Competing on Information in Selection Markets: Evidence from Auto Insurance*

Marco Cosconati — IVASS
Yi Xin — Caltech
Fan Wu — Caltech
Yizhou Jin — Toronto

June 7, 2024

Abstract

This paper develops a novel method for empirically analyzing competitive equilibrium in selection markets where firms offer differentiated products and have different cost structures and *information precision*. We apply this method to study a representative sample of Italian auto insurance contracts and associated claims from 2013 to 2021. We find substantial differences in the precision of risk rating across insurers, and insurers with less accurate risk-rating algorithms tend to have more efficient cost structures. If all insurers were to counterfactually adopt the least advanced information technology (e.g., under more stringent privacy regulations), the average consumer surplus would increase by 4.5%, where the gain is enjoyed primarily by high-risk consumers. Conversely, allowing all firms to share the best risk-rating technology improves the efficiency of the insurer–insuree match: high-risk consumers are more likely to be covered by insurers with lower claim processing costs, and the cost to insure consumers decreases by 4 euros per contract.

*Cosconati: IVASS, via del Quirinale, 21, 00187, Roma. Xin: Division of the Humanities and Social Sciences, California Institute of Technology, 1200 East California Blvd, MC 228-77, Pasadena, CA 91125. Wu: Division of the Humanities and Social Sciences, California Institute of Technology, 1200 East California Blvd, MC 228-77, Pasadena, CA 91125. Jin: Rotman School of Management, University of Toronto, 105 St. George Street, Toronto, Ontario M5S 3E6.

The opinions and errors in this paper are solely those of the authors and do not reflect the views of IVASS and the Bank of Italy.

1 Introduction

Following the seminal works of Akerlof (1970) and Rothschild and Stiglitz (1976), the modern information economics literature has remained largely theoretical until the recent renaissance in empirical studies of selection markets (such as credit or insurance markets), where "consumers vary not only in how much they are willing to pay for a product but also in how costly they are to the seller" (Einav et al., 2021). The interaction of market power and selection has garnered great interest in recent empirical work (see Cabral et al.,2018; Crawford et al., 2018; Nelson, 2018; Cuesta and Sepúlveda, 2021). However, these studies tend to assume that firms have *symmetric knowledge* about consumers and thus compete subject to the same degree of information asymmetry.

This paper develops a novel empirical framework to study imperfect competition in selection markets when firms have *heterogeneous information* about consumers. In selection markets, firms' profitability hinges not only on prices and quantities but also on their capacity to identify and attract low-cost (or high-value) consumers. Understanding how competition and markups change when some firms gain superior information on consumers is crucial for welfare analysis and policy evaluations. With the rapid development of big data, artificial intelligence, and machine learning, the importance of heterogeneous information has become even more relevant. Competing firms may have access to different amounts of data and use various pricing algorithms.[1] This phenomenon has raised important questions for competition policies beyond the traditional perspectives.

Empirical research on the equilibrium consequences of heterogeneous information has lagged behind for several reasons.[2] First, modeling equilibrium requires market-wide data, which are hard to obtain in general.[3] Collecting market-wide data on prices and consumer characteristics becomes even more challenging when firms compete on information, as firms with informational advantages often have strong incentives not to share their data (Jin and Vasserman, 2021; Jin and Wagman, 2021). Second, it is not straightforward how to

[1]Hendricks et al. (1987), Hendricks and Porter (1988), Hendricks et al. (1994), and Porter (1995) use data from oil and gas drainage lease auctions and show that "neighbour firms" are more informed about lease values compared to "non-neighbour firms," and their equilibrium participation and bidding strategies are different. More recently, Chen and Tsai (2023) find evidence that platform owner has an information advantage over third-party sellers on a vertically integrated e-commerce platform. Byrne et al. (2023) document asymmetric access to competitors' price information in retail gasoline markets.

[2]The impact of heterogeneous information on equilibrium outcomes has long interested theoretical scholars. Wilson (1967) explains how information asymmetries among bidders can shape equilibrium bidding in common-value auctions. Subsequent work, including Wilson (1977), Milgrom (1979), Milgrom (1981), Engelbrecht-Wiggans et al. (1983), and Maskin and Riley (2000), find that several theoretical results are sensitive to bidder asymmetries.

[3]As pointed out by Salanié (2017), it is relatively easy to obtain high-quality data from a single firm, but to understand and model equilibrium, "the most crying need is for market-wide data."

empirically measure information heterogeneity from equilibrium market outcomes (such as prices). For instance, an auto insurer with a more accurate risk-rating algorithm can mimic its low-accuracy competitors by charging a high premium to a safe driver if it leads to an increase in profits. Moreover, heterogeneous information may interact with cost structures and other supply-side factors in complex ways. Therefore, a rich and tractable empirical framework is needed to incorporate heterogeneous information to model competition in selection markets.

In this paper, we provide a new method to empirically analyze competitive equilibrium in selection markets where firms offer differentiated products, have different cost structures, and possess different *information precision*. This empirical framework allows us to (1) quantify the degree of supply-side heterogeneity (in information precision and cost structures), and (2) evaluate the equilibrium impact of policies that eliminate information heterogeneity. We apply the framework to a unique market-level micro dataset that contains a representative sample of Italian auto insurance contracts from *all* insurers between 2013 and 2021. The Italian auto insurance industry is well suited for our analysis. Liability insurance is mandatory for all drivers in Italy, and similar contract features are shared by all insurers. Furthermore, the market is highly segmented and displays wide price dispersion.[4] Our data include demographic information, vehicle characteristics, contract features, transaction prices, and claim records for each customer over multiple years. Importantly, policyholders are followed through changes of insurers.

Motivated by the institutional setting, we develop a static model of price competition among insurers. Each firm draws a private signal about the consumer's risk type and infers the expected cost to insure the consumer based on the signal. The dispersion of the signal distribution is determined by the firm's information precision. Firms' pricing strategies are functions of their risk evaluations. In equilibrium, firms optimally choose pricing coefficients to maximize their profits, taking into account their opponents' strategies and demand-side responses.[5] Our model allows for rich and flexible supply-side heterogeneity: insurers can differ in their benefits of contracting with a customer, efficiency at processing claims, and unobserved product attributes (e.g., quality and brand loyalty). The demand side of our model follows the standard differentiated product framework as in Berry (1994) and Berry et al. (1995).

[4]Extensive price dispersion has also been documented in consumer loans (e.g., Edelberg, 2006; Einav et al., 2013; Nelson,2018), mortgage markets (e.g., Allen et al., 2014), and health care (e.g., Cooper et al., 2019; Brown, 2019).

[5]Matcham (2023) studies lenders' credit limit choices, allowing them to have different screening technologies. The paper assumes that signal precision affects lender's optimal credit limit choices only non-competitively. In our model, signal precision plays a key role in firms' competition and it endogenously determines the equilibrium pricing strategies and consumer sorting.

Our paper provides novel econometric techniques to identify and estimate demand parameters, insurer-specific information precision, and cost structures when only transaction prices are available. We first estimate the individual-specific cost to the insurers (i.e., the consumer's risk type) using a panel of claim records realized after the issuance of insurance policies.[6] Next, we nonparametrically estimate the premium distribution conditional on each risk type *after* consumers are sorted into different insurers. This price distribution is different from the one consumers face *ex ante*, which is not observed in the data because, as in many other empirical settings, we have access to transaction prices but not to offered prices. One key observation we highlight is that the price distributions before and after selection occurs are connected through the consumers' choice probabilities. In addition to the classic loop over mean utilities that matches aggregate market shares (Berry et al., 1995), our method introduces an innovative outer loop that iterates over consumers' sorting propensities. This new approach allows us to jointly estimate consumers' choice probabilities, firm-specific offered price distributions, and demand parameters without imposing any parametric restrictions on the price distributions.

The offered price distribution for each firm plays an important role in identifying the signal distribution. The key insight is that the offered price is a monotonically increasing function of the firm's signal, similar to auction models in which bids are monotonically increasing in bidders' valuations. Using the one-to-one mapping between the offered price and the signal, we derive the model-implied joint distribution of premiums and risks to match the data. We show that the correlation between premiums and risk types within a firm monotonically decreases with the signal variance. Intuitively, when the signal distribution is very informative, consumers' risks are better reflected in their premiums, and vice versa.[7] Finally, we estimate firms' cost parameters (i.e., cost of managing contracts and efficiency at processing claims) using the first-order conditions derived from their profit maximization problem. We also show that given the linear structure imposed in firms' pricing strategies, the equilibrium pricing coefficients can be estimated using only the first and second moments of transaction prices and consumers' risks, which significantly reduces the computational burden.

We apply the model and the estimation strategy to the Italian auto insurance industry, focusing on a sample of new customers in Rome.[8] Our structural estimation suggests that

[6]Similarly, Porter (1995) estimates the ex-post value of oil leases (which is ex-ante unknown to bidders) based on amounts of extracted oil, industry costs, and the price of oil, etc. Biglaiser et al. (2020) use the ex-post likelihood of resale as a measure of the quality of used cars and they find that dealers have better information about car quality.

[7]A similar intuition appeared in Panetta et al. (2009), who argue that banks with superior screening abilities should charge an interest rate that is more "sensitive" to the firm's default risk.

[8]Focusing on new customers is common in empirical analysis of asymmetric information to avoid com-

there is a significant amount of firm heterogeneity along *all* dimensions. In particular, we find substantial differences in the precision of risk rating across insurers, and the ones with lower information precision tend to attract more high-risk customers.[9] We also find that insurers enjoy large and different net benefits from contracting with new customers (potentially due to inertia and dynamic pricing strategies) ranging from 298 to 883 euros per contract, and their efficiency at processing claims also differs significantly. Another interesting finding to highlight is that insurers with less accurate risk-rating algorithms tend to have more efficient cost structures, leading to distinct types of comparative advantages.

To understand the equilibrium consequences of heterogeneous information among competing firms, we consider information policies that require insurers to adopt the same—either the least or the most advanced—risk-rating technology. We find that if all insurers were to counterfactually adopt the least advanced information technology in risk rating (e.g., under more stringent privacy regulations), average consumer surplus would increase by 4.5% due to a price reduction of 5.1%. The gain in consumer surplus, however, is mainly enjoyed by high-risk drivers. On the other extreme, if all insurers were to share the most advanced risk-rating technology (e.g., under antitrust laws enforcing data/algorithm sharing) the surplus for low-risk consumers would increase by 3.9%, and the firms' average profit would increase by 5.4%.

Homogenizing information precision across insurers has an important impact on consumers' sorting patterns (i.e., high- and low-risk consumers are selected into different firms). When firms adopt equally bad information technology, risk sorting across insurers becomes much less significant compared to the data. When firms share the best risk-rating technology, we observe a significant amount of consumer sorting, which is driven primarily by specialization. Insurers more efficient at processing claims adjust their pricing strategies to target high-risk consumers; overall, the cost to insure consumers decreases by 4 euros per contract. This result is illuminating: when firms have heterogeneous cost structures, even if insurance is mandatory, reducing information asymmetry could still be welfare improving as it would result in a more efficient matching between insurers and insurees.

Lastly, we want to emphasize that information policies can have a differential impact on competing firms. Firms with less advanced risk-rating technology benefit the most, regardless of whether the policies improve or reduce the overall risk-rating precision in the market. This finding is striking and has broad implications for competition policies in the digital

plicated issues of dynamic pricing, and consumer/firm learning. See the discussions in Chiappori and Salanie (2000) and Crawford et al. (2018).

