Firm Learning in a Selection Market

Claudio Lucarelli[1] and Evan Saltzman*[2]

[1]University of Pennsylvania
[2]Emory University

August 1, 2023

Abstract

Creating new markets is a prevalent approach for implementing large social programs. Assuming firms have full information about the relevant parameters upon market inception is commonplace in the literature. In contrast, we develop an adaptive learning model with selection to study how firms' knowledge of demand and cost affects the market equilibrium. We estimate alternative learning models with data from the California ACA exchange and assess their external validity using novel data on firms' predicted costs from insurer rate filings. The learning models provide statistically significant improvements in fit relative to the standard model that assumes firms have full information. Most of the improvement results from allowing firms to learn about the relationship between demand and cost. Firms with full information can increase profit, but at taxpayers' expense. Regulation that prohibits firms from using certain consumer information to set premiums makes them react more to the information they can use.

Keywords: Adaptive learning, adverse selection, health insurance.
JEL Codes: I11, I13, L51, L88, H51

*We thank Ulrich Doraszelski, Ben Handel, Neale Mahoney, Ian McCarthy, Francesca Molinari, Anh Nguyen, Mark Pauly, Dan Polsky, Juan Rubio-Ramirez, Ashley Swanson, Pietro Tebaldi, and Bob Town for advice and helpful feedback. We also thank seminar participants at the University of Georgia, Cornell University, the Annual Health Econometrics Workshop, ASHE Annual Conference, and Caribbean Health Economics Symposium. Yisroel Cahn and Sam Wunderly provided excellent research assistance.

1 Introduction

Large-scale social programs are increasingly being delivered through the private sector. Prominent examples occur in education, child care, and health insurance. An increasing share of U.S. citizens obtain health insurance through publicly-supported private health insurance markets, a trend expected to continue for the foreseeable future (Gruber, 2017). As of 2020, there are 53.9 million enrollees in the Medicaid Managed Care program for the low-income population, 24.1 million enrollees in the Medicare Advantage program for the elderly, 46.5 million enrollees in the Medicare Part D prescription drug program, and 11.4 million enrollees in the Affordable Care Act (ACA) exchanges (Kaiser Family Foundation, 2020).

Implementation of these programs often involves establishing new markets. Firms participating in these new markets may initially have little knowledge of consumer preferences and their competitors' strategic behavior. In markets with selection, firms face the additional challenges of forecasting cost and understanding how cost is correlated with demand (Einav et al., 2010). However, analyses of these new markets in the empirical industrial organization (IO) literature and by policymakers largely assume firms have full information and the market is in equilibrium upon market inception (Doraszelski et al., 2018). This assumption is a significant shortcoming of the literature given the increasing prevalence of social programs delivered through new private markets.

In this paper, we study how firms' knowledge about their demand and cost affects the market equilibrium and consider potential implications for market design. The impact of firm information is theoretically ambiguous and depends on how well firms forecast key market features such as consumer price sensitivity, plan switching costs, and the distribution of risk across consumer types. For example, underestimating consumer price sensitivity might lead firms to overestimate their market power and set premiums above the true profit-maximizing levels. Conversely, firms that underestimate the risk of their projected enrollees could set premiums below the true profit-maximizing levels. These different perceptions about the true demand and cost parameters, and how they correlate, could generate very different predictions about the performance of these markets and the distribution of welfare.

Because of these theoretical ambiguities, we construct an empirical framework that incorporates these potential gaps in firm knowledge and allows us to distinguish between them. Similar to Doraszelski et al. (2018), we estimate an adaptive learning model that allows firms to progressively learn in a new market. We apply our framework to the state-based health insurance exchanges created in 2014 under the ACA, where eligible consumers can receive government subsidies for purchasing private plans. The ACA setting has two important features that make it particularly appealing for studying firm learning. First, firms faced considerable uncertainty in predicting consumer preferences for health insurance and the cost of insuring their enrollees. Potential enrollees came from two very distinct sources: those with coverage in the pre-ACA individual market (i.e., the market where consumers buy insurance directly from an insurer) and those without insurance (Gruber, 2017). The ACA's modified community rating regulations that prohibit the use of health status and other consumer characteristics to set premiums created additional sources of uncertainty (Pauly et al., 2015, 2020). Second, there are rich data on consumer plan choices, firm costs, and firms' own predictions about cost from the establishment of the exchanges in 2014. We obtain consumer-level administrative data on consumer plan choices from the California ACA exchange. Our California data account for approximately 13% of nationwide enrollment in the ACA exchanges (Kaiser Family Foundation, 2020) and contain nearly 10 million consumer plan choices between 2014 and 2019. We use data on firms' predictions about their costs from insurer rate filings to assess the external validity of our adaptive learning model compared to the full information approach. The rate review process requires firms to provide actuarial justification for their proposed premiums, including a detailed explanation of their cost forecast. The availability of credible data on firms' own cost predictions is a particularly novel feature of our setting.

We make four primary contributions to the literature: (1) we extend the empirical IO literature on firm learning to a selection market, where firms need to learn about the correlation of demand and cost; (2) we evaluate the external validity of alternative models and find the adaptive learning model fits our data better than the standard IO model that assumes firms have full information; (3) we find firms with full information can increase profits at the expense of taxpayers; and (4) we show community rating regulation that prohibits firms from using certain consumer information to price makes them react more to the information they do have available, exacerbating the impact of uncertainty in the market's initial years.

Our paper contributes to the empirical IO literature on firm learning in oligopoly markets (see Aguirregabiria and Jeon (2020) for a recent review of this literature), which mostly focuses on how consumers learn about their demand (Ackerberg, 2003; Dickstein, 2018) or how firms learn about their cost (Benkard, 2000; Zhang, 2010; Conley and Udry, 2010; Newberry, 2016). Jeon (2020) studies how firms in the container shipping industry learn about their demand. Several papers study whether the market converges to an equilibrium. Joskow et al. (1998) study the market for sulfur dioxide emissions following passage of 1990 Clean Air Act and find the market had become reasonably efficient by mid-1994. Hortaçsu and Puller (2008) analyze the bidding behavior of firms in the Texas electricity spot market from 2001 to 2003, finding that large firms made bids that were close to optimal. Hortaçsu et al. (2019) extend this work by examining the impact of large firms' superior strategic ability on market efficiency. Huang et al. (2021) study how firms learn about consumer demand in the Washington state liquor market following deregulation in 2012, finding prices converge to levels consistent with profit maximization. Doraszelski et al. (2018) use adaptive learning and fictitious play models to study how firms learn about their demand and competitors' behavior in the U.K. electricity market. They find that it takes several years before firms' behavior is consistent with a complete information Nash equilibrium and convergence to equilibrium is better described with learning models than standard IO models. We extend this literature by applying adaptive learning to a selection market where firms not only need to learn about their demand and cost, but also how demand and cost are correlated. To the best of our knowledge, our paper is the first to empirically study firm learning and the convergence to equilibrium in a selection market, where uncertainty is particularly acute. In our model, firms use only the *available* information on demand and cost to form expectations about the future.[1] Our framework explicitly accounts for adverse selection and moral hazard, as well as firm market power and consumer choice frictions. We also explicitly incorporate the key ACA policies that are essential for computing equilibria in this market, including the individual mandate, endogenous premium subsidies, community rating regulation, reinsurance, and risk adjustment.

We estimate alternative adaptive learning model specifications with our data on consumer choices, plan risk, and *realized* costs. For both the full information and alternative adaptive learning models, we form predictions of plan costs and compare them to the firms' predictions of plan costs as reported in their rate filings. Our validation approach uses the firms' predictions of plan costs only to validate each model and, importantly, does not use the firms' predictions in estimation. Hence, a novel feature of our work is that we can assess the external validity of each model specification, extending the work of Doraszelski et al. (2018). We find the adaptive learning model that allows firms to learn about all model parameters yields a statistically significant superior fit of our data compared to the standard IO model that assumes firms have full information. Models that allow firms to learn about the relationship between demand and cost, but assume firms know the other model parameters, also yield statistically significant improvements in fit. Conversely, models that assume firms know the relationship between demand and cost, but allow firms to learn the other model parameters, fare no better than the standard model. This result suggests it is particularly important to allow firms to learn about the relationship between demand and cost. We also find the benefits of using an adaptive learning approach are (1) reduced over time, and (2) largest for predicting the cost of the most generous plans.

[1] The field of macroeconomics has a long history of including adaptive learning in dynamic general equilibrium models (Sargent, 1993; Evans and Honkapohja, 2001)

We use our estimated learning model to simulate the impact of firm information on the market equilibrium. Assuming full information is commonplace in the previous literature evaluating the design of government-created health insurance markets, including the ACA exchanges (Tebaldi, 2022; Saltzman, 2021; Polyakova and Ryan, 2021; Einav et al., 2019), Medicare Advantage (Town and Liu, 2003; Lustig, 2009; Curto et al., 2020; Miller et al., 2019), Medicare Part D (Abaluck and Gruber, 2011, 2016; Ketcham et al., 2015; Decarolis et al., 2020; Fleitas, 2017; Lucarelli et al., 2012), Medigap (Starc, 2014), and the pre-ACA Massachusetts exchange (Ericson and Starc, 2015; Geruso et al., 2019; Hackmann et al., 2015; Finkelstein et al., 2019; Jaffe and Shepard, 2020). Our simulations indicate that firms set premiums below profit-maximizing levels on average. When firms have full information, average premiums are 6.9% higher in 2016, 3.4% higher in 2017, and only 0.1% higher in 2018 compared to the observed equilibrium. Average premiums are higher because firms overestimated premium sensitivity, underestimated switching costs, and underestimated claims risk (particularly for adults under age 45). With full information, firms increase annual per-capita profit by $347 in 2016, $209 in 2017, and $1 in 2018. Higher profits come primarily at taxpayers' expense because ACA subsidies shield consumers from higher premiums; annual per-capita net government spending increases $274 in 2016 and $147 in 2017, and decreases $37 in 2018.

The final part of the paper considers the interaction of learning with community rating regulation that limits the information firms can use to price discriminate. Relative to the baseline setting where price discrimination is partially restricted, a complete prohibition on price discrimination (i.e., pure community rating) increases average claims by nearly 10% in 2016, resulting in higher premiums and lower enrollment. Premium reductions realized by the winners (older adults) are smaller than the premium increases realized by the losers (young adults). The impact of information is generally largest in the pure community rating setting. Learning the full information model parameters increases total exchange enrollment by 1.2% with pure community rating, but has a negligible impact on enrollment in the baseline setting. Hence, prohibiting firms from using consumer information to set premiums makes them react more to the information they can use.

Our results have a number of important policy implications. Modeling firm learning may improve forecasts of proposed social programs, such as those conducted regularly by the Congressional Budget Office (CBO). In a retrospective assessment of its ACA forecast, the CBO concluded that it had substantially overestimated exchange premiums and federal spending on premium subsidies (Congressional Budget Office, 2017). Given our findings on the equilibrium impact of firm information, the CBO's forecast would have projected lower (and more accurate) premiums and spending on subsidies if it had accounted for firm uncertainty. Our study also has implications for which policies to adopt in new social programs. We find creating a new market with community rating exacerbates the effects of firm uncertainty. Instead of implementing community rating in a new market, policymakers could rely on expanded premium subsidies or reinsurance to protect high-risk consumers from price discrimination.[2]

The remainder of this paper is organized as follows. Section 2 describes the data and setting. Section 3 develops an adaptive learning model. Section 4 discusses estimation. Section 5 presents the model estimates and validation results. Section 6 uses the model to simulate the impact of learning. Section 7 uses the model to simulate policy counterfactuals. Section 8 concludes.

2 Data and Policy Background

The Affordable Care Act (ACA) seeks to expand health care access coverage by providing subsidized access to health insurance. A key mechanism for accomplishing this objective was the establishment of state-based health insurance exchanges in 2014. Eligible exchange consumers can receive subsidies to purchase health insurance from private insurance firms. Firms must comply with numerous regulations, including limitations on price discrimination. To study these exchanges, we use two primary sets of data: (1) 2014-2018 plan-market-level data on firm costs and predictions about cost from insurer rate filings and (2) 2014-2019 consumer-level data on enrollee choices from the California ACA exchange. We describe these data sources in the following two subsections.

2.1 Data on Firm Costs and Predicted Cost

We obtain data on firm costs and predictions about cost from insurer rate filings. All participating California exchange insurers must submit their proposed premiums for actuarial review at the Department of Managed Health Care (DMHC). Insurers are required to include detailed supporting data justifying premium increases, including past medical claims and expected trends. DMHC does not have the authority to reject premium increases, but can find the insurer's rate filing "unreasonable" if the supporting data do not support the rate increase and the insurer refuses to adjust their rates accordingly. Insurers must notify enrollees of an unreasonable finding. As part of the rate filing, insurers must include an independent actuarial certification which confirms its actuarial

[2] The ACA included a temporary reinsurance program that helped insurers offset the cost of covering consumers with unexpectedly high claims during the first three years of the exchanges. Since the program's expiration, fifteen states have reinstated reinsurance programs to reduce uncertainty and constrain premium growth, including Alaska, Colorado, Delaware, Georgia, Maine, Maryland, Minnesota, Montana, New Hampshire, New Jersey, North Dakota, Oregon, Pennsylvania, Rhode Island, and Wisconsin.

methodologies were audited by an independent firm. Because rate filings are subject to extensive scrutiny by both DHMC and independent auditors, we assume insurers truthfully report their projected costs and cannot strategically misreport in order to gain a competitive advantage.

The DMHC rate review process usually begins the summer before the new plan year when the proposed premiums take effect and can last several months. Firms submit their premiums for plan year t in the summer of year $t-1$ using experience data (i.e., supporting data) from plan year $t-2$. For example, rate filings for 2016 are submitted in the summer of 2015 and report experience from 2014, the most recent complete year of experience. The new premiums for 2016 take effect on January 1, 2016. Firms cannot adjust premiums in the middle of the plan year. Similarly, consumers can only switch exchange plans once a year during a period called "open enrollment."

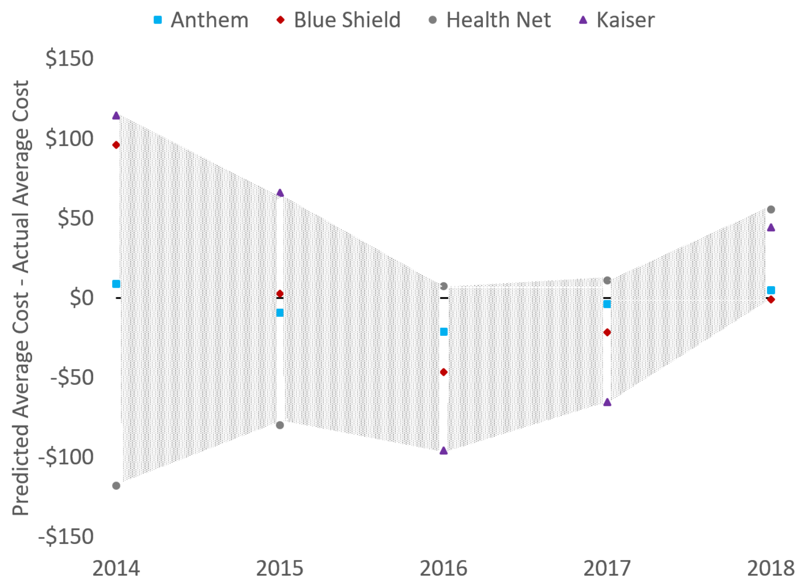
Insurers did not have any experience data from the exchanges in 2014 to make projections. Most insurers developed their 2014 premium rates using experience from other lines of business. Among the four dominant, statewide insurers in the California exchange, two insurers (Anthem and Health Net) used experience data from the small group market and the other two insurers (Blue Shield and Kaiser) used experience data from the pre-ACA individual market (Department of Managed Health Care, 2016). Although these were useful starting points, a substantial portion of the potential exchange population consisted of consumers who were uninsured. Insurers had to estimate both the size and health status of the uninsured population that would enroll. As part of its rate filing, Blue Shield indicated that it used the U.S. Census Bureau's Current Population Survey (CPS) to estimate the size of the uninsured population by age, income, and geography. Blue Shield estimated the uninsured population's take-up of insurance by calibrating premium sensitivity factors with its experience data for each age-income group. The firm assumed for each age group that the health status distribution of the uninsured population was the same as the health status distribution in its experience data.

The insurer rate filings provide key plan-market-level financial information, including data on enrollee medical claims and two important ACA risk mitigation programs – reinsurance and risk adjustment. Reinsurance was a temporary ACA program in effect from 2014 until 2016 that provided "insurance to insurers" for any enrollees with very high medical claims. The federal government served as the reinsurer and funded the program through a tax on all private insurance plans, including employer-sponsored plans. Risk adjustment is a permanent program where plans with lower-than-average risk make transfer payments to plans with higher-than-average risk. ACA risk adjustment transfers sum to zero, whereas the reinsurance program provides an inflow of funds to the ACA exchanges. Other risk adjustment programs, such as the one used in Medicare Advantage, may benchmark risk adjustment payments to the risk of those choosing the outside option (e.g., traditional Medicare) and also provide an inflow of funds to the market. The objective of risk adjustment is to disincentivize firms from cherry-picking the lowest-risk consumers to reduce cost (Handel et al., 2015; Layton, 2017; Mahoney and Weyl, 2017). Cherry-picking may result in the unraveling of the most generous, high-cost plans. Risk adjustment discourages strategic variation in premiums by plan generosity, but does not explicitly restrict such variation. In the next section, we discuss the calculation of ACA risk adjustment transfers.

