln(1 - exp(-\frac{\alpha_i \underline{c}}{x_j + \omega_i x_j^2})) + ln\left(\hat{G}(\underline{c}|x_j)\right) \right] \\ + d_{i,j} \mathbf{1}\{j \neq 1\} \mathbf{1}\{c_{i,j} \neq 0\} \left[ ln\left(\frac{\alpha_i}{x_j + \omega_i x_j^2}\right) - c_{i,j} * \left(\frac{\alpha_i}{x_j + \omega_i x_j^2}\right) \right] \end{aligned}$$

The gradient of this likelihood with respect to $\theta$ remains unchanged from the specification where $\underline{c}$ is assumed. The derivative of the likelihood with respect to $\underline{c}$ is:

$$\frac{\partial L(\theta, \underline{c})}{\partial \underline{c}} = \sum_{i=1}^{N} \sum_{j=1}^{J} d_{i,j} \mathbf{1}\{j \neq 1\} \mathbf{1}\{c_{i,j} = 0\} \left[ \frac{\alpha_i}{x_j + \omega_i x_j^2} \frac{exp(-\frac{\alpha_i \underline{c}}{x_j + \omega_i x_j^2})}{1 - exp(-\frac{\alpha_i \underline{c}}{x_j + \omega_i x_j^2})} + \frac{\hat{g}(\underline{c}|x_j)}{\hat{G}(\underline{c}|x_j)} \right]$$

To estimate $\theta$ and $\underline{c}$, we use a penalized maximum likelihood approach. A standard maximum likelihood approach would overfit the data by setting $\underline{c}$ very high. That is, a high enough $\underline{c}$ ensures that the likelihood of observed *any* cost $c_{i,j}$ is 1. To avoid overfitting, we penalize the magnitude of $\underline{c}$, maximizing the penalized likelihood function:

$$PL(\theta, \underline{c}) = L(\theta, \underline{c}) - \Psi \left[ log(\underline{c}\sigma\sqrt{2\pi}) + \frac{(log(\underline{c}) - \mu)^2}{2\sigma^2} \right]$$

Where the penalization parameter $\Psi$, and hyperparameters $\mu, \sigma^2$ are specified before estimation. This method can be derived from the assumption that $\underline{c}$ has a log normal prior distribution.

$$\underline{c} \sim lognormal(\mu, \sigma^2)$$

Then, the likelihood of observing the data and $\underline{c}$ under the assumed prior is equal to the penalized likelihood function:

$$PL(\theta, \underline{c}) = log(P(\text{observe data}|\theta, \underline{c})) + \Psi log(P(\text{observe } \underline{c}|\mu, \sigma^2)$$

$$PL(\theta, \underline{c}) = L(\theta, \underline{c}) - \Psi\left[log(\underline{c}) + \frac{(log(\underline{c}) - \mu)^2}{2\sigma^2}\right]$$

# F Counterfactual Algorithm

In this appendix section, we provide more detail on the algorithm we use to compute our counterfactual equilibrium. The approach mirrors that of Azevedo and Gottlieb (2017) for a competitive insurance market. We adjust the algorithm for the specific regulatory environment in Oregon and for our specification of expected utility.

We begin by collecting the parameters from both our supply and demand side estimation. From our maximum likelihood routine, which relies on both household plan choices and observed health care spending to identify the parameters of demand, we collect $\theta$:

$$\theta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}$$

where the derived parameters of utility depend on $\theta$ under our specification:

$$\begin{aligned} ln(\alpha_i) &= W_{1,i}\beta_1 \\ ln(\omega_i) &= W_{2,i}\beta_2 \\ ln(\psi_i) &= W_{3,i}\beta_3 \end{aligned} \tag{F.1}$$

Here, $\beta_0$ is the parameter vector multiplying the plan-specific indicator variables in the utility specification.

We also collect $(\beta_4, \beta_5)$, the parameters of our supply-side price-setting equation, Equation 6.6, where under perfect competition insurance carriers set total premium revenue equal to the sum of the portion of health care costs the insurer bears and the insurer's allocated administrative costs for operating the plan:

$$\widehat{R}_{jmt} = \sum_{i \in \{m,t\}}^{N_{mt}} \left(\widehat{s}_{ijmt} * \widehat{\beta}_4 \kappa_{ij} \bar{c}_{ijmt}\right) + \widehat{\beta}_5 A_{jmt} \tag{F.2}$$

In the equilibrium search, our goal is to find $p_{jmt}$, the plan-market-year baseline premium the insurer sets to equate revenue and costs in equilibrium for all plans and markets. We can translate the normalized baseline premium to (a) gross premiums for household i, $p_{ijmt}$, and (b) net (subsidized) premiums for household i, $p^s_{ijmt}$, using regulated age-rating factors in effect in Oregon in our sample period. We label these factors $\gamma_{k,i}$ for household member $k$ in household $i$:

$$
\begin{aligned}
p_{ijmt} &= p_{jmt} \sum_{k \in i}^{K_i} \gamma_{ki} \\
p^s_{ijmt} &= max\{p_{ijmt} * (1 - subs^p_{ijmt}) - subs^l_{ijmt}, 0\}
\end{aligned}
\tag{F.3}
$$

We simplify the premium subsidy formula for the purpose of searching for counterfactual premiums. In the true formula, premium subsidies for individual market households depend both on household income relative to regulated thresholds and the premium for the second cheapest silver plan offered in a market and year. To simplify our equilibrium search, while still capturing the key features of premium subsidies, we fix subsidy levels before the equilibrium search algorithm. We allow for two types of subsidies. The first is a proportional subsidy $subs^p_{ijmt}$ to be applied to the gross premium. The second is a lump sum subsidy $subs^l_{ijmt}$.

In the main specification, we set $subs^l_{ijmt} = 0$ for individual market households (or small group market households in counterfactuals where the small group market closes) and set the proportional subsidy $subs^p_{ijmt}$ to equal the subsidized share of premiums a household would face in the observed market. Under this method, we capture the differential elasticity of subsidized consumers but still allow a degree of price-sharing to constrain insurers from raising price. We omit the specific demand shift for an insurer that sets its premium to rank as the second cheapest silver plan. For small group market households in the extended ICHRA counterfactuals, we set $subs^p_{ijmt} = 0$ and set $subs^l_{ijmt}$ equal to the household's tax benefit and employer contribution for their enrolled plan in the observed data. This allows for the employer contribution to remain constant as plan purchasing shifts to the individual market. The tax benefit is equal to the household's average tax rate times the gross premium. The employer contribution is equal to 65% of the remaining (post-tax) premium.

## F.1 Steps in the equilibrium search

With these components, our algorithm proceeds as follows:

1. Compute each household's expected utility for plan $j$ given the $l$th guess at the household's effective premium, $p^{l,s}_{ijmt}$.[19]

$$\hat{V}_{ijmt} = x_{jt} + \frac{1}{2}\hat{\omega}_i x^2_{jt} - (\hat{\alpha}_i - \hat{\psi}_i) p^{l,s}_{ijmt} + \hat{\beta}_0 Z_{jmt}$$

2. Compute predicted shares, $\hat{s}_{ijmt}$:

$$\hat{s}_{ijmt} = \frac{exp(\hat{V}_{ijmt})}{\sum_{k=0}^{J_{mt}} exp(\hat{V}_{ikmt})}$$

   where $k = 0$ represents the outside good, with $\hat{V}_{i0mt} = 0$. Here $J_{mt}$ equals the number of insurance options in market $m$ in year $t$.