[9]Relatedly, Einav et al. (2013) document that high-performing dealers are better at assessing borrowers' risks in auto loans. Grodzicki (2023) shows that the computing and risk-rating capacities of banks have increased on average and become more dispersed over the past 20 years.

era. Information policies that aim at protecting consumer privacy or promoting competition for the benefit of consumers may in fact generate anti-competitive effects, especially when the larger firms in the market do not have the more advanced information technology as we observe from the data.

Related Literature Our paper contributes to the large empirical literature on selection markets (for an excellent review, see Einav et al. (2021)). In particular, our paper develops the first tractable empirical framework for analyzing imperfect competition when firms have heterogeneous information about consumers. The seminal works of Einav et al. (2010a) and Einav and Finkelstein (2011) laid the foundation for analyzing perfectly competitive selection markets. Mahoney and Weyl (2017) present a theoretical model of symmetric oligopoly competition. Recent empirical studies focus on the interaction between asymmetric information and market power using data from credit and insurance markets (e.g., Cabral et al., 2018; Crawford et al., 2018; Nelson, 2018; Cuesta and Sepúlveda, 2021), all under the assumption that information is symmetrically distributed among competing firms.[10] Importantly, our paper also incorporates multidimensional cost heterogeneity on the supply side and studies how it interacts with heterogeneous information.[11]

From a methodological point of view, our paper extends the classic demand estimation techniques by introducing a "selection loop" on top of the contraction mapping for mean utilities pioneered by Berry (1994) and Berry et al. (1995). We provide a new econometric method to recover firms' equilibrium pricing coefficients when only transaction prices are available. Our method for identifying and estimating information precision for competing firms leverages the one-to-one mapping between the offered price and the signal, which is related to the empirical nonparametric auction literature (Guerre et al., 2000).

Our paper also contributes to the recent discussions on antitrust policies and consumer protection with the rise of big data (Jin and Wagman, 2021). We provide new empirical evidence suggesting that regulatory attempts to protect consumer privacy may affect the competitive structure of the industry and disproportionately benefit firms with less advanced information technology (see related discussions in Campbell et al., 2015; Johnson et al., 2023; Alcobendas et al., 2023). Data/algorithm-sharing policies could lead to a more cost-efficient matching between consumers and competing firms, but these policies may hurt consumer welfare as they allow firms to better segment markets and more effectively

[10] A recent paper by Friedrich et al. (2023) studies participants' strategic responses to information asymmetry in a two-sided matching market (between students and medical school programs). They find that one of the programs was significantly better at learning about students' unobserved talents after adding a knowledge test and a personal interview.

[11] Existing papers allowing for multidimensional cost heterogeneity include Salanié (2017), Serna (2023), and Nelson (2018).

price discriminate (Taylor, 2004).

The rest of the paper is organized as follows. In Section 2 we introduce the institutional background of the Italian auto insurance market and briefly summarize the data. We then present the model and estimation strategies in Sections 3 and 4, respectively. Section 5 reports the empirical estimation results and model fit. The details of our counterfactual experiments are provided in Section 6. Section 7 concludes.

2 Institutional Background

Background and Data Our paper focuses on the market for mandatory third-party liability insurance for motor vehicles (i.e., *Responsabilità Civile Auto*) in Italy. This insurance covers damage to third parties' health and property in accidents where one is "not at fault."[12] The policies last for one year and are exclusive. Consumers decide whether to renew the contract or change insurers at the end of the year. The Italian mandatory liability insurance market is large; for example, in 2018, about 31 million contracts were underwritten, and over 4 million claims were filed. There are more than 50 competing firms nationwide.

The data we use for this research come from a micro dataset, IPER (*Indagine sui Prezzi Effettivi RC Auto*), collected by IVASS, the Italian insurance supervising authority. The IPER dataset covers a nationally representative sample of matched insurer–insuree panel with rich information on observable risk factors, premiums, coverage, and contractual clauses. The data also include information on the frequency and severity of claims for each sampled consumer in each contract year. Importantly, policyholders are tracked through changes of insurers.

Contract Design and Premium Two important features of the contract design stand out in the Italian auto insurance market. First, only a negligible fraction of contracts feature a deductible, and therefore consumers do not face complex deductible choices. Second, the law establishes a mandatory minimum of liability coverage.[13] The liability limits are 1 million and 6 million euros for property and health damage, respectively. In practice, claim payouts almost never exceed the mandatory minimum liability limits, and therefore consumers essentially enjoy full coverage.

The premiums of the contracts are determined by insurers' actuarial algorithms based

[12]The Italian system follows the common fault-based system. In the event of an accident, all individuals who experience physical damages, except for the liable driver, are compensated by the company covering the liable vehicle.

[13]That is, if the damage to third parties exceeds this limit, the policyholder is responsible for any amount exceeding the limit.

on various risk customization factors such as age, residence, and driving record. In Italy (as in many other countries), a uniform experience rating system relates the history of accidents to a class of risk, known as bonus-malus (BM) class. There are 18 BM classes with class 1 being the best. Young drivers with no driving history are assigned to BM class 14; if no accident occurs, the BM class decreases by one each year. The experience rating system suffers from the well-known saturation problem. The BM class, although publicly available, is not very informative for purposes of risk evaluations: about 80% of policyholders are in class 1. This provides strong incentives for firms to collect additional information and develop pricing algorithms to more accurately price consumers' risks.

In the data, we observe a set of commonly used factors for motor insurance pricing, including consumer characteristics (e.g., age, BM class, number of accidents over the previous five years), vehicle features (e.g., power, years since registry), province of residence, and city-level geographic characteristics. These observed factors explain around 50% of the price variations in the data, suggesting that premiums also depend on other information unobservable to researchers.[14] Moreover, the set of variables and algorithms used for pricing varies across firms. For example, some firms take into account the presence of safety devices, zip code, occupation, and socioeconomic variables, while others do not. Sales agents in some firms may have more experience and better knowledge about consumers' driving habits. These factors contribute to the heterogeneity of firms' risk-rating precision.

Premiums can also vary due to optional contractual clauses and the number of installments chosen by consumers.[15] In this paper, we abstract away from consumers' choices of additional contract features and payment method; instead, we treat them as consumer characteristics that can directly affect premiums and risks. We include these variables as controls when we predict individual expected claim payouts (i.e., consumers' risk types), as detailed in Section 4.1. We model additional contractual clauses in a reduced-form way, instead of a full-fledged structural analysis, mainly because our paper focuses on price competition when insurers have heterogeneous risk-rating technology and cost structures. Introducing an additional layer of endogenous contract design and consumer sorting into different types of contracts would be even more challenging, but it is an interesting avenue for future research.[16]

[14]The unobservable pricing variables include consumer/vehicle characteristics that directly enter firms' actuarial algorithms but are not available in the dataset, as well as "soft information" (e.g., a sales agent's perception of a consumer's risk type).

[15]For example, the "Exclusive Driving" clause lowers the premium, but the vehicle can be used only by the driver indicated in the contract. The "Protected Bonus" clause, which increases the premium, prevents the future premium from going up as much if an at-fault accident occurs. Not paying the entire amount upfront (i.e., in one installment) also implies a surcharge.

[16]Empirical work endogenizing contract designs in selection markets is scarce. Einav et al. (2012) analyze the optimal design of minimum down payment and interest rate in auto loans. In our context, there are more

Summary Statistics In our empirical analysis, we focus on a sample of new customers in Rome's metropolitan area.[17] Our sample includes 124,592 contracts sold between 2013–2021. Table 1 reports the summary statistics for contract premiums, claim frequency and severity, and consumer and vehicle characteristics. In our sample, the average annual premium paid by the consumers is 478 euros. Accidents are rare events: on average, the number of claims filed in each contract year is around 0.08. Around 56% of the sampled consumers are male. The average ages of drivers and vehicles are 48 and 8.3, respectively.

Table 1: Summary statistics

Variables	Mean	Std. Dev.	Min	Max	N
Premium (€)	478.06	209.03	134.26	1,334.70	124,592
Claim size (€)	264.90	10,213.62	0	2,521,014	124,592
No. of claims (within contract year)	0.08	0.30	0	6	124,592
No. of accidents in last 5 years	0.81	1.22	0	3	124,592
BM class	2.06	2.51	1	15	124,592
Age	48.24	14.11	18	99	124,592
Man	0.56	0.50	0	1	124,592
Median city	0.10	0.30	0	1	124,592
Big city	0.62	0.49	0	1	124,592
Car age	8.30	5.27	0	19	124,592
Horsepower	66.88	26.84	0	493	124,592
Petrol vehicle	0.52	0.50	0	1	124,592
One installment	0.67	0.47	0	1	124,592

We plot the average premium and claim payouts for each firm separately in Figure 1, where the size of the circle represents the market share of each firm. We observe a substantial cross-firm price dispersion in this market. For two insurers with similar market size and average claim payouts, the difference in their average premiums could exceed 100 euros. Moreover, firms differ significantly in their average claim payouts, suggesting that different types of consumers are sorted into different insurers. What are the main drivers of price variation and risk sorting observed in this market? To better understand firms' equilibrium pricing strategies and consumers' sorting patterns, a tractable empirical framework that incorporates rich heterogeneity in firms' risk-rating precision, cost structures, as well as product differentiation is needed.

than 10 optional contractual clauses offered by the insurers. Allowing endogenous choices in this extremely large space of contract features is exceptionally challenging. Several recent papers incorporating endogenous product attributes in conventional markets include Sweeting (2013), Crawford et al. (2019), Fan and Yang (2020), and Barwick et al. (2023).

[17]We define customer tenure as the number of years since the customer initiated her contract with the insurer. Our sample includes all customers with tenure equal to 0 in a given contract year for a given insurer (i.e., the customers are new to the insurer).

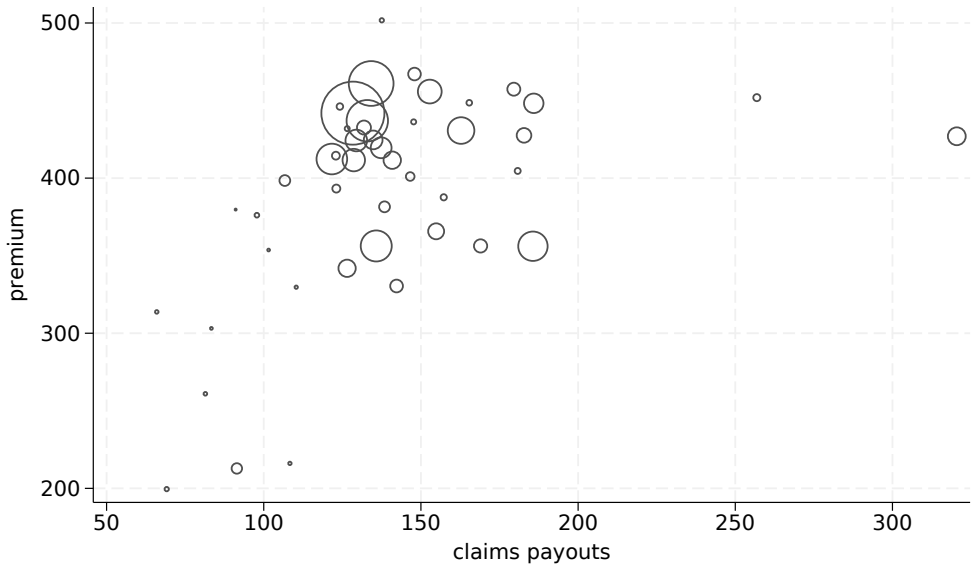


Figure 1: Heterogeneity across firms: Average premium and claim payouts (in euros). The size of the circle represents the market share of each firm.

