Firm Learning in a Selection Market

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

[1]University of Pennsylvania
[2]Emory University

December 4, 2022

Abstract

Creating new markets is a prevalent mechanism 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 conclusions about the market equilibrium. We estimate our learning model with data from the California ACA exchange and find it provides a statistically significant improvement 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. Assuming firms have full information leads to a less favorable conclusion about welfare in equilibrium because firms initially overestimated premium sensitivity and underestimated inertia. Premium regulation that prohibits firms from using certain consumer information to price 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, Anh Nguyen, Mark Pauly, Dan Polsky, Juan Rubio-Ramirez, Ashley Swanson, Pietro Tebaldi, and Bob Town for advice and helpful feedback. We also thank participants at 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 in the United States are increasingly being delivered through the private sector. Prominent examples occur in education, child care, health care services, and health insurance. An increasing share of Americans obtain health insurance through publicly-supported private health insurance markets. As of 2020, approximately 53.9 million people participate in the Medicaid Managed Care program for the low-income population, 24.1 million people are enrolled in the Medicare Advantage program for the elderly, 46.5 million people have a Medicare Part D prescription drug plan, and 11.4 million get health coverage through the Affordable Care Act (ACA) exchanges (Kaiser Family Foundation, 2020). The current trend of delivering social programs through private markets is expected to continue for the foreseeable future (Gruber, 2017).

Implementation of these social programs is usually accomplished by establishing new markets. A significant challenge for participating firms is that they initially have little knowledge of the relevant market characteristics such as consumer preferences and competitors' behavior for making optimal decisions. In markets with selection, firms face the additional challenges of forecasting cost and understanding how cost is correlated with demand (Einav et al., 2010). Suboptimal decisions while firms are learning may have significant welfare implications. Therefore, potential sources of firm uncertainty should be considered when designing and evaluating social programs. However, the empirical industrial organization (IO) literature and policy-makers usually assume firms have full information and the market is in equilibrium upon market inception, as noted by 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 conclusions about the market equilibrium and the potential welfare implications for market design. Similar to Doraszelski et al. (2018), we estimate an adaptive learning model that allows firms to progressively learn about demand preferences in a new market. We extend Doraszelski et al. (2018)'s adaptive learning model by allowing firms to learn about cost and the correlation between demand and cost. 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 insurance from private firms. The ACA setting has two important features that make it particularly appealing for studying firm learning. First, the ACA setting provides 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 also use data on firms' predictions about their costs and actual costs from insurer rate filings. 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 document that firms' predictions about their costs converged to their realized costs over the first several years of the exchange, suggesting that firms were learning. Second, 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 health insurance market (i.e., the market where consumers buy insurance directly from an insurer) and those without insurance (Gruber, 2017). New ACA restrictions that prohibited firms from using health status to set premiums created additional sources of uncertainty (Pauly et al., 2015, 2020).

We make four primary contributions to the literature: (1) we extend the empirical IO literature on firm learning to a market with selection, where firms need to learn about the correlation of demand and cost; (2) we find the adaptive learning model provides a statistically significant superior fit of the data than the standard IO model that assumes firms have full information upon market inception; (3) we demonstrate that in our setting, assuming firms have full information leads to a less favorable conclusion about welfare in market equilibrium; and (4) we show that price discrimination regulation that prohibits firms from using certain consumer information to price (known as community rating) makes firms react more to the information they do have available.

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 that 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 following deregulation. 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 with 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 is a straightforward extension of standard approaches in the empirical IO literature. The model accounts for adverse selection and moral hazard and endogenizes consumer choices, premiums, plan risk, and medical claims.

We estimate the adaptive learning model using our data on firm cost and consumer-level plan choices. Our parameter estimates indicate firms initially overestimated premium elasticities of demand and underestimated inertia. We assess how well alternative models fit our data by comparing each model's predictions of cost with the firms' predictions of cost, as reported in their rate filings. We find the adaptive learning model yields a statistically significant superior fit of our data compared to assuming firms have full information, reducing the mean absolute error by 24.0% and the root mean square error by 18.0%. 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 approach. This result suggests it is particularly important to allow firms to learn about the relationship between demand and cost. We also find that the benefits of using an adaptive learning approach compared to the standard approach are reduced over time.

We next use our estimated learning model to determine how much assuming firms have full information affects conclusions about the market equilibrium. The full information assumption is commonplace in the previous literature evaluating the design of government-created health insurance markets. In previous ACA exchange studies, Tebaldi (2022), Saltzman (2021), Polyakova and Ryan (2021), and Einav et al. (2019) assume full information. The full information assumption is also made in studies of 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.,

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

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). Given the incidence of government-created health insurance markets over the last two decades, it is especially important to assess the validity of assuming full information.

Assuming firms have full information in the California ACA exchange setting, compared to modeling firm learning, leads to a less favorable conclusion about the equilibrium; premiums are higher and social welfare is lower. The detrimental effect of full information provides empirical evidence of recent theoretical predictions in information design (Roesler and Szentes, 2017). The differences between the equilibrium results using full information and using only the available information generally decline over time, an indication that firms are learning. Relative to the observed equilibrium, average premiums are 4.2% higher in 2016, 1.9% higher in 2017, and only 0.3% higher in 2018. Annual per-capita social welfare is $34 lower and annual total social welfare is $70 million lower in 2016. The equilibrium is worse under full information because firms initially overestimated premium sensitivity and underestimated inertia, leading them to set lower premiums than if they had full information.

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 (ACA) setting where price discrimination is partially restricted, a complete prohibition on price discrimination (i.e., pure community rating) increases the cost of the marginal consumer by 22.5% 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 setting with pure community rating. 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 price makes them react more to the information they can use.

Our results have a number of important implications, both for researchers and policymakers. Firm learning may be relevant for analyses that forecast the impact of social programs, such as those conducted regularly by the Congressional Budget Office (CBO), and studies that retrospectively analyze program impact and the efficacy of certain program design features. Researchers should consider whether firm uncertainty is relevant for their specific setting and whether it is likely to affect estimates of key model parameters. Our study also has implications for whether policymakers should adopt policies that promote information sharing between firms to reduce uncertainty.

The remainder of this paper is organized as follows. Section 2 describes the data and ACA setting. Section 3 develops a model of the ACA exchanges. Section 4 discusses estimation. Section 5 presents the model parameter estimates. 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 promoting 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 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.

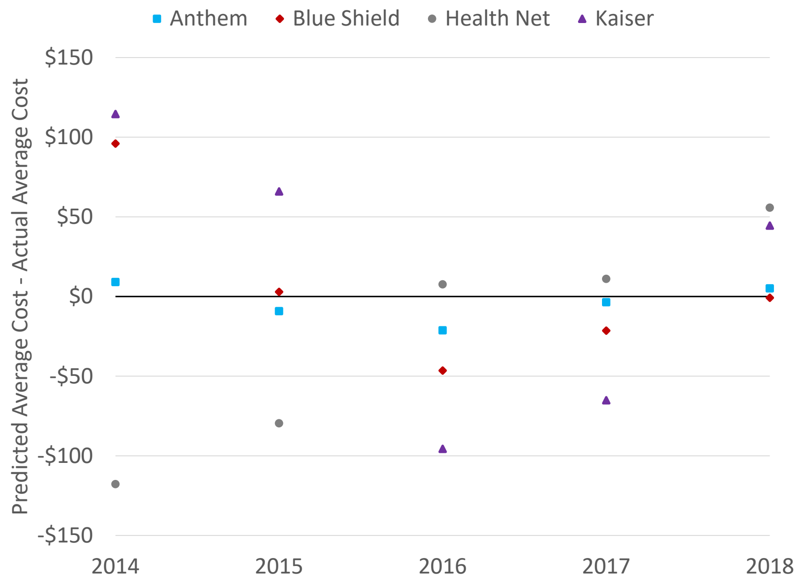
A unique feature of the rate filing data is the ability to compare firms' predictions about their costs with their realized costs. We refer to the difference between the firm's predicted and realized average costs as the *cost prediction error*, where cost is the sum of claims, risk adjustment, and reinsurance. 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. 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.

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

Figure 1: Cost Prediction Error By Year



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.

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.[2] Anthem's

[2]These firms include Chinese Community Health Plan, Contra Costa, L.A. Care Health Plan, Molina Healthcare,

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.

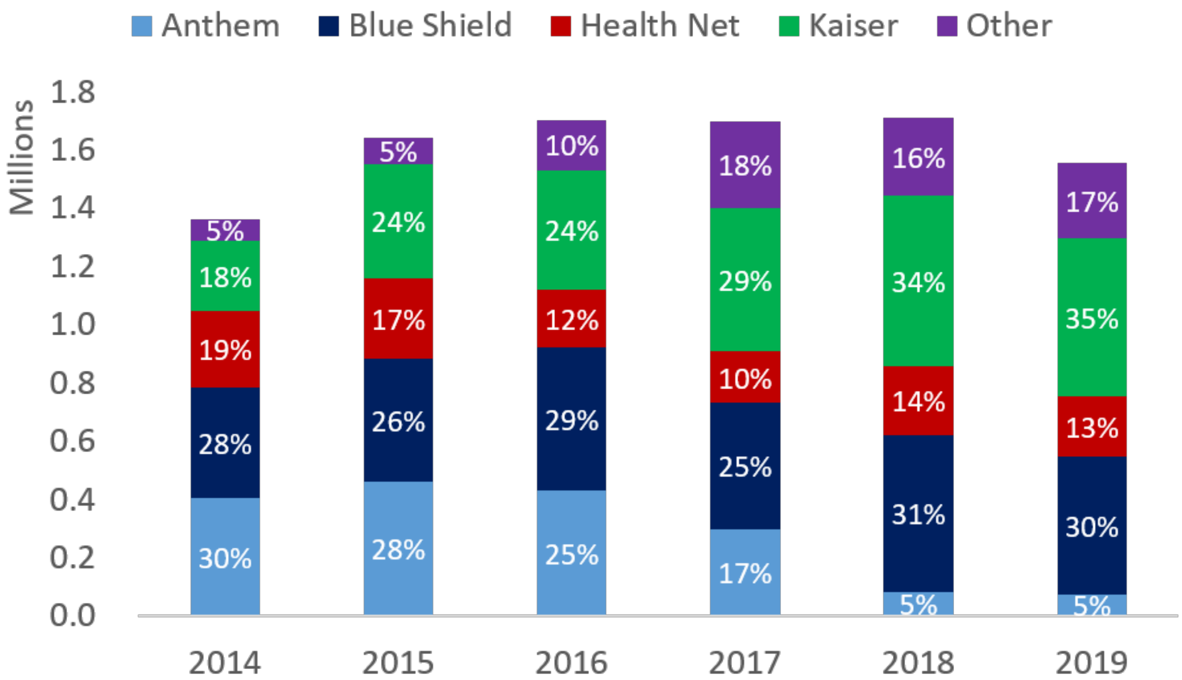
Consumers also have an outside option to forgo insurance. To construct the outside option population, we use consumer-level survey data on the uninsured from the American Community Survey (ACS) between 2014 and 2019 (Ruggles et al., 2022), consistent with the previous ACA IO literature (Tebaldi, 2022). 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 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).