[19] In the notation below, we leave off an $i$ subscript from $x_{jt}$, the actuarial value of plan $j$ at $t$. In our implementation, we allow $x_{jt}$ to vary by household for those low income households eligible for cost sharing subsidies.

3. Compute expected costs for household $i$ under plan $j$:

$$\bar{c}_{ijmt} = E[c_{ijmt}|\hat{\alpha}_i, \hat{\omega}_i] = \frac{x_{jt} + \hat{\omega}_i x_{jt}^2}{\hat{\alpha}_i}$$

4. Compute total costs to the insurer, the right-hand side of the price-setting equation:

$$\widehat{R}_{jmt} = \sum_{i \in \{m,t\}}^{N_{mt}} \left( \widehat{s}_{ijmt} * \widehat{\beta}_4 \kappa_{ij} \bar{c}_{ijmt} \right) + \widehat{\beta_5} A_{jmt}$$

We can also compute the left-hand side of the pricing equation using prices:

$$\widehat{R}_{jmt} = p_{jmt}^{l} \sum_{i \in (m,t)}^{N_{mt}} \left( \hat{s}_{ijmt} \left( \sum_{k \in i}^{K_i} \gamma_{ki} \right) \right)$$

Thus, to find the next guess at the normalized premium vector, $p_{jmt}^{l+1}$, we can combine the two sides of the equation as:

$$p_{jmt}^{l+1} = \frac{\sum_{i \in \{m,t\}}^{N_{mt}} \left( \widehat{s}_{ijmt} * \widehat{\beta}_4 \kappa_{ij} \bar{c}_{ijmt} \right) + \widehat{\beta}_5 A_{jmt}}{\sum_{i \in (m,t)}^{N_{mt}} \left( \hat{s}_{ijmt} \left( \sum_{k \in i}^{K_i} \gamma_{ki} \right) \right)}$$

5. We test for convergence by comparing the plan-specific price vector $p_{jmt}^{l+1}$ from the $(l+1)$st iteration against the $l$th iteration across all plans. We compute the mean across plans $j = 1, ...J_{mt}$ in all markets $m$ and $t$, which total $J$ possible plans:

$$\frac{1}{J} \sum_{\forall (j,m,t)}^{J} \left( \frac{p_{jmt}^{l+1} - p_{jmt}^{l}}{p_{jmt}^{l}} \right)$$

If the mean percentage difference in prices (excluding the outside good) is less than .001, we stop the search algorithm and define the equilibrium price vector, $p^{eql}$, of length $J$. An entry in $p^{eql}$ is equal to $p_{jmt}^{eql} = p_{jmt}^{l+1}$. If our condition on the mean percentage difference in premiums is not satisfied, we return to Step 1, using $p_{jmt}^{l+1}$ in place of $p_{jmt}^{l}$.

## F.2 Counterfactual sample

For each counterfactual scenario we examine, we define the set of households able to choose a plan in the individual insurance market in 2016. We include households in the data who purchase an individual market or small group plan plus the set of uninsured households in that year. In addition to this sample, we also add a group of households we label "behavioral types," as in Azevedo and Gottlieb (2017). In each market-year pair $(m, t)$, we add a total number of additional behavior types equal to 1% of households in $(m, t)$. These households incur no health care costs and choose each of the $J_{mt}$ plans in the market-year with equal probability: $s_{ijmt}^{behav} = 1/J_{mt}$. We also assume there is no additional administrative cost to serve these households; that is, $A_{jmt}$ remains unchanged. Following Azevedo and Gottlieb (2017), we must include behavioral types in the counterfactual sample to guarantee existence of the equilibrium.

### F.3 Counterfactual Standard Errors

We estimate standard errors for the main counterfactual results using a bootstrap methodology. These standard errors are reported in Tables 7 and 8. We use 20 independent bootstrap draws in this estimation. For each draw, we sample households with replacement from the full individual market sample and from the small group switchers sample. Demand parameters, $\beta_0$, $\beta_1$, $\beta_2$, and $\beta_3$, are estimated for each bootstrap sample. Following our empirical methodology, we then use the demand parameters estimated for individual market households and characteristics of households in the individual market bootstrap sample to predict household plan choice probabilities. Using these predicted choice probabilities and expected household costs, we re-estimate the premium setting equation. In this exercise, we treat premiums, $p_j$, as fixed across bootstrap samples. Lastly, using the bootstrapped demand and supply parameters, along with the (non-bootstrap) individual market and small group market sample, we simulate our main counterfactuals. Our reported standard errors are the standard deviation of counterfactual outcomes across these simulations.

## G Premium Comparison

In this section, we describe how we compute and compare household premiums in the small group and individual insurance markets.

### G.1 Premiums in the individual market

For the individual market, we observe each each household's composition in terms of the number of household members and each household member's age. For each plan in the household's choice set in its rating area and in a given year, we collect the standardized premium for those plans. The premium calculation involves the following steps:

1. For each member of a subscriber-HH, collect the age-rating factor according to the standardized age curve for Oregon.[20]

2. Multiply the person-specific age-rating factor against the standardized base premium for the plan.

3. Sum the age-based premium for household members to find the household premium for that plan. Following ACA regulations, for households with more than three dependents we include only the age-specific premiums for the three eldest dependents when computing the household's total premium.

### G.2 Premiums in the small group market

Computing household-level premiums in the small group market requires knowing not only the household's size and the ages of household members, but also information on all households in the observed small group. We compute a household's premium for each possible plan as follows:

[20] Oregon follows a state-specific age-rating curve. The exact rating factor or multiplier for each age can be found here: https://www.cms.gov/CCIIO/Programs-and-Initiatives/Health-Insurance-Market-Reforms/Downloads/StateSpecAgeCrv053117.pdf

1. For each member of a subscriber-HH, collect the age-rating factor according to the standardized age curve for Oregon.

2. Multiply the person-specific age-rating factor against the standardized base premium for the plan.

3. Sum the age-based premium for household members to find the household premium for that plan. All dependents' age-specific premiums contribute to the household's total premium.

4. Save the total premium computed above by household for *all* households in the small group. Sum these premiums to find the total "plan side premium"– that is, the total amount of premium dollars the small group pays the insurer to cover the households in the group. Call this sum $TotPrem_{plan}$.

5. To find the premium that each household faces, we divide up the total plan-side premium by weights:

   (a) Assign to each household a family rating factor: single = 1, single with children = 1.85, married without children = 2, and married with children = 2.85.

   (b) Sum these family rating factors across all households in a small group. For example, if a group consisted of two single households and one household with two parents and children, the groups rating factor would equal 1+1+2.85 = 4.85. Call this sum $TotFRF$.

   (c) Define one of four possible household premiums as: single = $1 * \frac{TotPrem_{plan}}{TotFRF}$, single with children = $1.85 * \frac{TotPrem_{plan}}{TotFRF}$, married without children = $2 * \frac{TotPrem_{plan}}{TotFRF}$, and married with children = $2.85 * \frac{TotPrem_{plan}}{TotFRF}$. Households thus don't pay their specific premium contribution, but only one of these four premium levels that divide the entire group premium pool by these fixed factors.