3 Model

We take as given that there are J insurers in the market, indexed by $j = 1, 2, \cdots, J$. Each of them offers a standardized insurance product that shares the same contract features. The insurance contract is exclusive and mandatory for all drivers. In other words, consumers have to purchase the product from one of the companies and there is no outside option. Let θ denote a consumer's true risk type, which can be measured by the expected cost of insuring a consumer within a contract year. The risk type θ is *not* observed by insurers ex ante. The population density of θ is denoted by $f_0(\theta)$, which is common knowledge among all firms.

Signal Structure When a consumer of type θ arrives, firm j draws a signal $\hat{\theta}_j$ from a cumulative distribution function $F_j(\hat{\theta}_j|\theta)$, with its density denoted by $f_j(\hat{\theta}_j|\theta)$. Intuitively, the signal $\hat{\theta}_j$ can be viewed as a risk score assigned to a consumer by firm j based on information it has access to. We assume that the signals are private and independent across firms conditional on θ. This setting is similar to that of common value auctions, where firms' signals can be interpreted as their noisy estimates of the true but unknown common value, i.e., the expected claim payouts to the consumer.

The signal distribution $F_j(\cdot|\theta)$ captures firm j's information technology for risk evaluation. For example, the signal can admit a normal distribution $N(\theta, \sigma_j^2)$, where σ_j measures

the dispersion of the signal distribution around the true risk type, and therefore $1/\sigma_j$ captures firm j's information precision. When $\sigma_j = 0$, firm j perfectly observes θ; as σ_j increases, the signal becomes less precise. In reality, the set of variables and algorithms used in the risk-rating process may differ across firms, leading to heterogeneity in the signal distribution $F_j(\cdot|\theta)$.

Risk Rating Firm j infers the risk type of the consumer upon observing the signal $\hat{\theta}_j$. Let $f(\theta|\hat{\theta}_j, j)$ denote the posterior density of θ conditional on the firm receiving a signal $\hat{\theta}_j$ and the consumer being *selected* into firm j. The posterior mean $E(\theta|\hat{\theta}_j, j)$ represents firm j's risk rating for a consumer conditional on signal $\hat{\theta}_j$ and her choosing j. Specifically,

$$E(\theta|\hat{\theta}_j, j) = \int_\theta \theta f(\theta|\hat{\theta}_j, j) d\theta = \frac{\int_\theta \theta Pr(j|\hat{\theta}_j, \theta) f_j(\hat{\theta}_j|\theta) f_0(\theta) d\theta}{\int_\theta Pr(j|\hat{\theta}_j, \theta) f_j(\hat{\theta}_j|\theta) f_0(\theta) d\theta}, \tag{3.1}$$

where $Pr(j|\hat{\theta}_j, \theta)$ is the probability that a consumer of type θ chooses firm j given the signal $\hat{\theta}_j$. Note that $Pr(j|\hat{\theta}_j, \theta)$ is an equilibrium object because it depends on all firms' signal precision and pricing strategies, as well as consumers' demand responses. Consequently, $E(\theta|\hat{\theta}_j, j)$ is an equilibrium object.

Pricing Strategy Firm j sets the price for a consumer based on their risk evaluation. We assume that the firm's pricing strategy takes the following form:

$$p_j(\hat{\theta}_j) = \alpha_j + \beta_j \underbrace{E(\theta|\hat{\theta}_j, j)}_{\text{risk rating}}, \tag{3.2}$$

where α_j and β_j are pricing coefficients optimally chosen by the firm. Specifically, α_j reflects the firm's baseline markup and β_j relates to the elasticity of price with respect to risk rating.

The price in our model is a function of the signal that the firm receives. Prices may vary with the signals in a complicated nonlinear way, even if we assume that price is a linear function of risk rating. This is simply because $E(\theta|\hat{\theta}_j, j)$ itself is a complex function of the signal due to self-selection into insurers, as shown in Equation (3.1). The linear structure imposed in Equation (3.2) is *not essential* for our model. More generally, we can assume that prices are monotonically increasing functions of risk rating, i.e., $p_j(\hat{\theta}_j) = \tau_j(E(\theta|\hat{\theta}_j, j))$. Imposing a linear pricing strategy makes our identification argument more transparent and significantly reduces the computational burden, as we will show in Section 4.

Demand We now describe the consumer's choice problem. We follow the standard differentiated product framework and assume that the level of utility that a consumer derives from a product is a function of the price and product characteristics. The insurance plans offered in this market are homogeneous in terms of observable characteristics, but may have unobserved (by the econometrician) product attributes such as service quality or brand loyalty, which we denote by ξ.[18] The utility derived by consumer i from a product of firm j is given by the scalar value

$$U_{ij} = -\gamma p_j(\hat{\theta}_j) + \xi_j + \varepsilon_{ij}. \tag{3.3}$$

We assume that ε_{ij} follows a type I extreme value distribution and is independent across all firms and individuals. The probability that a consumer chooses firm j given a vector of signals $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_J)$ drawn by all firms is

$$Pr(j|\hat{\boldsymbol{\theta}}) = \frac{\exp(-\gamma p_j(\hat{\theta}_j) + \xi_j)}{\sum_{j'=1}^{J} \exp(-\gamma p_{j'}(\hat{\theta}_{j'}) + \xi_{j'})}. \tag{3.4}$$

Although θ does not directly enter consumer's utility, it affects the choice probabilities through the signal distribution.

Profit Maximization Each firm is assumed to choose its pricing coefficients (α_j, β_j) to maximize its profit given the pricing strategies of competing firms. We assume that this is a static, complete information game, and therefore the primitives of each firm (e.g., their signal distributions, cost parameters, and unobserved product attributes) are common knowledge.

We use $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_J)$ and $\boldsymbol{\beta} = (\beta_1, \beta_2, \cdots, \beta_J)$ to denote the pricing coefficients of all firms. The profits of firm j are

$$\pi_j(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \int_{\theta} \int_{\hat{\boldsymbol{\theta}}} \Bigg(\underbrace{p_j(\hat{\theta}_j) + c_j - k_j\theta}_{\text{net profit}} \Bigg) \underbrace{Pr(j|\hat{\boldsymbol{\theta}})}_{\text{choice prob.}} \Bigg(\underbrace{\prod_{j'=1}^{J} f_{j'}(\hat{\theta}_{j'}|\theta)}_{\text{signal dist.}} \Bigg) \underbrace{f_0(\theta)}_{\text{type dist.}} d\hat{\boldsymbol{\theta}} d\theta. \tag{3.5}$$

The first term in Equation (3.5) represents firm j's net profit from servicing a consumer with type θ given the price $p_j(\hat{\theta}_j)$. We use c_j to denote the "net benefit" from contracting with a customer irrespective of her risk type. Insurers need to pay administrative costs (e.g., setting up the profile) when managing contracts, but may also receive future benefits from

[18] It is straightforward to extend our demand model to allow for observable product attributes. Alternatively, we can interpret ξ_j as firm-specific mean utilities, which depend on both observed and unobserved product attributes.

contracting with new customers due to inertia and dynamic pricing strategies.[19] Our paper focuses on a static price competition for new customers and does not provide a full-fledged dynamic analysis of firms' pricing strategies and consumers' switching behavior. We treat c_j as a primitive of our model. In the estimation section, we calibrate the expected future benefits from contracting with a new customer using the data and then isolate the marginal cost for managing contracts from c_j.

Another cost component in our model is $k_j\theta$, which represents the cost related to claim payouts. In addition to the indemnity paid to the consumers following an accident, which is measured by θ directly, insurers may incur extra costs of processing the claims. For instance, the insurer may need to hire adjusters to inspect the vehicle damage. The scalar k_j essentially measures the efficiency of the firm in processing claims.

The term $Pr(j|\hat{\boldsymbol{\theta}})$ in Equation (3.5) is the probability that a consumer is selected into firm j given the prices and product attributes across all firms. The third term in Equation (3.5) represents the joint distribution of signals from all firms, which, under our conditional independence assumption, can be written as the product of conditional densities. The last term $f_0(\theta)$ denotes the population distribution of consumers' risk type. Each firm integrates over the joint distribution of types and signals $(\theta, \hat{\boldsymbol{\theta}})$ to compute its expected profits.

To summarize, our model allows for rich supply-side heterogeneity. Insurers are differentiated by their cost structures (c_j, k_j), information technology (F_j), and unobserved product attributes (ξ_j). All of these are primitives of the firms and are allowed to be arbitrarily correlated. Throughout the paper, we maintain the assumption that firms' primitives cannot be adjusted within a relatively short period of time. Insurers optimally choose their pricing coefficients (α_j, β_j) to maximize profits, taking into account their opponents' strategies and demand-side responses.[20]

[19]For example, Handel (2013) identifies substantial inertia in health insurance markets. Honka (2014) documents that in the US auto insurance industry, about 70% of consumers stay with their insurance provider after the first year, and this percentage increases with tenure. High switching costs provide incentives for firms to use a "bargain-then-ripoff" pricing strategy (Farrell and Klemperer, 2007; Ericson, 2014). Cosconati (2023) documents that in Italian auto insurance markets, new customers on average receive a discount of 91 euros relative to old customers; premiums increase rapidly in the first five years and gradually become inelastic to tenure.

[20]The payoff functions π_j in Equation (3.5) are continuous in $(\boldsymbol{\alpha}, \boldsymbol{\beta})$. We verify numerically that π_j are concave in (α_j, β_j). Since the strategy spaces are nonempty compact convex subsets of a Euclidean space, a pure-strategy Nash equilibrium of our pricing game exists following Fudenberg and Tirole (1991, Theorem 1.2).

4 Estimation

There are three sets of parameters we need to estimate for the model. We first estimate consumers' risk types (i.e., the expected claim payouts to consumers) using a panel of claim records. Second, on the demand side, we estimate consumers' price sensitivity and the unobserved product attributes for each firm using a novel fixed-point algorithm. Finally, the supply-side parameters include the firms' pricing coefficients, signal distributions, and cost parameters. We exploit (1) the joint distribution of premiums and risk types for each firm and (2) the first-order conditions derived from each firm's profit maximization problem to identify and estimate the supply-side parameters. The details of each step are provided in this section.

4.1 Risk Type

A consumer's true risk type θ is ex-ante unknown by the insurer. However, the type is eventually revealed after the consumer is observed over multiple periods. In other words, the consumer's risk type is ex-post estimable given realized claim data. Existing literature on auto insurance markets often has access to contract and claim data from a single insurer (e.g., Cohen, 2005; Cohen and Einav, 2007; Jeziorski et al., 2017). Estimating consumers' risks using data from one firm is susceptible to selective censoring issues. This is because in the auto insurance markets, high-risk drivers are more likely to switch, especially after being involved in an accident (Cohen, 2005; Cosconati, 2023). The advantage of our dataset is that we are able to track individuals across different firms over a long period of time.

Given a panel dataset of claim records, which includes the number of accidents and the amount paid to the consumers following an accident in each contract year, we estimate an individual-specific riskiness measure using a two-part model. Specifically, we estimate claim frequency and severity separately and multiply the two to represent consumers' risk level. This two-part approach is widely adopted in the auto insurance industry. Actuaries rely heavily on the generalized linear models (GLM) and they typically model claim frequency and severity separately (see more details and discussions in Goldburd et al., 2016).