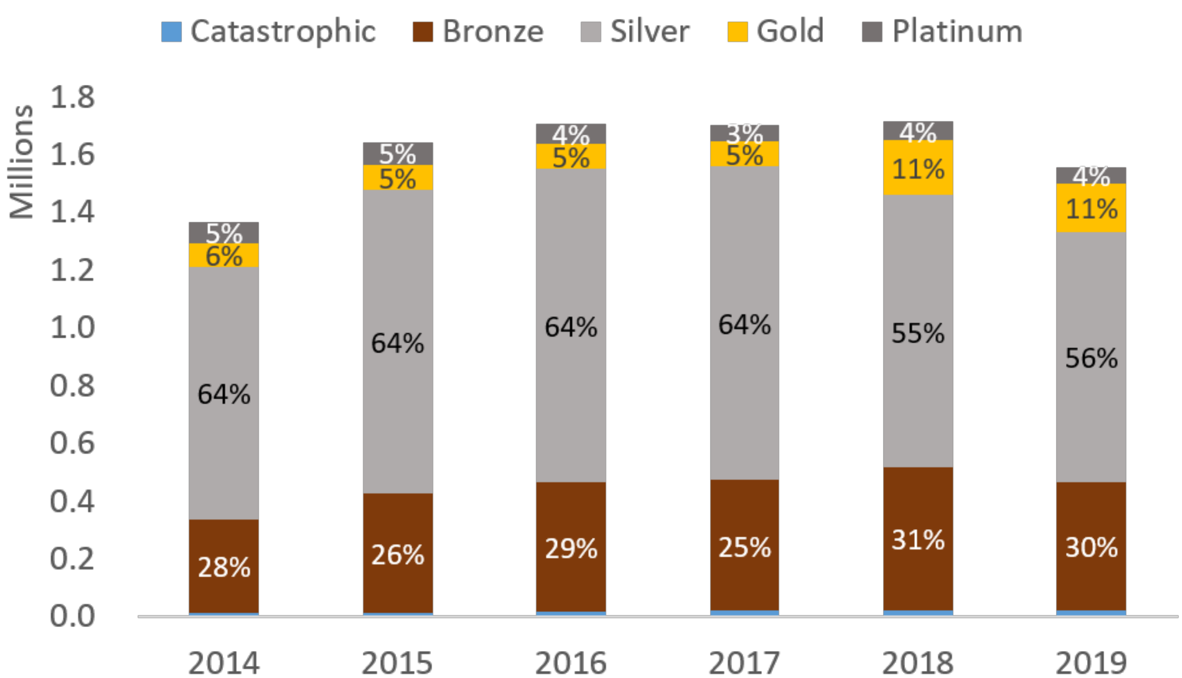
Although our focus is firm learning, a natural concern is whether consumers learn and adjust their plan choices accordingly. In related work, Saltzman et al. (2021) identify two significant features of this market that mitigate the concern of consumer learning: (1) annual churn into and out of the market 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

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

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 Model

Consider a two-stage game where in each period t (1) insurers set premiums simultaneously and (2) consumers choose plans. Below we detail how we model each of these two stages, beginning with consumer plan choice. Our model relaxes the assumption that firms have full information in estimation as discussed in Section 4.

3.1 Consumer Plan Choice

Households select the plan that maximizes their (indirect) utility function

$$U_{ijt}(p;\boldsymbol{\beta}) \quad \equiv \quad \beta_i^p p_{ijt}(p) + \beta_i^y y_{ij(t-1)} + x_{ij}'\beta^x + \xi_j + \epsilon_{ijt}^d \tag{1}$$

where $p = \mathbf{p}_t$ is the vector of plan base premiums set by all insurers in each market in year t, $p_{ijt}(p)$ 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, ξ_j is a vector of unobserved product characteristics, and ϵ_{ijt}^d is an error term. We allow the household's premium parameter $\beta_i^p = \beta^p + w_{it}'\phi$ to vary with household characteristics w_{it} and the household's inertia parameter $\beta_i^y = \beta^y + x_{ij}'\kappa + w_{it}'\nu$ to vary with household and product characteristics. Premium subsidies reduce the household's premium $p_{ijt}(p)$ as discussed below. CSRs increase the AV of silver plans in equation (1). The utility of the outside option $U_{i0t} = \beta_i^p \rho_{it} + \epsilon_{i0t}$, 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 $p_{ijt}(p)$ is calculated as

$$p_{ijt}(p) \quad = \quad \max\left\{ \underbrace{\sigma_{it} p_{jmt}}_{\substack{\text{full}\\ \text{premium}}} - \underbrace{\max\{\sigma_{it} p_{bmt} - \zeta_{it}, 0\}}_{\text{premium subsidy}}, 0 \right\} \tag{2}$$

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 ζ_{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.[3] Figure 3a compares the age rating curve in effect between 2014-2017[4] 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 ζ_{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 (1).

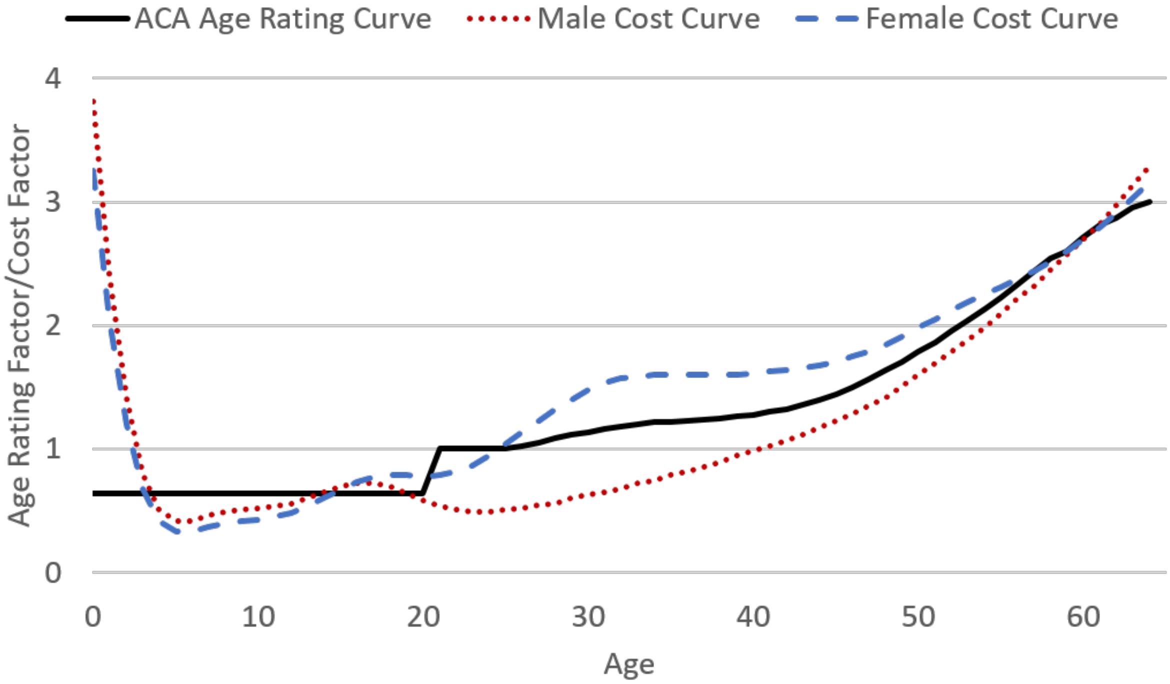
3.1.2 Calculating Demand

We assume that the vector of error terms ϵ_i has the generalized extreme value distribution so that equation (1) is a nested logit model with two nests. The first nest contains all exchange plans and

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

[4]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).

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 probabilities are

$$q_{ijt}(p;\boldsymbol{\beta}) \quad = \quad \frac{e^{V_{ijt}(p;\boldsymbol{\beta})/\lambda}\left(\sum_j e^{V_{ijt}(p;\boldsymbol{\beta})/\lambda}\right)^{\lambda-1}}{1+\left(\sum_j e^{V_{ijt}(p;\boldsymbol{\beta})/\lambda}\right)^{\lambda}} \tag{3}$$

where $V_{ijt}(p;\boldsymbol{\beta}) \equiv \beta_i^p p_{ijt}(p)+\beta_i^y y_{ij(t-1)}+x_{ij}'\beta^x+\xi_j$ and λ is the nesting parameter. The household choice probabilities in equation (3) converge to the standard logit choice probabilities when $\lambda \to 1$.

The sensitivity of a subsidized consumer's demand $q_{ijt}(p)$ to a premium change is

$$\frac{\partial q_{ikt}(p)}{\partial p_{jmt}} \quad = \quad \sum_{l\in J_{mt}} \frac{\partial q_{ikt}(p)}{\partial p_{ilt}(p)}\frac{\partial p_{ilt}(p)}{\partial p_{jmt}}$$

for all plans j,k, where J_{mt} is the set of available plans. If the subsidy does not exceed the full premium, then the sensitivity of the consumer's premium to the insurer's base premium is

$$\frac{\partial p_{ilt}(p)}{\partial p_{jmt}} = \begin{cases} 0 & l=j, j=b \\ \sigma_{it} & l=j, j\neq b \\ -\sigma_{it} & l\neq j, j=b \\ 0 & l\neq j, j\neq b \end{cases} \tag{4}$$

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 for estimation and simulation. 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.

3.2 Firm Premium-Setting

Assume firms are risk-neutral and maximize expected profit. A firm sets the vector of base premiums for the plans that it sells to maximize

$$\begin{aligned} \pi_{ft}(p;\boldsymbol{\theta}) \quad &= \quad R_{ft}(p;\boldsymbol{\beta}) - C_{ft}(p;\boldsymbol{\theta}) + RA_{ft}(p;\boldsymbol{\theta}) + RI_{ft}(p;\boldsymbol{\theta}) - V_{ft}(p;\boldsymbol{\beta}) - FC_{ft} \\ &= \quad R_{ft}(p;\boldsymbol{\beta}) - (1-\iota_{ft})C_{ft}(p;\boldsymbol{\theta}) + RA_{ft}(p;\boldsymbol{\theta}) - V_{ft}(p;\boldsymbol{\beta}) - FC_{ft} \end{aligned} \tag{5}$$

where $R_{ft}(\cdot)$ is total premium revenue, $C_{ft}(\cdot)$ is total claims, $RA_{ft}(\cdot)$ is risk adjustment received, $RI_{ft}(\cdot)$ is reinsurance received, $V_{ft}(\cdot)$ is variable administrative cost (e.g., commissions or fees), FC_{ft} is fixed cost, and ι_{ft} indicates the AV of the reinsurance contract (i.e., the expected percentage of claims paid by the reinsurer). The model parameters $\boldsymbol{\theta} \equiv (\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\mu})$, where $\boldsymbol{\beta}$ are the demand parameters (as defined above), $\boldsymbol{\gamma}$ are the risk score parameters, and $\boldsymbol{\mu}$ are the average claims parameters. As discussed 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. Formulas for firm revenue, claims, and administrative costs are straightforward and provided in Appendix A. The next two subsections discuss calculation of the risk adjustment transfer and the model equilibrium.