### G.3 Comparing premiums in the individual and small group markets

In Figure A1, we illustrate the differences in annual premiums comparing plans in the small group and individual insurance market by year. To do so, for each constructed plan available in the small group market, we identify an identical constructed plan in the individual market that shares the same carrier, plan type, metal tier, and geographic region. Our goal is to find the difference in premiums between small group and individual coverage for an example household that contains a 40-year-old single subscriber who is not eligible for federal premium subsidies in the individual market.

We carry out this analysis in three steps. In our first step, we compare the premiums without any tax adjustments or employer subsidies. To find our household's premium in the small group market, we apply the tiered-composite pricing system as if our household were in a small group with other households. We simulate the other members of the small group using the empirical distribution of household size and age composition that we observe in Oregon:

1. We loop through all observed small groups in Oregon

2. For each group we observe, we replace one random household in the group with our example 40-year old single adult.

3. We use the steps we describe in Appendix Section G.2 to compute our example household's required premium

4. We average these required premiums over the empirical distribution of small groups.

In our plot, we show the distribution of both the small group and individual market premiums that the household would face for all possible plans in a given year.

In our second step, we assume our single household has the median income for single adults in Oregon. We compute the average tax rate this household would face and adjust the small group premium to reflect the household's tax savings from using money for non-taxed insurance premiums rather than taxable income. There is no similar tax advantage in the individual market.

Finally, third, we add, in addition to the tax subsidy, a subsidy from the employer. We assume a 50% employer subsidy, which is the minimum employer's premium contribution that allows smaller businesses to earn the Small Business Health Care Tax Credit.

## G.4 Comparing premiums in the individual and small group markets outside Oregon

To compare premiums in the small group and individual insurance markets outside Oregon, we exploit the 'HIX Compare' data (hereafter 'HIX') compiled by the Robert Wood Johnson Foundation. The dataset provides plan-level detail on cost-sharing terms, premiums, and deductibles for nearly all plans offered in the health insurance marketplaces beginning in 2014. We divide our analyses on plans available in the individual insurance and small group insurance markets into two periods, 2014-2016 and 2017-2018.

In 2014, HIX data contains plans from 33 states with both small group and individual insurance plan data. In 2015-2018, all states and the District of Columbia appear with at least one plan in the silver and gold metal tiers in both the small group and individual markets.

In the HIX data, we save raw premiums for a 27-year-old single insurance enrollee for each plan offered in the individual and small group markets. We define a plan as a unique combination of year, state, rating area, carrier, metal tier, plan type, and individual deductible level. We compare the small group and individual markets by averaging the premiums we observe in each business segment over a specific subset of observations; we then match equivalent subsets between business lines.

We run our comparisons between the small group and individual markets using three subset definitions. First, in our most granular comparison, we compare the two markets at the plan level, matching plans by year, state, rating area, carrier, metal tier, plan type, and individual deductible amount. We match many plans across the two business lines. In 2015, for example, we find 2,411 bronze plans, 2,114 silver plans, and 1,378 gold plans that are offered in both the small group and individual market. At the year-state-metal level, we observe an average of 27 matched plans and a median of 16. In effect, under this definition, we must limit our analysis to carriers that offer nearly identical plan designs in both business lines.

Our second definition compares plans across the business lines in a state, year, and rating region, where the plans share only the same carrier and metal tier. That is, we find the mean at the carrier-metal level first in a geographic market and year and then compare these means across the business lines. Because we aggregate to the carrier-metal tier level, we can now use data in our calculations from carriers that offer slightly different types of plans within a metal tier across the business lines.[21] In a typical year-state-metal combination, we observe 29 matched plans on average, with a median of 18.

Finally, our third definition aggregates to a higher level before comparing across business lines. Specifically, we aggregate at the year, state, rating region, and metal tier level. For example, we aggregate all 52 individual market silver plans that carriers offered in rating area 1 in Alaska in 2014. We similarly aggregate the 8 small group silver plans carriers offered Alaska's region 1 in 2014. We then compare these means across business lines. In 2014, we observe 390 year-state-rating region combinations per metal tier. In 2015-2018, we observe roughly 500 such combinations.

Our results for years 2014-2016 appear in Table A10. Using our most granular definition (definition 1), reported in Panel A of Table A10, we see that in 2014-2015, across metal tiers, the small group premiums tend to be \$20-30 higher per month than the equivalent individual market plan. As a share of the small group premiums, these differences amount to an increase of roughly 7-12%. In 2016, the small group premiums again exceed the individual market premiums, but by a smaller amount: premiums are on the order of 1-3% higher.

In Panels B and C, we show premium differences using definitions 2 and 3 above, respectively. The results appear similar to Panel A, but slightly higher. Recall, in the later panels, we use more of the observed plan premiums because we do not require the carrier to be operating the same plan across both markets. Instead, we compare mean premiums at either the year-state-region-metal-carrier level or the year-state-region-metal level across the two business lines. In these comparisons, we find premiums are higher by approximately \$20-40 per month, amounting to roughly 10-15% higher premiums in the small group vs. individual insurance market for the same actuarial value plan in 2014-2016.[22]

We repeat this analysis in the HIX data for Oregon alone. HIX contains premium information for Oregon only in years 2015-2016. As in our main data, we again find that small group premiums exceed comparable premiums in the individual marketplaces. In Oregon, under the plan-level and carrier-metal definitions in the years we study, the gap between small group and individual insurance premiums is slightly larger than the national average.

Finally, as a robustness, we repeat our premium comparison for years 2017-2018. We omit an extensive discussion of these years in our main analysis both because we do not have claims data for these years and because there were two important regulatory changes in 2017. First, prior to 2017, the federal government paid compensation to insurers who enrolled consumers eligible for cost-sharing subsidies. Eligible consumers who purchase silver plans face reduced cost sharing (higher actuarial value plans) by statute, leading to added costs borne by the insurer; beginning in 2017, the federal government no longer reimbursed insurers for these added costs. Many insurers responded to this anticipated change by "silver loading" in 2017, meaning they raised the silver plan premiums to compensate for the added costs. Second, the temporary risk corridors program, administered under Section 1342 of the ACA, ended after the 2016 plan year. The program limited both the gains and losses to insurers offering plans in the early years of the individual marketplaces, with the goal of encouraging insurer participation.

[21] By requiring the plans to have the same metal designation, all the plans we aggregate together share the same actuarial value. However, the plans we aggregate may differ by PPO vs. HMO plan type, and may have a different mix of deductible and cost-sharing terms to achieve the same actuarial value.

[22] As a robustness, we also compute the minimum across plans that fit in a business line under definitions 2 and 3. We then compare the minimums across the small group and individual markets. That is, rather than average the premiums at the year and market level within a metal tier, we find the minimum across all plans or across all plans within a carrier. The qualitative findings are similar when using the minimum order statistic in place of the mean.

In part reflecting these regulatory changes, we find the unsubsidized individual market premiums exceed the comparable small group premiums in 2017-2018, across all plan tiers. Using the same definitions we apply to 2014-2016, we find higher baseline premiums on the order of 10-15% in the individual market. These differences appear both in Oregon and nationally.