Closely related to the reinsurance and risk adjustment programs was the ACA's implementation of medical loss ratio (MLR) requirements and a temporary risk corridor program. The MLR is the share of premiums spent on medical claims or efforts to improve quality of care (i.e., not profit distributions or plan administrative costs). ACA insurers must send rebates to their enrollees if the MLR falls below 80%. The MLR requirement does not appear to be binding in the California ACA exchange; only once did a California exchange insurer (out of 13) make MLR rebate payments across 5 years of data. We therefore omit MLR constraints in the model developed in the next section. The ACA's risk corridor program, in place between 2014 and 2016, reduced both insurer gains and losses. Insurers with substantial gains paid into the program, whereas insurers with substantial losses drew from the program. Profit and loss reduction were symmetric such that the risk corridor program had no impact on expected profit. Because the model developed in the next section assumes insurers are risk-neutral profit-maximizers and entry decisions are exogenous, risk corridors have no impact in our model.

A unique feature of the rate filing data is the ability to compare firms' predictions about their costs with their realized costs. We refer to the difference between the firm's predicted and realized average costs as the *cost prediction error*, where cost is the sum of claims, risk adjustment, and reinsurance. The cost prediction error for year t is the difference between the predicted average cost reported in the year t rate filing and the realized average cost reported in the year $t+2$ rate filing. For example, the 2016 cost prediction error uses predicted cost data from the 2016 rate filing and the realized cost data for 2016 as reported two years later in the 2018 rate filing.

Figure 1 reports the firms' cost prediction error. In the first year of the ACA exchanges, Blue Shield and Kaiser over-predicted average monthly costs by \$96 and \$115, respectively, whereas Health Net under-predicted its average monthly cost by \$118. The firms' prediction error narrowed considerably over the first five years of the exchanges. During this period, the direction of the prediction error reversed for all four firms, most strikingly for Kaiser. This reversal suggests that the firms were not strategically misleading regulators with their predictions, nor deliberately implementing a dynamic pricing strategy or invest and harvest. By 2018, Anthem and Blue Shield were able to predict their average costs to within \$5 of their actual costs. Kaiser also had its smallest cost prediction error in 2018. Health Net reduced its prediction error by more then half from 2014 to 2018. We interpret this convergence of predicted and actual costs as evidence of firm learning. Morrisey et al. (2017) also find anecdotal evidence of substantial initial uncertainty and firm learning in interviews with insurance firm representatives from 5 states, including California.



Notes: Figure shows the evolution of the average cost prediction error for the four large firms. Average cost equals average claims minus the average risk adjustment transfer received and average reinsurance received.

2.2 Data on Enrollee Choices

We obtain consumer-level enrollment data from the California ACA exchange. There are approximately 10 million records in our enrollment data between 2014 and 2019. Our enrollment data include every enrollee's chosen plan and key enrollee characteristics, but not enrollee utilization. The data provide sufficient information define every household's complete choice set and the household-specific premium paid for each plan in its choice set.

Appendix Table A1 summarizes enrollee characteristics by plan year. About 90% of exchange enrollees are eligible for premium subsidies. Premium subsidies are available to consumers who (1) have income between 100% and 400% of the federal poverty line (FPL); (2) are citizens or legal residents; (3) are ineligible for public insurance such as Medicare or Medicaid; and (4) lack access to an "affordable plan offer" through employer-sponsored insurance. Most households in California with income below 138% of FPL are eligible for Medicaid and therefore ineligible for premium subsidies. A plan is defined as "affordable" if the employee's contribution to the employer's single coverage plan is less than 9.5% of the employee's household income in the 2014 plan year. This percentage increases slightly each year. The next section discusses the complex ACA formula used to calculate premium subsidies.

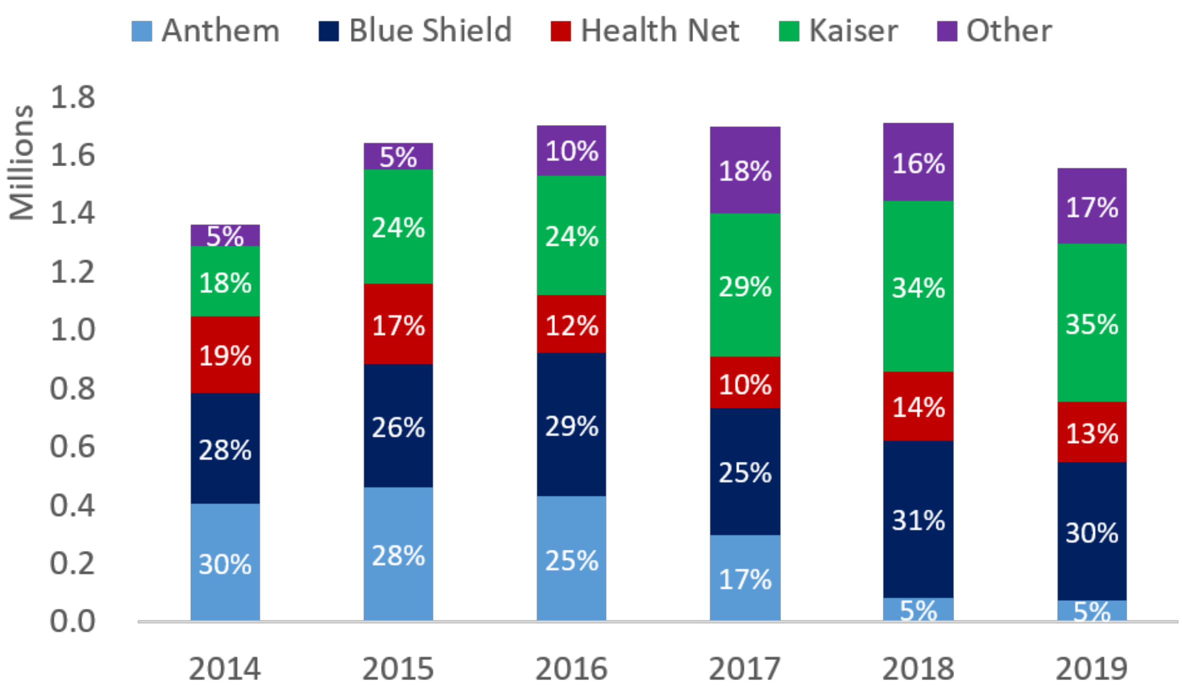
Exchange consumers have access to a diverse set of plans that varies by geographic market and age. Figure 2a shows that 4 firms – Anthem, Blue Shield, Health Net, and Kaiser – dominate the California exchange. There are also 9 regional firms that offer exchange plans.[3] Anthem's market share declined substantially in 2018 when it exited most of the state. Consumers can select a plan from one of the four actuarial value (AV) or "metal" tiers: bronze (with 60% AV), silver (with 70% AV), gold (with 80% AV), and platinum (with 90% AV). Individuals under age 30 can buy a basic catastrophic plan, but premium subsidies cannot be used to purchase catastrophic plans. Consequently, Figure 2b indicates that only 1% of consumers select a catastrophic plan. In contrast, about 60% of consumers choose a plan from the silver tier because eligible consumers must choose silver to receive cost sharing reductions (CSRs) that reduce that reduce deductibles, copays, etc. CSRs increase the AV of the silver plan from 70% to (1) 94% for consumers with income below 150% of the federal poverty level (FPL); (2) 87% for consumers with income between 150% and 200% of FPL; and (3) 73% for consumers with income between 200% and 250% of FPL. Consumers with income above 250% of FPL are ineligible for CSRs. Approximately two-thirds of California consumers are eligible for CSRs.

Consumers also have an outside option to forgo insurance. To construct the outside option population, we use consumer-level survey data on the uninsured from the American Community Survey (ACS) between 2014 and 2019 (Ruggles et al., 2022), consistent with the previous ACA IO literature (Tebaldi, 2022). Our uninsured sample from the ACS excludes consumers who are explicitly or de facto ineligible for the exchanges, such as consumers with access to another source of coverage (e.g., employer-sponsored insurance or Medicaid). We combine the administrative data from Covered California with the survey data from the ACS to form the universe of consumers in our market setting.

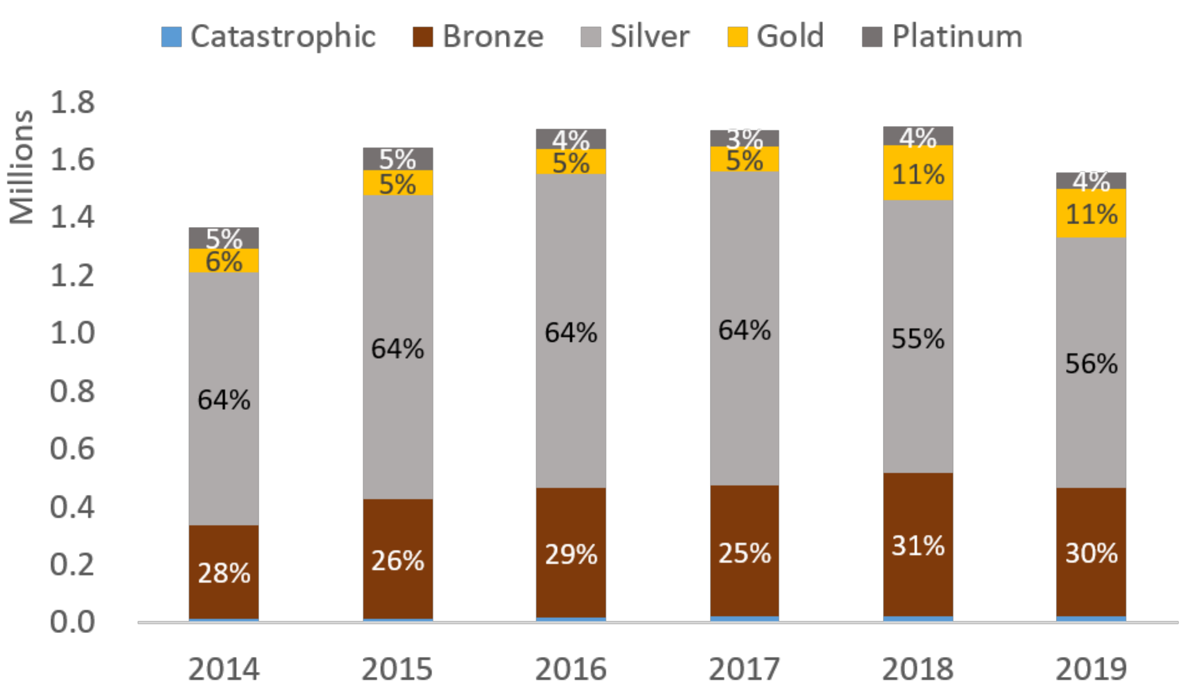
Consumers without insurance may be subject to a penalty under the ACA's individual mandate. The individual mandate penalty was phased in between 2014 and 2016. The penalty for a single person was the greater of $95 and 1% of income (exceeding the tax filing threshold) in 2014 and the greater of $695 and 2.5% of income in 2016. After passage of the Tax Cuts and Jobs Act of 2017, the penalty was set to 0 starting in 2019. Exemptions from the ACA's individual mandate

[3]These firms include Chinese Community Health Plan, Contra Costa, L.A. Care Health Plan, Molina Healthcare, Oscar, Sharp Health Plan, United Healthcare, Valley Health Plan, and Western Health Advantage.

Figure 2: Market Share By Year



(a) By Insurer



(b) By Metal Tier

are made for certain groups, including (1) those with income below the tax filing threshold and (2) individuals who lack access to a health insurance plan that is less than 8% of their income in 2014 (this percentage changes slightly each year).

Although our focus is firm learning, a natural concern is whether consumers learn and adjust their plan choices accordingly. In related work, Saltzman et al. (2021) identify two significant features of this market that mitigate the concern of consumer learning: (1) annual churn into and out of the market is substantial and (2) switching between plans is minimal despite highly volatile premiums during the study timeframe. High levels of churn suggest limited opportunities for consumers to learn and low levels of switching indicate consumers are not adjusting their plan choices over time. Hence, we do not model consumer learning.

3 Adaptive Learning Model

We now develop an adaptive learning model of the ACA exchanges that accommodates firm learning about demand, cost, and the correlation between demand and cost. The previous literature usually adopts the rational expectations approach, which assumes (1) firms can optimally forecast future demand and cost and (2) future forecasts of demand and cost are consistent with dynamic equilibrium play between firms (Evans and Honkapohja, 2001). Given the unique features and complexities of health insurance markets, these assumptions require an unrealistic level of firm knowledge. We instead adopt an adaptive learning approach (Sargent, 1993). This approach assumes firms form expectations of key variables, such as demand and cost, by using only the information currently available to them.

Define the data tuple $D_t \equiv (\mathbf{p}_t, \mathbf{q}_t, \mathbf{c}_t, \mathbf{d}_t)$, where $\mathbf{p}_t$ are the plan base premiums set by all insurers in year t, $\mathbf{q}_t \equiv \{q_{ijt}\}_{i \in I, j \in J}$ are indicator variables that equal 1 if household i chooses plan j at time t, $\mathbf{c}_t \equiv \{c_{jmt}\}_{j \in J, m \in M}$ are average plan medical claims for plan j in market m at time t, and $\mathbf{d}_t \equiv \{d_{ijmt}\}_{i \in I, j \in J, m \in M}$ are other relevant variables that influence optimal agent decisions. Let $\mathcal{I}_t \equiv \{D_\tau\}_{\tau=1,\ldots,t-1}$ for $t > 1$ be the set of information available to firms at time t. Firm forecasts in period t are based on data available in year $1, \ldots, t-1$.

We consider a two-stage game where in each year t (1) firms set premiums simultaneously to maximize expected profit, given their forecasts of demand, cost, and the correlation between demand and cost and (2) households choose plans to maximize utility. Define U_{ijt} as household i's (indirect) utility for plan j at time t. Households solve the discrete choice problem

$$\max_{\mathbf{q}_{it}} U_{ijt} \tag{1}$$

where $\mathbf{q}_{it} = \{q_{ijt}\}_{j\in J}$ is household i's choice vector at time t (i.e., $q_{ijt} \in \{0,1\}$ and $\sum_{i\in I} q_{ijt} = 1$). Firms do not know consumer utilities and must forecast household plan choices using past information $\mathcal{I}_t$. The firm's forecast of household plan choices $E[q_{ijt}|\mathbf{p}_t, \mathcal{I}_t]$ is a function of current year premiums $\mathbf{p}_t$ and available past information $\mathcal{I}_t$. In addition to household choices, firms must forecast plan costs. The firm's forecast of plan cost $E[c_{jmt}|\{E[q_{ijt}|\mathbf{p}_t, \mathcal{I}_t]\}_{i\in I, j\in J}, \mathcal{I}_t]$ for $t > 1$ depends on both past information and the forecast of demand, allowing for correlation between demand and cost. Define the vector of demand forecasts $\widehat{\mathbf{q}}_t \equiv \{E[q_{ijt}|\mathbf{p}_t, \mathcal{I}_t]\}_{i\in I, j\in J}$ and the vector of cost forecasts $\widehat{\mathbf{c}}_t \equiv \{E[c_{jmt}|\{E[q_{ijt}|\mathbf{p}_t, \mathcal{I}_t]\}_{i\in I, j\in J}, \mathcal{I}_t]\}_{m\in M, j\in J}$. Each firm f's premiums must be consistent with its beliefs about demand and cost and expected profit maximization such that its premiums $\mathbf{p}_{ft}$ solve

$$\max_{\mathbf{p}_{ft}} E\left[\pi_{ft}(\widehat{\mathbf{q}}_t, \widehat{\mathbf{c}}_t)\right] \tag{2}$$

The vector of premiums $\mathbf{p}_t^*$ satisfying equation (2) for all firms constitutes an adaptive learning equilibrium. At the adaptive learning equilibrium, firms' plan premiums must be optimal conditional on their competitors' plan premiums and their beliefs about plan costs and optimal household choices in equation (1).

To maintain tractability, we do not model other potential sources of uncertainty that are less relevant for our setting. Our model does not allow for structural or strategic uncertainty that arises when firms have private information about their demand and cost primitives. In this market, firms have ample access to their competitors' rate filings and the regulatory rate review process occurs over several months, providing firms numerous opportunities to learn about their competitors' proposed rates. We also assume consumers are myopic and do not learn over time. As discussed in Section 2, evidence of consumer learning appears to be minimal in our setting.

To operationalize our adaptive learning model of the ACA exchanges, we specify parametric forms of (1) and (2). Least squares learning is a common approach where the econometrician specifies reduced-form parametric relationships between the relevant variables and allows firms to learn about the reduced-form parameters over time (Evans and Honkapohja, 2001). We specify parametric forms of (1) and (2) using standard structural approaches from the IO literature and allow firms to learn about the structural parameters over time. Section 3.1 details how we specify consumer preferences in (1) and Section 3.2 discusses how we specify firm premium-setting in (2). Section 4 explains how we estimate the structural parameters.