Let i be the index of each customer and t be the index of a contract year. For each consumer in the dataset, we observe a vector of characteristics (such as age, vehicle features, geographic locations, and contractual clauses), which we denote by $\boldsymbol{X}_{it}$. We first estimate a Poisson regression of claim counts on observables with an individual fixed effect. Specif-

ically,

$$E[ClaimCount_{it}|\boldsymbol{X}_{it}, \lambda_i] = \exp(\boldsymbol{X}_{it}\delta_c + \omega_i). \tag{4.1}$$

We then estimate a log-normal regression of claim size conditional on the consumer being involved in an accident:[21]

$$\log(ClaimSize_{it}) = \delta_0 + \boldsymbol{X}_{it}\delta_s + \eta_{it}. \tag{4.2}$$

With the estimated regression coefficients δ and the individual fixed effects ω from Equations (4.1) and (4.2), we predict the expected number of accidents and the expected indemnity conditional on involvement in an accident for each consumer. Multiplying the two yields an estimate for the expected cost of insuring consumer i in a year, which is the risk type θ in our model. Since we control for contract features in these regressions, the impact of consumers self-selecting into different clauses has been factored into our estimate of θ.[22] In the rest of the analysis, we take θ as given and known by the econometricians.

4.2 Demand

In classic demand models, the price of a product is often assumed to be the same for all consumers (e.g., the list price of a vehicle). The premiums for insurance contracts, however, depend on various consumer characteristics and discounts are highly likely to vary across individuals. The key challenge in estimating the demand model is that our data do not include the full price menu faced by consumers. Our data include only contracts sold in the market, and therefore we do not have access to the premiums of the unchosen products. This is a common issue in many empirical studies (e.g., Goldberg, 1996; Cicala, 2015; Crawford et al., 2018; Allen et al., 2019; Salz, 2022; Sagl, 2023) and the standard approach is predicting the alternative prices using regressions.

This standard approach unfortunately does not directly apply to our setting for two reasons. First, we do not have access to all pricing variables insurers may use. For each consumer, we observe a set of demographic characteristics and their vehicle features, but insurers may collect many more variables for risk rating and pricing (such as GPA for students, marital status, socioeconomic status, etc.). Second and more importantly, different firms may use different sets of pricing variables and algorithms. This phenomenon further complicates the task of predicting alternative prices using regression models. To provide

[21] We assume that claim size is independent of the individual fixed effect following Jeziorski et al. (2017). This assumption is motivated by the arguments in the actuarial literature that accident severity is more random and less related to the individual's driving ability.

[22] Our estimates of θ also take into account the moral hazard effect in a reduced-form way. That is, consumers may change their risky behavior after choosing different contracts.

a more concrete example, Crawford et al. (2018) face a similar problem as we do (not observing the full price menu). They estimate the following linear regression model to predict prices:

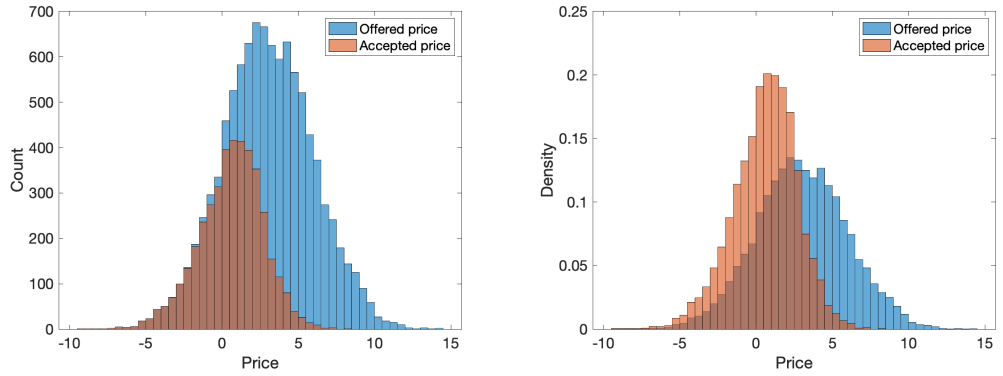
$$p_{ij} = \gamma_0 + \boldsymbol{X}_{ij}\gamma_1 + \lambda_j + \omega_i + \varepsilon_{ij}, \tag{4.3}$$

where λ_j is the firm fixed effect and ω_i is the individual fixed effect. The term ω_i includes factors that are observed by all firms in the market, but are unobserved by econometricians (i.e., "soft information" that firms may have about consumers). The regression approach in Crawford et al. (2018) requires observations of consumers contracting with multiple firms, and more importantly, assumes that all firms observe the same set of soft information about consumers (ω_i). This assumption is restrictive particularly in our setting because it essentially erases any informational asymmetry among competing firms.

Our paper adopts a novel fixed-point approach that jointly estimates consumers' sorting probabilities, price distributions, and demand parameters when only transaction prices are available. This approach extends the classic demand estimation techniques developed in the seminal paper of Berry et al. (1995). In addition to matching observed aggregate market share of each firm, our method introduces an outer loop that iterates over consumers' sorting propensities.

Offered vs. Accepted Prices We first highlight that there are two kinds of price distributions in our setting. The price distribution we observe in the data for each firm is the *accepted* price distribution after selection occurs. It is different from the *offered* price distribution that consumers face when making their enrollment decisions. To draw an analogy, the offered prices correspond to submitted bids in auctions and potential wages in Roy models (Roy, 1951), while the accepted prices correspond to winning bids and observed wages. The key observation is that the offered and accepted price distributions are linked through the demand system (or, more generally, the selection rules). Figure 2 provides a graphical illustration of the relationship between offered and accepted price distributions from a simulation exercise. Figure 2a shows that if firm j charges a lower price, the consumer is more likely to stay with j. Normalizing the histograms in Figure 2a to plot the density of offered and accepted prices, we observe in Figure 2b that after selection, the density of the price distribution shifts to the left. Consequently, lower prices are over-represented, while higher prices are under-represented in the accepted price distribution observed by econometricians.

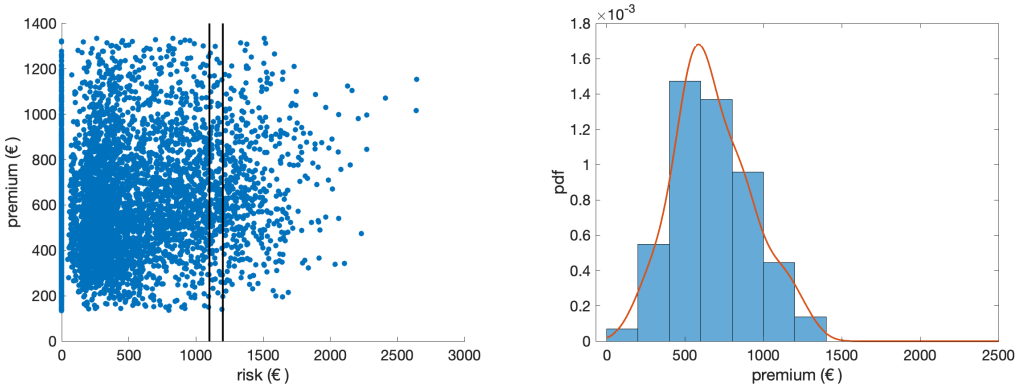
We derive firms' offered price distributions from their accepted price distributions and the choice probabilities using the demand model. Consider a firm j and a particular risk type θ. We use $\tilde{g}_j(p|\theta)$ to denote the density of the accepted price in firm j conditional on



(a) Count (b) Density

Figure 2: Comparing offered and accepted price distributions for a firm using simulated data. We draw offered prices from $N(3,3)$, and the probability that the consumer given price p stays with the firm is given by $\exp(-p)/(0.1+\exp(-p))$. Panels (a) and (b) plot the histograms of the offered and accepted prices, normalized by count and density, respectively.

the true risk type being θ. Note that $\tilde{g}_j(p|\theta)$ can be directly and nonparametrically estimated from the data. For concreteness, Figure 3a shows a scatter plot of premiums and risk types for a firm using real data. Since our measure of risks is continuous, we discretize θ and estimate the distribution of premiums conditional on θ falling into the bin (highlighted by the two vertical lines). The kernel density estimate of the premiums is shown in Figure 3b.



(a) Scatter plot of premiums and risks (b) Kernel density estimates

Figure 3: Estimating ex-post price distribution conditional on a certain risk type

Let $g_j(p|\theta)$ denote the density of the offered price in firm j. The densities of the offered and accepted prices are linked through the probability that consumers are selected into firm

j. Specifically,

$$g_j(p|\theta) = \frac{\tilde{g}_j(p|\theta)/Pr(j|p_j = p, \theta)}{\int_{p'} \tilde{g}_j(p'|\theta)/Pr(j|p_j = p', \theta)dp'}, \tag{4.4}$$

where the denominator is for normalization purposes. In Equation (4.4), $Pr(j|p_j = p, \theta)$ represents the probability that consumers of type θ choose firm j given a particular price p. The term $Pr(j|p_j = p, \theta)$ can be derived using the demand model and the distributions of offered prices of other firms. Let $\boldsymbol{p}_{-j} = (p_1, \cdots, p_{j-1}, p_{j+1}, \cdots, p_J)$ denote the vector of prices excluding firm j's price.

$$Pr(j|p_j = p, \theta) = \int_{\boldsymbol{p}_{-j}} \frac{\exp(-\gamma p + \xi_j)}{\exp(-\gamma p + \xi_j) + \sum_{j' \neq j} \exp(-\gamma p_{j'} + \xi_{j'})} \left(\prod_{j' \neq j} g_{j'}(p_{j'}|\theta) \right) d\boldsymbol{p}_{-j}. \tag{4.5}$$

Combining Equations (4.4) and (4.5) for all firms yields a system that characterizes the *offered* price distributions as its fixed point for any given value of demand parameters $(\gamma, \boldsymbol{\xi})$, where $\boldsymbol{\xi} = (\xi_1, \xi_2, \cdots, \xi_J)$. Wu and Xin (in process) show that for fixed $(\gamma, \boldsymbol{\xi})$, the operator in the space of the offered price distributions is a contraction under mild conditions.

Iterative Algorithm Given the fixed point problem defined in Equations (4.4) and (4.5), we propose an iterative procedure to solve for the *offered* price distributions for all firms as functions of demand parameters. We nest the contraction mapping for mean utilities in Berry et al. (1995) in our algorithm to obtain $\boldsymbol{\xi}$ that matches aggregate market shares. Our algorithm works as follows:

1. Fix the value of γ.
2. Set the initial value: $\xi_j^1 = 0$, and $Pr^1(j|p_j = p, \theta) = \exp(-\gamma p)$ for all j.[23]
3. At the r-th iteration, denote the vector of unobserved product attributes by $\boldsymbol{\xi}^r = (\xi_1^r, \xi_2^r, \cdots, \xi_J^r)$. Let $Pr^r(j|p_j = p, \theta)$ denote the probability of consumers being sorted into firm j given price p and type θ for $j = 1, 2, \cdots, J$.
4. Compute the offered price distribution for the r-th iteration using Equation (4.4):

$$g_j^r(p|\theta) = \frac{\tilde{g}_j(p|\theta)/Pr^r(j|p_j = p, \theta)}{\int_p \tilde{g}_j(p|\theta)/Pr^r(j|p_j = p, \theta)dp}. \tag{4.6}$$

[23]The initial value for choice probabilities is the numerator in the demand model (when $\xi_j = 0$). It can be viewed as a coarse approximation to the selection probability, assuming changing firm j's price would only impact the numerator, but not the denominator. The fact that it is not a well-defined probability measure does not matter, because it only serves as a weight and will be normalized when we update the offered price distribution using Equation (4.4). Since the operator is a contraction, the choices of initial values are less important. In our estimation, we try using different initial values for choice probabilities, and they all converge to the same results.