# H Risk-related programs

We describe three risk-related programs in effect during the time period of our data. The first, called risk adjustment, begin in 2014 and continues to operate in both the small group and individual insurance market segments. The remaining two programs, risk corridors and reinsurance, were temporary and operated only from 2014 through 2016. Both programs were intended to stabilize premiums in the initial years of the ACA marketplaces; the risk corridors program collected penalties or made payments to insurers in both the individual and small group markets, while the reinsurance program made transfers only to individual insurance plans.[23]

## H.1 Risk adjustment

Under the risk adjustment design for the plan years 2014 through 2016, all non-grandfathered plans in the small group and individual insurance market were part of the risk adjustment program. Both plans offered on marketplaces and those offered outside the marketplaces ("off exchange") participated in risk adjustment. Apart from states that have mandated common premiums in the individual and small group markets—in our time period, Massachusetts, Vermont, and the District of Columbia—the risk adjustment system operates separately in each market segment in a state.

The payments are made or collected at the plan level; as reported by Cox et al. (2016), the risk adjustment methodology generates an average risk score at the plan level, which is a weighted average of all enrollees' individual risk scores. Plans with relatively low average risk scores pay into the system while those with relatively high average risk scores receive payments. In an HHS summary report for the 2015 benefit year (Centers for Medicare & Medicaid Services, 2016), the 2014 risk adjustment transfers averaged 6% of premiums in the small group market and 10% of premiums in the individual market. In this risk adjustment design, payments to plans net to zero against the payments collected from plans with lower average risk scores. Beginning in the 2017 and 2018 benefit years, after the span of our analysis data, the risk adjustment design also accounts for the potential added costs of partial-year enrollees and for prescription drug utilization.

## H.2 Reinsurance

Of the two temporary programs included in the ACA, reinsurance only applied to the individual insurance market. The goal of the program was to stabilize premiums: individual insurance plans with higher-cost enrollees would receive transfers, reducing the need for insurers to charge high

[23] An alternative to pooling the markets directly is to regulate the markets via premium-setting rules or by instituting risk-sharing programs. In the time period of our data, Massachusetts, Vermont, and the District of Columbia regulated premiums such that, for identical plans offered by the same carrier in the small group and individual markets, the premiums had to be equal. Alternatively, regulators could design risk-related programs, like risk adjustment or risk corridors, to pool patient costs across market segments for the purpose of assigning penalties or issuing payments. In the time period of our data, these programs operated separately by market segment. Moreover, even with pooling risks, premiums may remain distinct in the markets depending on the degree of competition and the split of risks across markets.

premiums out of fear that guaranteed issue policies would bring in high-cost enrollees. The reinsurance payments were set according to realized health costs rather than expected or predicted costs based on an individual's risk score. Further, the pool of funds used to make these transfers came from insurers offering fully insured plans in any segment, including the individual, small group, and large group markets.

Eligible plans received transfers if an enrollee's costs exceeded an "attachment point", which equaled $45,000 in 2014 and 2015 and $90,000 in 2016. The plan would receive a payment of 80% of the costs above the attachment point to a cap of $250,000. In 2014, the funds collected from insurers exceeded the claims for transfers, allowing 100% of the costs above the attachment point to be paid out to the insurers (Cox et al., 2016).

### H.3 Risk corridors

Finally, for 2014-2016, a third risk-related program, known as risk corridors, operated in both the individual and small group insurance markets. Like reinsurance, the program was intended to reduce the uncertainty that insurers faced when setting premiums in a new marketplace environment with guaranteed issue requirements. All qualifying health plans were judged against a target level of 80% of their premium dollars collected (coinciding with MLR regulations on allowable costs); if a plan's health care costs plus the costs of qualifying quality improvement programs fell below 97% or above 103% of the target threshold, the plan was subject to a penalty paid to the program or a transfer paid out to the plan, respectively.[24]

Under the original language of the ACA, the risk corridors program was not required to net to zero–that is, payments and penalties need not equal each other. However, subsequent Congressional bills required that the risk corridors program be paid out to insurers in a revenue neutral way (Cox et al., 2016). Because the payments owed far exceeded transfers collected under the program, actual payments made were prorated. Nationally, insurers with payments owed to them received 12.6% of the amount due to them in both market segments. In Oregon in 2014 in the individual insurance market, the risk corridors program collected $899,831 from insurers and owed $102.23 million. The program paid out $12.90 million (12.6% of the amount owed). In the small group market in Oregon in 2014, the risk corridors program collected $171,755 and owed $5.23 million. The program paid out $660,326.

The balance of pro-rated payments, under the statute, were supposed to be paid out with any excess funds collected in the program in the 2015 and 2016 plan years. However, by the end of 2016, the program owed $12 billion in payments to insurers under the program. Multiple insurers sued the federal government; the US Supreme Court ruled in April 2020 that the government did have an obligation to make risk corridors payments under Section 1342 of the ACA (Keith, 2020). That there was uncertainty in the amounts insurers might receive from the program could have affected the level of premiums set in the plan years observed in our analysis sample. We describe how this uncertainty might affect our specifications in Section 6.

[24]As described in Cox et al. (2016), a qualified health plan with costs of 92% – 97% of the target level paid into the corridors program an amount equal to 50% of the difference between its actual claims and 97% of its target amount. If the plan had observed claims less than 92% of the target, the plan paid 2.5% of the target amount plus 80% of the difference between their actual claims and 92% of its target. A payment to the plan followed a symmetric schedule for observed costs greater than 103% of the target level.

# I Appendix Tables and Figures

**Table A1**: Full sample of forced switchers

| | Total |
|---|---|
| Individual market | 5,321 |
| Coverage through spouse | 1,148 |
| Large group market | 687 |
| Other | 233 |
| Uninsured | 9,242 |
| Total | 16,631 |

Note: This table describes the destination of small group enrollees whose plans where closed in 2014 or 2015. “Other” includes coverage as a dependent or coverage through Medicare, another small group plan, an association plan, or other miscellaneous types of plan. “Uninsured” comprises enrollees who do not fall into any other category and are therefore untracked.

**Table A2**: Administrative costs

| **Variable** | **Mean** | **S.D.** | **Median** |
|---|---|---|---|
| Monthly admin costs in the individual market per enrollee ($) | | | |
| Total | 71.53 | 28.24 | 62.86 |
| Taxes and fees | 21.23 | 25.64 | 11.96 |
| Wellness activities | 4.14 | 2.99 | 3.64 |
| General administrative expenses | 46.17 | 16.22 | 42.85 |
| Number of annual subscribers in Oregon | | | |
| Individual market | 25,931 | 31,702 | 16,778 |
| Small group market | 21,883 | 13,197 | 19,812 |
| Large group market | 84,770 | 92,063 | 42,990 |
| Other markets | 40,772 | 68,557 | 6,543 |
| Number of states where the payer is active | 7.5 | 14.8 | 2 |

Note: This table displays measures of insurers' administrative costs. We define all cost measures from publicly available data collected by the Centers for Medicare and Medicaid Services (CMS). We collect this insurer-state level data for the individual market for the years 2014-2016. In the table, we report total per-enrollee monthly administrative costs and the constituent elements of administrative costs: (1) taxes and fees, which include federal and state taxes deductible from premiums, community benefit expenditures, contribution to the Federal Transition Reinsurance Program, and various other regulatory fees; (2) wellness activities, which include activities to improve and promote health outcomes and prevent hospital readmission as well as expenses related to health information technology; and, (3) general administrative expenses, which include direct sales salaries and benefits, agents and broker fees, and all other non-medical general and administrative expenses. In addition, in the table we report other measures available in the CMS data. We compute statistics on the count of annual subscribers in Oregon that each insurer covers in each business line. Finally, we present statistics on the number of states, including Oregon, where payers that operate in Oregon are active. We classify an insurer as active in a state when it files a report with CMS in that state and year.