3.1 Household Plan Choice

Our specification of household preferences follows standard approaches in the discrete choice literature (Train, 2009). Firms form expectations of household choices using the parameterized utility function

$$U_{ijt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t) \equiv \beta^p_{it} p_{ijt}(\mathbf{p}_t) + \beta^y_{it} y_{ij(t-1)} + x'_{ij}\beta^x_t + f'_{ijt}\zeta^f_t + \xi_j + \epsilon^d_{ijt} \tag{3}$$

where $p_{ijt}(\mathbf{p}_t)$ is household i's premium for plan j in year t, $y_{ij(t-1)}$ indicates whether household i chose plan j in the previous year, x_{ij} is a vector of observed product characteristics including the plan AV, f_{ijt} is a vector of year fixed effects, market fixed effects and insurer-market fixed effects, ξ_j is a vector of unobserved product characteristics, and ϵ^d_{ijt} is an error term. Each year, firms obtain new information to update their estimates $\boldsymbol{\beta}_t$ of the full information parameter vector $\boldsymbol{\beta}$ and form new forecasts of demand. We allow the household's premium parameter $\beta^p_{it} = \beta^p_t + w'_{it}\phi_t$ to vary with household characteristics w_{it} and over time (i.e., we interact the premium variable with household characteristics and with year dummies). The household's switching cost parameter $\beta^y_{it} = \beta^y_t + x'_{ij}\kappa_t + w'_{it}\nu_t$ varies with household characteristics, product characteristics, and over time. Premium subsidies reduce the household's premium $p_{ijt}(\mathbf{p}_t)$ as discussed below. CSRs increase the AV of silver plans in equation (3). The utility of the outside option $U_{i0t} = \beta^p_{it}\rho_{it} + \epsilon_{i0t}$, where ρ_{it} is the household's penalty for not purchasing insurance in year t.

3.1.1 Calculating Household Premiums

The household's premium $p_{ijt}(\mathbf{p}_t)$ is calculated as

$$p_{ijt}(\mathbf{p}_t) = \max\left\{ \underbrace{\sigma_{it} p_{jmt}}_{\substack{\text{full}\\ \text{premium}}} - \underbrace{\max\{\sigma_{it} p_{bmt} - \varphi_{it}, 0\}}_{\text{premium subsidy}}, 0 \right\} \tag{4}$$

where σ_{it} is the household's rating factor, p_{jmt} is the base premium of plan j in market m and year t, p_{bmt} is the base premium of the benchmark plan, and φ_{it} is the household's income contribution cap. The product of the rating factor and the plan's base premium equals the household's full or unsubsidized premium.

Household rating factors are subject to the ACA's "modified community rating" regulations. California insurers cannot use health status to rate plan premiums and are only permitted to use age and geographic residence of the household's members.[4] Figure 3a compares the age rating curve

[4]The ACA also permits rating by tobacco usage, but California prohibits tobacco rating.

in effect between 2014-2017[5] with average cost differences by age and gender (Yamamoto, 2013). Insurers are able to charge a 64-year-old up to 3 times as much as a 21-year-old (i.e., the age rating factor in Figure 3a is 3 for a 64-year-old and 1 for a 21-year-old). However, Figure 3a indicates that 64-year-old females cost insurers an average of 4 times as much as 21-year-old females and 64-year-old males cost insurers an average of 6 times as much as 21-year-old males. Insurers therefore undercharge older adults (particularly females) and must overcharge younger adults (particularly males) relative to their expected cost, creating the potential for adverse selection. Females tend to have higher medical costs during their child-bearing years, whereas males have higher medical costs over age 60. Figure 3b shows the partition of California's 58 counties into 19 rating areas. An insurer's premium must be the same for all consumers of the same age within a rating area.

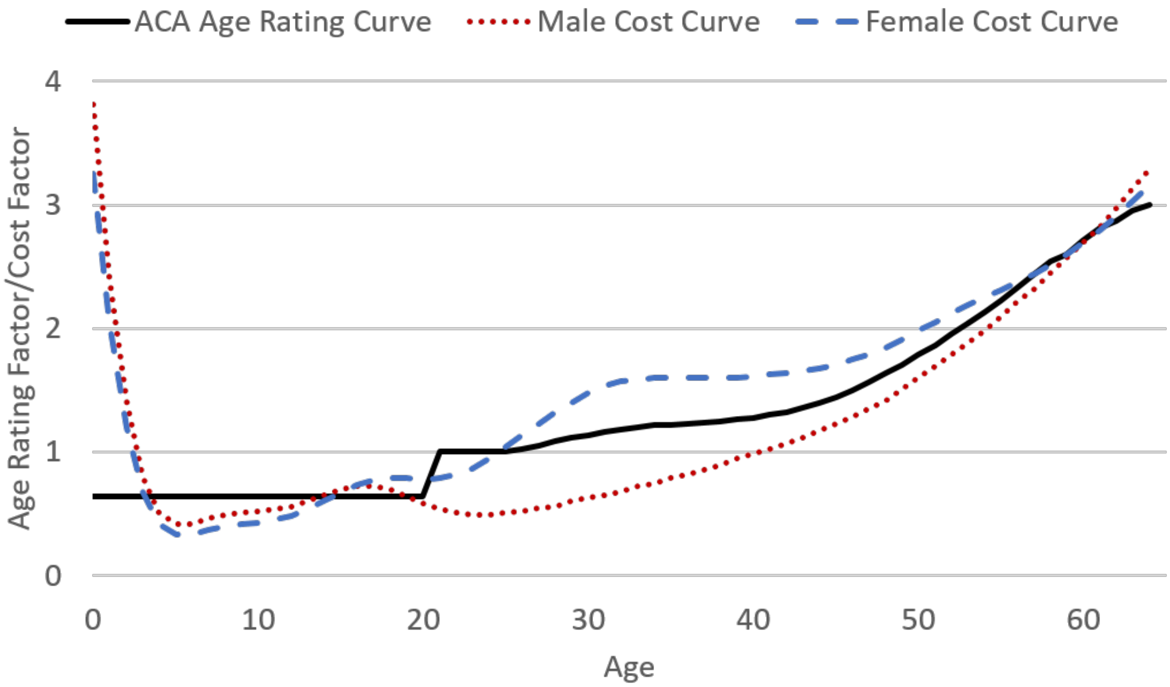
Premium subsidies are calculated as the difference the household's unsubsidized premium for the benchmark plan ($\sigma_{it}p_{bmt}$) and the household's income contribution cap φ_{it} as specified by the ACA. The ACA's premium subsidy is endogenous because it depends on the benchmark plan premium. The ACA defines the benchmark plan as the second-cheapest silver plan available to the household. The benchmark plan varies across households because of heterogeneous firm entry across markets. The income contribution cap ranged from 2% of annual income for consumers earning 100% of the federal poverty level (FPL) and 9.5% of annual income for consumers earning 400% of FPL in 2014. The contribution caps were set initially by the ACA and are updated annually by the Internal Revenue Service (IRS). Because the ACA's subsidy formula uses the second-cheapest silver plan premium as the benchmark, the premium subsidy may exceed the full premium of some bronze plans; the subsidy is reduced in these cases to ensure the premium is nonnegative. As discussed in the next section, this nonlinearity in the ACA's subsidy formula creates exogenous variation in relative premiums that we use to identify the premium parameter in equation (3).

3.1.2 Calculating Demand

We assume that the vector of error terms ϵ_i has the generalized extreme value distribution so that equation (3) is a nested logit model with two nests. The first nest contains all exchange plans and the second nest contains the outside option. This nest structure captures the primary substitution channel between silver plans (which must be selected to receive CSRs) and the outside option. Under this nest structure, the household choice probability forecasts are

[5]The age rating curve for 2018 used slightly higher age rating factors for children under 21

Figure 3: Modified Community Rating in the California Exchange



(a) ACA Age Rating Curve vs. Observed Age Cost Curves By Gender

(b) Premium Rating Regions in California

Notes: Panel (a) compares the ACA's age rating curve with the observed age cost curves by age and gender (Yamamoto, 2013). By design, a 21-year-old is assigned a rating factor of 1 and a 64-year-old is assigned a rating factor of 3. A 64-year-old can therefore be charged 3 times as much as a 21-year-old. Panel (b) shows the partition of California's 58 counties into 19 rating areas (Department of Managed Health Care, 2016).

$$q_{ijt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t) \equiv E[q_{ijt}|\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t] \;\; = \;\; \frac{e^{V_{ijt}(\mathbf{p}_t,\mathcal{I}_t;\boldsymbol{\beta}_t)/\lambda_t}\left(\sum_j e^{V_{ijt}(\mathbf{p}_t,\mathcal{I}_t;\boldsymbol{\beta}_t)/\lambda_t}\right)^{\lambda_t-1}}{1+\left(\sum_j e^{V_{ijt}(\mathbf{p}_t,\mathcal{I}_t;\boldsymbol{\beta}_t)/\lambda_t}\right)^{\lambda_t}} \tag{5}$$

where $V_{ijt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t) \equiv \beta^p_{it} p_{ijt}(\mathbf{p}_t) + \beta^y_{it} y_{ij(t-1)} + x'_{ij}\beta^x_t + f'_{ijt}\zeta^f_t + \xi_j$ and λ_t is the firm's estimate of the full information nesting parameter λ. The household choice probabilities in equation (5) converge to the standard logit choice probabilities when $\lambda \to 1$. Define J_{mt} as the set of available plans in market m at time t. The sensitivity of a subsidized consumer's demand to a base plan premium change is

$$\frac{\partial q_{ikt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)}{\partial p_{jmt}} \;\; = \;\; \sum_{l \in J_{mt}} \frac{\partial q_{ikt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)}{\partial p_{ilt}(\mathbf{p}_t)} \frac{\partial p_{ilt}(\mathbf{p}_t)}{\partial p_{jmt}} \tag{6}$$

for all plans j, k, where

$$\frac{\partial q_{ikt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)}{\partial p_{ilt}} = \begin{cases} \beta^p_i q_{ilt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)\left[\frac{1}{\lambda_t} + \frac{\lambda_t - 1}{\lambda_t} q'_{ilt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t) - q_{ilt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)\right] & k = l \\ \beta^p_i q_{ilt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)\left[\frac{\lambda_t - 1}{\lambda_t} q'_{ilt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t) - q_{ilt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)\right] & k \neq l \end{cases} \tag{7}$$

such that $q'_{ijt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)$ is the probability of choosing j, conditional on choosing a plan. Assuming the subsidy does not exceed the full premium,[6] we have from equation (4) that

$$\frac{\partial p_{ilt}(\mathbf{p}_t)}{\partial p_{jmt}} = \begin{cases} 0 & l = j, j = b \\ \sigma_{it} & l = j, j \neq b \\ -\sigma_{it} & l \neq j, j = b \\ 0 & l \neq j, j \neq b \end{cases} \tag{8}$$

Equation 8 indicates that an increase in a plan's base premium results in consumers paying more for that plan, unless it is the benchmark plan. A small increase in the benchmark plan base premium increases the subsidy by the same amount. Hence, the consumer's contribution to the benchmark plan premium remains constant, but the larger subsidy reduces what consumers pay for all other plans. Modeling this endogenous subsidy design poses substantial computational issues. We model the ACA's endogenous subsidy because of the key role premium subsidies play in determining the extent to which consumers, firms, or taxpayers assume the cost of learning.

Define total plan demand $q_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\beta}_t) \equiv \sum_{i \in I} (\mathbb{I}_{i,m,t})\, q_{ijt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)$ and total firm demand $q_{ft}(\widehat{\mathbf{q}}_t; \boldsymbol{\beta}_t) \equiv \sum_{k \in J_{ft}, m \in M} q_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\beta}_t) = \sum_{i \in I, k \in J_{ft}} q_{ikt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)$, where $\widehat{\mathbf{q}}_t \equiv \{q_{ijt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)\}_{i \in I, j \in J} =$

[6] If the subsidy does exceed the full premium, then $\frac{\partial p_{ilt}(\mathbf{p}_t)}{\partial p_{jmt}} = 0$.

$\{E[q_{ijt}|\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t]\}_{i\in I, j\in J}$ is the vector of demand forecasts, J_{ft} is the set of plans sold by firm f at time t, and $\mathbb{I}_{i,m,t}$ indicates whether household i lives in market m at time t. The sensitivities of plan demand and firm demand to a base plan premium change are

$$\frac{\partial q_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\beta}_t)}{\partial p_{jmt}} = \sum_{i\in I} (\mathbb{I}_{i,m,t}) \frac{\partial q_{ikt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)}{\partial p_{jmt}} \tag{9}$$

$$\frac{\partial q_{ft}(\widehat{\mathbf{q}}_t; \boldsymbol{\beta}_t)}{\partial p_{jmt}} = \sum_{k\in J_{ft}, m\in M} \frac{\partial q_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\beta}_t)}{\partial p_{jmt}} \tag{10}$$

3.2 Firm Premium-Setting

A risk-neutral firm sets its base plan premiums $\mathbf{p}_{ft}$ to maximize expected profit

$$E\left[\pi_{ft}(\widehat{\mathbf{q}}_t, \widehat{\mathbf{c}}_t; \boldsymbol{\theta}_t)\right] = R_{ft}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) + RA_{ft}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) - (1-\iota_{ft})C_{ft}(\widehat{\mathbf{q}}_t, \widehat{\mathbf{c}}_t; \boldsymbol{\theta}_t) - V_{ft}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) - FC_{ft} \tag{11}$$

where $R_{ft}(\cdot)$ is total premium revenue, $RA_{ft}(\cdot)$ is risk adjustment received, $C_{ft}(\cdot)$ is total claims, $V_{ft}(\cdot)$ is variable administrative cost (e.g., commissions or fees), FC_{ft} is fixed cost, ι_{ft} indicates the AV of the reinsurance contract (i.e., the expected percentage of claims paid by the reinsurer), $\widehat{\mathbf{c}}_t \equiv \{c_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)\}_{m\in M, j\in J} = \{E[c_{jmt}|\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t]\}_{m\in M, j\in J}$ is the vector of cost forecasts, and $\widehat{\mathbf{q}}_t \equiv \{q_{ijt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)\}_{i\in I, j\in J} = \{E[q_{ijt}|\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t]\}_{i\in I, j\in J}$ is the vector of demand forecasts as defined above. The risk corridor program makes a positive monotonic transformation of firm profit and hence does not affect the optimal solution, assuming firms are risk-neutral and maximize expected profit. We also ignore MLR constraints because the empirical evidence suggests that they are not binding. The vector $\boldsymbol{\theta}_t \equiv (\boldsymbol{\beta}_t, \boldsymbol{\gamma}_t, \boldsymbol{\mu}_t)$ is an estimate of the full information parameter vector $\boldsymbol{\theta} \equiv (\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\mu})$, where $\boldsymbol{\beta}$ are the demand parameters, $\boldsymbol{\gamma}$ are the risk score parameters, and $\boldsymbol{\mu}$ are the average claims parameters. We define the parameterization of risk scores and average claims below. Differentiating equation (11) yields the Nash equilibrium conditions

$$\frac{\partial R_{ft}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}{\partial p_{jmt}} + \frac{\partial RA_{ft}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}{\partial p_{jmt}} = (1-\iota_{ft})\frac{\partial C_{ft}(\widehat{\mathbf{q}}_t, \widehat{\mathbf{c}}_t; \boldsymbol{\theta}_t)}{\partial p_{jmt}} + \frac{\partial V_{ft}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}{\partial p_{jmt}} \tag{12}$$

for all markets m in which plan j is offered by the firm in year t. Equation (12) accounts for potential intra-firm cannibalization between plans (i.e., a decrease in plan j's premium may reduce demand for the firm's other plans).

The equilibrium conditions in equation (12) are used for estimating the model parameters and computing new equilibria in counterfactuals. In the following four subsections, we show how to take these equilibrium conditions to our data. We write the revenue, risk adjustment, claims, and administrative cost variables in equations (11) and (12) in terms of three estimable variables: (1) household choice probabilities $q_{ijt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)$; (2) plan risk scores $r_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)$; and (3) average claims $c_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)$. Household choice probabilities are defined above in equation (5), plan risk scores are defined below in equation (15), and average claims are defined below in equation (22).

3.2.1 Revenue

The firm's total expected premium revenue equals the expected (unsubsidized) premium collected from each household. That is, $R_{ft}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) = \sum_{i \in I, m \in M, k \in J_{fmt}} (\mathbb{I}_{i,m,t})\, \sigma_{it} p_{kmt} q_{ikt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)$, where J_{fmt} is the set of plans sold by firm f in market m at time t. The sensitivity of firm revenue to a base plan premium change is

$$\frac{\partial R_{ft}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}{\partial p_{jmt}} = \sum_{i \in I, k \in J_{fmt}} \mathbb{I}_{i,m,t} \sigma_{it} \left(q_{ijt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t) + p_{kmt} \frac{\partial q_{ikt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)}{\partial p_{jmt}} \right) \tag{13}$$

We can compute $\frac{\partial R_{ft}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}{\partial p_{jmt}}$ using $q_{ijt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)$ from equation (5) and $\frac{\partial q_{ikt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)}{\partial p_{jmt}}$ from equation (6).