5. Update $\boldsymbol{\xi}^r$ using aggregate market shares. Given γ and the offered price distributions $\boldsymbol{g}^r = (g_1^r, g_2^r, \cdots, g_J^r)$, $\boldsymbol{\xi}^{r+1}$ solves the following system of equations:

$$\begin{aligned}
\hat{s}_1 &= s_1(\boldsymbol{\xi}^{r+1}, \gamma, \boldsymbol{g}^r), \\
\hat{s}_2 &= s_2(\boldsymbol{\xi}^{r+1}, \gamma, \boldsymbol{g}^r), \\
&\vdots \\
\hat{s}_J &= s_J(\boldsymbol{\xi}^{r+1}, \gamma, \boldsymbol{g}^r),
\end{aligned}$$

where $\hat{s}_j$ represents the observed market share for firm j. We denote by $s_j(\boldsymbol{\xi}^{r+1}, \gamma, \boldsymbol{g}^r)$ the model-implied market share given by the following equation:

$$s_j(\boldsymbol{\xi}^{r+1}, \gamma, \boldsymbol{g}^r) = \int_\theta \int_{\boldsymbol{p}} \frac{\exp(-\gamma p_j + \xi_j^{r+1})}{\sum_{j'} \exp(-\gamma p_{j'} + \xi_{j'}^{r+1})} \left(\prod_{j'} g_{j'}^r(p_{j'}|\theta) \right) f_0(\theta) d\boldsymbol{p} d\theta. \quad (4.7)$$

This step is similar to the contraction mapping for the mean utilities in Berry et al. (1995). Using the same iterative procedure as in Berry et al. (1995), we solve for the vector of unobserved product attributes $\boldsymbol{\xi}^{r+1}$ from observed market shares (ξ_1 is normalized to 0 without loss).

6. Given $\boldsymbol{\xi}^{r+1}$, we update the selection probabilities for all firms using Equation (4.5):

$$\begin{aligned}
&Pr^{r+1}(j|p_j = p, \theta) \quad (4.8) \\
&= \int_{\boldsymbol{p}_{-j}} \frac{\exp(-\gamma p + \xi_j^{r+1})}{\exp(-\gamma p + \xi_j^{r+1}) + \sum_{j' \neq j} \exp(-\gamma p_{j'} + \xi_{j'}^{r+1})} \left(\prod_{j' \neq j} g_{j'}^r(p_{j'}|\theta) \right) d\boldsymbol{p}_{-j}.
\end{aligned}$$

7. Advance to the next iteration with $\boldsymbol{\xi}^{r+1}$ and $Pr^{r+1}(j|p_j = p, \theta)$ for all j.

8. Iterate until $\boldsymbol{g}^r$ converges. In the limit, for a fixed price sensitivity parameter γ, we obtain (1) the offered price distribution faced by consumers $g_j(p|\theta)$ for all firms, (2) the probability that consumers are selected into firm j, $Pr(j|p_j = p, \theta)$, and (3) the vector of unobserved product attributes $\boldsymbol{\xi}$ that match the market shares.

Price Sensitivity We estimate the price sensitivity parameter γ using the sorting pattern observed in the data. To see which moments in the data identify γ, consider the likelihood that a consumer of type θ chooses firm j:

$$Pr(j|\theta; \gamma) = \int Pr(j|p_j = p, \theta; \gamma) g_j(p|\theta; \gamma) dp, \quad (4.9)$$

where $Pr(j|p_j = p, \theta; \gamma)$ and $g_j(p|\theta; \gamma)$ are obtained from the iterative process given γ. The probability that different types of consumers are sorted into different firms is governed by γ. Consequently, the average risk levels in different firms impose identifying restrictions on γ. Intuitively, in an extreme case where γ is equal to 0, consumers do not respond to price variations across firms, and therefore different types of consumers are randomly assigned to firms (so the average risk levels would be the same across firms).

For a sample of consumers indexed by $i = 1, 2, \cdots, N$, let θ_i denote their risk types and $y_i \in \{1, 2, \cdots, J\}$ denote their choices. In the outer loop of the iterative process, we search for γ to maximize the following log-likelihood function:

$$LL_d(\gamma) = \sum_{i=1}^{N} \log \Big(Pr(y_i|\theta_i; \gamma) \Big). \tag{4.10}$$

Our approach provides several important insights for the demand estimation literature. First, our method solves the problem of missing full price menu commonly encountered in empirical studies. We introduce a new equilibrium approach that combines insights from the seminal work of Berry et al. (1995) with a novel observation that the accepted and offered price distributions are linked through the demand system. Second, our estimation of the price distribution does not rely on reduced-form regressions and imposes no parametric restrictions. This makes our estimation method very flexible and robust to misspecifications. Third, our estimation strategy fully allows for heterogeneity across firms. We do not require that the offered price distributions are the same across firms. This is particularly useful for analyzing markets with rich supply-side heterogeneity. Lastly, we propose an easy-to-implement iterative procedure to numerically solve for the offered price distributions and sorting probabilities given demand parameters, which is very transparent and performs well numerically.

4.3 Pricing Coefficients

We now turn to the estimation of the supply-side parameters. We first focus on estimating the pricing coefficients (α_j, β_j), which are equilibrium objects optimally chosen by the firm. For each consumer in firm j, we observe her risk type θ and the premium charged by firm j. The difficulty in estimating the pricing coefficients is due to the fact that we do not observe the equilibrium risk rating ($E(\theta|\hat{\theta}_j, j)$) assigned to the consumer by the firm given their private signal $\hat{\theta}_j$.

We show that under the linear structure imposed on the pricing strategy in Equation (3.2), α_j and β_j are identified from the first and second moments of the joint distribution of

premiums and risks. To see this, we derive the within-firm mean and covariance of p and θ:

$$E(p_j(\hat{\theta}_j)) = \alpha_j + \beta_j E(E(\theta|\hat{\theta}_j, j)) = \alpha_j + \beta_j \mu_j, \tag{4.11}$$

$$cov(p_j(\hat{\theta}_j), \theta) = \beta_j var(E(\theta|\hat{\theta}_j, j)) = \frac{var(p_j(\hat{\theta}_j))}{\beta_j}, \tag{4.12}$$

where μ_j denotes the expectation of θ among consumers who choose firm j. Equations (4.11) and (4.12) uniquely determine (α_j, β_j) as the solution to a system of two linear equations. Note that all the first- and second-order moments of premiums and risks in firm j can be directly estimated from the data. Plugging the moment estimates back into these equations, we easily obtain the estimators for (α_j, β_j). An important advantage of this approach is that it relies only on data moments and imposes no parametric constraints on the type and signal distributions.

We highlight that given the linear structure imposed on firms' pricing strategies, the equilibrium pricing coefficients can be identified and estimated separately from other parts of the model. This significantly reduces the computational burden. If we instead assume that the premium is a nonlinear function of the risk rating, we then need to estimate the pricing coefficients together with other supply-side primitives, such as firms' signal distributions and cost parameters, which is feasible but computationally much more demanding.

4.4 Signal Distribution

Thus far, we have recovered the offered price distribution $g_j(p|\theta)$ faced by consumers and the pricing coefficients (α_j, β_j) for all j. The next step is to pin down the firm-specific signal distribution that generates the observed price dispersion.

The key insight to identify the signal distribution is that there is a one-to-one mapping between the offered price and the signal. This concept is analogous to the auction models, where bids are monotonically increasing in bidders' valuations. The nonparametrically identified bid distribution therefore identifies the underlying valuation distribution under model restrictions (Guerre et al., 2000).

Following similar ideas in auction models, when price is a monotonically increasing function of the signal, the change of variables formula yields

$$G_j(p_j(\hat{\theta}_j)|\theta) = F_j(\hat{\theta}_j|\theta), \tag{4.13}$$

where $G_j(p|\theta)$ denotes the cumulative distribution function of the offered price distribution of firm j for a given risk type θ. Note that $G_j(p|\theta)$ has been recovered from previous steps

and can thus be treated as known. Equation (4.13) implies that for a given θ, if we know the signal distribution F_j, we can pin down the corresponding price for a signal $\hat{\theta}_j$ using the inverse of the observed offered price distribution G_j. Specifically,

$$p_j^o(\hat{\theta}_j; F_j, \theta) = G_j^{-1}\Big(F_j(\hat{\theta}_j|\theta)\Big). \tag{4.14}$$

Equation (4.14) identifies the relationship between the price and the signal without directly invoking model restrictions other than that there is a one-to-one mapping between the two.

We now derive the price as a function of the signal based on our model. For a fixed signal distribution F_j, we first derive the risk rating $E(\theta|\hat{\theta}_j, j; F_j)$. Recall from Equation (3.1) that

$$\begin{aligned} E(\theta|\hat{\theta}_j, j; F_j) &= \frac{\int_\theta \theta Pr(j|\hat{\theta}_j, \theta) f_j(\hat{\theta}_j|\theta) f_0(\theta) d\theta}{\int_\theta Pr(j|\hat{\theta}_j, \theta) f_j(\hat{\theta}_j|\theta) f_0(\theta) d\theta} \\ &= \frac{\int_\theta \theta Pr(j|p_j^o(\hat{\theta}_j; F_j, \theta), \theta) f_j(\hat{\theta}_j|\theta) f_0(\theta) d\theta}{\int_\theta Pr(j|p_j^o(\hat{\theta}_j; F_j, \theta), \theta) f_j(\hat{\theta}_j|\theta) f_0(\theta) d\theta}. \end{aligned} \tag{4.15}$$

The second equality in Equation (4.15) holds because the signal $\hat{\theta}_j$ affects the choice probabilities only through the price, and therefore $Pr(j|\hat{\theta}_j, \theta) = Pr(j|p_j^o(\hat{\theta}_j; F_j, \theta), \theta)$. Notice that in the previous step, we have obtained the selection probability conditional on the price and consumer's risk type, that is $Pr_j(j|p_j = p, \theta)$. This allows us to evaluate the risk rating in Equation (4.15) for any signal distribution F_j.

Once we have obtained $E(\theta|\hat{\theta}_j, j; F_j)$, we derive the model-generated price as a function of the signal, which we denote by $gp_j(\hat{\theta}_j; F_j)$. Specifically,

$$gp_j(\hat{\theta}_j; F_j) = \alpha_j + \beta_j E(\theta|\hat{\theta}_j, j; F_j). \tag{4.16}$$

Note that the pricing coefficients (α_j, β_j) have been recovered from the previous step. Equation (4.16) implies that we can generate the firm's price as a function of $\hat{\theta}_j$ for any fixed signal distribution F_j. This is the key for us to obtain a model-implied joint distribution of premiums and risk types for any fixed F_j. To see this, we first derive the posterior joint distribution of signals and consumers' risk types for any given signal distribution F_j:

$$f(\hat{\theta}_j, \theta|j; F_j) = \frac{Pr(j|p_j^o(\hat{\theta}_j; F_j, \theta), \theta) f_j(\hat{\theta}_j|\theta) f_0(\theta)}{s_j}. \tag{4.17}$$

We can then construct the posterior joint distribution of the premiums and risk types using

the change of variables formula:

$$h(p, \theta|j; F_j) = \frac{f(gp_j^{-1}(p; F_j), \theta|j; F_j)}{gp'(gp_j^{-1}(p; F_j); F_j)}, \tag{4.18}$$

as $p = gp_j(\hat{\theta}_j; F_j)$ and $\hat{\theta}_j = gp_j^{-1}(p; F_j)$. Here, gp_j^{-1} is the inverse function to back out the signal corresponding to an observed premium p. Matching the joint distribution of the premiums and risk types derived from the model (i.e., $h(p, \theta|j; F_j)$ in Equation (4.18)) with the data pattern provides the key identifying restrictions for the signal distribution.