**Table A3**: Demographics of the switchers

| | Small-Group | | Switchers | | Compliers | |
|---|---|---|---|---|---|---|
| **Demographic variable** | **Mean** | **S.D.** | **Mean** | **S.D.** | **Mean** | **S.D.** |
| Single-membered | 0.75 | | 0.73 | | 0.73 | |
| Married, no dependent | 0.08 | | 0.08 | | 0.08 | |
| Not married, with dependent(s) | 0.07 | | 0.09 | | 0.09 | |
| Married, with dependent(s) | 0.10 | | 0.10 | | 0.10 | |
| Number of dependents | 2.21 | 1.25 | 2.30 | 1.32 | 2.27 | 1.29 |
| HH health severity score | 1.25 | 2.20 | 1.33 | 2.56 | 1.31 | 2.59 |
| Income (share of FPL) | 2.38 | 0.29 | 2.39 | 0.28 | 2.39 | 0.28 |
| Age | 42.62 | 11.28 | 45.56 | 11.10 | 45.43 | 11.10 |
| Over 50 | 0.28 | | 0.38 | | 0.38 | |
| Living in rating areas 1, 2, or 3 | 0.78 | | 0.67 | | 0.67 | |
| Size of group | 13.85 | 11.51 | 7.17 | 8.47 | 7.20 | 8.49 |
| Chose gold or platinum plan [0/1] | 0.42 | | 0.44 | | 0.44 | |
| Chose silver plan [0/1] | 0.43 | | 0.35 | | 0.35 | |
| Number of subscriber-year observations | 383,036 | | 16,631 | | 14,563 | |

Note: This table compares the demographic characteristics of (a) the population of "forced switchers" who lose small group coverage to (b) the larger population of small group enrollees in years 2014 - 2016. The first column reports summary statistics for the full small group sample. The second column includes statistics for all households in the switcher sample in the year prior to their exit from small group coverage. We compute the household health severity score as the sum of health status scores for all members of a household, where we predict each member's score using the Johns Hopkins' ACG software. Number of dependents is calculated for the subset of households who have dependents. Rating areas 1-3 include the urban areas of Portland, Eugene, and Salem, respectively. Rating areas 4-7 include largely rural areas of the state.

**Table A4**: Demographics of the uninsured

| Demographic variable | Uninsured Households Mean | Uninsured Households S.D. | Insured Households Mean | Insured Households S.D. |
|---|---|---|---|---|
| Single-membered | 0.51 | | 0.70 | |
| Married, no dependent | 0.16 | | 0.15 | |
| Not married, with dependent(s) | 0.13 | | 0.06 | |
| Married, with dependent(s) | 0.20 | | 0.09 | |
| Number of dependents | 1.97 | 1.34 | 1.91 | 1.10 |
| HH health severity score | 1.90 | 3.34 | 1.44 | 2.49 |
| Income (share of FPL) | 2.37 | 0.29 | 2.46 | 0.29 |
| Age | 37.65 | 12.41 | 47.00 | 11.80 |
| Over 50 | 0.22 | | 0.45 | |
| Living in rating areas 1, 2, or 3 | 0.68 | | 0.69 | |
| Number of subscriber-year observations | 600,487 | | 444,255 | |

Note: This table compares the demographic characteristics of uninsured households and households who chose plans in the individual market in years 2014 - 2016. We compute the household severity score as the sum of household health status scores for all members of a household, where we predict each member's score using the Johns Hopkins' ACG software. Number of dependents is calculated with the subset of households who have dependents. Rating areas 1-3 include the urban areas of Portland, Eugene, and Salem, respectively. Rating areas 4-7 include largely rural areas of the state.

**Table A5**: Estimated model of small group fees

| | Specifications | | | | |
|---|---|---|---|---|---|
| | **(1)** | **(2)** | **(3)** | **(4)** | **(5)** |
| Constant | | | | | |
| Bronze | 1.449 | 2.042 | 2.069 | 1.930 | 1.906 |
| | (0.022) | (0.111) | (0.111) | (0.112) | (0.121) |
| Gold | 1.276 | 0.252 | 0.262 | 0.380 | 0.334 |
| | (0.022) | (0.104) | (0.104) | (0.105) | (0.114) |
| Platinum | 2.003 | 1.515 | 1.303 | 1.061 | 1.260 |
| | (0.029) | (0.130) | (0.133) | (0.134) | (0.143) |
| Age | | | | | |
| Bronze | | -0.015 | -0.015 | -0.015 | -0.016 |
| | | (0.003) | (0.003) | (0.003) | (0.003) |
| Gold | | 0.027 | 0.029 | 0.030 | 0.030 |
| | | (0.003) | (0.003) | (0.003) | (0.003) |
| Platinum | | 0.013 | 0.025 | 0.024 | 0.025 |
| | | (0.003) | (0.004) | (0.004) | (0.004) |
| HH health severity score | | | | | |
| Bronze | | | -0.335 | -0.338 | -0.343 |
| | | | (0.026) | (0.026) | (0.027) |
| Gold | | | -0.380 | -0.385 | -0.395 |
| | | | (0.023) | (0.023) | (0.025) |
| Platinum | | | -0.690 | -0.705 | -0.656 |
| | | | (0.030) | (0.030) | (0.031) |
| Group size | | | | | |
| Bronze | | | | 0.008 | 0.008 |
| | | | | (0.001) | (0.001) |
| Gold | | | | -0.005 | -0.005 |
| | | | | (0.000) | (0.000) |
| Platinum | | | | 0.014 | 0.015 |
| | | | | (0.001) | (0.001) |
| Family size | | | | | |
| Bronze | | | | | 0.025 |
| | | | | | (0.049) |
| Gold | | | | | 0.053 |
| | | | | | (0.050) |
| Platinum | | | | | -0.228 |
| | | | | | (0.053) |
| Number of observations | 46,310 | 46,310 | 46,310 | 46,310 | 46,310 |

Note: This table reports the maximum likelihood estimates from the employer preferences specification in Equation D.13. In this analysis, we omit households who choose grandfathered or other plans for which we do not observe premiums or metal tier. All variables are interactions between metal tier indicators (silver-tier excluded) and household characteristics. The household characteristics that we use include: the primary subscriber's age; the household heath severity score, which sums the health scores of all individuals in the household; the number of other households in the focal household's small group contract; and the number of individuals in the household.