3.2.2 Risk Adjustment

Using the notation of our model, the firm's risk adjustment transfer $RA_{ft}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)$ using the official ACA risk adjustment formula (Pope et al., 2014) equals[7]

$$\begin{aligned} RA_{ft}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) &= \sum_{m \in M, k \in J_{kmt}} \left[\widehat{c}_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) - \widetilde{c}_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) \right] q_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) \\ &= \sum_{m \in M, k \in J_{kmt}} \left[\frac{\widehat{h}_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}{\sum_{l \in J_t} \widehat{h}_{lmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) s_{lmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)} \nu \overline{p} - \frac{\widetilde{h}_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}{\sum_{l \in J_t} \widetilde{h}_{lmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) s_{lmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)} \nu \overline{p} \right] q_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) \end{aligned} \tag{14}$$

where $\widehat{c}_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)$ is expected plan average claims with adverse selection and $\widetilde{c}_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)$ is expected plan average claims without adverse selection. Transfer formula (14) redistributes money so that each firm faces the same (unobserved) enrollee health risk, which firms are prohibited from considering when determining premiums. Firms are not compensated for observable differences in age, geography, moral hazard, or plan AV that can be considered when determining premiums. Through the cost factor $\widehat{h}_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)$, the variable $\widehat{c}_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)$ accounts for plan differences in enrollee health risk, age, geography, moral hazard, and plan AV. Through the cost factor

[7] Saltzman (2021) uses a simplified version of the risk adjustment formula in Pope et al. (2014)'s Appendix A1. In this analysis, we use the full version of the risk adjustment formula in Pope et al. (2014)'s Appendix A2.

$\widetilde{h}_{jmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)$, the variable $\widetilde{c}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)$ accounts for plan differences in age, geography, moral hazard, and plan AV, but not enrollee health risk. Therefore, the difference $\widehat{c}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)-\widetilde{c}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)$ compensates plans for differences in enrollee health risk only (i.e., the plan's relative risk due to adverse selection only). Plans with higher-than-average risk will receive a risk adjustment transfer ($\widehat{c}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)-\widetilde{c}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)>0$), whereas plans with lower-than-average risk will pay a risk adjustment transfer ($\widehat{c}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)-\widetilde{c}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)<0$).

Our model fully endogenizes all terms in the second line of formula (14). The cost factor $\widehat{h}_{jmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \equiv \mathrm{IDF}_j\mathrm{GCF}_{mt}r_{jmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)$ is the product of the plan's induced demand factor (or moral hazard factor), geographic cost factor, and risk score. The induced demand factors and geographic cost factors are set by CMS. Each year, CMS estimates plan risk scores as function of the plan AV, enrollee characteristics including age, and diagnosed medical conditions using a regression-based procedure (Pope et al., 2014). We endogenize plan risk scores using the estimating equation

$$\ln r_{jmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \quad = \quad \sum_{d\in D}\gamma_t^d s_{djmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)+MT_j'\gamma_t^{MT}+\epsilon_{jmt}^r \tag{15}$$

The predicted demographic share $s_{djmt}(\cdot)=\frac{q_{djmt}(\cdot)}{q_{jmt}(\cdot)}$ is the share of plan j's enrollment in market m and year t with demographic characteristic d, MT_j is a vector metal (or AV) tier fixed effects, ϵ_{jmt}^r is an error term, and $\boldsymbol{\gamma}_t=(\gamma_t^d,\gamma_t^{MT},\gamma_t^n)$ is the firm's estimate of the full information risk score parameter vector $\boldsymbol{\gamma}$. We compute demographic shares by aggregating the choice probabilities in equation (5). The cost factor $\widetilde{h}_{jmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \equiv \mathrm{AV}_j\mathrm{IDF}_j\mathrm{GCF}_{mt}a_{jmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)$ is the product of the plan's AV, induced demand factor, geographic cost factor, and average ACA age rating factor $a_{jmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)=\frac{\sum_{i\in I}(\mathbb{I}_{i,m,t})a_{it}q_{ijt}(\widehat{\mathbf{p}}_t,\mathcal{I}_t;\boldsymbol{\theta}_t)}{q_{jmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}$ across the plan's enrollees, where a_{it} is the household's CMS age rating factor. The plan's market share is $s_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)=\frac{q_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\sum_{j\in J_t}q_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}$, where J_t is the set of all plans offered in the market in year t. Importantly, this summation occurs over all metal plans in the market at the state level because of the ACA's single risk pool provisions.[8] The average statewide premium is denoted $\overline{p}$. Finally, the expected percentage of collected premiums that is spent on claims is denoted as ν. CMS set ν to 100% from 2014-2017 and then reduced it to 86% starting in 2018.

Now we calculate the sensitivity of the firm's risk adjustment transfer to a base plan premium change. Define the risk share $rs_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \equiv \frac{\widehat{h}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)q_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\sum_{l\in J_t}\widehat{h}_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)q_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}$ and the utilization share $us_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \equiv \frac{\widetilde{h}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)q_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\sum_{l\in J_t}\widetilde{h}_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)q_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}$. Let $R_t(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \equiv \sum_{f\in F}R_{ft}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)$ be total market premium revenue in year t. Then

[8] Catastrophic plans have a separate risk adjustment pool.

$$\frac{\partial RA_{ft}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} = \nu \sum_{k\in J_{fmt}} \left[\frac{\partial R_t(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} \left(rs_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) - us_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \right) + R_t(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \left(\frac{\partial rs_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} - \frac{\partial us_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} \right) \right] \quad (16)$$

where the partial $\frac{\partial R_t(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} = \sum_{f\in F} \frac{\partial R_{ft}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}}$ and the partials for the risk and utilization shares are

$$\frac{\partial rs_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} = \left(\sum_{m'\in M, l\in J_{mt}} \widehat{h}_{lm't}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) q_{lm't}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \right)^{-1} \left[\left(\widehat{h}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \frac{\partial q_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} + q_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \frac{\partial \widehat{h}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} \right) - rs_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \sum_{l\in J_{mt}} \left(\widehat{h}_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \frac{\partial q_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} + q_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \frac{\partial \widehat{h}_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} \right) \right] \quad (17)$$

$$\frac{\partial us_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} = \left(\sum_{m'\in M, l\in J_{m't}} \widetilde{h}_{lm't}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) q_{lm't}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \right)^{-1} \left[\left(\widetilde{h}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \frac{\partial q_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} + q_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \frac{\partial \widetilde{h}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} \right) - us_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \sum_{l\in J_{mt}} \left(\widetilde{h}_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \frac{\partial q_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} + q_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \frac{\partial \widetilde{h}_{lmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} \right) \right] \quad (18)$$

Computing the partials for the risk and utilization shares requires computing partials for the cost factors, which are given by

$$\frac{\partial \widehat{h}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} = \text{IDF}_k \text{GCF}_{mt} \left[\frac{r_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{q_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)} \sum_{d\in D} \gamma_t^d \left[\frac{\partial q_{dkmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} - s_{dkmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \frac{\partial q_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} \right] \right] \quad (19)$$

$$\frac{\partial \widetilde{h}_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} = \text{AV}_k \text{IDF}_k \text{GCF}_{mt} \frac{\left[\sum_{i\in I} \left(\mathbb{I}_{i,m,t} \right) a_{it} \frac{\partial q_{ikt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)}{\partial p_{jmt}} - a_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \frac{\partial q_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} \right]}{q_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)} \quad (20)$$

where $q_{dkmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)$ is the number of enrollees in plan k with demographic characteristic d.

Using equations (16)-(20), we can compute the partial derivative $\frac{\partial RA_{ft}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}}$ with estimates of the choice probabilities and risk scores. Only the choice probabilities are needed to compute total premium revenue $R_t(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)$ and its partial $\frac{\partial R_t(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} = \sum_{f\in F} \frac{\partial R_{ft}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}}$ using equation (13). Computation of the risk share $rs_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)$ and its partial derivative $\frac{\partial rs_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}}$ using equations (17) and (19) requires the choice probabilities and plan risk scores. Computation of the utilization share $us_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)$ and its partial derivative $\frac{\partial us_{kmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}}$ using equations (18) and (20) only requires the choice probabilities.

3.2.3 Claims

Total expected claims paid by the firm equal the expected claims paid out for each plan. That is, $C_{ft}(\widehat{\mathbf{q}}_t, \widehat{\mathbf{c}}_t; \boldsymbol{\theta}_t) = \sum_{m \in M, k \in J_{fmt}} c_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) q_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)$, where $c_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)$ is plan average claims. We calculate the sensitivity of total expected claims to a base plan premium change as

$$\frac{\partial C_{ft}(\widehat{\mathbf{q}}_t, \widehat{\mathbf{c}}_t; \boldsymbol{\theta}_t)}{\partial p_{jmt}} = \sum_{k \in J_{fmt}} \left(q_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) \frac{\partial c_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}{\partial p_{jmt}} + c_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) \frac{\partial q_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}{\partial p_{jmt}} \right) \quad (21)$$

We calculate plan average claims as a function of the plan risk score using the estimating equation

$$\begin{aligned} \ln c_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) &= \mu_t^r \ln r_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) + x_j' \mu_t^x + \mu_t^l l_t + n_m' \mu_t^n + \epsilon_{jmt}^{c'} \\ &= \mu_t^r \left(\sum_{d \in D} \gamma_t^d s_{djmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) + MT_j' \gamma^{MT} \right) + x_j' \mu_t^x + \mu_t^l l_t + n_m' \mu_t^n + \epsilon_{jmt}^c \quad (22) \end{aligned}$$

where $r_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)$ is the predicted risk score computed using equation (15), x_j are product characteristics (not including plan AV), l_t is a linear trend, n_m' are market fixed effects, $\epsilon_{jmt}^c = \epsilon_{jmt}^{c'} + \mu_t^r \epsilon_{jmt}^r$ is an error term, and $\boldsymbol{\mu}_t = (\mu_t^r, \mu_t^x, \mu_t^l, \mu_t^n)$ is an estimate of the full information claims parameter vector $\boldsymbol{\mu}$. The inclusion of a trend term in (22) incorporates cost growth that is a common feature of health insurance markets. The sensitivity of average claims to a base plan premium change is

$$\frac{\partial c_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}{\partial p_{jmt}} = \mu_t^r \frac{c_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}{q_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)} \sum_{d \in D} \gamma_t^d \left[\frac{\partial q_{dkmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}{\partial p_{jmt}} - s_{dkmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) \frac{\partial q_{kmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}{\partial p_{jmt}} \right] \quad (23)$$

Hence, we can compute the partial derivative $\frac{\partial C_{ft}(\widehat{\mathbf{q}}_t, \widehat{\mathbf{c}}_t; \boldsymbol{\theta}_t)}{\partial p_{jmt}}$ with estimates of the choice probabilities, plan risk scores, and average claims using equations (21) and (23). To gain insight into how adverse selection manifests in the model, define marginal claims $MC_{jmt}(\widehat{\mathbf{q}}_t, \widehat{\mathbf{c}}_t; \boldsymbol{\theta}_t) \equiv \frac{\partial C_{ft}(\widehat{\mathbf{q}}_t, \widehat{\mathbf{c}}_t; \boldsymbol{\theta}_t)}{\partial q_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)}$. Then

$$MC_{jmt}(\widehat{\mathbf{q}}_t, \widehat{\mathbf{c}}_t; \boldsymbol{\theta}_t) = c_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) \left(1 + \mu_t^r \sum_{d \in D} \gamma_t^d s_{djmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t) \right) \quad (24)$$

The presence of adverse selection depends on the sign of the "selection term" $\mu_t^r \sum_{d \in D} \gamma_t^d s_{djmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)$. The selection term equals zero in a market without selection when marginal claims and average claims are equal. The selection term is positive in a market with advantageous selection when marginal claims are greater than average claims. The selection term is negative in a market with adverse selection when marginal claims are less than average claims.

3.2.4 Variable Administrative Costs and Fixed Costs

Variable administrative costs include expenses such as commissions and fees, whereas fixed costs include expenses such as building overhead. Data on both come from the CMS MLR reports. We calculate variable administrative costs as $V_{ft}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)=v_{ft}q_{ft}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)$, where v_{ft} is the variable administrative cost per-member per-month calculated from the MLR data. The sensitivity of variable administrative costs to a base plan premium change is

$$\frac{\partial V_{ft}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} = v_{ft}\frac{\partial q_{ft}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}} \tag{25}$$

We only require the choice probabilities and equations (10) and (25) to compute the partial $\frac{\partial V_{ft}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jmt}}$.

4 Estimation

In this section, we explain how we estimate the adaptive learning parameter vector $\boldsymbol{\theta}_t$. To estimate $\boldsymbol{\theta}_t$, we create four sets of moment conditions: (1) demand moments that match observed choices and predicted household choice probabilities; (2) risk score moments that match observed and predicted risk scores; (3) average claims moments that match observed and predicted average claims; and (4) the first-order conditions for profit maximization in equation (12). Define the previous period set $T_t \equiv \{2014,\ldots,t-1\}$. Denote N^{IJT_t} as the number of plans available to all households in year t, N^{JMT_t} as the number of plans available in all markets in year t, and N^M_{jt} as the number of markets where plan j is offered in year t. Let q_{ijt} be an indicator of whether household i chose plan j at time t, r_{jmt} be the observed plan risk score, and c_{jmt} be the observed plan average claims. Define the risk score covariates $\mathbf{z}^r_{jmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \equiv (s_{djmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t), MT_j)$ and the average claims covariates $\mathbf{z}^c_{jmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \equiv (\ln r_{jmt}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t), x_j, u_t, n_m)$. To estimate $\boldsymbol{\theta}_t$, we form the moment conditions

$$\begin{aligned}
\frac{1}{N^{IJT_t}}\sum_{i\in I, j\in J, \tau\in T_t}\frac{q_{ij\tau}\partial \ln q_{ij\tau}(\mathbf{p}_t,\mathcal{I}_t;\boldsymbol{\beta}_t)}{\partial\boldsymbol{\theta}_t} &= \mathbf{0}\\
\frac{1}{N^{JMT_t}}\sum_{j\in J, m\in M, \tau\in T_t}\mathbf{z}^r_{jm\tau}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)\left(\ln r_{jm\tau}-\boldsymbol{\gamma}_t'\mathbf{z}^r_{jm\tau}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)\right) &= \mathbf{0}\\
\frac{1}{N^{JMT_t}}\sum_{j\in J, m\in M, \tau\in T_t}\mathbf{z}^c_{jm\tau}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)\left(\ln c_{jm\tau}-\boldsymbol{\mu}_t'\mathbf{z}^c_{jm\tau}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)\right) &= \mathbf{0}\\
\frac{1}{N^M_{j\tau}}\sum_{m\in M}\overline{g}_{jm\tau}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) &= 0, \quad \forall j\in J_\tau, \tau\in\{2014,\ldots,t\}
\end{aligned} \tag{26}$$

where the first-order condition values

$$\overline{g}_{jm\tau}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t) \equiv \frac{\partial R_{f\tau}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jm\tau}} - (1-\iota_{f\tau})\frac{\partial C_{f\tau}(\widehat{\mathbf{q}}_t,\widehat{\mathbf{c}}_t;\boldsymbol{\theta}_t)}{\partial p_{jm\tau}} + \frac{\partial RA_{f\tau}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jm\tau}} - \frac{\partial V_{f\tau}(\widehat{\mathbf{q}}_t;\boldsymbol{\theta}_t)}{\partial p_{jm\tau}}$$

Only data through year $t-1$ is used to estimate $\boldsymbol{\theta}_t$ in the first three sets of moment conditions. The fourth set of moment conditions are the firms' first-order conditions through year t. We include the first-order conditions for year t because the estimated parameters $\boldsymbol{\theta}_t$ should be consistent with profit-maximizing behavior in year t (even though firms only have access to data through year $t-1$).

Because model (26) over-identifies the model parameters, we use two-step feasible GMM to find the values of $\boldsymbol{\theta}_t$ that minimize the GMM objective $[\mathbf{m}(\boldsymbol{\theta}_t)]'\mathbf{W}^{-1}[\mathbf{m}(\boldsymbol{\theta}_t)]$, where $\mathbf{m}(\boldsymbol{\theta}_t)$ is the vector of moment values in model (26) and the optimal weight matrix $\mathbf{W}$ is a consistent estimate of the variance-covariance matrix of the moment values.

The primary estimation challenge is to identify the effect of premiums on household choices (i.e., the parameter β^p_{it}). We observe most plan characteristics, but some time-invariant plan characteristics that vary at the insurer-market level such as plan formularies or customer service are unobserved. We include insurer-market fixed effects to control for these unobservables. Using the institutional detail of the ACA setting, we can also exploit several other sources of plausibly exogenous variation in premiums. First, kinks in the household premium formula (4) create exogenous variation in relative premiums (i.e., between plans). As discussed above, some bronze plans may be "free" to low-income consumers if the subsidy exceeds the full premium (i.e., the second-cheapest silver plan available to the consumer may exceed the premium of some bronze plans). The set of free plans varies by market, time, and exogenously-determined household characteristics, including age, income, and household composition. Second, the phasing-in of the mandate penalty between 2014 and 2016 and elimination of the penalty in 2019 creates exogenous variation in absolute premiums (i.e., relative to the outside option). One caveat with using this time-varying source of variation is that it is limited in the initial years of the exchanges. Third, age rating factors set by CMS are applied to the insurers' premiums to obtain household premiums, creating variation in premiums across plans and households. We allow the premium parameter to vary by age group to leverage this variation for identification, assuming premium sensitivity does not vary within each age group (but premiums do vary within age groups). Saltzman (2019) also estimates the demand parameters using the control function approach of Petrin and Train (2010) with geographic cost factors as instruments and finds minimal differences in the estimated parameters. We do not use the control function approach here because the first stage imposes a hedonic pricing model.