To provide a more concrete example, we consider a special case where the signals are drawn from a normal distribution in Appendix B. This is also the empirical specification we use for our application. We show that when the risk rating $E(\theta|\hat{\theta}_j, j; F_j)$ can be approximated by a linear function of the signal $\hat{\theta}_j$, the correlation between the premiums and risk types within a firm (which is directly estimable from the data) decreases monotonically with the variance of the signal distribution. This result is intuitive. When the signal distribution is very informative, consumers' risks are better reflected in their premiums, and vice versa.

Given the identification argument, we estimate the firm-specific signal distribution by matching the model-implied posterior distribution of the premiums and risk types with the data. Since the identification of the signal distribution relies on within-firm variations, we can estimate it for each firm separately, which significantly reduces the computational burden. Again, we use θ_i, p_i, and y_i to denote consumer i's risk type, premium, and choice of product. We estimate F_j for each firm separately by maximizing the following likelihood function:

$$LL_p(F_j) = \sum_{i=1}^{N} \mathbf{1}\{y_i = j\} \log\Big(h(p_i, \theta_i|j; F_j)\Big). \tag{4.19}$$

4.5 Cost

The final step of our estimation is to recover the cost parameters from the first-order conditions of firms' profit maximization problem. Recall that the profit function of firm j given

the vector of pricing coefficients $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is

$$\begin{aligned}\pi_j(\boldsymbol{\alpha}, \boldsymbol{\beta}) &= \int_{\theta}\int_{\hat{\boldsymbol{\theta}}} \Big(\alpha_j + \beta_j E(\theta|\hat{\theta}_j, j) + c_j - k_j\theta\Big) Pr(j|\hat{\boldsymbol{\theta}}) f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}} d\theta. \\ &= \int_{\theta}\int_{\hat{\boldsymbol{\theta}}} (\alpha_j + c_j - k_j\theta) Pr(j|\hat{\boldsymbol{\theta}}) f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}} d\theta \\ &+ \quad \beta_j \int_{\theta}\int_{\hat{\boldsymbol{\theta}}} \theta Pr(j|\hat{\boldsymbol{\theta}}) f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}} d\theta, \end{aligned} \tag{4.20}$$

where to simplify notation we use $f(\hat{\boldsymbol{\theta}}|\theta)$ to represent the joint density of signals conditional on the type θ, i.e., $f(\hat{\boldsymbol{\theta}}|\theta) = \prod_{j'=1}^{J} f_{j'}(\hat{\theta}_{j'}|\theta)$. Taking the first-order derivatives of the profit function with respect to (α_j, β_j) yields

$$\begin{aligned}\frac{\partial \pi_j(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \alpha_j} &= \int_{\theta}\int_{\hat{\boldsymbol{\theta}}} Pr(j|\hat{\boldsymbol{\theta}}) f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}} d\theta \\ &+ \quad \int_{\theta}\int_{\hat{\boldsymbol{\theta}}} (\alpha_j + \beta_j\theta + c_j - k_j\theta) \frac{\partial Pr(j|\hat{\boldsymbol{\theta}})}{\partial \alpha_j} f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}} d\theta, \end{aligned} \tag{4.21}$$

$$\begin{aligned}\frac{\partial \pi_j(\boldsymbol{\alpha}, \boldsymbol{\beta})}{\partial \beta_j} &= \int_{\theta}\int_{\hat{\boldsymbol{\theta}}} \theta Pr(j|\hat{\boldsymbol{\theta}}) f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}} d\theta \\ &+ \quad \int_{\theta}\int_{\hat{\boldsymbol{\theta}}} (\alpha_j + \beta_j\theta + c_j - k_j\theta) \frac{\partial Pr(j|\hat{\boldsymbol{\theta}})}{\partial \beta_j} f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}} d\theta. \end{aligned} \tag{4.22}$$

In Equations (4.21) and (4.22), the first terms capture the direct effects of changing the pricing coefficients on profit, while the second terms quantify the marginal changes to the profit through indirect sorting effects. We approximate the derivatives of sorting probabilities with respect to α_j and β_j in the following equations:

$$\frac{\partial Pr(j|\hat{\boldsymbol{\theta}})}{\partial \alpha_j} \approx -\gamma Pr(j|\hat{\boldsymbol{\theta}})(1 - Pr(j|\hat{\boldsymbol{\theta}})), \tag{4.23}$$

$$\frac{\partial Pr(j|\hat{\boldsymbol{\theta}})}{\partial \beta_j} \approx -\gamma E(\theta|\hat{\theta}_j, j) Pr(j|\hat{\boldsymbol{\theta}})(1 - Pr(j|\hat{\boldsymbol{\theta}})). \tag{4.24}$$

The approximations in Equations (4.23) and (4.24) ignore the impact of (α_j, β_j) on risk rating $E(\theta|\hat{\theta}_j, j)$. Given the equilibrium concept we use (i.e., Nash equilibrium), changing firm j's pricing coefficients does not affect other firms' ($j' \neq j$) pricing strategy $p_{j'}(\hat{\theta}_{j'})$. In other words, other firms will keep using the same pricing strategies they currently use. Changing α_j and β_j will affect firm j's own expected risk level (i.e., $E(\theta|\hat{\theta}_j, j)$), but the impact of this single term on the selection probability could be small.

With the approximations in Equations (4.23) and (4.24), the first-order conditions in Equation (4.21) and (4.22) are reduced to a system of two linear equations involving (c_j, k_j).

Specifically,

$$\hat{s}_j = \gamma \int_{\theta} \int_{\hat{\boldsymbol{\theta}}} (\alpha_j + \beta_j \theta + c_j - k_j \theta) Pr(j|\hat{\boldsymbol{\theta}})(1 - Pr(j|\hat{\boldsymbol{\theta}})) f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}} d\theta, \tag{4.25}$$

$$\hat{\mu}_j \hat{s}_j = \gamma \int_{\theta} \int_{\hat{\boldsymbol{\theta}}} (\alpha_j + \beta_j \theta + c_j - k_j \theta) E(\theta|\hat{\theta}_j, j) Pr(j|\hat{\boldsymbol{\theta}})(1 - Pr(j|\hat{\boldsymbol{\theta}})) f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}} d\theta, \tag{4.26}$$

where $\hat{s}_j$ and $\hat{\mu}_j$ represent the estimated market share and the average risk for firm j, respectively. These two equations uniquely pin down (c_j, k_j) as the solution to the system of linear equations. All other terms in these two equations are either directly estimable from the data (for example, $\hat{s}_j$ and $\hat{\mu}_j$) or have been recovered in previous steps (e.g., α_j, β_j, σ_j).

5 Results

We apply our model and estimation strategy to the Italian auto insurance industry. Our sample includes 124,592 liability insurance contracts for new customers sold in Rome between 2013-2021. This market has more than 50 insurance companies. To reduce computational burden, we focus on the top 10 largest firms and the group of remaining fringe firms (which we denote as Firm 11). In the estimation, we assume that the firms receive Gaussian signals from $N(\theta, \sigma_j^2)$ and we estimate the standard deviation of the signal distribution σ_j. Note that $1/\sigma_j$ measures the precision of risk-rating technology in firm j.

5.1 Data Patterns

Using the two-part model described in Section 4.1, we estimate the individual-specific risk type for all consumers in the sample. Table 2 reports the means and standard deviations of risks and premiums (both measured in euros) across firms and the last column reports the market share of each firm. On average, the expected cost to insure a consumer per year is around 331 euros. What is striking to observe is the substantial amount of risk sorting across firms. The average risk level among consumers ranges from 286 euros to 481 euros and it does not seem to have a direct relationship with the market share. There is also a significant price dispersion in this market, even if we take into account price adjustments due to risk sorting.

Using the data on consumers' risks and premiums, we run a simple linear regression of price on risk for each firm separately. The linear regression coefficients and their standard errors are reported in Table 3. In Figure 4, we provide scatter plots of premiums versus risks for Firms 7 and 8 separately as two illustrative examples (the two firms have similar

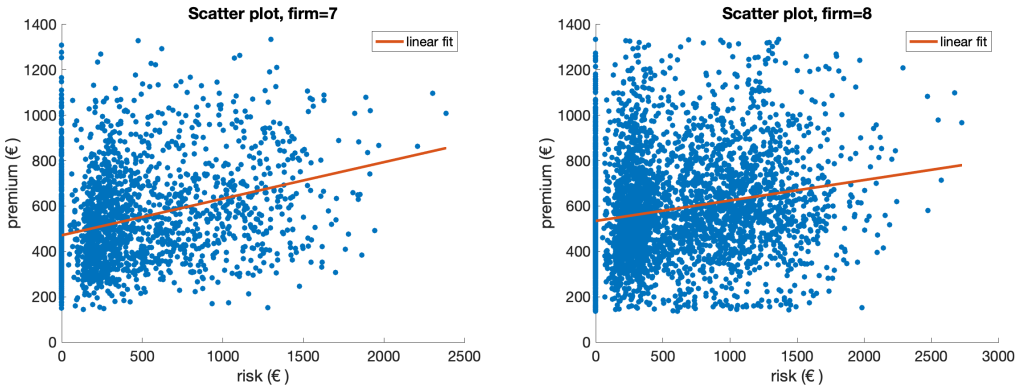
Table 2: Risk sorting and price dispersion across firms

Firm ID	Risk (€)		Premium (€)		Market Share
	Mean	Std. Dev.	Mean	Std. Dev.	
1	352.71	433.72	519.81	209.46	0.17
2	318.53	416.51	557.12	221.44	0.06
3	320.07	432.11	568.59	216.07	0.06
4	363.97	461.81	427.78	199.55	0.04
5	285.56	401.07	506.97	201.39	0.03
6	326.86	423.73	448.57	173.50	0.13
7	346.70	426.08	526.93	209.00	0.02
8	481.29	511.30	577.66	227.70	0.03
9	324.17	397.91	552.28	225.59	0.01
10	322.36	435.80	523.75	212.33	0.02
11	316.25	412.78	431.04	198.01	0.41
All	331.79	426.83	478.06	209.03	1.00

market shares). The red line in the figure shows the linear fit of the regression model. Firm 7 charges a premium that is more sensitive to the risk, indicating that Firm 7 is potentially better at evaluating the risk when a new customer arrives. Of course, the patterns shown in Figure 4 are complex equilibrium outcomes and the differences in slopes could be due to other factors. For example, if Firm 8 is more efficient at processing claims (i.e., k_j is smaller), in equilibrium it may design a pricing algorithm to target high-risk consumers. We therefore need the structural model to better understand to what extent firms are differentiated in their cost structures and information technology, and which factors drive the equilibrium sorting patterns and pricing strategies.