**Table A6**: Estimated fees paid by small group households

| | Specifications | | | | |
|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) |
| *Overall* | 71.24 | 72.90 | 50.43 | 48.15 | 47.88 |
| | (0.77) | (0.88) | (1.28) | (1.28) | (1.40) |
| *Household type* | | | | | |
| No dependent, under 50 | 70.59 | 69.92 | 57.91 | 55.81 | 56.05 |
| | (0.77) | (0.78) | (0.92) | (0.92) | (0.94) |
| With dependent(s), under 50 | 78.81 | 80.45 | 27.54 | 23.93 | 20.22 |
| | (0.85) | (0.89) | (2.30) | (2.31) | (5.34) |
| No dependent, over 50 | 67.69 | 74.62 | 46.56 | 44.76 | 45.27 |
| | (0.74) | (1.94) | (2.27) | (2.27) | (2.29) |
| With dependent(s), over 50 | 78.85 | 88.79 | 21.83 | 17.94 | 15.74 |
| | (0.86) | (2.09) | (3.50) | (3.51) | (5.52) |
| *Household plan choice* | | | | | |
| Bronze | 144.88 | 137.75 | 112.23 | 108.89 | 109.40 |
| | (2.20) | (2.57) | (3.14) | (3.15) | (3.40) |
| Silver | — | — | — | — | — |
| | ( — ) | ( — ) | ( — ) | ( — ) | ( — ) |
| Gold/Platinum | 154.11 | 163.08 | 104.14 | 98.55 | 97.42 |
| | (1.91) | (2.15) | (3.26) | (3.27) | (3.53) |

Note: This table reports average estimated "fees" paid by households in the small group market for their chosen plans. The columns in this table correspond with the model specifications described in Table A5. In this analysis, we omit households who choose grandfathered or other plans for which we do not observe premiums or metal tier. The sample used includes 46,310 households. The estimated fees decrease on average as we allow for more heterogeneity in fees by household type in the model. The fees are identified by the difference between observed plan choices and predicted plan choices using the preferences we estimate from the population of forced switchers.

**Table A7**: Variation within constructed plans

| Panel A: Individual Market (N = 1196) | | | | | | |
|---|---|---|---|---|---|---|
| | | Percentile | | | | |
| | Min | 25th | 50th | 75th | Max | Mean |
| *Number of* | | | | | | |
| Plans | 1.0 | 2.0 | 3.0 | 3.0 | 15.0 | 2.9 |
| Networks | 1.0 | 1.0 | 1.0 | 2.0 | 4.0 | 1.3 |
| Service areas | 1.0 | 1.0 | 1.0 | 2.0 | 4.0 | 1.4 |
| Plan types | 1.0 | 1.0 | 1.0 | 1.0 | 2.0 | 1.0 |
| National network types | 1.0 | 1.0 | 1.0 | 1.0 | 2.0 | 1.2 |

| Panel B: Small Group Market (N = 1772) | | | | | | |
|---|---|---|---|---|---|---|
| | | Percentile | | | | |
| | Min | 25th | 50th | 75th | Max | Mean |
| *Number of* | | | | | | |
| Plans | 1.0 | 3.0 | 5.0 | 10.0 | 96.0 | 8.6 |
| Networks | 1.0 | 1.0 | 1.0 | 2.0 | 4.0 | 1.4 |
| Service areas | 1.0 | 1.0 | 1.0 | 2.0 | 4.0 | 1.4 |
| Plan types | 1.0 | 1.0 | 1.0 | 1.0 | 2.0 | 1.1 |
| National network types | 1.0 | 1.0 | 1.0 | 1.0 | 2.0 | 1.1 |

Note: In order to simplify our analysis, we bin similar plans together to create constructed plans. A constructed plan consists of all plans with the same insurer, rating area, metal tier, on/off exchange status, and managed care status. This table uses data on 2014-2016 insurance plans from the National Association of Insurance Commissioners' SERFF database to describe the variation of plan characteristics within a constructed plan. Plan type describes whether a plan is an HMO, EPO, POS, or PPO plan. National network type is an indicator for whether a plan has a national network.

**Table A8**: Change in consumer surplus by household type

| | Markup = 0% | | Markup = 25% | |
|---|---|---|---|---|
| | Individual | Small-Group | Individual | Small-Group |
| *Overall* | 44 | 57 | 18 | 28 |
| *Household income* | | | | |
| 1st quartile | 21 | 60 | 7 | 39 |
| 2nd quartile | 27 | 57 | 10 | 30 |
| 3rd quartile | 52 | 32 | 22 | -14 |
| 4th quartile | 72 | 26 | 31 | -40 |
| *HH health status score* | | | | |
| 1st quartile | 32 | 67 | 12 | 38 |
| 2nd quartile | 33 | 52 | 12 | 20 |
| 3rd quartile | 56 | 21 | 24 | -20 |
| 4th quartile | 84 | 8 | 37 | -34 |

Note: This table shows the effects of merging the individual and the small group markets in 2016 under the mandated insurance counterfactual assumption for different types of households. All reported numbers are the difference between consumer surplus in the mandated insurance merged market counterfactual and the base counterfactual. Each row corresponds to a set of households. Each number is an average over households in the corresponding set of households. These averages are reported in $s at the monthly level. Quartiles are determined using all households in both the individual and small group markets. To account for outliers, consumer surplus is reported for the set of households where $\alpha - \psi > 0.05$.

**Table A9**: Comparing out-of-sample model predictions to observed data

| | Estimated | Observed |
|---|---|---|
| *Choice shares* | | |
| Uninsured | 56 | 49 |
| Bronze | 16 | 16 |
| Silver | 21 | 27 |
| Gold | 7 | 8 |
| *Average cost* | | |
| Full sample | 242 | 207 |
| Uninsured HHs | 0 | 0 |
| Bronze HHs | 282 | 222 |
| Silver HHs | 496 | 413 |
| Gold HHs | 771 | 740 |

Note: This table assesses model fit by comparing out-of-sample estimates of choice shares and average costs to the corresponding observed objects. To construct the estimates in column 1, we estimate the preferred plan choice with a randomly-drawn 80% sample of individual market households in the 2014-2016 data. Using the obtained demand estimates and the supply estimates described in Table 6, we apply the base counterfactual algorithm to the entire population of individual market households in the 2016 data. We then determine the choice shares and average costs of the 20% out-of-sample households in the counterfactual results. Column 2 presents summary statistics of observed choices and costs related to the 20% out-of-sample households. In both columns, costs include only costs borne by insurers and exclude patient out-of-pocket costs. The costs reported are censored below at $20 and winsorized above at the 99th percentile of individual market cost. Estimated household costs are calculated as an average of 100 draws from the household's cost distribution.

**Table A10**: National comparison of small group and individual market premiums, 2014-2016

Panel A: Plan-level definition

| | | | SG-IM ($) | | % Diff |
|---|---|---|---|---|---|
| Year | Metal | Count | Mean | Std | Mean |
| 2014 | Bronze | 451 | 21.97 | 21.48 | 11.70 |
| 2014 | Gold | 319 | 21.32 | 33.40 | 7.23 |
| 2014 | Silver | 363 | 20.83 | 25.31 | 8.45 |
| 2015 | Bronze | 2411 | 22.24 | 30.71 | 10.80 |
| 2015 | Gold | 1378 | 25.84 | 44.24 | 8.54 |
| 2015 | Silver | 2114 | 31.23 | 38.79 | 12.22 |
| 2016 | Bronze | 1573 | 2.95 | 36.24 | 2.12 |
| 2016 | Gold | 792 | 1.57 | 51.48 | 1.38 |
| 2016 | Silver | 1400 | 4.70 | 42.46 | 2.85 |

Panel B: Carrier-metal definition

| | | | SG-IM ($) | | % Diff |
|---|---|---|---|---|---|
| Year | Metal | Count | Mean | Std | Mean |
| 2014 | Bronze | 559 | 13.76 | 88.57 | 11.28 |
| 2014 | Gold | 741 | 21.24 | 90.37 | 9.76 |
| 2014 | Silver | 741 | 21.32 | 87.87 | 11.67 |
| 2015 | Bronze | 2432 | 32.21 | 34.53 | 15.06 |
| 2015 | Gold | 2244 | 47.27 | 57.51 | 16.24 |
| 2015 | Silver | 2417 | 44.74 | 48.58 | 17.25 |
| 2016 | Bronze | 1932 | 10.99 | 37.82 | 5.46 |
| 2016 | Gold | 1831 | 16.14 | 56.19 | 5.81 |
| 2016 | Silver | 1885 | 18.24 | 45.76 | 7.05 |