Another identification challenge is that we do not observe patient medical conditions that are used to predict plan risk scores. Estimates of the risk score parameter γ^d_t may be biased by omitting patient medical conditions. We address this potential source of bias by computing predicted demographic shares using the estimated consumer-level choice probabilities from equation (5) instead of the observed demographic shares, which may be endogenous. The identifying assumption is that the predicted demographic shares are based on exogenous determinants of consumer plan demand. Choice model (5) can be interpreted as the first-stage of an IV regression for computing unbiased estimates of plan risk scores. A similar empirical strategy is widely used in the hospital choice literature to compute measures of hospital market concentration (e.g., Kessler and McClellan (2000)).

A third identification challenge is to compute an unbiased estimate of the average claims parameter μ_t^r. We compute predicted plan risk scores using equation (15) instead of the observed plan risk scores, which may be endogenous. Enrollee characteristics should affect average claims through the plan risk score only and not directly affect average claims. This may not be the case if the ACA risk score is an imprecise measure of plan claims risk.

One of the primary goals of our analysis is to compare the fit of the adaptive learning model with the fit of the standard (or full information) model. Using the full information approach, the econometrician pools data from all years to estimate the model parameters. In contrast, the adaptive learning approach only exploits the data available to firms to estimate the model parameters. Denote the full information parameters $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\mu})$. To estimate $\boldsymbol{\theta}$, we form the moment conditions

$$
\begin{aligned}
\frac{1}{N^{IJT}} \sum_{i\in I, j\in J, t\in T} \frac{\chi_{ijt}\partial \ln q_{ijt}(\mathbf{p}_t; \boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= \mathbf{0} \\
\frac{1}{N^{JMT}} \sum_{j\in J, m\in M, t\in T} \mathbf{z}^r_{jmt}(\mathbf{p}_t; \boldsymbol{\theta}) \left(\ln r_{jmt} - \boldsymbol{\gamma}'\mathbf{z}^r_{jmt}(\mathbf{p}_t; \boldsymbol{\theta})\right) &= \mathbf{0} \\
\frac{1}{N^{JMT}} \sum_{j\in J, m\in M, t\in T} \mathbf{z}^c_{jmt}(\mathbf{p}_t; \boldsymbol{\theta}) \left(\ln c_{jmt} - \boldsymbol{\mu}'\mathbf{z}^c_{jmt}(\mathbf{p}_t; \boldsymbol{\theta})\right) &= \mathbf{0} \\
\frac{1}{N^M_{jt}} \sum_{m\in M} g_{jmt}(\mathbf{p}_t; \boldsymbol{\theta}) &= 0, \qquad \forall j \in J_t, t \in T \qquad (27)
\end{aligned}
$$

where the first-order condition values

$$
g_{jmt}(\mathbf{p}_t; \boldsymbol{\theta}) \equiv \frac{\partial R_{ft}(\mathbf{p}_t; \boldsymbol{\theta})}{\partial p_{jmt}} - (1 - \iota_{ft}) \frac{\partial C_{ft}(\mathbf{p}_t; \boldsymbol{\theta})}{\partial p_{jmt}} + \frac{\partial RA_{ft}(\mathbf{p}_t; \boldsymbol{\theta})}{\partial p_{jmt}} - \frac{\partial V_{ft}(\mathbf{p}_t; \boldsymbol{\theta})}{\partial p_{jmt}}
$$

Importantly, the model variables depend on premiums $\mathbf{p}_t$, but no longer depend on the demand forecasts $\widehat{\mathbf{q}}_t$ or cost forecasts $\widehat{\mathbf{c}}_t$.

5 Estimation Results

5.1 Parameter Estimates

Table I summarizes the adaptive learning estimates $\widehat{\boldsymbol{\theta}}_t$ for $t \in \{2016, 2017, 2018\}$ and the full information estimates $\widehat{\boldsymbol{\theta}}$. Detailed parameter estimates for the adaptive learning and full information models are provided in Table A2 in Appendix B. We also estimate six "intermediate" specifications that assume firms know a certain subset of the full information parameters and estimate the remaining parameters conditional on this knowledge. We do not include the adaptive learning estimates $\widehat{\boldsymbol{\theta}}_{2015}$ because the one year of available data from 2014 is insufficient to identify switching costs.[9]

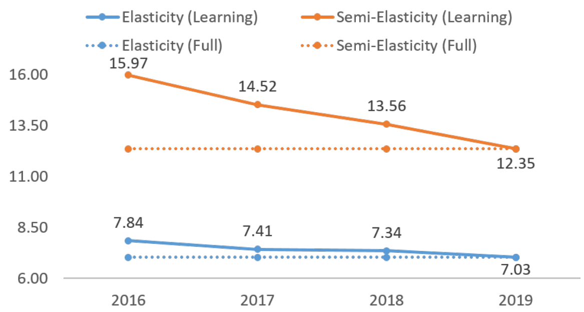
Our results indicate that firms overestimated premium sensitivity in the ACA's initial years. The premium parameter increases slightly from 2016 and 2017, before falling in 2018. Figure 4 shows the mean own-premium elasticities and exchange coverage elasticities of demand implied by these parameter estimates. Firms overestimated the sensitivity of a plan's demand to its own premium in the ACA's initial years; firms also overestimated the sensitivity of total exchange enrollment to a change in exchange plan premiums. Firms substantially underestimated switching costs (i.e., the previous choice parameter). The previous choice parameter estimate in 2016 is only 77% of the previous choice parameter estimate using the full information approach. The trend for the plan generosity parameter is similar to the trend for the premium parameter. Firms initially underestimated the effect of plan generosity. We also find firms initially underestimated the nesting parameter, expecting less substitution between exchange plans and the outside option.[10]

Learning estimates of the supply-side parameters generally converge over time toward the full information estimates. Estimates of the silver, gold, and platinum parameters in the risk score regression converge non-monotonically toward the standard approach estimates. Platinum plans have the greatest exposure to claims risk. Estimates of the young adult (under age 45) share parameters are negative as expected. Firms substantially underestimated the claims risk of adults under age 45, particularly in 2016. The parameter for the "Share Ages 18 to 25" variable is nearly double in magnitude for the 2016 estimates than for the full information estimates. Consumers of Hispanic origin have less claims risk than other racial and ethnic groups, but firms underestimated this effect initially. Firms slightly overestimated the relationship between average claims and the plan risk score; the estimated parameter decreases from 1.074 using data available in 2016 to 1.038 using all

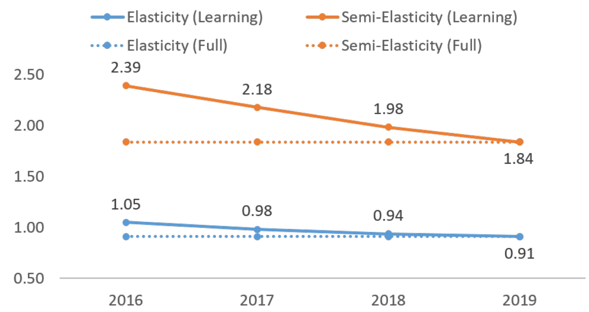
[9]Parameter estimates for the intermediate models are available upon request. We estimated the adaptive learning model for all years without the lagged choice variables $y_{ij(t-1)}$, but obtained very different estimates. We find our model suffers from omitted variable bias when the lagged choice variables are absent and therefore do not include estimates $\widehat{\boldsymbol{\theta}}_{2015}$.

[10]Recall that our nested logit choice model converges to the standard logit as $\lambda \to 1$.

Figure 4: Estimated Premium Elasticities of Demand By Year



(a) Mean Own-Premium Elasticities



(b) Mean Coverage Elasticities

Notes: Figure shows the premium elasticities of demand implied by the learning parameter estimates $\widehat{\theta}_t$ and the full information estimates $\widehat{\theta}$. Panel (a) shows how a plan's demand responds to a change in its own premium. Panel (b) shows how total exchange enrollment responds to a change in all exchange premiums. Semi-elasticities are calculated for a \$100 change in annual premiums.

data. Estimates of the HMO parameter in the average claims regression have the wrong (positive) sign in 2016 and 2017, but are small and not statistically significant. The estimated HMO parameter is negative and statistically significant in 2018. The estimated time trend is negative in 2016 and 2017, but reverses to the correct (positive) sign in 2018.

5.2 Assessing External Validity

Now we assess how well the alternative adaptive learning models fit the data compared to the standard model. We calculate goodness-of-fit measures by comparing the plan average costs implied by our alternative models with the firms' **predictions** of plan average costs as reported in their rate filings. In other words, we quantify how closely each model's estimates of cost match with the firms' predictions of cost. Data from the rate filings on predicted plan costs are only used in this section to calculate model fit; we used the available data on **realized** plan average costs for estimation. Hence, we can assess the external validity of the alternative models.

We compute six common goodness-of-fit measures: (1) mean absolute error (MAE); (2) root mean square error (RMSE); (3) R^2; (4) Adjusted R^2; (5) Akaike information criterion (AIC); and (6) Bayesian information criterion (BIC). Overfitting is a potential concern in assessing goodness-of-fit, particularly because the adaptive learning models have many more parameters. The latter three measures – Adjusted R^2, AIC, and BIC – address this concern by imposing a penalty for includ-

Table I: Summary of Parameter Estimates

	$\widehat{\boldsymbol{\theta}}_{2016}$	$\widehat{\boldsymbol{\theta}}_{2017}$	$\widehat{\boldsymbol{\theta}}_{2018}$	$\widehat{\boldsymbol{\theta}}$
Demand Parameters ($\widehat{\boldsymbol{\beta}}_t$)				
Monthly Premium ($100)	-1.139^{***}	-1.178^{***}	-1.151^{***}	-1.096^{***}
	(0.008)	(0.008)	(0.007)	(0.006)
Previous Choice	1.915^{***}	2.278^{***}	2.407^{***}	2.479^{***}
	(0.088)	(0.070)	(0.058)	(0.052)
AV	3.193^{***}	3.222^{***}	3.216^{***}	3.188^{***}
	(0.028)	(0.025)	(0.022)	(0.020)
Nesting Parameter	0.548^{***}	0.623^{***}	0.649^{***}	0.680^{***}
	(0.005)	(0.005)	(0.004)	(0.004)
Risk Score Parameters ($\widehat{\boldsymbol{\gamma}}_t$)				
Silver	0.814^{***}	0.824^{***}	0.785^{***}	0.761^{***}
	(0.062)	(0.042)	(0.033)	(0.028)
Gold	0.880^{***}	0.913^{***}	0.859^{***}	0.851^{***}
	(0.071)	(0.044)	(0.034)	(0.029)
Platinum	1.083^{***}	1.247^{***}	1.278^{***}	1.287^{***}
	(0.077)	(0.047)	(0.036)	(0.031)
Share Ages 18 to 25	-1.661^{**}	-1.602^{***}	-0.965^{***}	-0.975^{***}
	(0.804)	(0.495)	(0.365)	(0.313)
Share Ages 26 to 44	-1.334^{***}	-0.977^{***}	-1.125^{***}	-1.061^{***}
	(0.395)	(0.219)	(0.161)	(0.141)
Share Male	-0.355	-0.088	0.035	-0.340^{*}
	(0.653)	(0.305)	(0.220)	(0.197)
Share Hispanic	-0.233	-0.334^{**}	-0.642^{***}	-0.744^{***}
	(0.184)	(0.130)	(0.097)	(0.084)
Average Claims Parameters ($\widehat{\boldsymbol{\mu}}_t$)				
Log Risk Score	1.074^{***}	1.049^{***}	1.038^{***}	1.038^{***}
	(0.009)	(0.005)	(0.004)	(0.004)
HMO	0.010	0.037^{*}	-0.155^{***}	-0.187^{***}
	(0.064)	(0.023)	(0.011)	(0.010)
Trend	-0.027^{***}	-0.010^{***}	0.015^{***}	0.019^{***}
	(0.008)	(0.003)	(0.002)	(0.002)

Notes: Table summarizes the adaptive learning parameter estimates $\widehat{\boldsymbol{\theta}}_t$ for $t \in \{2016, 2017, 2018\}$ and the full information estimates $\widehat{\boldsymbol{\theta}}$. Robust standard errors are in parentheses (*** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level). We compute the household-specific monthly premium and previous choice parameters for each household using the demographic interaction terms and report an average across all households in this table. The raw parameter estimates are available in Table A2.

ing additional parameters. Table II compares these measures for eight alternative specifications, depending on firm knowledge of the relevant parameters. Specification (1) is the adaptive learning model where none of the parameters are known; we use the adaptive learning estimates $\widehat{\theta}_{2016}, \widehat{\theta}_{2017}$, and $\widehat{\theta}_{2018}$ to compute plan average costs for 2016, 2017, and 2018, respectively. Specification (8) is the full information model where firms know all parameters; we use the full information estimates $\widehat{\theta}$ to compute plan average costs for 2016, 2017, and 2018. Specifications (2)-(7) are the intermediate models where we assume firms know a subset of the parameters and estimate the other model parameters conditional on this knowledge. For example, specification (4) assumes firms know the full information demand parameters $\boldsymbol{\beta}$ and learn about the risk score parameters $\boldsymbol{\gamma}$ and claims parameters $\boldsymbol{\mu}$, conditional on knowing $\boldsymbol{\beta}$.

For nearly all measures, the model that allows firms to learn about all model parameters (Specification 1) performs better than the model that assumes firms have full information (Specification 8). The MAE declines by 19.2% from 158 to 128 and the RMSE declines by 14.1% from 217 to 186. There is also a substantial improvement in the R^2, which increases from 0.18 to 0.40. Although specification (1) has triple the number of parameters as the full information model, the adjusted R^2 is still considerably higher and the AIC lower than in the full information model. However, the BIC which imposes the largest penalty for adding parameters is lower for the full information model.

Improvements in fit relative to the full information model are also substantial for Specifications (2)-(5), which allow firms to learn about the claims parameters $\boldsymbol{\mu}$. Relative to the full information model, all four specifications have lower MAE, RMSE, and AIC values than the full information model, as well as higher R^2 and Adjusted R^2 values. Specifications (4) and (5) also have lower BIC values than the full information model. In contrast, specifications (6) and (7) which assume firms know the full information claims parameters have a worse fit than the full information model. In sum, these results indicate that it is particularly important to allow firms to learn about the relationship between demand and cost in equation (22).

The results in Table II suggest the adaptive learning model provides a better fit of the data than the full information model. To formalize this finding, we construct a statistical test similar to Doraszelski et al. (2018). Define the difference in absolute errors (DAE_{jmt}) as

$$DAE_{jmt} \quad = \quad \left|\zeta_{jmt}(\widehat{\mathbf{q}}_t; \widehat{\boldsymbol{\theta}}_t) - \zeta_{jmt}\right| - \left|\zeta_{jmt}(\mathbf{p}_t; \widehat{\boldsymbol{\theta}}) - \zeta_{jmt}\right| \tag{28}$$

where $\zeta_{jmt}(\widehat{\mathbf{q}}_t; \widehat{\boldsymbol{\theta}}_t) \equiv (1 - \iota_{ft})c_{jmt}(\widehat{\mathbf{q}}_t; \widehat{\boldsymbol{\theta}}_t) - ra_{jmt}(\widehat{\mathbf{q}}_t; \widehat{\boldsymbol{\theta}}_t)$ is the adaptive learning estimate of plan cost, $\zeta_{jmt}(\mathbf{p}_t; \widehat{\boldsymbol{\theta}}) \equiv (1 - \iota_{ft})c_{jmt}(\mathbf{p}_t; \widehat{\boldsymbol{\theta}}) - ra_{jmt}(\mathbf{p}_t; \widehat{\boldsymbol{\theta}})$ is the full information estimate of plan cost, and ζ_{jmt} is the average cost forecast in the firms' rate filings. In Table III, we show the results of regressing DAE_{jmt} on a constant for Specifications (1)-(5) in Table II. The -40.1 point

Table II: Model Goodness-of-Fit

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Known Parameters								
Switching Cost (β_i^y)		✓		✓	✓			✓
Premium (β_i^p)			✓	✓	✓			✓
Other Demand (β_i^x)				✓	✓			✓
Risk ($\boldsymbol{\gamma}$)					✓		✓	✓
Claims ($\boldsymbol{\mu}$)						✓	✓	✓
Fit Measure								
MAE	127.9	126.9	129.9	127.0	125.9	164.5	166.0	158.4
RMSE	186.1	182.5	189.2	187.6	187.0	217.8	227.2	216.6
R2	0.40	0.42	0.37	0.39	0.39	0.17	0.10	0.18
Adjusted R^2	0.27	0.32	0.27	0.32	0.33	0.04	-0.03	0.13
AIC	19252	19098	19251	19021	18977	19709	19828	19398
BIC	20882	20498	20716	19938	19806	21053	21084	19941

Notes: Table shows how well the alternative models' predictions of plan cost fit the firms' predictions of plan cost. Six goodness-of-fit measures are shown: (1) mean absolute error (MAE); (2) root mean square error (RMSE); (3) R^2; (4) Adjusted R^2; (5) Akaike information criterion (AIC); and (6) Bayesian information criterion (BIC). To compute the six measures, we use the plan shares as weights. The first panel indicates which parameters are known to the firm for each of the eight models. Specification (1) corresponds to the model where firms must learn all model parameters, whereas specification (8) corresponds to the full information model.

estimate for the constant parameter in Specification (1) is exactly the difference between the MAE for Specifications (1) and (8) in Table II (i.e., $127.9 - 158.4 = -30.5$). We find that this point estimate is highly statistically significant, indicating the substantial improved fit of the adaptive learning model relative to the full information model. Specifications (2)-(5) also yield statistically significant improvements in fit compared to the full information model.