3.2.1 Calculating Risk Adjustment Transfers

Because risk adjustment plays an important role in the ACA exchanges, we endogenize the program in our model. Under the ACA's single risk pool provisions, risk adjustment occurs at the state level for all metal plans (i.e., platinum, gold, silver, and bronze) in the individual market. Catastrophic plans have a separate risk adjustment pool. Pope et al. (2014) derive the official risk adjustment transfer formula that is used in the ACA exchanges.[5] The average transfer received by a plan is

$$\begin{aligned} ra_{jmt}(p;\boldsymbol{\theta}) &= \widehat{c}_{jmt}(p;\boldsymbol{\theta}) - \widetilde{c}_{jmt}(p;\boldsymbol{\theta}) \\ &= \frac{\widehat{h}_{jmt}(p;\boldsymbol{\theta})}{\sum_{l\in J_t} \widehat{h}_{lmt}(p;\boldsymbol{\theta}) s_{lmt}(p;\boldsymbol{\theta})} \nu\overline{p} - \frac{\widetilde{h}_{jmt}(p;\boldsymbol{\theta})}{\sum_{l\in J_t} \widetilde{h}_{lmt}(p;\boldsymbol{\theta}) s_{lmt}(p;\boldsymbol{\theta})} \nu\overline{p} \end{aligned} \tag{6}$$

where $\widehat{c}_{jmt}(p;\boldsymbol{\theta})$ is the plan's expected average claims with adverse selection and $\widetilde{c}_{jmt}(p;\boldsymbol{\theta})$ is the plan's expected average claims without adverse selection. The cost factor $\widehat{h}_{jmt}(p;\boldsymbol{\theta}) \equiv \mathrm{IDF}_j \mathrm{GCF}_{mt} r_{jmt}(p;\boldsymbol{\theta})$ is the product of the plan's induced demand factor (or moral hazard factor), geographic cost factor, and risk score. The cost factor $\widetilde{h}_{jmt}(p;\boldsymbol{\theta}) \equiv \mathrm{AV}_j \mathrm{IDF}_j \mathrm{GCF}_{mt} a_{jmt}(p;\boldsymbol{\theta})$ is the product of the plan's AV, induced demand factor, geographic cost factor, and average ACA age rating factor $a_{jmt}(p;\boldsymbol{\theta})$ across the plan's enrollees. We denote the plan's market share as $s_{lmt}(p;\boldsymbol{\theta})$, the average statewide premium as $\overline{p}$, and the expected percentage of collected premiums that is spent on claims as ν. CMS set ν to 100% from 2014-2017 and then reduced it to 86% starting in 2018.

The plan's total risk adjustment transfer $RA_{jmt}(p;\boldsymbol{\theta})$ equals

[5] 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.

$$\begin{aligned} RA_{jmt}(p;\boldsymbol{\theta}) &= ra_{jmt}(p;\boldsymbol{\theta})q_{jmt}(p;\boldsymbol{\theta}) \\ &= \left[\frac{\widehat{h}_{jmt}(p;\boldsymbol{\theta})q_{jmt}(p;\boldsymbol{\theta})}{\sum_{l\in J_t}\widehat{h}_{lmt}(p;\boldsymbol{\theta})q_{lmt}(p;\boldsymbol{\theta})} - \frac{\widetilde{h}_{jmt}(p;\boldsymbol{\theta})q_{jmt}(p;\boldsymbol{\theta})}{\sum_{l\in J_t}\widetilde{h}_{lmt}(p;\boldsymbol{\theta})q_{lmt}(p;\boldsymbol{\theta})}\right]\nu R_t(p;\boldsymbol{\theta}) \end{aligned} \tag{7}$$

where $R_t(p;\boldsymbol{\theta})$ is total market premium revenue. Formula (7) 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. The cost factor $\widehat{h}_{jmt}(p;\boldsymbol{\theta})$ in the firm term of formula (7) accounts for differences in moral hazard and geography, as well as age, plan AV, and enrollee health risk in the plan risk score. The cost factor $\widetilde{h}_{jmt}(p;\boldsymbol{\theta})$ in the second term of formula (7) accounts for differences in moral hazard, geography, age, and plan AV, but <u>not</u> enrollee health risk. Therefore, the difference between the first and second terms of formula (7) compensates firms for differences in enrollee health risk only (i.e., the firm's relative risk due to adverse selection only). Firms with higher-than-average risk will receive a risk adjustment transfer ($RA_{jmt}(p;\boldsymbol{\theta}) > 0$), whereas firms with lower-than-average risk will pay a risk adjustment transfer ($RA_{jmt}(p;\boldsymbol{\theta}) < 0$).

3.2.2 Equilibrium

Now we find the first-order necessary conditions for a Nash equilibrium. Differentiating equation (5) yields the first-order conditions

$$0 = \frac{\partial \pi_{ft}(p;\boldsymbol{\theta})}{\partial p_{jmt}} = \frac{\partial R_{ft}(p;\boldsymbol{\beta})}{\partial p_{jmt}} - (1-\iota_{ft})\frac{\partial C_{ft}(p;\boldsymbol{\theta})}{\partial p_{jmt}} + \frac{\partial RA_{ft}(p;\boldsymbol{\theta})}{\partial p_{jmt}} - \frac{\partial V_{ft}(p;\boldsymbol{\beta})}{\partial p_{jmt}} \tag{8}$$

for all markets m in which plan j is offered by the firm in year t. Equation 8 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). Appendix A shows that every variable in equations (5) and (8) can be written in terms of three estimable variables: (1) household choice probabilities $q_{ijt}(p;\boldsymbol{\beta})$; (2) plan risk scores $r_{jmt}(p;\boldsymbol{\theta})$; and (3) average claims $c_{jmt}(p;\boldsymbol{\theta})$. Household choice probabilities are computed using equation (3). We calculate plan risk scores as a function of observable enrollee characteristics and the plan generosity using the estimating equation

$$\ln r_{jmt}(p;\boldsymbol{\theta}) = \sum_{d\in D}\gamma^d s_{djmt}(p;\boldsymbol{\beta}) + MT_j'\gamma^{MT} + \epsilon^r_{jmt} \tag{9}$$

The predicted demographic share $s_{djmt}(\cdot)$ is the share of plan j's enrollment in market m and year

t with demographic characteristic d, MT_j is a vector metal tier fixed effects, ϵ^r_{jmt} is an error term, and the vector of risk score parameters $\boldsymbol{\gamma} = (\gamma^d, \gamma^{MT}, \gamma^n)$. The demographic shares are computed by aggregating the household choice probabilities. We calculate plan average claims as a function of the plan risk score using the estimating equation

$$\ln c_{jmt}(p; \boldsymbol{\theta}) = \mu^r \ln r_{jmt}(p; \boldsymbol{\theta}) + x'_j \mu^x + \mu^l l_t + n'_m \mu^n + \epsilon^c_{jmt} \tag{10}$$

where $r_{jmt}(\cdot)$ is the predicted risk score computed using equation (9), x_j are product characteristics (not including plan AV), l_t is a linear trend, n'_m are market fixed effects, ϵ^c_{jmt} is an error term, and $\boldsymbol{\mu} = (\mu^r, \mu^x, \mu^l, \mu^n)$ are the claims parameters. Equation (18) in Appendix A provides a formula for how average claims respond to a change in premiums. If a plan is adversely selected, then $\frac{\partial c_{jmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}} > 0$.

4 Estimation

In this section, we explain how we use the generalized method of moments (GMM) to estimate the parameter vector $\boldsymbol{\theta}$. We first review a standard approach in the IO literature that assumes firms have full information. We then relax the full information assumption to accommodate firm learning.

4.1 Full Information Approach

In the standard or full information approach, the econometrician pools data from all years to estimate the demand parameters $\boldsymbol{\beta}$, the risk score parameters $\boldsymbol{\gamma}$, and the average claims parameters $\boldsymbol{\mu}$. To estimate these parameters, 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 (8). Denote N^{IJT} as the number of plans available to all households in all years, N^{JMT} as the number of plans available in all markets and years, and N^M_{jt} as the number of markets where plan j is offered in year t. Let χ_{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 $\mathbf{z}^r_{jmt}(p; \boldsymbol{\beta}) \equiv (s_{djmt}(p, \boldsymbol{\beta}), MT_j)$ and the average claims covariates $\mathbf{z}^c_{jmt}(p; \boldsymbol{\theta}) \equiv (\ln r_{jmt}(p; \boldsymbol{\theta}), x_j, u_t, n_m)$. The moment conditions are

$$
\begin{aligned}
\frac{1}{N^{IJT}} \sum_{i\in I, j\in J, t\in T} \frac{\chi_{ijt}\partial \ln q_{ijt}(p;\boldsymbol{\beta})}{\partial \boldsymbol{\beta}} &= \mathbf{0} \\
\frac{1}{N^{JMT}} \sum_{j\in J, m\in M, t\in T} \mathbf{z}^{r}_{jmt}(p;\boldsymbol{\theta})\left(\ln r_{jmt} - \boldsymbol{\gamma}'\mathbf{z}^{r}_{jmt}(p;\boldsymbol{\theta})\right) &= \mathbf{0} \\
\frac{1}{N^{JMT}} \sum_{j\in J, m\in M, t\in T} \mathbf{z}^{c}_{jmt}(p;\boldsymbol{\theta})\left(\ln c_{jmt} - \boldsymbol{\mu}'\mathbf{z}^{c}_{jmt}(p;\boldsymbol{\theta})\right) &= \mathbf{0} \\
\frac{1}{N^{M}_{jt}} \sum_{m\in M} g_{jmt}(p;\boldsymbol{\theta}) &= 0, \qquad \forall j \in J_t, t \in T \qquad (11)
\end{aligned}
$$

where the first-order condition values

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

Because model (11) over-identifies the model parameters, we use two-step feasible GMM to find the values of $\boldsymbol{\theta}$ that minimize the GMM objective $[\mathbf{m}(\boldsymbol{\theta})]'\mathbf{W}^{-1}[\mathbf{m}(\boldsymbol{\theta})]$, where $\mathbf{m}(\boldsymbol{\theta})$ is the vector of moment values in model (11) 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_i^p). Two primary sources of exogenous variation are used to identify the premium parameter β_i^p, including: (1) exogenous variation in absolute premiums (i.e., relative to the outside option) that results from the phasing-in of the mandate penalty between 2014 and 2016 and elimination of the penalty in 2019; (2) exogenous variation in relative premiums (i.e., between plans) that results from kinks in the household premium formula (2). 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 household characteristics, including age, income, and household composition. While we believe these sources of exogenous variation cleanly identify the premium parameter, we also estimate equation (1) with insurer-market fixed effects. These fixed effects control for unobservables at the insurer's discretion such as provider networks and formularies that could be correlated with premiums. Ho and Pakes (2014) and Tebaldi (2022) follow a similar approach. Our estimates are similar when including insurer-market fixed effects.

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 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 (3) 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 (3) 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^r. We compute predicted plan risk scores using equation (9) 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.

4.2 Adaptive Learning

In this section, we allow firms to learn about the parameter vector $\boldsymbol{\theta}$. Firm participation in a new insurance market involves two principal sources of uncertainty: demand and cost uncertainty. Demand uncertainty arises when firms do not know consumer preferences. Cost uncertainty arises when firms do not know the cost of insuring their enrollees. A defining feature of a selection market such as insurance is that demand and cost uncertainty are correlated. Our model accommodates this correlation by estimating equations (9) and (10). With adaptive learning estimates of $\boldsymbol{\theta}$, we can compute the partial derivative $\frac{\partial c_{jmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}}$ using equation (18) to quantify the firm's estimate of adverse selection.