Panel B: Rating area-level definition

| | | | SG-IM ($) | | % Diff |
|---|---|---|---|---|---|
| Year | Metal | Count | Mean | Std | Mean |
| 2014 | Bronze | 322 | 12.11 | 75.03 | 9.99 |
| 2014 | Gold | 390 | 25.44 | 71.23 | 10.56 |
| 2014 | Silver | 390 | 25.83 | 58.76 | 12.11 |
| 2015 | Bronze | 498 | 30.85 | 24.14 | 13.06 |
| 2015 | Gold | 499 | 39.52 | 36.40 | 12.34 |
| 2015 | Silver | 499 | 48.77 | 30.19 | 17.58 |
| 2016 | Bronze | 498 | 6.26 | 33.67 | 3.25 |
| 2016 | Gold | 499 | 27.29 | 50.10 | 8.68 |
| 2016 | Silver | 499 | 32.32 | 40.43 | 11.74 |

Note: This table reports the difference in premiums between the small group and individual insurance markets from HIX Compare data. We collect baseline premiums for a 27-year-old single enrollee in a plan for 33 US states in 2014 and all US States and DC in 2015-2016; premiums for the plan are specific to the year, state, rating region, carrier, metal tier, plan type, and deductible level. Panel A reports statistics for a plan-level comparison. Panel B reports statistics after aggregating plans to the carrier-metal-market. Panel C reports statistics after aggregating plans to the metal-market in each business segment.

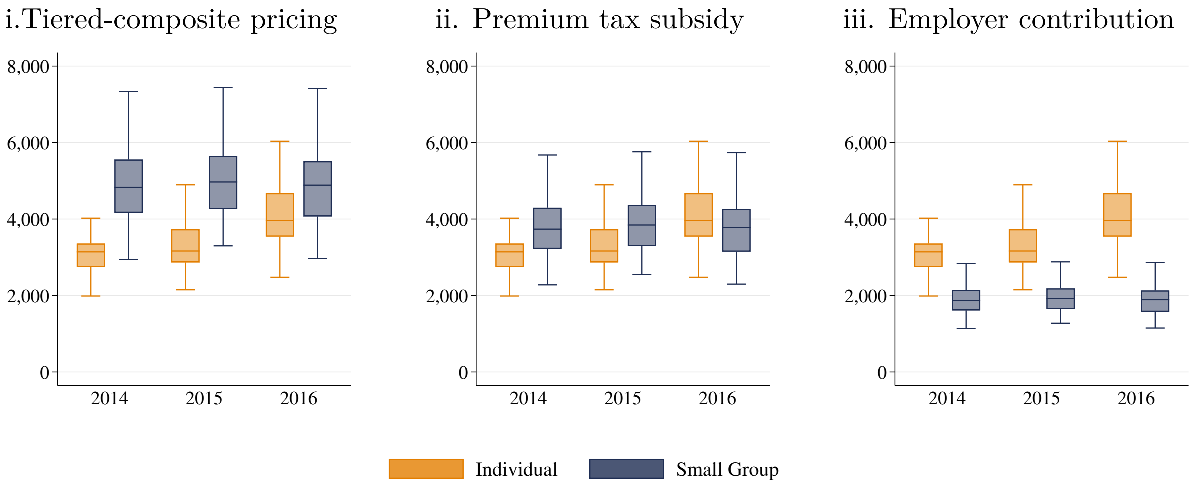


**Figure A1**: Comparison of subsidy schemes

Note: This figure compares premiums in the individual and small group market after accounting for different subsidies households may receive in the small group market. For each constructed plan available in the small group market, we identify an identical constructed plan in the individual market. We then simulate for each plan the average premium that a 40-year-old single subscriber would pay in the two markets under different subsidy schemes. The plot on the left assumes that the only difference in pricing is in the tiered-composite pricing in the small group market. The plot in the middle combines tiered-composite pricing with the implicit premium tax subsidy in the small-group market. The plot on the right combines tiered-composite pricing, the premium tax subsidy, as well as the employer's premium contribution. We set the contribution at 50% for illustration.