In Specifications (1')-(5') in Table III, we regress DAE_{jmt} on metal tier dummies and a time trend. We find the improvement in fit of the adaptive learning model relative to the full information model is increasing in plan generosity. This result is consistent with firms having greater difficulty in predicting who will select into the more generous plans and how much they will cost. The extent of moral hazard also presents a challenge for predicting cost for the more generous plans. Another key finding is that the improvement in fit declines over time, as indicated by the positive parameter estimate on the time trend term. The advantages of using the adaptive learning model are therefore largest in the ACA's initial years when firms had more limited data available for making predictions.

Table III: Comparison of Alternative Learning Models with Full Information Model

Model	(1)	(1′)	(2)	(2′)	(3)	(3′)
Known Parameters						
Switching Cost (β_i^y)			✓	✓		
Premium (β_i^p)					✓	✓
Other Demand (β_i^x)						
Risk (γ)						
Claims ($\boldsymbol{\mu}$)						
Constant	-30.515^{***}	-22.395^{**}	-31.511^{***}	-20.102	-28.472^{***}	-18.789^{*}
	(2.871)	(10.611)	(3.242)	(12.204)	(2.861)	(10.745)
Silver		-39.724^{***}		-37.886^{***}		-39.207^{***}
		(5.906)		(6.793)		(5.981)
Gold		-49.447^{***}		-51.341^{***}		-42.789^{***}
		(9.697)		(11.153)		(9.820)
Platinum		-92.261^{***}		-100.260^{***}		-85.285^{***}
		(10.489)		(12.064)		(10.622)
Time Trend		8.298^{***}		7.129^{*}		7.295^{**}
		(3.165)		(3.640)		(3.205)

Model (Cont.)	(4)	(4′)	(5)	(5′)
Known Parameters				
Switching Cost	✓	✓	✓	✓
Premium	✓	✓	✓	✓
Other Demand	✓	✓	✓	✓
Risk			✓	✓
Cost				
Constant	-31.426^{***}	-25.775^{***}	-32.509^{***}	-15.366^{*}
	(2.617)	(9.398)	(2.548)	(9.192)
Silver		-42.693^{***}		-46.371^{***}
		(5.231)		(5.116)
Gold		-49.841^{***}		-55.540^{***}
		(8.589)		(8.400)
Platinum		-88.561^{***}		-79.110^{***}
		(9.291)		(9.087)
Time Trend		9.576^{***}		6.319^{**}
		(2.803)		(2.741)

Notes: Table shows regression results that compare the fit of the first five alternative adaptive learning models in Table II to the fit of the full information model (Specification 8 in Table II). In all regressions, the dependent variable is the difference in absolute errors as defined in equation (28). A negative coefficient indicates the alternative adaptive learning model has a lower mean absolute error than the full information model. For each of the five models, we consider two specifications: (1) regressing the difference in absolute errors on a constant and (2) regressing the difference in absolute errors on a constant, metal tier, and linear time trend. Plans are weighted by their market shares.

6 Impact of Information

6.1 Simulation Methodology

In this section, we simulate the impact of information on the estimated model equilibrium. We do this by (1) replacing the firms' learning estimates $\widehat{\theta}_t$ for $t \in \{2016, 2017, 2018\}$ with the full information parameters (i.e., the last column in Table I); (2) solving for the new vector of premiums that satisfy the firms' first-order conditions in equation (12); and (3) computing several measures of the new equilibrium outcome, including average premiums, enrollment, and social welfare. Social welfare consists of consumer surplus CS_t, total firm profit π_t, and net government spending GS_t. for every year t. We compute total consumer surplus

$$CS_t \quad = \quad -\sum_{i \in I} \frac{1}{\beta^p_{it}} \ln \left(\sum_{j \in J} \exp\left(V_{ijt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t)/\lambda_t\right)^{\lambda_t} + \exp\left(\beta^p_{it}\rho_{it}\right) \right) + \sum_{j \in J} \left[q_{ijt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\beta}_t) * \frac{\beta^y_{ijt} * y_{ij(t-1)}}{\beta^p_{it}} \right] \qquad (29)$$

where the first term of equation (29) is the standard nested logit formula for consumer surplus and the second term "corrects" the first term to remove gains in welfare that result from inertia. Total firm profit is $\pi_t = \sum_{t \in T} E[\pi_{ft}(\widehat{\mathbf{q}}_t, \widehat{\mathbf{c}}_t; \boldsymbol{\theta}_t)]$, where $E[\pi_{ft}(\widehat{\mathbf{q}}_t, \widehat{\mathbf{c}}_t; \boldsymbol{\theta}_t)]$ is defined in equation (11). Net government spending GS_t equals the sum of spending on premium subsidies, CSRs, and uncompensated care for the uninsured minus revenue collected from the mandate penalty. Premium subsidy spending is the sum of subsidies received by each consumer in equation (4). Spending on CSRs is computed as

$$CSR_t \quad = \quad \sum_{i \in I, j \in J} s^g_j q_{ijt}(\mathbf{p}_t, \mathcal{I}_t; \boldsymbol{\theta}_t) c_{jmt}(\widehat{\mathbf{q}}_t; \boldsymbol{\theta}_t)$$

where s^g_j is the expected share of claims paid by the government for plan j.[11] We calculate spending on uncompensated care by multiplying the number of uninsured that we estimate in each scenario by \$2,025, the estimated annual uncompensated care cost per uninsured[12], and a factor accounting for the change in the uninsured population's risk score. Penalty revenue collected by the government equals $\sum_{i \in I} q_{i0t}\rho_{it}$, where q_{i0t} is the household's probability of choosing the outside option.

[11] Ignoring moral hazard, the government's expected outlay is $94 - 70 = 24\%$ of claims for the 94% CSR plan, $87 - 70 = 17\%$ of claims for the 87% CSR plan, and $73 - 70 = 3\%$ of claims for the 73% CSR plan. To account for moral hazard, we follow Pope et al. (2014) and assume there is no moral hazard for consumers in the 73% plan, while consumers in the 87% and 94% plans increase consumption by 12%. Including moral hazard, the $s^g_j = 26.88\%$ for the 94% CSR plan, $s^g_j = 19.04\%$ for the 87% CSR plan, and $s^g_j = 3\%$ for the 73% CSR plan.

[12] We multiply the per-capita amount of medical costs that are paid on behalf of the nonelderly uninsured as estimated by Coughlin et al. (2014) by an inflation factor using data from the National Health Expenditure Accounts to adjust the estimates to the timeframe of this study (Centers for Medicare and Medicaid Services, 2018).

6.2 Impact of Information

Figure 5 summarizes the impact of information over time. We report the impact of information on average unsubsidized premiums (Figure 5a), average subsidized premiums (Figure 5b), the change in average unsubsidized premiums by metal tier (Figure 5c), the change in plan market share by metal tier (Figure 5d), total exchange enrollment (Figure 5e), and the change in annual per-capita social welfare (Figure 5f).

Overall, our results indicate that firms can benefit with full information at taxpayers' expense. Using the full information parameters increases average unsubsidized premiums in 2016 by 6.9% from $396 to $424. The percentage increase in average unsubsidized premiums falls to 3.4% in 2017 and only 0.1% in 2018. Conversely, using the full information parameters reduces average subsidized premiums by 1.6% in 2016 and by 0.7% in 2017, but slightly increases average subsidized premiums by 2.6% in 2018. The opposite impact of full information on unsubsidized and subsidized premiums is the result of the ACA's unique usage of silver premiums to determine subsidies. Figure 5c shows full information increases average silver premiums by slightly more than the overall unsubsidized premium, leading to a larger subsidy that can be applied towards the purchase of any plan. The impact of full information on total exchange enrollment is negative, but small.

Figure 5f shows the impact of full information on annual per-capita social welfare. Firms earn substantially higher annual per-capita profit of $347 in 2016 when firms know the full information parameters. However, producer surplus gains are mostly offset by an increase of $274 in annual per-capita government spending, primarily on premium subsidies. Consumer surplus is largely unchanged because premium subsidies increase by the same amount as premiums increase, as shown in equation (4).[13] Over time, the increases in both profit and government spending that result from learning the full information parameters moderate.

Our finding that firms set premiums below profit-maximizing levels during the ACA's initial years is consistent with firms overestimating premium sensitivity, underestimating switching costs, and underestimating claims risk. Markups are lower in a market where profit-maximizing firms believe that consumers are more premium sensitive than they actually are. Similarly, switching costs are an important source of market power that firms exploit in this market to set higher premiums (Saltzman et al., 2021). Underestimating switching costs leads firms to set lower premiums than they otherwise would. Similarly, underestimating claims risk results in firms setting lower premiums than if they had known their true higher costs. If firms had instead underestimated premium sensitivity, overestimated switching costs, and overestimated claims risk, they would have been

[13]This assumes the subsidy is strictly positive before the premium increase.

incentivized to set higher premiums.

7 Policy Simulations

In this section, we examine the interaction of firm knowledge with policy. We focus on the interaction of information with community rating because it explicitly restricts firms from using certain consumer information to set premiums. Firms cannot consider a consumer's health history or gender, and limitations are placed on using age according to the ACA's modified community rating rules. Community rating prevents firms from accurately pricing risk and exacerbates adverse selection. ACA policies such as the individual mandate and premium subsidies attempt to mitigate the effects of adverse selection that result from community rating.

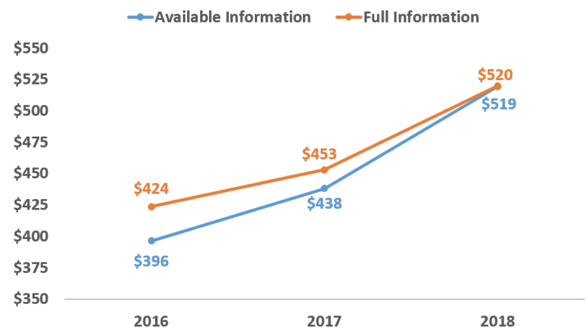
The ACA's modified community rating rules allow firms to set premiums by age and geography. We simulate two changes to modified community rating: (1) relaxing restrictions on age and gender rating and (2) requiring firms to charge all consumers the same premium, regardless of age and geography (i.e., pure community rating). We allow firms to rate consumers by both age and gender without restriction by replacing the ACA age rating factors with the age-gender cost factors in Figure 3a; we then solve for the new vector of premiums that satisfy the firms' first-order conditions. We simulate pure community rating by setting all household rating factors to 1 and solving for the new vector of premiums that satisfy the firms' first-order conditions. We run all simulations using both the 2016 parameters and the full information parameters.

Figure 6 summarizes our results. We report the impact on average unsubsidized premiums (panel a), average subsidized premiums (panel b), total exchange enrollment (panel c), average claims (panel d), annual per-capita profit (panel e) and annual per-capita net government spending (panel f). Our results indicate: (1) the equilibrium outcome improves when community rating is relaxed and worsens when pure community rating is required and (2) prohibiting firms from using certain information to set premiums makes them react more to the information they can use. Hence, the impact of information is generally largest in a setting with pure community rating where firms must charge all consumers the same premium. We discuss each of these two main findings in detail below.

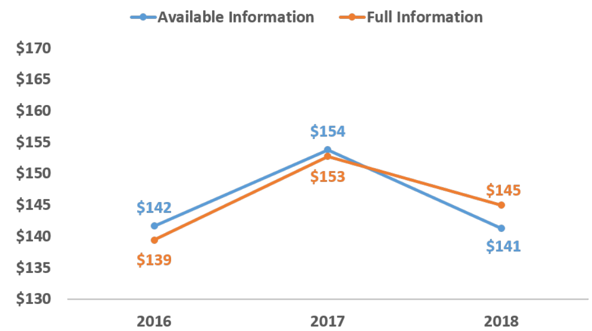
7.1 Effect of Community Rating

We assess the impact of community rating under 2016 information using the blue columns in Figure 6. Consistent with theory, we find relaxing community rating reduces average premiums and

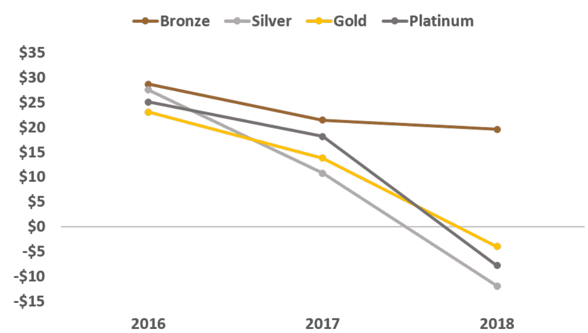
Figure 5: Impact of Assuming Full Information By Year



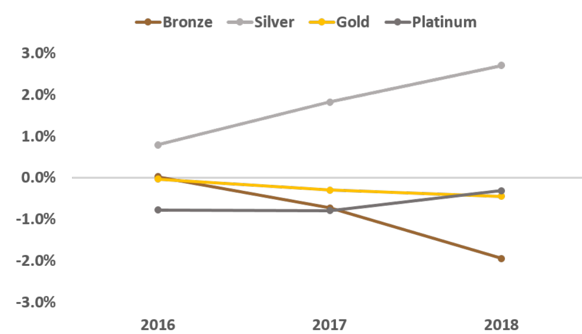
(a) Average Unsubsidized Premiums



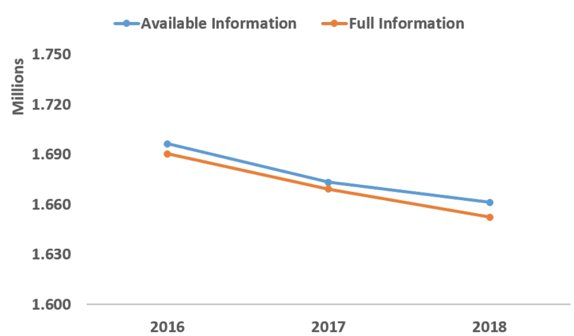
(b) Average Subsidized Premiums



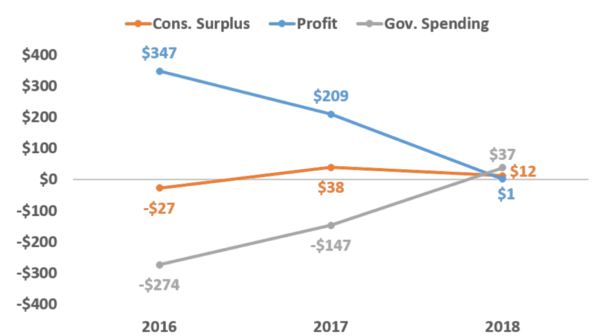
(c) Change in Unsubsidized Premiums By Metal



(d) Change in Market Share By Metal



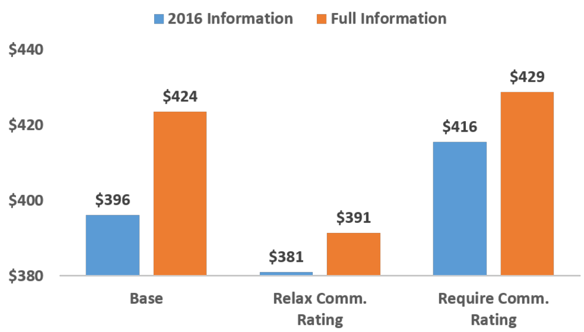
(e) Total Exchange Enrollment



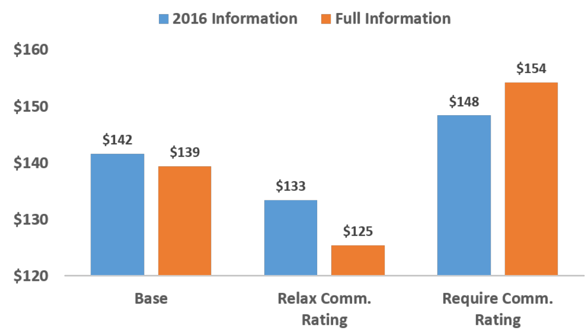
(f) Change in Annual Per-Capita Welfare

Notes: Figure shows the equilibrium impact of using the full information parameters instead of the learning parameters by year. Panel (a) shows the impact on average unsubsidized premiums, panel (b) shows the impact on average subsidized premiums, panel (c) shows the change in average unsubsidized premiums by metal tier, panel (d) shows the change in market share by metal tier, panel (e) shows the impact on total exchange enrollment, and panel (f) shows the change in annual per-capita welfare when switching to the full information parameters. An increase in government spending is shown having a negative welfare cost.