We assume that in any plan year t, firms use data from years $2014, \ldots, t-1$ to estimate $\boldsymbol{\theta}$. This estimate of $\boldsymbol{\theta}$ is used to predict cost in year t. Define the previous period set $T_t \equiv \{2014, \ldots, t-1\}$ and denote $\boldsymbol{\theta}_t$ as the firm's estimate of $\boldsymbol{\theta}$ at time t. To estimate $\boldsymbol{\theta}_t$, we form the moment conditions

$$\begin{aligned}
\frac{1}{N^{IJT_t}} \sum_{i\in I, j\in J, \tau\in T_t} \frac{\chi_{ij\tau}\partial \ln q_{ij\tau}(p;\boldsymbol{\beta}_t)}{\partial \boldsymbol{\beta}_t} &= \mathbf{0} \\
\frac{1}{N^{JMT_t}} \sum_{j\in J, m\in M, \tau\in T_t} \mathbf{z}^r_{jm\tau}(p;\boldsymbol{\beta}_t)\left(\ln r_{jm\tau} - \boldsymbol{\gamma}_t' \mathbf{z}^r_{jm\tau}(p;\boldsymbol{\beta}_t)\right) &= \mathbf{0} \\
\frac{1}{N^{JMT_t}} \sum_{j\in J, m\in M, \tau\in T_t} \mathbf{z}^c_{jm\tau}(p;\boldsymbol{\theta}_t)\left(\ln c_{jm\tau} - \boldsymbol{\mu}_t' \mathbf{z}^c_{jm\tau}(p;\boldsymbol{\theta}_t)\right) &= \mathbf{0} \\
\frac{1}{N^M_{j\tau}} \sum_{m\in M} \overline{g}_{jm\tau}(p;\boldsymbol{\theta}_t) &= 0, \quad \forall j \in J_\tau, \tau \in \{2014, \ldots, t\} \qquad (12)
\end{aligned}$$

where the first-order condition values

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

The first three moment conditions in (12) are identical to the first three moment conditions in (11), except that we only use data through year $t-1$ instead of all the available data. 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 $t-1$).

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.

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 C. 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 inertia.[6] The equilibrium implications of the adaptive learning estimates compared to the estimates where firms are assumed to have full information are discussed in Section 6.

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

[6]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 the model suffers from omitted variable bias when the lagged choice variables are absent and therefore do not include estimates $\widehat{\boldsymbol{\theta}}_{2015}$.

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 inertia (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.[7]

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 55) share parameters are negative as expected, but quite volatile. The standard errors are large initially, but decline over time because firms have more data available to learn the relevant parameters. Consumers of Hispanic origin have considerably less claims risk than other racial and ethnic groups, but firms underestimate this effect initially. Firms initially overestimated the relationship between average claims and the plan risk score; the estimated parameter decreases from 1.075 using data available in 2016 to 1.059 using all 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 Comparing Goodness-of-Fit

One of the central goals of our study is to assess how well the adaptive learning estimates fit the data relative to the fit of the full information approach estimates. We assess model goodness-of-fit by comparing the plan average costs implied by our adaptive learning parameter estimates with the firms' **predictions** of plan average costs, as reported in their rate filings. In other words, we quantify how closely our adaptive learning model's estimates of cost match with the firms' predictions of cost. Data from the rate filings on predicted plan average costs is only used in this section to assess model fit (we used the available data on **realized** plan average costs to estimate the model parameters in Section 4). We compute two goodness-of-fit measures, the mean absolute error (MAE) and the

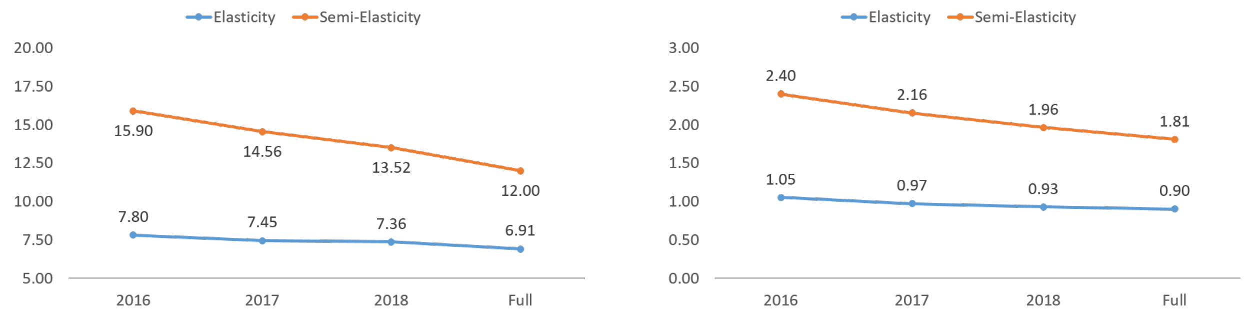
[7] Recall that our nested logit choice model converges to the standard logit as $\lambda \to 1$.

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 ($100)	-1.146^{***}	-1.171^{***}	-1.145^{***}	-1.090^{***}
	(0.008)	(0.007)	(0.006)	(0.005)
Previous Choice	1.941^{***}	2.257^{***}	2.396^{***}	2.513^{***}
	(0.089)	(0.068)	(0.056)	(0.050)
AV	3.202^{***}	3.190^{***}	3.150^{***}	3.088^{***}
	(0.028)	(0.025)	(0.022)	(0.020)
Nesting Parameter	0.554^{***}	0.618^{***}	0.649^{***}	0.694^{***}
	(0.005)	(0.005)	(0.004)	(0.004)
Risk Score Parameters ($\widehat{\boldsymbol{\gamma}}_t$)				
Silver	0.814^{***}	0.825^{***}	0.789^{***}	0.764^{***}
	(0.062)	(0.042)	(0.033)	(0.028)
Gold	0.882^{***}	0.915^{***}	0.863^{***}	0.852^{***}
	(0.071)	(0.044)	(0.034)	(0.029)
Platinum	1.084^{***}	1.252^{***}	1.288^{***}	1.293^{***}
	(0.077)	(0.047)	(0.036)	(0.031)
Share Ages 18 to 25	-1.647^{**}	-1.347^{***}	-0.666^{*}	-0.903^{***}
	(0.802)	(0.488)	(0.366)	(0.336)
Share Ages 26 to 44	-1.330^{***}	-0.873^{***}	-0.963^{***}	-0.913^{***}
	(0.395)	(0.217)	(0.160)	(0.143)
Share Male	-0.350	-0.047	0.106	-0.339^{*}
	(0.653)	(0.299)	(0.214)	(0.193)
Share Hispanic	-0.234	-0.348^{***}	-0.652^{***}	-0.741^{***}
	(0.184)	(0.130)	(0.097)	(0.085)
Average Claims Parameters ($\widehat{\boldsymbol{\mu}}_t$)				
Log Risk Score	1.075^{***}	1.053^{***}	1.045^{***}	1.059^{***}
	(0.009)	(0.004)	(0.003)	(0.004)
HMO	0.013	0.036	-0.130^{***}	-0.160^{***}
	(0.064)	(0.022)	(0.010)	(0.010)
Trend	-0.026^{***}	-0.006^{**}	0.022^{***}	0.021^{***}
	(0.007)	(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.

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 $\widehat{\theta}_t$ for $t \in \{2016, 2017, 2018\}$ 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.

root mean square error (RMSE).

Table II compares the goodness-of-fit 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 $\boldsymbol{\beta}$.

Relative to the model that assumes firms have full information (Specification 8), allowing firms to learn about all model parameters (Specification 1) reduces the MAE by 24.0% from 167 to 127 and the RMSE by 18.0% from 227 to 186. 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 $\boldsymbol{\mu}$. However, Specifications (6) and (7) which assume firms know the full information claims parameters have a slightly worse fit than the full information model. These results indicate that it is particularly important to allow firms to learn about the relationship between demand and cost in equation (10).

Table II: Model Goodness-of-Fit

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Known Parameters								
Inertia (β_i^y)		✓		✓	✓			✓
Premium (β_i^p)			✓	✓	✓			✓
Other Demand (β_i^x)				✓	✓			✓
Risk ($\boldsymbol{\gamma}$)					✓		✓	✓
Claims ($\boldsymbol{\mu}$)						✓	✓	✓
Fit Measure								
Mean absolute error	127.1	130.2	127.6	128.4	128.3	173.1	175.4	167.2
Root mean square error	186.4	187.9	186.7	189.8	192.3	228.2	238.7	227.3

Notes: Table shows how well the alternative models' predictions of plan cost fit the firms' predictions of plan cost. Two goodness-of-fit measures are shown: (1) mean absolute error (MAE) and (2) root mean square error (RMSE). To compute MAE and RMSE, 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.

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

$$DAE_{jmt} = \left|\zeta_{jmt}(p;\widehat{\boldsymbol{\theta}}_t) - \zeta_{jmt}\right| - \left|\zeta_{jmt}(p;\widehat{\boldsymbol{\theta}}) - \zeta_{jmt}\right| \tag{13}$$

where $\zeta_{jmt}(p;\widehat{\boldsymbol{\theta}}_t) \equiv (1-\iota_{ft})c_{jmt}(p;\widehat{\boldsymbol{\theta}}_t) - ra_{jmt}(p;\widehat{\boldsymbol{\theta}}_t)$ is the adaptive learning estimate of plan cost, $\zeta_{jmt}(p;\widehat{\boldsymbol{\theta}}) \equiv (1-\iota_{ft})c_{jmt}(p;\widehat{\boldsymbol{\theta}}) - ra_{jmt}(p;\widehat{\boldsymbol{\theta}})$ 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 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., $127.1 - 167.2 = -40.1$). 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.

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, which we assume are the estimates $\widehat{\theta}$ obtained using all available data (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 (8); and (3) computing several measures of the new equilibrium outcome, including average premiums, enrollment, and social welfare.

We compute total social welfare in year t as

$$SW_t \quad = \quad CS_t + \pi_t - \delta GS_t$$

where CS_t is total consumer surplus, π_t is total firm profit, and GS_t is net government spending adjusted by the deadweight loss of taxation δ that results from distortions in prices and consumer behavior. The deadweight loss of taxation corresponds to the additional compensation consumers need in order to obtain their original utility levels (i.e., before government spending on the premium subsidies, CSRs, etc.) at the distorted prices (Hausman and Poterba, 1987). Following Hausman and Poterba (1987) and Decarolis et al. (2020), we multiply government spending by a factor of 1.3 to account for the deadweight loss of taxation. We compute total consumer surplus

$$CS_t \quad = \quad -\sum_{i \in I} \frac{1}{\beta_i^p} \ln \left(\sum_{j \in J} \exp\left(V_{ijt}(p;\boldsymbol{\beta}_t)/\lambda\right)^{\lambda} + \exp\left(\beta_{it}^p \rho_{it}\right) \right) + \sum_{j \in J} \left[q_{ijt}(p) * \frac{\beta_{ij}^y * y_{ij(t-1)}}{\beta_i^p} \right] \qquad (14)$$

where the first term of equation (14) 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 $\pi_t = \sum_{t \in T} \pi_{ft}(p;\boldsymbol{\theta}_t)$, where $\pi_{ft}(p;\boldsymbol{\theta}_t)$ is defined in equation (5). 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 (2). Spending on CSRs is computed as

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

Model	(1)	(1′)	(2)	(2′)	(3)	(3′)
Known Parameters						
Inertia (β_i^y)			✓	✓		
Premium (β_i^p)					✓	✓
Other Demand (β_i^x)						
Risk (γ)						
Claims ($\boldsymbol{\mu}$)						
Constant	-40.146^{***}	-21.726^{**}	-37.044^{***}	-35.083^{***}	-39.662^{***}	-24.004^{**}
	(3.032)	(10.800)	(3.463)	(12.778)	(2.991)	(10.619)
Silver		-48.112^{***}		-41.683^{***}		-48.231^{***}
		(6.012)		(7.112)		(5.911)
Gold		-63.251^{***}		-59.718^{***}		-62.338^{***}
		(9.871)		(11.678)		(9.705)
Platinum		-111.623^{***}		-114.005^{***}		-109.935^{***}
		(10.677)		(12.632)		(10.498)
Time Trend		7.274^{**}		11.588^{***}		8.149^{**}
		(3.221)		(3.811)		(3.167)

Model (Cont.)	(4)	(4′)	(5)	(5′)
Known Parameters				
Inertia	✓	✓	✓	✓
Premium	✓	✓	✓	✓
Other Demand	✓	✓	✓	✓
Risk			✓	✓
Cost				
Constant	-38.793^{***}	-30.218^{***}	-38.926^{***}	-30.397^{***}
	(2.878)	(10.087)	(3.004)	(10.950)
Silver		-46.947^{***}		-48.940^{***}
		(5.614)		(6.095)
Gold		-61.201^{***}		-64.284^{***}
		(9.218)		(10.007)
Platinum		-108.060^{***}		-90.335^{***}
		(9.971)		(10.825)
Time Trend		10.210^{***}		10.227^{***}
		(3.008)		(3.266)

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 (13). 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.

$$CSR_t = \sum_{i \in I, j \in J} s_j^g q_{ijt}(p; \boldsymbol{\theta}_t) c_{jmt}(p; \boldsymbol{\theta}_t)$$

where s_j^g is the expected share of claims paid by the government for plan j.[8] 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[9], 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.