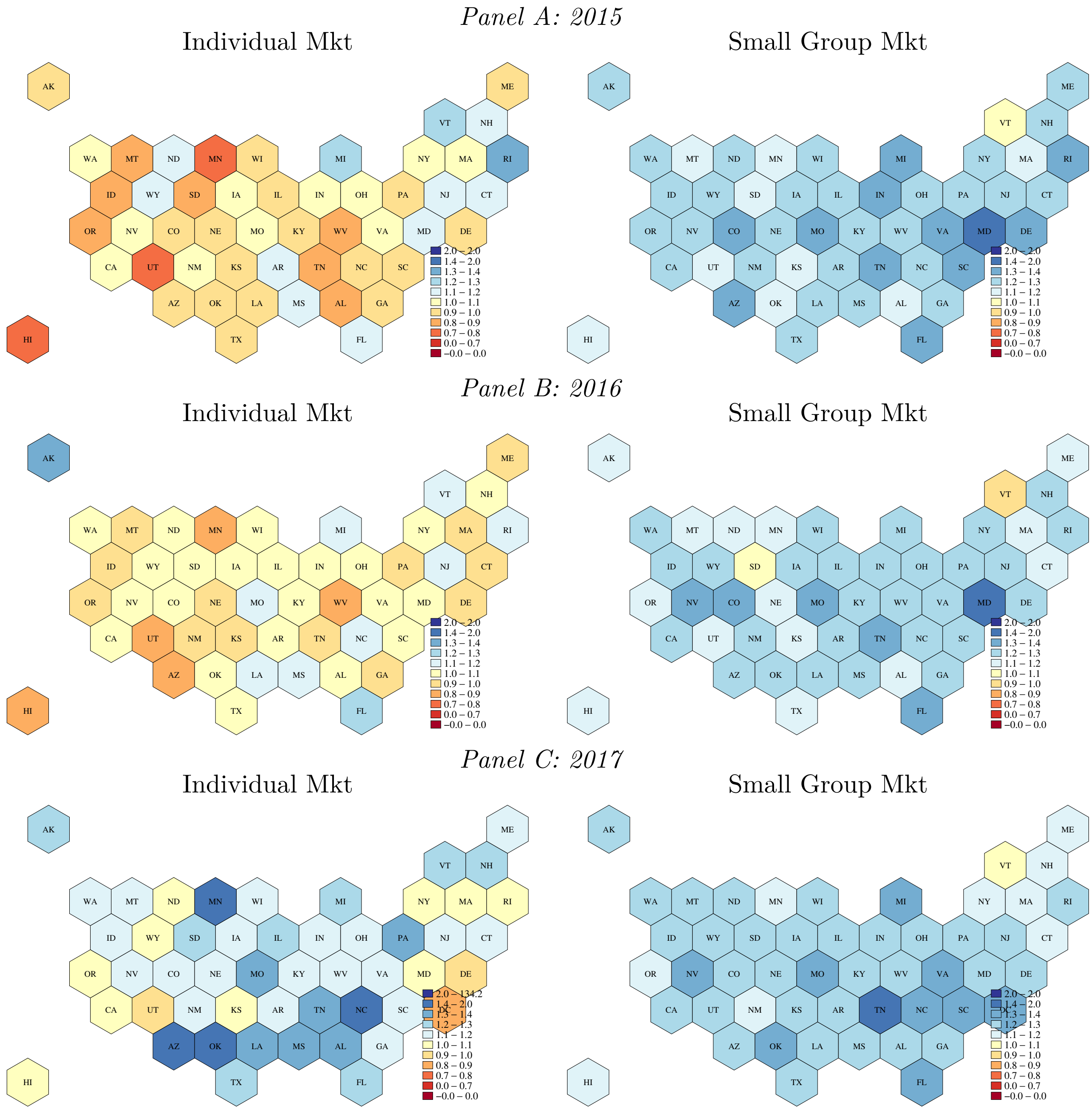


**Figure A2**: Average medical markups by state, year, and insurance segment

Note: This figure depicts the medical markup by state, year, and insurance market segment computed using Medical Loss Ratio data. The medical markup is calculated as the ratio of premium revenue divided by the total medical costs insurers incur. The cost in this ratio does not account for risk adjustment payments or other transfers to the insurer.

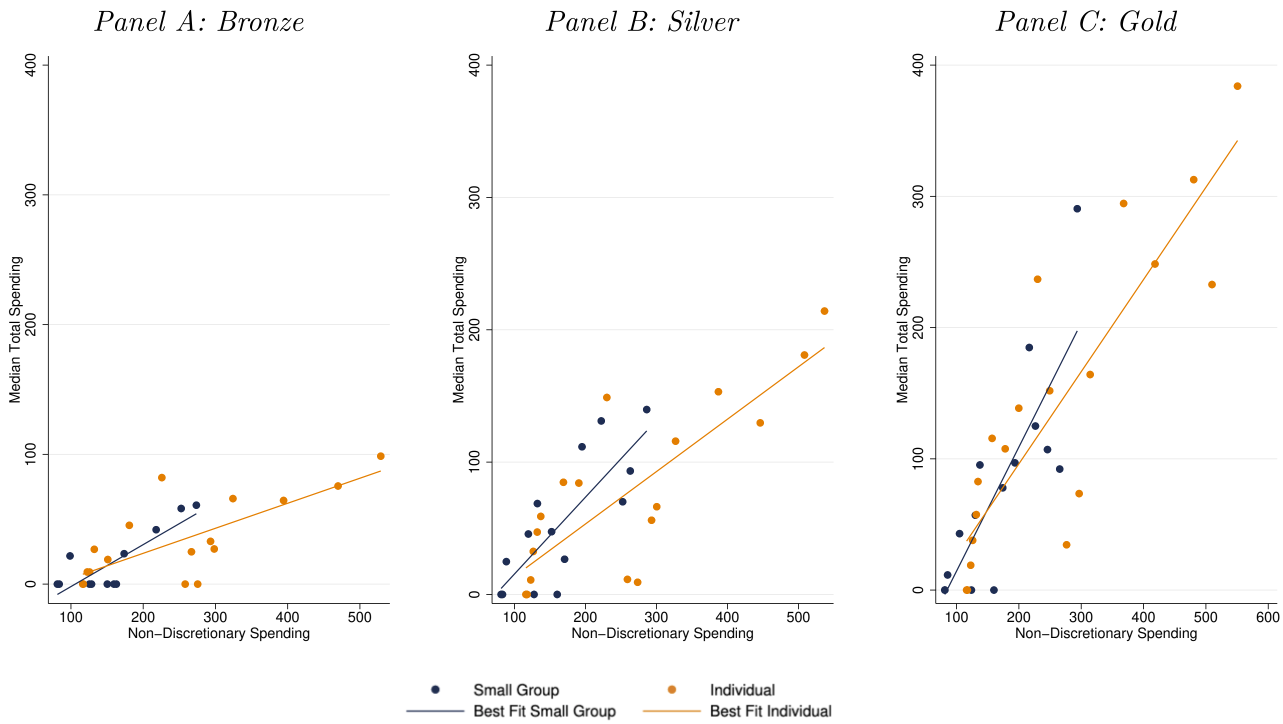


**Figure A3**: Identification of demand parameters: moral hazard

Note: Panels A through C compare total spending in the small group switcher and individual market populations for various expected non-discretionary spending levels, $\lambda$. The figure reports the relationship between $\lambda$ and total observed spending, distinguished by panel according to the actuarial value of the household's plan choice. All spending is reported in $s per month. We exclude outliers by focusing on households that satisfy $\alpha - \psi > .05$. Then, households within each market are divided into 20 buckets based on $\lambda$ levels; each point in the figure represents the median $\lambda$ and spending in a particular bucket.

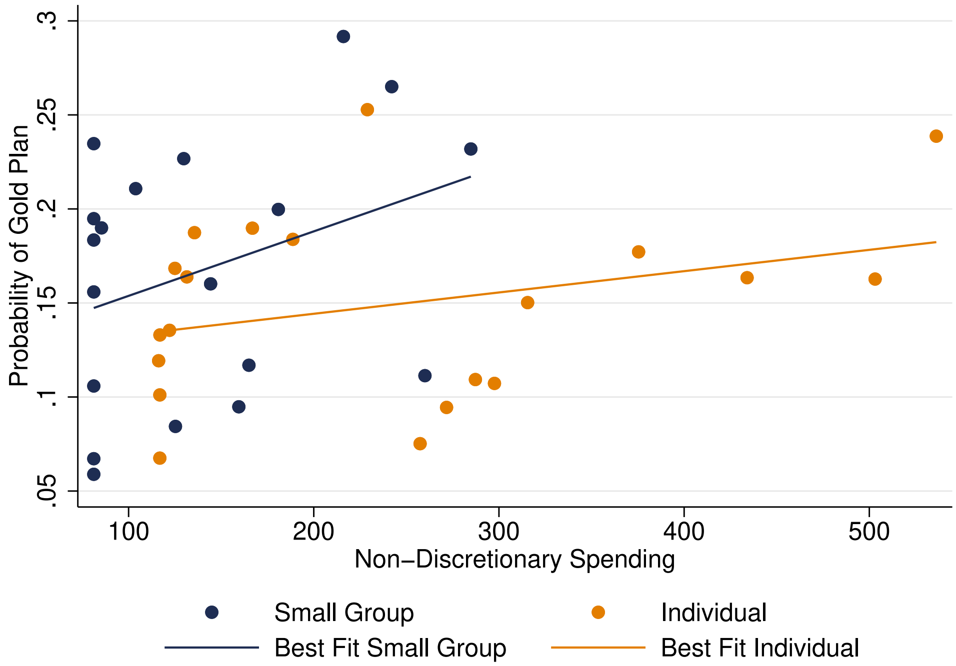


**Figure A4**: Identification of demand parameters: risk aversion

Note: This figure reports the probability of gold plans being chosen in the small group switcher and individual market populations for various expected non-discretionary spending levels, $\lambda$. Spending is reported in \$s per month. We exclude outliers by focusing on households that satisfy $\alpha - \psi > .05$. Then, households within each market are divided into 20 buckets based on $\lambda$ levels; each point in the figure represents the median $\lambda$ and spending in a particular bucket.