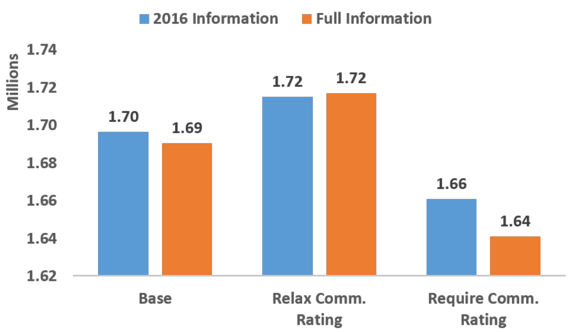
Figure 6: Policy Impact of Information



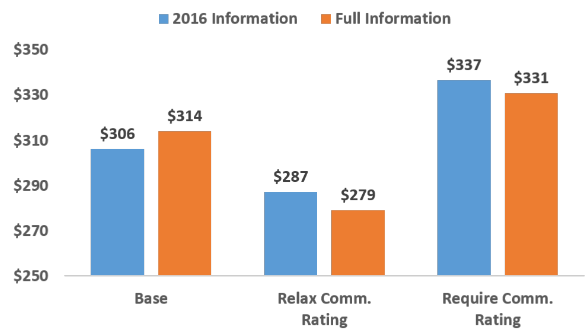
(a) Average Unsubsidized Premiums



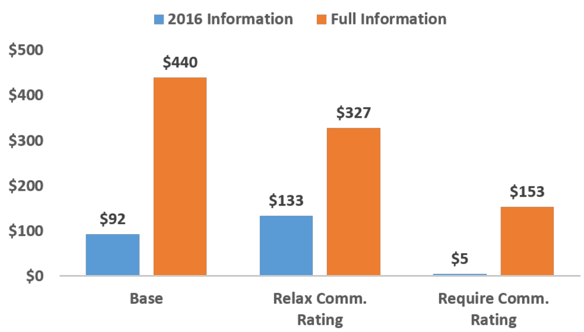
(b) Average Subsidized Premiums



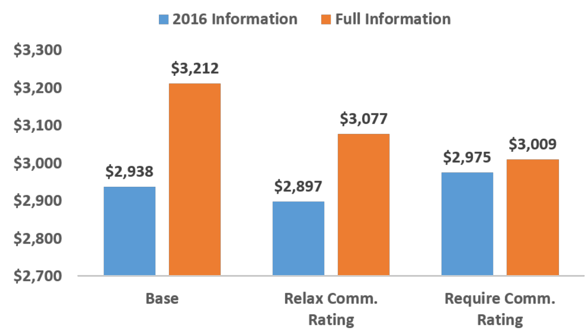
(c) Total Exchange Enrollment



(d) Average Claims



(e) Annual Per-Capita Profit



(f) Annual Per-Capita Gov. Spending

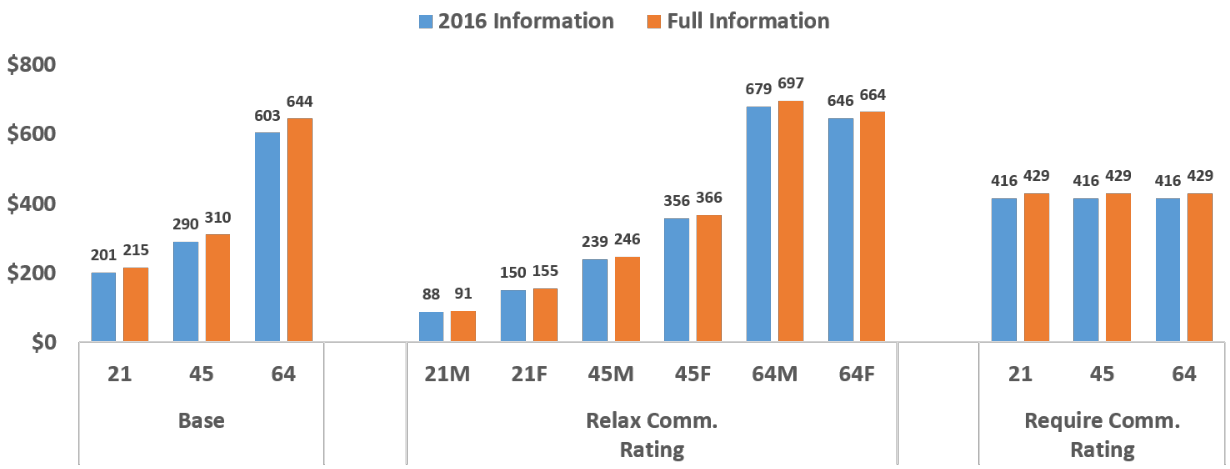
Notes: Figure shows the impact of using the full information parameters instead of the 2016 learning parameters in three different settings: (1) Base (modified community rating under the ACA); (2) relaxed community rating; and (3) pure community rating. The figures show the equilibrium impact on average unsubsidized premiums (panel a), average subsidized premiums (panel b), total exchange enrollment (panel c), average claims (panel d), annual per-capita profit (panel e), and annual per-capita net government spending (panel f).

requiring pure community rating increases average premiums. When community rating is relaxed, average unsubsidized premiums decrease 3.8%. Requiring pure community rating results in average unsubsidized premiums increasing 4.9%. These effects would likely be larger in a setting without policies such as the mandate and premium subsidies that are designed to mitigate selection. Relaxing community rating leads to average subsidized premiums decreasing 5.8%. Average subsidized premiums increase 4.8% when pure community rating is required. These premium changes have the expected impact on total exchange enrollment. When community rating is relaxed, total exchange enrollment increases 1.1%. Requiring pure community rating leads to a 2.1% reduction in enrollment. Figure 6d indicates that community rating has a substantial impact on the average enrollee. Relaxing community rating reduces average claims by 6.2%. Requiring pure community rating increases average claims by 9.9%. These results suggest that community rating exacerbates adverse selection.

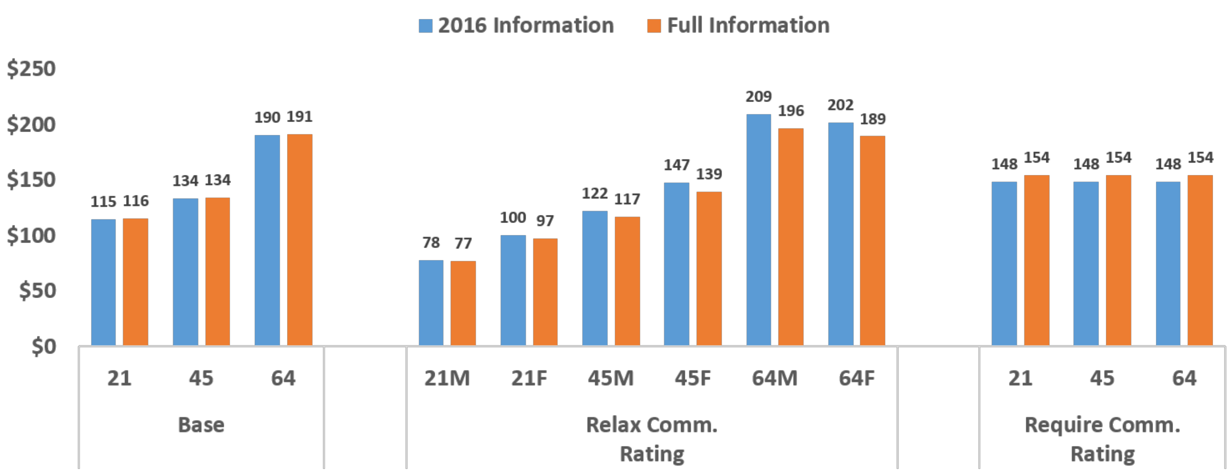
Figures 6e and 6f show the impact of community rating regulation on annual per-capita profit and annual per-capita net government spending, respectively. Annual per-capita profit increases $347 or 43.9% when community rating is relaxed, but decreases $147 or 96.2% when community rating is required. The effect of community rating regulation on government spending is smaller; annual per-capita net government spending decreases $40 or 1.4% when community rating is relaxed and increases $37 or 1.2% when community rating is required.

Figure 7 indicates community rating has a heterogeneous impact across age and gender. Relaxing community rating provides the most benefit to 21-year-old males, who pay 56.2% less in unsubsidized premiums and 32.2% less in subsidized premiums. Relaxing community rating is most detrimental to 64-year-old males, who pay 12.5% more in unsubsidized premiums and 9.8% more in subsidized premiums. Therefore, relaxing community rating creates winners (young adults, especially males) and losers (older adults, especially males). However, the premium reductions realized by the winners are larger in magnitude than the premium increases realized by the losers. Requiring community rating has the reverse impact. Relative to the Base scenario, 64-year-olds pay 31.1% less in average unsubsidized premiums and 22.1% less in average subsidized premiums. In contrast, 21-year-olds pay 106.7% more in average unsubsidized premiums and 29.4% more in average subsidized premiums. The winners under pure community rating are older adults and the losers are young adults, although the premium reductions realized by the winners are smaller in magnitude than the premium increases realized by the losers.

Figure 7: Effect of Information by Age and Gender



(a) Average Unsubsidized Premiums by Age and Gender



(b) Average Subsidized Premiums by Age and Gender

Notes: Figure shows the impact of using the full information parameters instead of the 2016 learning parameters on average unsubsidized premiums (panel a) and average subsidized premiums (panel b) in three different settings (1) base, where premiums only vary by age; (2) relaxed community rating, where premiums vary by both age and gender without restriction; and (3) pure community rating, where premiums do not vary with age or gender. We compute average premiums using enrollee weights for 21-year-olds, 45-year-olds, and 64-year-olds by gender (if applicable).

7.2 Impact of Information in Alternative Settings

Now we assess the impact of information by comparing the blue and orange bars in Figure 6 in three alternative settings: (1) modified community rating under the ACA (Base setting); (2) relaxed community rating; and (3) pure community rating. Figure 6a indicates that learning the full information parameters increases average unsubsidized premiums in all three settings. The premium increase is smaller in the setting with relaxed community rating (2.7%) than the setting with pure community rating (3.2%). Learning reduces average subsidized premiums in the base and relaxed community rating setting, but increases them in the pure community rating setting. Overall, the premium impacts are generally smallest with relaxed community rating and largest with pure community rating. The impact of information on enrollment is also largest with pure community rating. Figure 6c indicates that learning the full information parameters reduces total exchange enrollment by 1.20% when pure community rating is required. In contrast, learning very slightly increases total exchange enrollment by 0.11% in the relaxed community rating setting and decreases total exchange enrollment by 0.36% in the modified community rating setting.

Figures 6e and 6f show the impact of information on annual per-capita profit and annual per-capita net government spending, respectively. Under 2016 information, annual per-capita profit is nearly zero in the pure community rating setting. Hence, pure community rating restrictions and firm uncertainty combine to nearly eliminate firm profits. Learning the full information parameters in the pure community rating setting increases annual per-capita profit to $153, which is still substantially less than in the modified community rating or relaxed community rating settings. Two offsetting forces drive changes in annual per-capita net government spending: (1) higher average premiums and (2) lower exchange enrollment. Annual per-capita net government spending increases the least when firms learn the full information parameters in the pure community rating setting because total exchange enrollment falls so much in this setting.

Taken together, our results indicate that learning the full information parameters has the greatest welfare impact when pure community rating is in place and the smallest impact when community rating is relaxed. Pure community rating prohibits firms from using consumer-specific information to rate consumers, including health history, age, and gender. Barred from using this information, firms react more to the information they do have about consumer sensitivity to premiums, switching costs, and claims risk.

8 Conclusion

Large-scale social programs, including recent public health insurance expansions, are increasingly being implemented by creating new markets with private sector participation. In new markets, the standard IO assumptions of market equilibrium and complete information might be unrealistic (Doraszelski et al., 2018). We study the effects of relaxing these standard assumptions by estimating an adaptive learning model in a selection market using data from the California ACA exchange. Firms initially faced considerable uncertainty in predicting who would enroll and how much their enrollees would cost. Our setting is appealing because we observe the creation of a new market and can exploit data on firms' predictions about their costs, as well as their actual costs.

A novel feature of our study is the ability to assess the external validity of alternative models using data on the firms' predictions of costs. We find that the adaptive learning model provides a statistically significant improvement in fit relative to the standard IO model that assumes firms have full information. Most of the improvement in fit results from allowing firms to learn about the relationship between demand and cost. Hence, models of new markets may need to accommodate firm learning about the demand-cost relationship, particularly in the initial years of the market.

Because firms initially overestimated premium sensitivity, underestimated switching costs, and underestimated claims risk, they set premiums below the full information profit-maximizing levels. With full information, firms earn more profit, but at taxpayers' expense. We also study the interaction of learning with community rating regulation. We find the impact of information is largest in a setting with pure community rating. Prohibiting firms from using certain consumer information to rate consumers therefore makes them react more to the information they can use. Creating new insurance markets with community rating regulation increases the impact of firm uncertainty, especially in the initial years of the market.

Our study can be extended in several directions. One useful extension may consider both consumer and firm learning. Prior work in other settings suggests consumers may also learn and adjust their plan choices accordingly from year to year (Ketcham et al., 2012; Miravete, 2003; List, 2003, 2004, 2006; List and Millimet, 2008). A significant feature of the ACA exchanges is high consumer churn due to exogenous reasons, such as job status changes or substantial income shocks. We also find minimal evidence of consumers switching plans despite highly volatile premiums during our study period. Hence, consumers have limited opportunities and incentives to learn in our study setting. Similar to Doraszelski et al. (2018), our model is not fully dynamic and does not allow firms to be forward-looking when setting premiums. Given the existing empirical tools, it is unclear how to estimate time-specific learning parameters and consistent equilibrium beliefs in a model of learning. Even in models without learning, modeling forward-looking behavior in health insurance markets is particularly challenging and would require significant compromises on key institutional details (Fleitas, 2017; Miller, 2019).

Our study has important implications for policy. Forecasts of large-scale social programs by the CBO and similar organizations may improve when considering the impact of information. The CBO forecast of the ACA's impact substantially overestimated premiums and subsidy spending. Given our study findings, the CBO's forecast would have projected lower (and more accurate) premiums and subsidy spending if it had accounted for firm uncertainty. Policymakers might also consider alternatives to community rating, which we find exacerbates firm uncertainty, when starting a new market. Expanded premium subsidies or reinsurance are alternatives that would protect high-risk consumers from price discrimination.

We expect the establishment of new insurance markets to be an increasingly important mechanism for expanding access to health insurance and reducing health care costs, especially under recent proposals to transform Medicare into a premium support or defined contribution program. The methods used in this paper might be useful for analyzing the potential impact of these markets and the impact of proposed regulation while firms are still learning.

References

Abaluck, J. and J. Gruber (2011, May). Heterogeneity in choice inconsistencies among the elderly: Evidence from prescription drug choice. *The American Economic Review 101*(3), 377–381.

Abaluck, J. and J. Gruber (2016). Evolving choice inconsistencies in choice of prescription drug insurance. *American Economic Review 106*(8), 2145–2184.

Ackerberg, D. (2003). Advertising, learning, and consumer choice in experience good markets: An empirical examination. *International Economic Review 44*(3), 1007–1040.

Aguirregabiria, V. and J. Jeon (2020). Firms' beliefs and learning: Models, identification, and empirical evidence. *Review of Industrial Organization 56*, 203–235.

Benkard, C. L. (2000). Learning and forgetting: The dynamics of aircraft production. *The American Economic Review 90*(4), 1034–1054.

Centers for Medicare and Medicaid Services (2018, January). *National Health Expentire Data*. `https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html`.

Congressional Budget Office (2017, December). *CBO's Record of Projecting Subsidies for Health Insurance Under the Affordable Care Act: 2014 to 2016*. https://www.cbo.gov/system/files/115th-congress-2017-2018/reports/53094-acaprojections.pdf.

Conley, T. and C. Udry (2010). Learning about a new technology: Pineapple in ghana. *The American Economic Review 100*(1), 35–69.

Coughlin, T., J. Holahan, K. Caswell, and M. McGrath (2014, May). *Uncompensated Care for Uninsured in 2013: A Detailed Examination*. https://kaiserfamilyfoundation.files.wordpress.com/2014/05/8596-uncompensated-care-for-the-uninsured-in-2013.pdf.

Curto, V., L. Einav, J. Levin, and J. Bhattacharya (2020, June). Can health insurance competition work? evidence from medicare advantage. Accepted, Journal of Political Economy.

Decarolis, F., M. Polyakova, and S. Ryan (2020, May). Subsidy design in privately provided social insurance: Lessons from medicare part d. *Journal of Political Economy 128*(5).

Department of Managed Health Care (2016). *Premium Rate Review Filings*. http://wpso.dmhc.ca.gov/ratereview/.

Dickstein, M. (2018). Efficient provision of experience goods: Evidence from antidepressant choice.

Doraszelski, U., G. Lewis, and A. Pakes (2018). Just starting out: Learning and equilibrium in a new market. *The American Economic Review 108*(3), 565–615.

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

Einav, L., A. Finkelstein, and P. Tebaldi (2019). Market design in regulated health insurance markets: Risk adjustment vs. subsidies.

Ericson, K. and A. Starc (2015). Pricing regulation and imperfect competition on the massachusetts health insurance exchange. *Review of Economics and Statistics 97*(3), 667–682.

Evans, G. and S. Honkapohja (2001). *Learning and Expectations in Macroeconomics*. Princeton University Press.

Finkelstein, A., N. Hendren, and M. Shepard (2019). Subsidizing health insurance for low-income adults: Evidence from massachusetts. *American Economic Review 109*(4).

Fleitas, S. (2017, May). Dynamic competition and price regulation when consumers have inertia: Evidence from medicare part d. Working paper.