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).

Our results indicate that learning the full information parameters worsens the equilibrium outcome. The differences in the equilibrium outcome measures that result from using the full information parameters instead of the adaptive learning parameters are generally largest in 2016 and moderate over time. Using the full information parameters increases average unsubsidized premiums in 2016 by 4.2% from \$400 to \$417. The percentage increase in average unsubsidized premiums falls to 1.9% in 2017 and only 0.3% in 2018. Conversely, using the full information parameters reduces average subsidized premiums by 2.8% in 2016 and by 1.1% in 2017, but slightly increases average subsidized premiums by 0.5% 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. Using the full information parameters results in a fairly substantial \$33 in-

[8]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 $s_j^g = 26.88\%$ for the 94% CSR plan, $s_j^g = 19.04\%$ for the 87% CSR plan, and $s_j^g = 3\%$ for the 73% CSR plan.

[9]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).

crease in average platinum premiums in 2016 (Figure 5c). As a consequence, platinum plans lose about 1% market share, whereas bronze and silver plans gain market share (Figure 5d). The shift in enrollment from platinum to bronze and silver is a potential indication of full information exacerbating underinsurance. By 2018, the effect of information on the enrollment distribution across the metal tiers is small. The impact of full information on total exchange enrollment is negligible.

Figure 5f shows the impact of full information on annual per-capita social welfare. Firms earn substantially higher annual per-capita profit of $230 in 2016 when firms know the full information parameters. However, producer surplus gains are offset by increased 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 (2).[10] Overall, annual per-capita social welfare declines by $34 when firms learn the full information parameters. Over time, the increases in both profit and government spending that result from learning the full information parameters moderate.

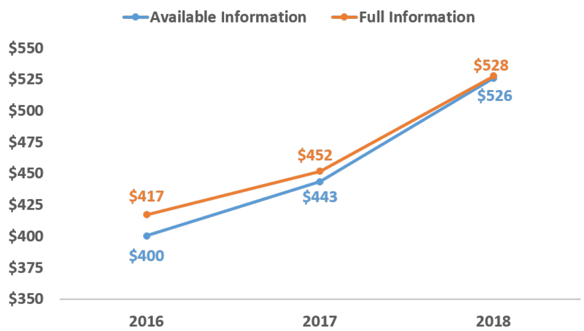
Our finding of a worse equilibrium outcome when firms learn the full information parameters is consistent with firms initially overestimating premium sensitivity and underestimating inertia. Markups are lower in a market where profit-maximizing firms believe that consumers are more premium sensitive than they actually are. Similarly, inertia is an important source of market power that firms exploit in this market to set higher premiums (Saltzman et al., 2021). Underestimating inertia leads firms to set lower premiums than they otherwise would. Learning the full information parameters therefore increases premiums and decreases welfare. If firms had instead underestimated premium sensitivity and overestimated inertia, they would have been incentivized to set higher premiums. Assuming firms have full information would then reduce premiums and increase social welfare.

7 Policy Simulations

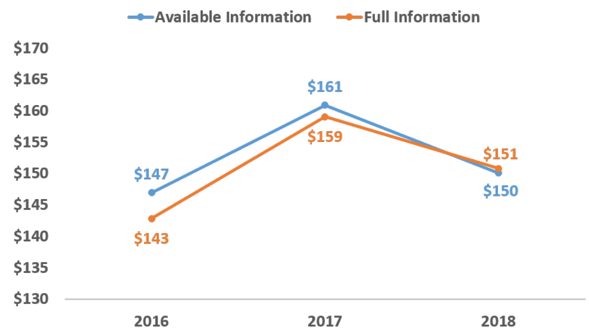
In this section, we examine whether firm knowledge interacts with policy. We studied a diverse set of prominent ACA policies and market features, including community rating regulations, inertia, endogenous subsidies, risk adjustment and the individual mandate. The exposition below focuses on the interaction 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

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

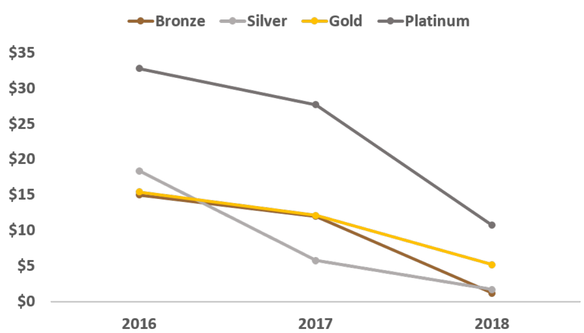
Figure 5: Impact of Assuming Full Information By Year



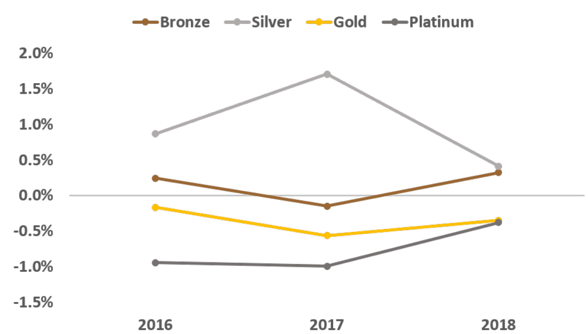
(a) Average Unsubsidized Premiums



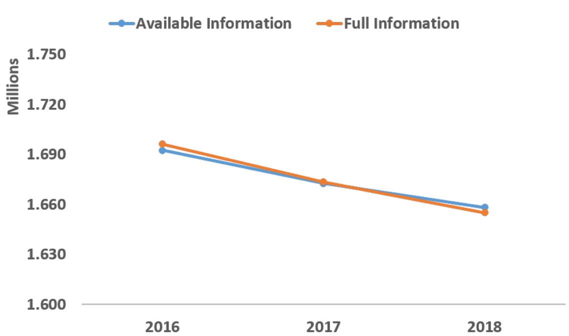
(b) Average Subsidized Premiums



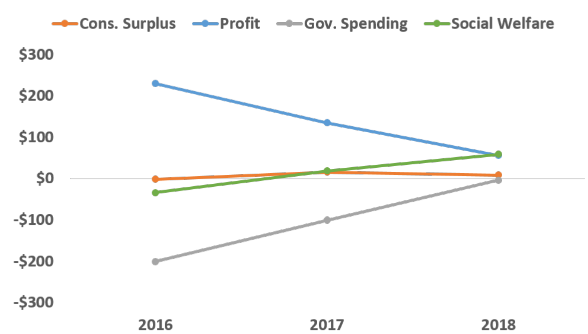
(c) Change in Unsubsidized Premiums By Metal



(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.

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.

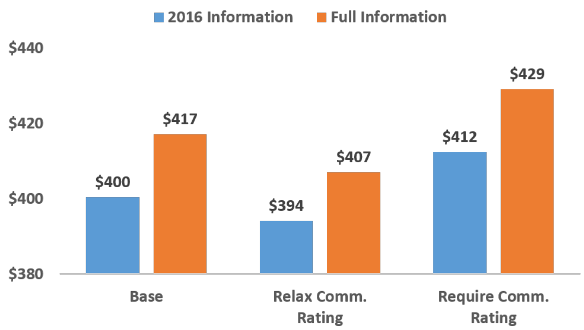
Figure 6 summarizes our results. We report the impact on average unsubsidized premiums (panel a), average subsidized premiums (panel b), total exchange enrollment (panel c), market share by metal tier (panel d), marginal claims using equation (22) in Appendix A (panel e), and the change in annual per-capita social welfare (panel f). From our results, we find: (1) the equilibrium outcome mostly improves when community rating is relaxed and worsens when pure community rating is required and (2) prohibiting firms from using certain information to price 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

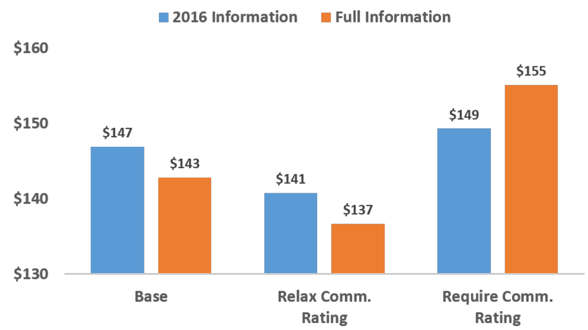
We assess the impact of community rating under 2016 information using the blue columns in Figure 6. Consistent with theory, we find relaxing community rating reduces average premiums and requiring pure community rating increases average premiums. When community rating is relaxed, average unsubsidized premiums decrease 1.6%. Requiring pure community rating results in average unsubsidized premiums increasing 3.0%. 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 4.2%. Average subsidized premiums increase 1.6% 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 0.8% reduction in enrollment. Figure 6d indicates that although pure community