Table 3: Estimates of linear projection coefficients by firm

Firm ID	Intercept		Slope	
	Estimates	Std. Err.	Estimates	Std. Err.
1	475.24	(1.81)	0.13	(0.00)
2	507.65	(2.97)	0.16	(0.01)
3	518.09	(2.84)	0.16	(0.01)
4	381.75	(3.27)	0.13	(0.01)
5	462.72	(3.74)	0.15	(0.01)
6	403.15	(1.63)	0.14	(0.00)
7	471.09	(5.06)	0.16	(0.01)
8	534.30	(4.67)	0.09	(0.01)
9	498.51	(7.56)	0.17	(0.01)
10	485.15	(4.68)	0.12	(0.01)
11	376.43	(1.03)	0.17	(0.00)



(a) Firm 7 (b) Firm 8

Figure 4: Scatter plots of premiums and risks with the linear fit

5.2 Estimation Results

We report the estimates for firms' pricing coefficients (α_j, β_j), standard deviations of the signal distributions σ_j, unobserved product attributes ξ_j, and cost parameters (c_j, k_j) in Table 4. The standard errors reported in parentheses are computed using bootstraps. Our estimated price sensitivity parameter γ is equal to 3.00 (with a standard error of 0.22).[24] From our estimates, we observe a huge amount of firm heterogeneity along *all* dimensions. Comparing pricing coefficients across the firms shown in Panel (A) of Table 4, we find that firms differ dramatically not only in their baseline premium (α_j), but also in how much they adjust premiums with respect to their risk ratings. For example, Firms 1, 8, and 10 charge a relatively lower baseline premium but their prices are much more sensitive to their risk ratings, while Firms 3, 6, and 11 tend to do the opposite. Pricing coefficients (α_j, β_j) are complex equilibrium objects chosen optimally by firms. The differences in pricing strategies could be due to heterogeneity in firms' cost structures and information technology, as well as competition in the market.

We observe substantial differences in the precision of risk rating across firms, as measured by the standard deviation of their signal distributions (see the first column of Panel (B) in Table 4). Based on our estimates, Firm 8 has the least advanced risk-rating technology. The standard deviation of its signal distribution is almost double that of Firm 5, which has the best risk-rating technology. The heterogeneity in information precision has an important impact on consumers' equilibrium sorting patterns. Figure 5 presents a scatter plot of the ranks for information precision versus average consumer risks across firms. It is clear that in equilibrium, insurers with low information precision (large σ) tend to attract

[24] In the estimation of the demand model, premiums are represented in thousands of euros.

Table 4: Parameter estimates

Panel (A): Pricing Coefficients

Firm ID	α_j		β_j	
	Estimates	Std. Err.	Estimates	Std. Err.
1	-131.26	(19.85)	1.85	(0.06)
2	-22.56	(24.93)	1.82	(0.08)
3	61.39	(20.16)	1.58	(0.06)
4	-109.61	(25.99)	1.48	(0.08)
5	42.38	(25.71)	1.63	(0.10)
6	54.16	(11.41)	1.21	(0.04)
7	8.95	(31.95)	1.49	(0.10)
8	-481.99	(88.68)	2.20	(0.18)
9	-75.90	(64.41)	1.94	(0.21)
10	-115.26	(51.82)	1.98	(0.17)
11	9.53	(5.69)	1.33	(0.02)

Panel (B): Firm Primitives

Firm ID	Information Precision σ_j		Latent Attributes ξ_j		Net Benefits c_j		Claim Efficiency k_j	
	Estimates	Std. Err.	Estimates	Std. Err.	Estimates	Std. Err.	Estimates	Std. Err.
1	1392.91	(56.27)	0.00	0.00	504.63	(74.27)	1.63	(0.16)
2	962.98	(57.07)	-0.85	(0.03)	487.12	(64.56)	2.16	(0.15)
3	868.88	(55.55)	-0.77	(0.03)	405.61	(56.95)	1.93	(0.13)
4	1252.67	(75.88)	-1.67	(0.04)	357.36	(90.39)	1.17	(0.17)
5	802.34	(56.43)	-1.71	(0.03)	428.28	(64.82)	2.09	(0.17)
6	977.38	(40.40)	-0.57	(0.03)	331.25	(60.80)	1.17	(0.11)
7	923.66	(81.85)	-2.09	(0.03)	396.07	(63.24)	1.69	(0.16)
8	1597.91	(118.20)	-1.47	(0.03)	882.67	(139.98)	2.28	(0.33)
9	1093.97	(131.70)	-2.56	(0.05)	297.97	(104.22)	1.57	(0.35)
10	1130.27	(91.63)	-1.93	(0.03)	412.44	(83.06)	1.83	(0.23)
11	1022.42	(29.55)	0.57	(0.03)	667.48	(86.60)	1.53	(0.13)

high-risk consumers. The correlation coefficient between σ and average risk within each firm is 0.85. Overall, our estimates of the information precision are consistent with the linear regression results shown in Table 3. For firms with better risk-rating technology (i.e., smaller σ_j), their premiums are more sensitive to consumers' risks.

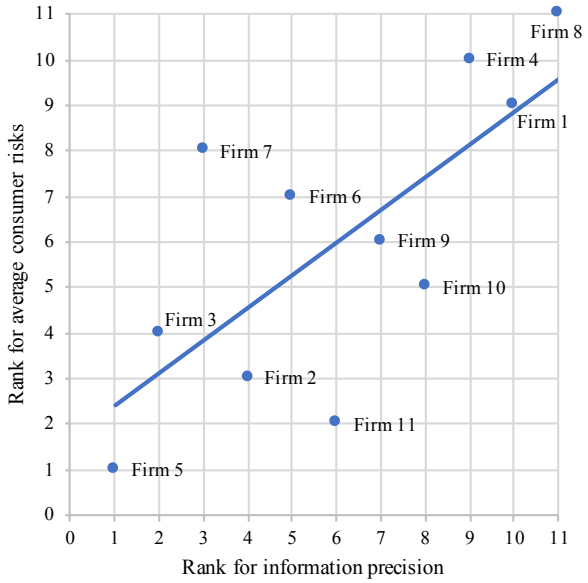


Figure 5: Who goes where: Rank correlation between information precision and average consumer risks across firms. Firms are ranked by information precision (highest to lowest on the x-axis) and average consumer risks (lowest to highest on the y-axis). The blue solid line represents a linear fit between the two rankings.

Another interesting finding from Table 4 is that firms enjoy large net benefits from contracting with new customers. Our estimates of these net benefits range from 298 to 883 euros per year and per customer across the competing firms. Although this paper does not provide a full-fledged dynamic model to account for consumers' switching behavior and firms' dynamic pricing strategies, we perform a back-of-the-envelope calculation of the future benefits firms may receive, based directly on the data. Let A_j denote the set of consumers who choose firm j in our sample. For each of these consumers, we observe the premium they pay, the actual claim costs, and whether they stay in firm j over the next T years. We can therefore compute the sum of the discounted future premiums averaged across all consumers in A_j, which we denote by $\overline{P}_j$. Specifically,

$$\overline{P}_j = \frac{\sum_{i \in A_j} \sum_t^T \delta^t P_{it} \mathbf{1}\{i \text{ stays in firm } j \text{ at } t\}}{|A_j|}, \tag{5.1}$$

where the discount factor δ is set to 0.95. We use a similar formula as in Equation (5.1) to

compute the sum of discounted claim costs and contract sales, which we denote by $\overline{CS}_j$ and $\overline{N}_j$, respectively. Let mc_j denote firm j's marginal cost of managing a contract. The following equation computes the net benefit firms receive from contracting with a new customer:

$$\underbrace{(\overline{P}_j - \overline{CS}_j k_j - \overline{N}_j mc_j)}_{\text{discounted value of future profits}} - mc_j = c_j. \tag{5.2}$$

Equation (5.2) isolates the firm's marginal cost mc_j from other factors. We report the estimates of marginal costs in Table 5.

Comparing our marginal cost estimates with the average premiums across firms, we find that the marginal cost is around 20%–40% of the premiums charged to consumers.[25] Another important finding from our estimation is that there is a moderate negative correlation between σ and mc (the correlation coefficient is -0.42). This indicates that firms with lower information precision tend to have lower costs. Firms in equilibrium seemed to have developed different comparative advantages. Some firms rely on their more precise risk-rating technology to cream-skim low-risk consumers, while other firms can afford to serve the high-risk segment of the market due to a more efficient cost structure.

Table 5: Estimates of marginal costs

Firm ID	Marginal Cost (€)	Perc. of Price (%)
1	210.63	40.52
2	180.57	32.41
3	197.94	34.81
4	160.38	37.49
5	190.23	37.52
6	174.85	38.98
7	214.46	40.70
8	91.68	15.87
9	235.88	42.71
10	210.77	40.24
11	111.40	25.85

5.3 Model Fit

Using the estimates in Table 4, we simulate the premiums offered by each firm and consumers' plan choices given the simulated price menu. Table 8 in the appendix presents the simulated post-selection means and standard deviations of risks and premiums across firms,

[25]This proportion aligns with findings in other markets; for instance, Chiappori et al. (2006) reported a 20% marginal administrative cost in the French auto insurance market. In addition, our estimate is close to the nationwide Italian average administrative cost obtained from balance sheet data of the insurance companies.

and the last column reports the simulated market share of each firm. Overall, our model matches the key moments of the data reasonably well. Our model-generated price distribution and market shares are very close to what we observe in the real dataset (see details in Table 2). As for the risk sorting pattern, we are able to match the mean and variance of consumers' risks for the majority of firms, and the ranking of average risk levels across firms is largely preserved. The only exception is Firm 8, which attracts consumers with extremely high risks. While our model under-predicts the average risk level of this firm, it correctly predicts that Firm 8 has the highest risk level compared to all other firms. This underprediction primarily stems from the limited sample size of consumers with extremely high risks. Consequently, for this subgroup, the price distribution and consumers' sorting patterns are estimated with much less precision.

Our estimation relies on the assumption that the premium is a monotonically increasing function of the firm's signal. We show that this assumption is self-consistent by validating the monotone relationship between the risk rating $E_j(\theta|\hat{\theta}_j, j)$—which is computed post-estimation—and the signal $\hat{\theta}_j$. In Figure 7 in the appendix, the risk rating, and therefore the premium, indeed monotonically increase with the signal of Firm 1 (similar patterns are observed for all other firms).

6 Counterfactuals

The empirical estimation results in Section 5 suggest that firms are not only heterogeneous in their cost structure and product attributes, but also differ substantially in their risk-rating technologies. Moreover, firms have developed different comparative advantages: those with lower information precision tend to have lower costs. From a policy perspective, it is therefore important to understand the equilibrium consequences of information heterogeneity among competing firms and how it affects competition and consumer welfare.

To address these questions, we conduct two counterfactual experiments, where firms are assumed to adopt the same—either the least or the most advanced—risk-rating technology existing in the market. The first counterfactual policy is motivated by the recent discussions on privacy protection.[26] With the rapid development of big data, machine learning, and AI technologies, firms now are able to collect a huge amount of information from con-

[26]See reviews by Acquisti et al. (2016) and Goldfarb and Que (2023) on the economics of privacy. There is also a growing literature on data protection policies; see Johnson (2022) for a review of studies on the economic consequences of the European Union's General Data Protection Regulation. A recent paper by Alcobendas et al. (2023) studies privacy protection in online advertising markets. They find that certain privacy regulations may generate anti-competitive effects by providing large tech companies with informational monopoly power.

sumers, which they can exploit to make additional profits. Researchers and policy makers have raised concerns regarding the legitimacy of using data from consumers for advertising, targeting, and price discrimination, etc. To simulate the consequences of potential privacy regulations where firms are required to use only the minimum amount of consumer characteristics, we assume that all firms are forced to use the least advanced information technology. That is, the standard deviation of their signal distributions is set to be the largest one currently observed in the market.