Geruso, M., T. Layton, G. McCormack, and M. Shepard (2019). The two margin problem in insurance markets.

Gruber, J. (2017). Delivering public health insurance through private plan choice in the united states. *Journal of Economic Perspectives 31*(4), 3–22.

Hackmann, M., J. Kolstad, and A. Kowalski (2015, March). Adverse selection and an individual mandate: When theory meets practice. *The American Economic Review 105*(3), 1030–1066.

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

Hortaçsu, A., F. Luco, S. L. Puller, and D. Zhu (2019). Does strategic ability affect efficiency? evidence from electricity markets. *American Economic Review 109*(12), 4302–42.

Hortaçsu, A. and S. L. Puller (2008). Understanding strategic bidding in multi-unit auctions: a case study of the texas electricity spot market. *The RAND Journal of Economics 39*(1), 86–114.

Huang, Y., P. Ellickson, and M. Lovett (2021). Learning to set prices. Working paper.

Jaffe, S. and M. Shepard (2020). Price-linked subsidies and imperfect competition in health insurance. *American Economic Journal: Economic Policy 12*(3), 279–311.

Jeon, J. (2020). Learning and investment under demand uncertainty in container shipping.

Joskow, P. L., R. Schmalensee, and E. M. Bailey (1998). The market for sulfur dioxide emissions. *American Economic Review*, 669–685.

Kaiser Family Foundation (2020, April). *State Health Facts: Health Reform*. `https://www.kff.org/statedata/`.

Kessler, D. and M. McClellan (2000). Is hospital competition socially wasteful? *Quarterly Journal of Economics 115*, 577–615.

Ketcham, J., C. Lucarelli, E. Miravete, and C. Roebuck (2012). Sinking, swimming, or learning to swim in medicare part d. *American Economic Review 102*(6), 2639–2673.

Ketcham, J., C. Lucarelli, and C. Power (2015). Paying attention or paying too much in medicare part d. *American Economic Review 105*(1), 204–233.

Layton, T. (2017). Imperfect risk adjustment, risk preferences, and sorting in competitive health insurance markets. *Journal of Health Economics*.

List, J. A. (2003). Does market experience eliminate market anomalies? *The Quarterly Journal of Economics 118*(1), 41–71.

List, J. A. (2004). Neoclassical theory versus prospect theory: Evidence from the marketplace. *Econometrica 72*(2), 615–625.

List, J. A. (2006). The behavioralist meets the market: Measuring social preferences and reputation effects in actual transactions. *Journal of political Economy 114*(1), 1–37.

List, J. A. and D. L. Millimet (2008). The market: Catalyst for rationality and filter of irrationality. *The BE Journal of Economic Analysis & Policy 8*(1).

Lucarelli, C., J. Prince, and K. Simon (2012). The welfare impact of reducing choice in medicare part d: A comparison of two regulation strategies. *International Economic Review 53*(4), 1155–1177.

Lustig, J. (2009). Measuring welfare losses from adverse selection and imperfect competition in privatized medicare.

Mahoney, N. and E. G. Weyl (2017). Imperfect competition in selection markets. *The Review of Economics and Statistics 99*(4), 637–651.

Miller, K. (2019, June). Estimating costs when consumers have inertia: Are private medicare insurers more efficient? Working paper.

Miller, K., A. Petrin, R. Town, and M. Chernew (2019, February). Optimal managed competition subsidies. Working paper.

Miravete, E. J. (2003, March). Choosing the wrong calling plan? ignorance and learning. *American Economic Review 93*(1), 297–310.

Morrisey, M. A., A. M. Rivlin, R. P. Nathan, and M. A. Hall (2017). Five-state study of aca marketplace competition: A summary report. *Risk Management and Insurance Review 20*(2), 153–172.

Newberry, P. (2016). An empirical study of observational learning. *RAND Journal of Economics 47*(2), 394–432.

Pauly, M., S. Harrington, and A. Leive (2015). "sticker shock" in individual insurance under health reform? *American Journal of Health Economics 1*(4), 494–514.

Pauly, M., A. Leive, and S. Harrington (2020). Losses (and gains) from health reform for non-medicaid uninsureds. *Journal of Risk and Insurance 87*(1), 41–66.

Petrin, A. and K. Train (2010). A control function approach to endogeneity in consumer choice models. *Journal of marketing research 47*(1), 3–13.

Polyakova, M. and S. Ryan (2021). Subsidy targeting with market power. *The National Bureau of Economic Research*.

Pope, G., H. Bachofer, A. Pearlman, J. Kautter, E. Hunter, D. Miller, and P. Keenan (2014). Risk transfer formula for individual and small group markets under the affordable care act. *Medicare and Medicaid Research Review 4*(3), 1–46.

Ruggles, S., S. Flood, R. Goeken, M. Schouweiler, and M. Sobek (2022, February). *IPUMS USA: Version 12.0 [dataset]*.

Saltzman, E. (2019). Demand for health insurance: Evidence from the california and washington aca exchanges. *Journal of Health Economics 63*, 197–222.

Saltzman, E. (2021). Managing adverse selection: Underinsurance vs. underenrollment. *The RAND Journal of Economics 52*(2), 359–381.

Saltzman, E., A. Swanson, and D. Polsky (2021, July). Inertia, market power, and adverse selection in health insurance: Evidence from the aca exchanges. Working Paper 29097, National Bureau of Economic Research.

Sargent, T. (1993). *Bounded Rationality in Macroeconomics*. Oxford University Press.

Starc, A. (2014). Insurer pricing and consumer welfare: Evidence from medigap. *RAND Journal of Economics 45*(1), 198–220.

Tebaldi, P. (2022). Estimating equilibrium in health insurance exchanges: Price competition and subsidy design under the aca.

Town, R. and S. Liu (2003). The welfare impact of medicare hmos. *The RAND Journal of Economics 34*(4), 719–736.

Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge: Cambridge University Press.

Yamamoto, D. (2013, June). Health care costs - from birth to death. *Society of Actuaries*. `https://www.cms.gov/CCIIO/Programs-and-Initiatives/Premium-Stabilization-Programs/Downloads/June-30-2016-RA-and-RI-Summary-Report-5CR-063016.pdf`.

Zhang, J. (2010). The sound of silence: Observational learning in the u.s. kidney market. *Marketing Science 29*(2), 315–335.

A Summary Statistics

Table A1: Demographic Distribution By Year

	2014	2015	2016	2017	2018	2019	Overall
Market Size	1,980,628	2,060,535	2,066,938	2,087,663	2,121,503	2,016,462	12,333,730
Total Enrollment	1,362,316	1,639,923	1,702,160	1,697,074	1,710,469	1,553,374	9,665,316
Income							
138% FPL or less	4.7%	3.5%	3.3%	4.0%	4.0%	3.5%	3.8%
138% FPL to 150% FPL	14.1%	14.3%	14.6%	14.7%	14.4%	14.0%	14.4%
150% FPL to 200% FPL	32.8%	32.8%	31.9%	30.3%	28.8%	28.4%	30.8%
200% FPL to 250% FPL	16.8%	16.7%	16.3%	16.3%	16.7%	16.7%	16.6%
250% FPL to 400% FPL	22.4%	23.4%	23.6%	23.6%	25.8%	27.4%	24.4%
400% FPL or greater	9.3%	9.3%	10.3%	11.0%	10.3%	9.9%	10.0%
Subsidy Status							
Subsidized	89.6%	88.8%	87.5%	86.5%	87.3%	87.7%	87.8%
Unsubsidized	10.4%	11.2%	12.5%	13.5%	12.7%	12.3%	12.2%
Age							
0-17	5.7%	6.0%	6.2%	6.7%	7.3%	7.3%	6.5%
18-25	11.1%	11.3%	11.1%	10.7%	10.5%	10.0%	10.8%
26-34	16.3%	16.9%	17.4%	17.6%	17.7%	17.3%	17.2%
35-44	16.6%	15.9%	15.3%	15.1%	15.2%	15.1%	15.5%
45-54	24.4%	23.5%	22.8%	22.2%	21.4%	21.0%	22.5%
55+	25.8%	26.3%	27.2%	27.8%	27.9%	29.3%	27.4%
Gender							
Female	52.6%	52.2%	51.9%	52.2%	52.5%	52.5%	52.3%
Male	47.4%	47.8%	48.1%	47.8%	47.5%	47.5%	47.7%
Race							
Asian	22.8%	21.8%	22.0%	22.6%	23.0%	23.4%	22.6%
Black/African American	2.7%	2.5%	2.4%	2.4%	2.4%	2.4%	2.5%
Hispanic	27.5%	28.2%	28.0%	28.3%	28.4%	27.8%	28.0%
Non-Hispanic White	39.4%	39.5%	39.6%	38.5%	37.1%	36.8%	38.5%
Other Race	7.7%	7.9%	7.9%	8.2%	9.1%	9.6%	8.4%

B Complete Parameter Estimates

Table A2: Estimated Parameters

Demand Parameters ($\widehat{\beta}_t$)

	$\widehat{\theta}_{2016}$	$\widehat{\theta}_{2017}$	$\widehat{\theta}_{2018}$	$\widehat{\theta}$
Monthly Premium ($100)	−0.539***	−0.604***	−0.625***	−0.642***
	(0.008)	(0.008)	(0.007)	(0.006)
250% to 400% of FPL	0.184***	0.219***	0.208***	0.181***
	(0.006)	(0.006)	(0.005)	(0.004)
> 400% of FPL	0.299***	0.358***	0.364***	0.344***
	(0.007)	(0.006)	(0.005)	(0.005)
Ages 0 to 17	−0.282***	−0.246***	−0.248***	−0.235***
	(0.017)	(0.015)	(0.013)	(0.011)
Ages 18 to 34	−0.854***	−0.866***	−0.838***	−0.793***
	(0.008)	(0.007)	(0.006)	(0.005)
Ages 35 to 54	−0.372***	−0.388***	−0.381***	−0.366***
	(0.006)	(0.005)	(0.005)	(0.004)
Male	−0.146***	−0.154***	−0.144***	−0.131***
	(0.006)	(0.005)	(0.004)	(0.004)
Family	0.009*	−0.011***	−0.026***	−0.028***
	(0.005)	(0.004)	(0.004)	(0.003)
Asian	−0.194***	−0.185***	−0.185***	−0.173***
	(0.007)	(0.006)	(0.006)	(0.005)
Black	−0.293***	−0.309***	−0.317***	−0.308***
	(0.015)	(0.014)	(0.012)	(0.010)
Hispanic	−0.541***	−0.547***	−0.523***	−0.472***
	(0.008)	(0.007)	(0.006)	(0.005)
Other race	0.064***	0.060***	0.048***	0.036***
	(0.010)	(0.009)	(0.008)	(0.007)
Year 2015	−0.000	0.001	0.004	0.004
	(0.005)	(0.006)	(0.006)	(0.006)
Year 2016		0.068***	0.082***	0.086***
		(0.006)	(0.006)	(0.006)
Year 2017			0.113***	0.114***
			(0.005)	(0.006)
Year 2018				0.226***
				(0.005)
AV	3.193***	3.222***	3.216***	3.188***
	(0.028)	(0.025)	(0.022)	(0.020)
Silver	0.571***	0.663***	0.716***	0.713***
	(0.008)	(0.008)	(0.007)	(0.006)
HMO	0.395***	0.513***	−0.044***	−0.143***
	(0.016)	(0.016)	(0.008)	(0.006)
Anthem	1.183***	1.241***	0.534***	0.412***
	(0.021)	(0.019)	(0.011)	(0.009)
Blue Shield	1.158***	1.261***	0.553***	0.449***
	(0.021)	(0.020)	(0.010)	(0.008)
Kaiser	0.846***	0.794***	0.674***	0.627***
	(0.011)	(0.009)	(0.007)	(0.006)
Health Net	0.515***	0.399***	0.142***	0.079***
	(0.010)	(0.009)	(0.007)	(0.006)
Anthem x HMO	−1.189***	−1.462***	−0.957***	−0.896***
	(0.022)	(0.023)	(0.015)	(0.014)
Nesting Parameter	0.548***	0.623***	0.649***	0.680***
	(0.005)	(0.005)	(0.004)	(0.004)

	$\widehat{\theta}_{2016}$	$\widehat{\theta}_{2017}$	$\widehat{\theta}_{2018}$	$\widehat{\theta}$
Previous Choice	1.948***	2.103***	1.996***	1.888***
	(0.088)	(0.070)	(0.058)	(0.052)
250% to 400% of FPL	0.301***	0.323***	0.309***	0.246***
	(0.026)	(0.019)	(0.015)	(0.013)
> 400% of FPL	0.624***	0.701***	0.653***	0.605***
	(0.040)	(0.029)	(0.023)	(0.020)
Ages 0 to 17	−0.136**	−0.105**	−0.108***	−0.174***
	(0.069)	(0.052)	(0.040)	(0.034)
Ages 18 to 34	0.004	0.018	0.082***	0.063***
	(0.029)	(0.022)	(0.017)	(0.015)
Ages 35 to 54	−0.006	−0.001	0.012	0.006
	(0.025)	(0.019)	(0.015)	(0.013)
Male	0.136***	0.172***	0.187***	0.191***
	(0.028)	(0.021)	(0.017)	(0.014)
Family	−0.206***	−0.279***	−0.311***	−0.298***
	(0.020)	(0.015)	(0.012)	(0.011)
Asian	−0.198***	−0.262***	−0.278***	−0.285***
	(0.025)	(0.019)	(0.015)	(0.013)
Black	0.035	−0.094*	0.008	0.037
	(0.077)	(0.055)	(0.048)	(0.042)
Hispanic	0.110***	0.030	0.021	0.027**
	(0.027)	(0.020)	(0.016)	(0.014)
Other race	−0.161***	−0.172***	−0.136***	−0.147***
	(0.039)	(0.030)	(0.025)	(0.022)
Year 2016		−0.024*	−0.005	0.009
		(0.014)	(0.015)	(0.015)
Year 2017			−0.166***	−0.167***
			(0.014)	(0.015)
Year 2018				−0.342***
				(0.015)
Anthem	−0.478***	−0.381***	−0.039	0.391***
	(0.059)	(0.047)	(0.038)	(0.033)
Blue Shield	−0.128**	−0.009	0.320***	0.889***
	(0.064)	(0.050)	(0.040)	(0.035)
Kaiser	−0.337***	−0.278***	−0.049*	0.182***
	(0.050)	(0.037)	(0.028)	(0.019)
Health Net	−0.839***	−0.911***	−0.525***	0.076***
	(0.050)	(0.038)	(0.028)	(0.020)
HMO	0.452***	0.440***	0.611***	0.700***
	(0.043)	(0.034)	(0.030)	(0.030)
AV	1.457***	1.754***	1.589***	1.406***
	(0.096)	(0.074)	(0.059)	(0.052)
Silver	−0.632***	−0.732***	−0.769***	−0.896***
	(0.022)	(0.018)	(0.015)	(0.013)

Risk Score Parameters ($\widehat{\gamma}_t$)

	$\widehat{\theta}_{2016}$	$\widehat{\theta}_{2017}$	$\widehat{\theta}_{2018}$	$\widehat{\theta}$
Silver	0.814***	0.824***	0.785***	0.761***
	(0.062)	(0.042)	(0.033)	(0.028)
Gold	0.880***	0.913***	0.859***	0.851***
	(0.071)	(0.044)	(0.034)	(0.029)
Platinum	1.083***	1.247***	1.278***	1.287***
	(0.077)	(0.047)	(0.036)	(0.031)
Share Ages 18 to 25	−1.661**	−1.602***	−0.965***	−0.975***
	(0.804)	(0.495)	(0.365)	(0.313)
Share Ages 26 to 44	−1.334***	−0.977***	−1.125***	−1.061***
	(0.395)	(0.219)	(0.161)	(0.141)
Share Male	−0.355	−0.088	0.035	−0.340*
	(0.653)	(0.305)	(0.220)	(0.197)
Share Hispanic	−0.233	−0.334**	−0.642***	−0.744***
	(0.184)	(0.130)	(0.097)	(0.084)

Average Claims Parameters ($\widehat{\mu}_t$)

	$\widehat{\theta}_{2016}$	$\widehat{\theta}_{2017}$	$\widehat{\theta}_{2018}$	$\widehat{\theta}$
HMO	0.010	0.037*	−0.155***	−0.187***
	(0.064)	(0.023)	(0.011)	(0.010)
Log risk score	1.074***	1.049***	1.038***	1.038***
	(0.009)	(0.005)	(0.004)	(0.004)
Trend	−0.027***	−0.010***	0.015***	0.019***
	(0.008)	(0.003)	(0.002)	(0.002)
Anthem	0.119*	0.093***	0.171***	0.145***
	(0.066)	(0.026)	(0.018)	(0.019)
Blue Shield	0.021	0.069**	−0.009	−0.081***
	(0.080)	(0.034)	(0.030)	(0.031)
Health Net	−0.029	−0.032	0.163***	0.176***
	(0.095)	(0.046)	(0.035)	(0.035)
Kaiser	−0.132***	−0.081***	0.047***	0.066***
	(0.044)	(0.016)	(0.016)	(0.016)

Notes: Robust standard errors are in parentheses (*** indicates statistical significance at the 1% level, ** at the 5% level, and * at the 10% level). Parameter estimates for the market fixed effects in equations (3) and (22) are omitted.