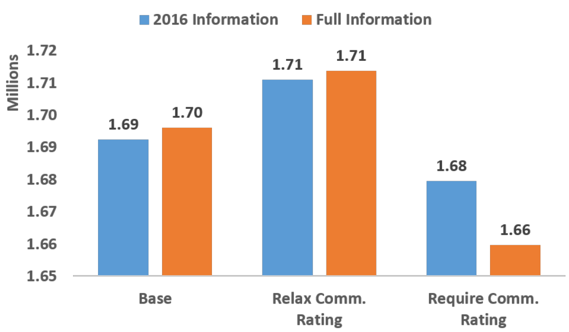
Figure 6: Policy Impact of Information



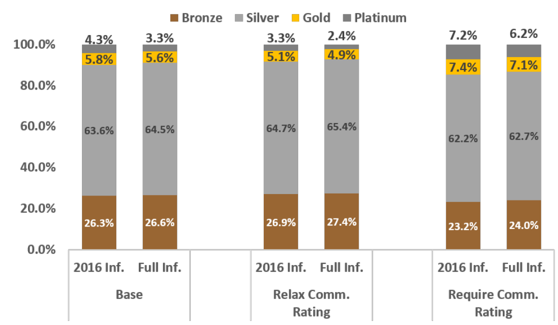
(a) Average Unsubsidized Premiums



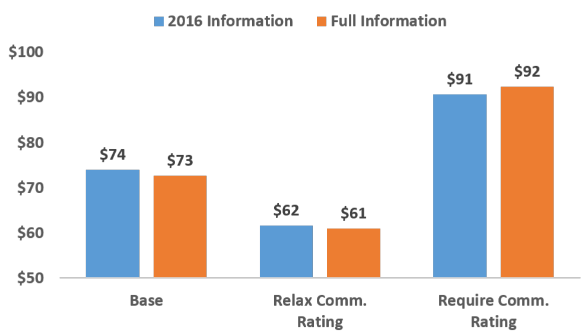
(b) Average Subsidized Premiums



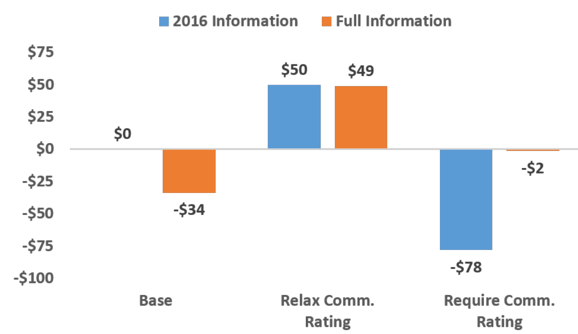
(c) Total Exchange Enrollment



(d) Market Share By Metal



(e) Marginal Claims



(f) Change in Annual Per-Capita Welfare

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), and market share by metal tier (panel d), marginal claims (panel e), the change in annual per-capita social welfare, relative to the base setting with partial information (panel f).

rating reduces total enrollment, it shifts the distribution of consumers towards the more generous gold and platinum plans. Requiring pure community rating increases the market share of gold and platinum plans by 4.5 percentage points. Conversely, relaxing community rating reduces the share of consumers in gold and platinum plans by 1.7 percentage points.

Figure 6e indicates that community rating has a substantial impact on the cost of the marginal enrollee. Relaxing community rating reduces marginal claims by 16.7%. Requiring pure community rating increases marginal claims by 22.5%. These results suggest that community rating exacerbates adverse selection. The welfare impact of community rating is shown in Figure 6f. Relaxing community rating increases annual per-capita social welfare by $50. Requiring pure community rating reduces annual per-capita social welfare by $78.[11]

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 50.4% less in unsubsidized premiums and 28.7% less in subsidized premiums. Relaxing community rating is most detrimental to 64-year-old males, who pay 13.6% more in unsubsidized premiums and 13.2% 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.

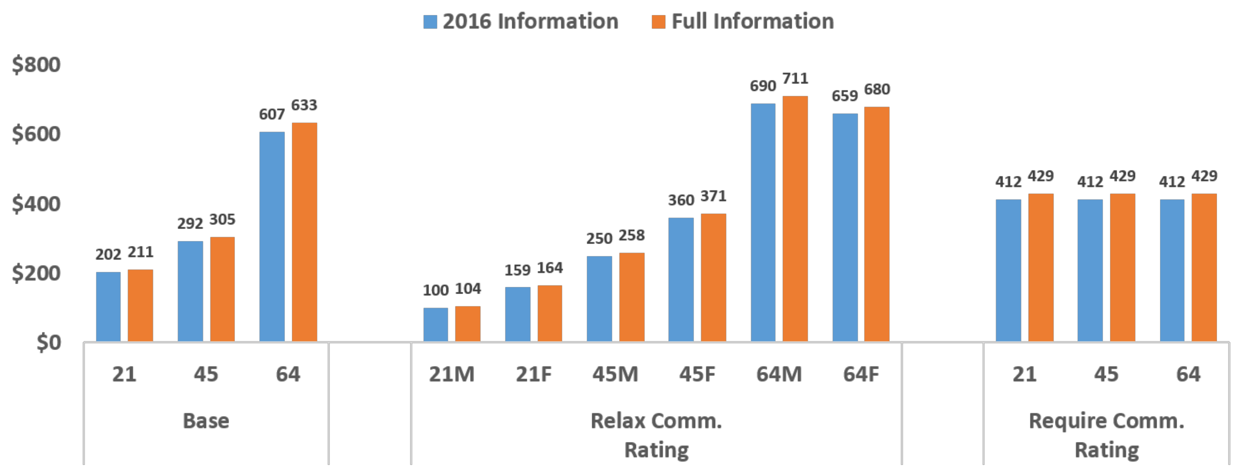
Requiring community rating has the reverse impact. Relative to the Base scenario, 64-year-olds pay 32.1% less in average unsubsidized premiums and 24.0% less in average subsidized premiums. In contrast, 21-year-olds pay 103.8% more in average unsubsidized premiums and 28.2% more in average subsidized 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.

7.2 Impact of Information in Alternative Settings

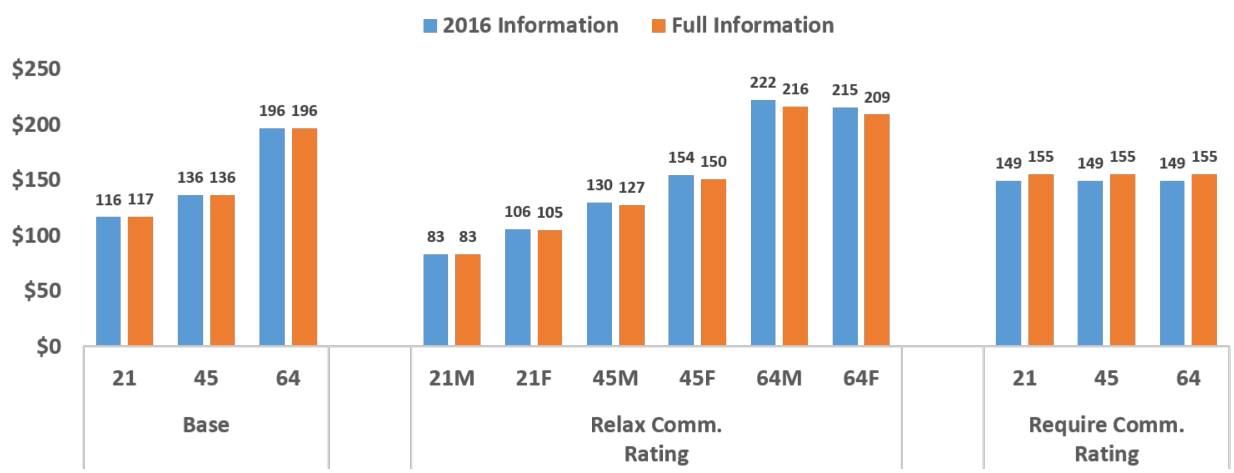
Now we compare the impact of information in the three alternative settings: (1) modified community rating under the ACA (Base setting); (2) relaxed community rating; and (3) pure community rating. Figure 6a indicates that learning the full information parameters increases average unsubsidized premiums in all three settings. The premium increase is smaller in the setting with relaxed community rating (3.3%) than the settings with modified community rating (4.2%) or pure community rating (4.1%). Learning reduces average subsidized premiums by 2.8% in the base setting

[11] When firms have full information, annual per-capita social welfare increases by $32. This small increase is the result of the 2.2% decline in total exchange enrollment, which substantially reduces federal spending on premium subsidies.

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).

and by 2.9% with relaxed community rating setting, but increases average subsidized premiums by 3.9% with pure community rating. 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 1.18% when pure community rating is required. In contrast, learning very slightly increases total exchange enrollment by 0.22% and 0.16% in the modified community rating and relaxed community rating settings, respectively. In all three settings, learning shifts the distribution of consumers from the more generous gold and platinum plans to the less generous bronze and silver plans. When firms learn the full information parameters, the market share of bronze plans increases 0.3 percentage points in the modified community rating setting, 0.5 percentage points in the relaxed community rating setting, and 0.8 percentage points in the pure community rating setting. Once again, the largest impact occurs in the pure community rating setting.

As shown in Figure 6e, the impact of information on marginal claims is also largest with pure community rating. Learning the full information parameters decreases marginal claims by 1.8% with modified community rating and by 0.9% with relaxed community rating. With pure community rating, marginal claims increase by 1.9% when firms learn the full information parameters.

Figure 6f shows the impact of information on annual per-capita social welfare is greatest in the pure community rating setting. Learning the full information parameters decreases social welfare by $34 in the modified community rating setting and has a negligible impact in the relaxed community rating setting. In contrast, learning increases annual per-capita social welfare by $76 with pure community rating.

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 premium sensitivity and inertia.

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.

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 and underestimated inertia, they initially set premiums lower than they would have with full information. Hence, assuming firms have full information leads to a less favorable conclusion about the market equilibrium. Although this finding does not necessarily generalize to other settings, it does suggest that researchers should consider whether firm uncertainty is relevant for their specific setting when analyzing a new or recently-created social program. Policymakers may need to consider whether to promote information sharing between firms to reduce uncertainty. An all-payer claims database or "active purchasing" model where insurance regulators actively solicit cost information from firms can help promote information sharing, although the U.S. Supreme Court's recent decision in *Gobeille v. Liberty Mutual* limits collection of health care claims data.

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.

Our study has some limitations. Although the primary focus of our study is firm learning, a natural concern is whether consumers 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. A useful extension of our study may consider both consumer and firm learning. 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).

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`.

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.

Hausman, J. and J. Poterba (1987). Household behavior and the tax reform act of 1986. *Journal of Economic Perspectives 1*(1), 101–119.

Ho, K. and A. Pakes (2014, December). Hospital choices, hospital prices, and financial incentives to physicians. *The American Economic Review 104*(12), 3841–3884.

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.

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.

Roesler, A.-K. and B. Szentes (2017). Buyer-optimal learning and monopoly pricing. *American Economic Review 107*(7), 2072–80.

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

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.

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 Mathematical Formulas in ACA Exchange Model

In this appendix, we write the variables in equations (5) and (8) in terms of three variables and associated partial derivatives: (1) the household choice probabilities $q_{ijt}(p;\boldsymbol{\beta})$; (2) the risk scores $r_{jmt}(p;\boldsymbol{\theta})$; and (3) plan average claims $c_{jmt}(p;\boldsymbol{\theta})$.

Household Choice Probabilities:
The household choice probabilities as defined in equation (3) are

$$q_{ijt}(p;\boldsymbol{\beta}) \quad = \quad \frac{e^{V_{ijt}(p;\boldsymbol{\beta})/\lambda}\left(\sum_j e^{V_{ijt}(p;\boldsymbol{\beta})/\lambda}\right)^{\lambda-1}}{1+\left(\sum_j e^{V_{ijt}(p;\boldsymbol{\beta})/\lambda}\right)^{\lambda}}$$

The (k,j) element of the Jacobian matrix of the household choice probability is

$$\frac{\partial q_{ikt}(p;\boldsymbol{\beta})}{\partial p_{ijt}} = \begin{cases} \beta_i^p q_{ijt}(p;\boldsymbol{\beta})\left[\frac{1}{\lambda}+\frac{\lambda-1}{\lambda}q'_{ijt}(p;\boldsymbol{\beta})-q_{ijt}(p;\boldsymbol{\beta})\right] & k=j \\ \beta_i^p q_{ijt}(p;\boldsymbol{\beta})\left[\frac{\lambda-1}{\lambda}q'_{ijt}(p;\boldsymbol{\beta})-q_{ijt}(p;\boldsymbol{\beta})\right] & k\neq j \end{cases} \tag{15}$$

where $q'_{ijt}(p;\boldsymbol{\beta})$ is the probability of choosing j, conditional on choosing a plan. Household i's demand partial derivative with respect to the firm's base plan premium p_{jmt} is

$$\frac{\partial q_{ikt}(p;\boldsymbol{\beta})}{\partial p_{jmt}} \quad = \quad \sum_{l\in J_{mt}} \frac{\partial q_{ikt}(p;\boldsymbol{\beta})}{\partial p_{ilt}(p)}\frac{\partial p_{ilt}(p)}{\partial p_{jmt}}$$

where $\frac{\partial p_{ilt}(p)}{\partial p_{jmt}}$ is given in equation (4). Total plan demand $q_{jmt}(p;\boldsymbol{\beta})$ and total firm demand $q_{ft}(p;\boldsymbol{\beta})$ equal

$$q_{jmt}(p;\boldsymbol{\beta}) = \sum_{i\in I} \left(\mathbb{I}_{i,m,t}\right) q_{ijt}(p;\boldsymbol{\beta})$$

$$q_{ft}(p;\boldsymbol{\beta}) = \sum_{i\in I, k\in J_f} q_{ikt}(p;\boldsymbol{\beta})$$

The plan and firm demand partial derivatives are

$$\frac{\partial q_{kmt}(p;\boldsymbol{\beta})}{\partial p_{jmt}} = \sum_{i\in I} \left(\mathbb{I}_{i,m,t}\right) \frac{\partial q_{ikt}(p;\boldsymbol{\beta})}{\partial p_{jmt}}$$

$$\frac{\partial q_{ft}(p;\boldsymbol{\beta})}{\partial p_{jmt}} = \sum_{i\in I, k\in J_f} \frac{\partial q_{kmt}(p;\boldsymbol{\beta})}{\partial p_{jmt}}$$

Plan Risk Scores:

We define the plan risk score in equation (9) as

$$\ln r_{jmt}(p;\boldsymbol{\theta}) \quad = \quad \sum_{d\in D} \gamma^d s_{djmt}(p;\boldsymbol{\beta}) + MT_j'\gamma^{MT} + \epsilon^r_{jmt}$$

The (k,j)-element of the Jacobian matrix of the plan risk score with respect to the premium equals

$$\frac{\partial r_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}} \quad = \quad \frac{r_{kmt}(p;\boldsymbol{\theta})}{q_{kmt}(p;\boldsymbol{\beta})} \sum_{d\in D} \gamma^d \left[\frac{\partial q_{dkmt}(p;\boldsymbol{\beta})}{\partial p_{jmt}} - s_{dkmt}(p;\boldsymbol{\beta}) \frac{\partial q_{kmt}(p;\boldsymbol{\beta})}{\partial p_{jmt}}\right] \tag{16}$$

The (k,j)-element of the Jacobian matrix of the plan risk score with respect to demand equals

$$\frac{\partial r_{kmt}(p;\boldsymbol{\theta})}{\partial q_{jmt}(p;\boldsymbol{\theta})} \quad = \quad \frac{r_{kmt}(p;\boldsymbol{\theta})}{q_{kmt}(p;\boldsymbol{\beta})} \sum_{d\in D} \gamma^d s_{dkmt}(p;\boldsymbol{\beta}) \tag{17}$$

for $k = j$ and 0 otherwise.

Plan Average Claims:

We define log average claims in equation (10) as

$$\ln c_{jmt}(p;\boldsymbol{\theta}) \quad = \quad \mu^r \ln r_{jmt}(p;\boldsymbol{\theta}) + x_j'\mu^x + \mu^l l_t + n_m'\mu^n + \epsilon^c_{jmt}$$

The (k,j)-element of the Jacobian matrix of plan average claims with respect to the premium equals

$$\frac{\partial c_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}} \quad = \quad \mu^r \frac{c_{kmt}(p;\boldsymbol{\theta})}{r_{kmt}(p;\boldsymbol{\theta})} \frac{\partial r_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}} \tag{18}$$

The (k,j)-element of the Jacobian matrix of plan average claims with respect to demand equals

$$\frac{\partial c_{kmt}(p;\boldsymbol{\theta})}{\partial q_{jmt}(p;\boldsymbol{\theta})} \quad = \quad \mu^r \frac{c_{kmt}(p;\boldsymbol{\theta})}{r_{kmt}(p;\boldsymbol{\theta})} \frac{\partial r_{kmt}(p;\boldsymbol{\theta})}{\partial q_{jmt}(p;\boldsymbol{\theta})} \tag{19}$$

Plan Age Rating Factors:

The average plan age rating factor is

$$a_{jmt}(p;\boldsymbol{\beta}) \quad = \quad \frac{\sum_{i\in I}\left(\mathbb{I}_{i,m,t}\right) a_{it} q_{ijt}(p;\boldsymbol{\beta})}{q_{jmt}(p;\boldsymbol{\beta})}$$

The (k,j)-element of the Jacobian matrix of the average plan age rating factor is

$$\frac{\partial a_{kmt}(p;\boldsymbol{\beta})}{\partial p_{jmt}} \quad = \quad \left(q_{kmt}(p;\boldsymbol{\beta})\right)^{-1}\left[\sum_{i\in I}\left(\mathbb{I}_{i,m,t}\right) a_{it}\frac{\partial q_{ikt}(p;\boldsymbol{\beta})}{\partial p_{jmt}} - a_{kmt}(p;\boldsymbol{\beta})\frac{\partial q_{kmt}(p;\boldsymbol{\beta})}{\partial p_{jmt}}\right]$$

Firm Revenue:

Total premium revenue earned by the firm is

$$R_{ft}(p;\boldsymbol{\beta}) = \sum_{i\in I, m\in M, k\in J_{fmt}} \mathbb{I}_{i,m,t}\sigma_{it} p_{kmt} q_{ikt}(p;\boldsymbol{\beta})$$

and the partial derivative with respect to the premium p_{jmt} is

$$\frac{\partial R_{ft}(p;\boldsymbol{\beta})}{\partial p_{jmt}} \quad = \quad \sum_{i\in I, k\in J_{fmt}} \mathbb{I}_{i,m,t}\sigma_{it}\left(q_{ijt}(p;\boldsymbol{\beta}) \; + p_{kmt}\frac{\partial q_{ikt}(p;\boldsymbol{\beta})}{\partial p_{jmt}}\right) \tag{20}$$

Firm Claims:

Total claims paid by the firm are

$$C_{ft}(p;\boldsymbol{\theta}) = \sum_{k\in J_{fmt}} c_{kmt}(p;\boldsymbol{\theta}) q_{kmt}(p;\boldsymbol{\beta})$$

and the partial derivative with respect to the premium p_{jmt} is

$$\frac{\partial C_{ft}(p;\boldsymbol{\theta})}{\partial p_{jmt}} \quad = \quad \sum_{k\in J_{fmt}}\left(q_{kmt}(p;\boldsymbol{\beta})\frac{\partial c_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}} + c_{kmt}(p;\boldsymbol{\theta})\frac{\partial q_{ikt}(p;\boldsymbol{\beta})}{\partial p_{jmt}}\right) \tag{21}$$

Marginal claims is

$$\begin{aligned} MC_{jmt}(p;\boldsymbol{\theta}) \equiv \frac{\partial C_{ft}(p;\boldsymbol{\theta})}{\partial q_{jmt}} \quad &= \quad c_{jmt}(p;\boldsymbol{\theta}) + \sum_{k\in J_{fmt}} q_{kmt}(p;\boldsymbol{\beta})\frac{\partial c_{kmt}(p;\boldsymbol{\theta})}{\partial q_{jmt}(p;\boldsymbol{\theta})} \\ &= \quad c_{jmt}(p;\boldsymbol{\theta}) + \mu^r c_{jmt}(p;\boldsymbol{\theta})\sum_{d\in D}\gamma^d s_{djmt}(p;\boldsymbol{\theta}) \end{aligned} \tag{22}$$

Firm Variable Administrative Cost:

Total variable administrative cost is

$$V_{ft}(p;\boldsymbol{\beta}) = v_{ft} q_{ft}(p;\boldsymbol{\beta})$$

where v_{ft} is the variable administrative cost per-member per-month. The partial derivative with respect to the premium p_{jmt} is

$$\frac{\partial V_{ft}(p;\boldsymbol{\beta})}{\partial p_{jmt}} = v_{ft}\frac{\partial q_{ft}(p;\boldsymbol{\beta})}{\partial p_{jmt}} \tag{23}$$

Firm Risk Adjustment:

The firm's risk adjustment transfer is

$$\begin{aligned}
RA_{ft}(p;\boldsymbol{\theta}) &= \sum_{m\in M,k\in J_{fmt}} q_{kmt}\left[\widehat{c}_{kmt}(p;\boldsymbol{\theta}) - \widetilde{c}_{kmt}(p;\boldsymbol{\theta})\right] \\
&= \sum_{m\in M,k\in J_{fmt}} \left(\frac{\widehat{h}_{kmt}(p;\boldsymbol{\theta})q_{kmt}(p;\boldsymbol{\beta})}{\sum_{l\in J_t}\widehat{h}_{lmt}(p;\boldsymbol{\theta})q_{lmt}(p;\boldsymbol{\beta})} - \frac{\widetilde{h}_{kmt}(p)q_{kmt}(p;\boldsymbol{\beta})}{\sum_{l\in J_t}\widetilde{h}_{lmt}(p;\boldsymbol{\theta})q_{lmt}(p;\boldsymbol{\beta})}\right)\nu R_t(p;\boldsymbol{\beta}) \\
&= \sum_{m\in M,k\in J_{fmt}} \left(rs_{kmt}(p;\boldsymbol{\theta}) - us_{kmt}(p;\boldsymbol{\theta})\right)\nu R_t(p;\boldsymbol{\beta})
\end{aligned}$$

and the partial derivative with respect to the premium p_{jmt} is

$$\begin{aligned}
\frac{\partial RA_{ft}(p)}{\partial p_{jmt}} = \nu \sum_{k\in J_{fmt}} &\left[\frac{\partial R_t(p;\boldsymbol{\beta})}{\partial p_{jmt}}\left(rs_{kmt}(p;\boldsymbol{\theta}) - us_{kmt}(p;\boldsymbol{\theta})\right)\right. \\
&+ \left. R_t(p;\boldsymbol{\beta})\left(\frac{\partial rs_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}} - \frac{\partial us_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}}\right)\right]
\end{aligned} \tag{24}$$

where

$$\frac{\partial R_t(p)}{\partial p_{jmt}} = \sum_{f\in F}\frac{\partial R_{ft}(p)}{\partial p_{jmt}} = \sum_{i\in I,f\in F,k\in J_{fmt}} \mathbb{I}_{i,m,t}\sigma_{it}\left(q_{ijt}(p;\boldsymbol{\beta}) + p_{kmt}\frac{\partial q_{ikt}(p;\boldsymbol{\beta})}{\partial p_{jmt}}\right)$$

$$\begin{aligned}
\frac{\partial rs_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}} = &\left(\sum_{m\in M,l\in J_{mt}}\widehat{h}_{lmt}(p;\boldsymbol{\theta})q_{lmt}(p;\boldsymbol{\beta})\right)^{-1}\left[\left(\widehat{h}_{kmt}(p;\boldsymbol{\theta})\frac{\partial q_{kmt}(p;\boldsymbol{\beta})}{\partial p_{jmt}} + q_{kmt}(p;\boldsymbol{\beta})\frac{\partial \widehat{h}_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}}\right)\right. \\
&- \left. rs_{kmt}(p;\boldsymbol{\theta})\sum_{l\in J_{mt}}\left(\widehat{h}_{lmt}(p;\boldsymbol{\theta})\frac{\partial q_{lmt}(p;\boldsymbol{\beta})}{\partial p_{jmt}} + q_{lmt}(p;\boldsymbol{\beta})\frac{\partial \widehat{h}_{lmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}}\right)\right]
\end{aligned}$$

$$\begin{aligned}
\frac{\partial us_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}} = &\left(\sum_{m\in M,l\in J_{mt}}\widetilde{h}_{lmt}(p;\boldsymbol{\theta})q_{lmt}(p;\boldsymbol{\beta})\right)^{-1}\left[\left(\widetilde{h}_{kmt}(p;\boldsymbol{\theta})\frac{\partial q_{kmt}(p;\boldsymbol{\beta})}{\partial p_{jmt}} + q_{kmt}(p;\boldsymbol{\beta})\frac{\partial \widetilde{h}_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}}\right)\right. \\
&- \left. us_{kmt}(p;\boldsymbol{\theta})\sum_{l\in J_{mt}}\left(\widetilde{h}_{lmt}(p;\boldsymbol{\theta})\frac{\partial q_{lmt}(p;\boldsymbol{\beta})}{\partial p_{jmt}} + q_{lmt}(p;\boldsymbol{\beta})\frac{\partial \widetilde{h}_{lmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}}\right)\right]
\end{aligned}$$

$$\frac{\partial \widehat{h}_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}} = \mathrm{IDF}_k\mathrm{GCF}_{mt}\frac{\partial r_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}}$$

$$\frac{\partial \widetilde{h}_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}} = \mathrm{AV}_k\mathrm{IDF}_k\mathrm{GCF}_{mt}\frac{\partial a_{kmt}(p;\boldsymbol{\theta})}{\partial p_{jmt}}$$

B 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%

C Complete Parameter Estimates

Table A2: Estimated Parameters

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

	$\widehat{\theta}_{2016}$	$\widehat{\theta}_{2017}$	$\widehat{\theta}_{2018}$	$\widehat{\theta}$
Monthly Premium (\$100) ×	-0.544^{***}	-0.574^{***}	-0.566^{***}	-0.544^{***}
	(0.008)	(0.007)	(0.006)	(0.005)
250% to 400% of FPL	0.188^{***}	0.213^{***}	0.201^{***}	0.183^{***}
	(0.006)	(0.006)	(0.005)	(0.004)
> 400% of FPL	0.303^{***}	0.351^{***}	0.360^{***}	0.357^{***}
	(0.007)	(0.006)	(0.005)	(0.005)
Ages 0 to 17	-0.279^{***}	-0.242^{***}	-0.236^{***}	-0.202^{***}
	(0.018)	(0.015)	(0.013)	(0.011)
Ages 18 to 34	-0.857^{***}	-0.861^{***}	-0.835^{***}	-0.794^{***}
	(0.008)	(0.007)	(0.006)	(0.005)
Ages 35 to 54	-0.374^{***}	-0.386^{***}	-0.382^{***}	-0.373^{***}
	(0.006)	(0.005)	(0.005)	(0.004)
Male	-0.147^{***}	-0.153^{***}	-0.144^{***}	-0.134^{***}
	(0.006)	(0.005)	(0.004)	(0.004)
Family	0.009	-0.011	-0.026^{***}	-0.032^{***}
	(0.005)	(0.004)	(0.004)	(0.003)
Asian	-0.195^{***}	-0.185^{***}	-0.187^{***}	-0.175^{***}
	(0.007)	(0.006)	(0.006)	(0.005)
Black	-0.295^{***}	-0.308^{***}	-0.319^{***}	-0.315^{***}
	(0.015)	(0.014)	(0.012)	(0.011)
Hispanic	-0.544^{***}	-0.544^{***}	-0.522^{***}	-0.478^{***}
	(0.008)	(0.007)	(0.006)	(0.005)
Other race	0.065^{***}	0.058^{***}	0.048^{***}	0.045^{***}
	(0.010)	(0.009)	(0.008)	(0.007)
AV	3.202^{***}	3.190^{***}	3.150^{***}	3.088^{***}
	(0.028)	(0.025)	(0.022)	(0.020)
Silver	0.578^{***}	0.658^{***}	0.720^{***}	0.739^{***}
	(0.008)	(0.007)	(0.007)	(0.006)
HMO	0.404^{***}	0.514^{***}	-0.015^{*}	-0.106^{***}
	(0.016)	(0.016)	(0.008)	(0.007)
Anthem	1.198^{***}	1.238^{***}	0.563^{***}	0.444^{***}
	(0.021)	(0.019)	(0.010)	(0.009)
Blue Shield	1.174^{***}	1.259^{***}	0.584^{***}	0.489^{***}
	(0.021)	(0.019)	(0.010)	(0.008)
Kaiser	0.854^{***}	0.790^{***}	0.674^{***}	0.629^{***}
	(0.011)	(0.008)	(0.007)	(0.006)
Health Net	0.522^{***}	0.403^{***}	0.159^{***}	0.102^{***}
	(0.010)	(0.008)	(0.007)	(0.006)
Anthem × HMO	-1.206^{***}	-1.452^{***}	-0.979^{***}	-0.955^{***}
	(0.022)	(0.022)	(0.015)	(0.014)
Nesting Parameter	0.554^{***}	0.618^{***}	0.649^{***}	0.694^{***}
	(0.005)	(0.005)	(0.004)	(0.004)

	$\widehat{\theta}_{2016}$	$\widehat{\theta}_{2017}$	$\widehat{\theta}_{2018}$	$\widehat{\theta}$
Previous Choice ×	1.979^{***}	2.006^{***}	1.804^{***}	1.608^{***}
	(0.089)	(0.068)	(0.056)	(0.050)
250% to 400% of FPL	0.310^{***}	0.331^{***}	0.329^{***}	0.266^{***}
	(0.026)	(0.019)	(0.015)	(0.014)
> 400% of FPL	0.642^{***}	0.713^{***}	0.680^{***}	0.642^{***}
	(0.040)	(0.029)	(0.023)	(0.020)
Ages 0 to 17	-0.143^{**}	-0.125^{**}	-0.149^{***}	-0.219^{***}
	(0.070)	(0.051)	(0.040)	(0.035)
Ages 18 to 34	0.007	0.011	0.072^{***}	0.064^{***}
	(0.029)	(0.022)	(0.017)	(0.015)
Ages 35 to 54	-0.005	-0.004	0.011	0.016
	(0.026)	(0.019)	(0.015)	(0.013)
Male	0.139^{***}	0.172^{***}	0.188^{***}	0.198^{***}
	(0.028)	(0.021)	(0.017)	(0.015)
Family	-0.211^{***}	-0.277^{***}	-0.310^{***}	-0.304^{***}
	(0.020)	(0.015)	(0.012)	(0.011)
Asian	-0.200^{***}	-0.259^{***}	-0.279^{***}	-0.290^{***}
	(0.025)	(0.019)	(0.015)	(0.013)
Black	0.038	-0.092^{*}	0.018	0.057
	(0.078)	(0.055)	(0.048)	(0.043)
Hispanic	0.114^{***}	0.033^{*}	0.027^{*}	0.037^{***}
	(0.027)	(0.019)	(0.016)	(0.014)
Other race	-0.164^{***}	-0.171^{***}	-0.135^{***}	-0.153^{***}
	(0.040)	(0.030)	(0.025)	(0.022)
Anthem	-0.483^{***}	-0.380^{***}	-0.005	0.568^{***}
	(0.060)	(0.046)	(0.037)	(0.033)
Blue Shield	-0.128^{**}	-0.009	0.349^{***}	1.043^{***}
	(0.065)	(0.050)	(0.040)	(0.035)
Kaiser	-0.340^{***}	-0.273^{***}	-0.023	0.267^{***}
	(0.050)	(0.037)	(0.028)	(0.019)
Health Net	-0.846^{***}	-0.914^{***}	-0.510^{***}	0.164^{***}
	(0.050)	(0.038)	(0.028)	(0.021)
HMO	0.458^{***}	0.429^{***}	0.615^{***}	0.746^{***}
	(0.043)	(0.034)	(0.030)	(0.030)
AV	1.461^{***}	1.840^{***}	1.729^{***}	1.519^{***}
	(0.096)	(0.073)	(0.059)	(0.053)
Silver	-0.637^{***}	-0.734^{***}	-0.781^{***}	-0.930^{***}
	(0.023)	(0.017)	(0.015)	(0.013)

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

	$\widehat{\theta}_{2016}$	$\widehat{\theta}_{2017}$	$\widehat{\theta}_{2018}$	$\widehat{\theta}$
Silver	0.814^{***}	0.825^{***}	0.789^{***}	0.764^{***}
	(0.062)	(0.042)	(0.033)	(0.028)
Gold	0.882^{***}	0.915^{***}	0.863^{***}	0.852^{***}
	(0.071)	(0.044)	(0.034)	(0.029)
Platinum	1.084^{***}	1.252^{***}	1.288^{***}	1.293^{***}
	(0.077)	(0.047)	(0.036)	(0.031)
Share Ages 18 to 25	-1.647^{**}	-1.347^{***}	-0.666^{*}	-0.903^{***}
	(0.802)	(0.488)	(0.366)	(0.336)
Share Ages 26 to 44	-1.330^{***}	-0.873^{***}	-0.963^{***}	-0.913^{***}
	(0.395)	(0.217)	(0.160)	(0.143)
Share Male	-0.350	-0.047	0.106	-0.339^{*}
	(0.653)	(0.299)	(0.214)	(0.193)
Share Hispanic	-0.234	-0.348^{***}	-0.652^{***}	-0.741^{***}
	(0.184)	(0.130)	(0.097)	(0.085)

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

	$\widehat{\theta}_{2016}$	$\widehat{\theta}_{2017}$	$\widehat{\theta}_{2018}$	$\widehat{\theta}$
HMO	0.013	0.036	-0.130^{***}	-0.160^{***}
	(0.064)	(0.022)	(0.010)	(0.010)
Log risk score	1.075^{***}	1.053^{***}	1.045^{***}	1.059^{***}
	(0.009)	(0.004)	(0.003)	(0.004)
Trend	-0.026^{***}	-0.006^{**}	0.022^{***}	0.021^{***}
	(0.007)	(0.003)	(0.002)	(0.002)
Anthem	0.125^{*}	0.087^{***}	0.117^{***}	0.116^{***}
	(0.069)	(0.025)	(0.018)	(0.020)
Blue Shield	0.027	0.060^{*}	-0.053^{*}	-0.092^{***}
	(0.083)	(0.034)	(0.029)	(0.033)
Health Net	-0.030	-0.044	0.085^{**}	0.134^{***}
	(0.095)	(0.047)	(0.036)	(0.038)
Kaiser	-0.128^{***}	-0.086^{***}	-0.010	0.060^{***}
	(0.045)	(0.016)	(0.015)	(0.017)

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 (1) and (10) are omitted.