At the other extreme, policy makers may enforce data portability or algorithm sharing to promote competition in the market. This in our model is equivalent to requiring all firms to adopt the most advanced information technology. That is, the standard deviation of their signal distributions is set to be the smallest one observed in the market.

For each counterfactual scenario, we re-solve for the market equilibrium using the iterative procedure described in Appendix C. We also re-solve for the market equilibrium given model estimates in Section 5.2 and use that as the benchmark case for the rest of the analysis. We discuss the impact of counterfactual policies on consumer surplus, firm profits, and equilibrium sorting pattern in the following subsections.

6.1 Consumer Surplus and Firm Profit

We report the average consumer surplus, premium, firm profit, and HHI under the baseline and two counterfactual scenarios (i.e., under privacy protection and data/algorithm sharing, respectively) in Table 6. To measure consumer surplus, we compute $-\gamma p_j + \xi_j$ for each individual, where j denotes the firm chosen by the consumer in equilibrium and p_j is the transaction price, and then we convert the utility into a monetary value expressed in euros. To evaluate the equilibrium consequence for consumers with different risk types, we report the average consumer surplus for high- and low-risk consumers separately.[27]

In Table 6, we can see that if all firms were to adopt the least advanced risk-rating technology, consumer surplus would increase by 4.5%, mainly due to lower prices (premiums on average would decrease by 5.1%). The gain in surplus is primarily enjoyed by the high-risk consumers. Since it is more difficult for firms with inferior risk-rating technology to identify high-risk consumers, the premiums these consumers pay significantly decrease. Under the privacy protection counterfactual scenario, we also observe that the market becomes slightly more concentrated (the HHI increases from 2091 to 2119). This is mainly driven by firms with lower marginal costs gaining larger market shares, when all firms are equally bad at selecting low-risk consumers.

[27] Consumers with $\theta < 1$ are classified as low risk. In the data, 41% of consumers had no accidents during the sample period, and therefore their individual-specific risk type estimate is extremely close to zero.

Table 6: Information policies: Privacy protection versus data/algorithm sharing

	Baseline	Privacy Protection (worst tech, max σ)	Data/Algorithm Sharing (best tech, min σ)
Avg. CS (€)	-546.87	-522.41	-564.22
Avg. CS: low risk	-498.18	-497.80	-478.85
Avg. CS: high risk	-580.14	-539.22	-622.54
Avg. premium (€)	451.35	428.53	465.05
Avg. profit (€)	462.46	446.87	487.58
HHI	2091.24	2119.78	2065.13
Avg. marginal cost (mc_j,€)	155.07	153.75	153.87
Avg. claim cost ($k_j\theta$,€)	537.34	539.88	534.51

When all firms share the same most advanced risk-rating technology, we observe the opposite effect on consumer surplus, again due to changes in premiums. Improvement in precision of risk rating increases the surplus for low-risk consumers by 3.9%, but lowers the surplus for high-risk consumers by 7.3%. Under this counterfactual scenario, firms' average profit increases by 5.4% compared to the baseline.

6.2 Sorting Pattern

To examine the effect of eliminating information heterogeneity on the equilibrium sorting pattern, we report the average risk across firms for the baseline and counterfactual scenarios in Table 7. To measure the degree of sorting, we compute the cross-firm standard deviation of the average risk, which is shown in the last row of Table 7. One interesting observation from this table is that consumer sorting (i.e., high- and low-risk consumers are selected into different firms) almost vanishes under the privacy protection counterfactual scenario. The risk levels of firms are much more alike than they are in the baseline. This result is reasonable because when firms are equally bad at risk rating, consumers are likely to be offered idiosyncratic prices and are then randomly assigned to firms.

When all firms adopt the most advanced risk-rating technology, there is a significant amount of consumer sorting, mainly driven by specialization. As indicated in the last column of Table 7, firms that attract high-risk consumers, such as Firms 4 and 6, are more efficient at processing claims (evidenced by smaller k_j). This result is very intuitive: when all firms are equally good at distinguishing high- and low-risk consumers, firms tailor their pricing strategies based on comparative advantages. Firms that are better at processing claims can offer more competitive prices to high-risk consumers.

Additionally, the data/algorithm sharing counterfactual scenario leads to a socially more cost-efficient matching between insurers and insurees. The last two rows in Table 6 indicate

Table 7: Eliminating heterogeneity in information precision: Sorting patterns

Firm ID	Baseline	Privacy Protection (worst tech, max σ)	Data/Algorithm Sharing (best tech, min σ)
1	353.12	324.79	319.50
2	280.42	312.13	294.58
3	277.49	319.56	302.01
4	383.94	334.33	397.72
5	261.69	322.81	324.25
6	341.55	347.36	381.08
7	309.11	334.47	346.01
8	365.95	316.29	284.30
9	317.76	342.75	328.04
10	336.85	320.63	317.71
11	332.83	339.16	336.54
Cross-firm Std. Dev.	38.67	11.61	34.52

Notes: The numbers reported in this table are the average risk levels among consumers (measured in euros) within each firm under the baseline and two counterfactual scenarios.

that the total cost to insure consumers (i.e., $mc_j + k_j\theta$) decreases by 4 euros per contract. Given the size of the Italian auto insurance market, the overall impact of this cost reduction is considerable. This result provides a compelling insight into insurance markets: even if insurance is mandatory, reducing information asymmetry (in our case, by allowing firms to use better information technology) can still improve welfare through a more efficient matching between insurers and insurees. This contrasts with the prediction of the canonical insurance model that mandatory full insurance can achieve the first-best outcome.[28]

6.3 Differential Impact

The last key finding we highlight is that policies that eliminate heterogeneity in information precision have a differential impact on competing firms. In Figure 6, we report percentage changes in profit under each counterfactual scenario (relative to the baseline) for the 11 firms separately. Firms are ranked from the highest to the lowest risk-rating precision. It is clear that firms with less advanced risk-rating technology (for example, Firms 1 and 8) benefit the most, regardless of whether the policies improve or reduce the overall risk-rating precision in the market. By contrast, firms with more advanced risk-rating precision (such as Firms 3 and 5) experience a significant profit loss under both counterfactual scenarios.

This finding has important implications for competition policies. Policies that aim at

[28] See more discussions related to this point in Einav et al. (2010b). Jeziorski et al. (2017) assess the welfare cost of mandatory liability insurance, focusing on the effect of moral hazard.

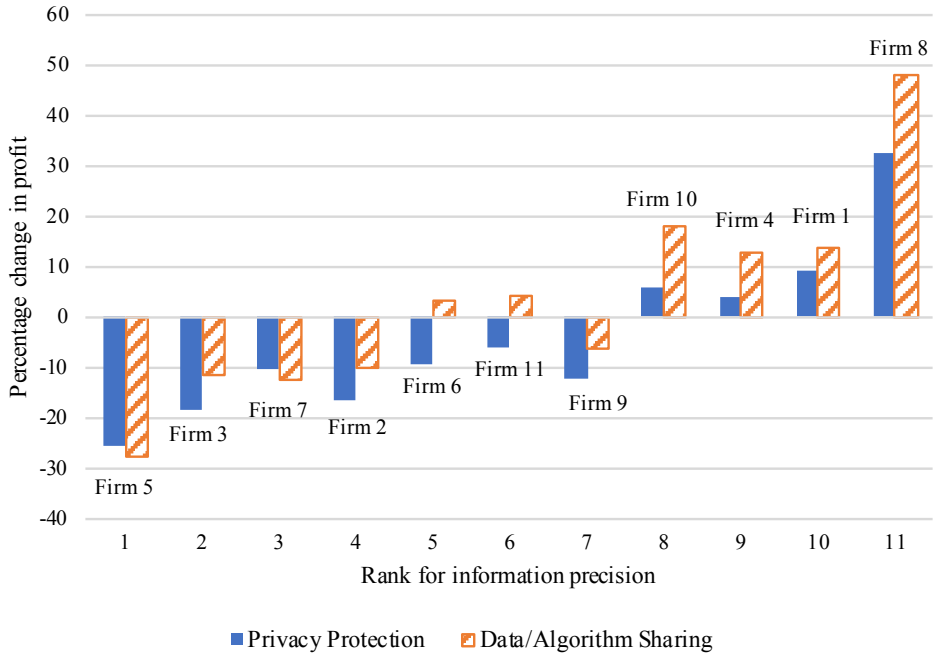


Figure 6: Differential impact on firms: Percentage change in profit relative to the baseline. Firms are ranked from the highest to the lowest risk-rating precision on the x-axis. The blue solid bars represent the percentage changes in profit under the privacy protection scenario, while the orange striped bars indicate changes under the data/algorithm sharing scenario.

protecting consumer privacy or promoting competition may in fact generate anti-competitive effects. Larger firms may not possess the most advanced information technology, as observed in this empirical setting; therefore, policies that eliminate heterogeneity in information technology might inadvertently grant more market power to these large players, leading to higher market concentration.

7 Conclusions

Our paper develops a novel empirical framework for studying imperfect competition in selection markets when firms have *heterogeneous information* about consumers, differ in cost structures, and offer differentiated products. We build a theoretical model where firms receive private signals about consumers' risk types and choose their optimal pricing strategies to maximize their profit, taking into account their opponents' strategies and demand-side responses. In our model, firms' inferences about consumers' risk types depend not only on their signals but also on equilibrium sorting patterns.

We introduce novel econometric techniques to identify demand parameters, as well as firm-specific information precision and cost structures. Applying the method to a representative sample of Italian auto insurance contracts with associated claims from 2013 to 2021, we find substantial differences in the precision of risk rating across firms; moreover, firms with lower information precision tend to have lower costs. This suggests that firms have developed different comparative advantages. Some firms focus on improving their risk-rating technology so that they can cream-skim low-risk consumers, while other firms are more efficient at underwriting contracts and/or liquidating claims, leading to a competitive equilibrium in which a variety of pricing strategies co-exist. The methodology we develop extends beyond auto insurance; it is applicable to a wide range of selection markets, including credit, health insurance, and labor markets.

We conduct counterfactual experiments to evaluate the equilibrium impact of information policies that aim at protecting consumer privacy or promoting competition (through data/algorithm sharing). If all firms were to adopt the least advanced information technology (e.g., under more stringent privacy regulation), consumer surplus would increase, but mainly for high-risk consumers. Conversely, allowing all firms to share the most advanced risk-rating technology leads to significant risk sorting, because firms that are more efficient at processing claims strategize to attract more high-risk consumers. This policy leads to a more efficient insurer–insuree matching, which reduces the social cost of servicing consumers in this market.

Our findings have broad implications for competition policies in the era of big data,

artificial intelligence, and machine learning. Information policies aimed at protecting consumer privacy or promoting competition may disproportionately benefit low-tech firms, which in our data include insurers with large market shares, and hence eventually create anti-competitive effects.

Finally, our paper mainly focuses on static price competition with information and cost heterogeneity in selection markets. Several other important aspects of competition, including dynamic pricing, endogenous contract designs, as well as more flexible demand models (e.g., allowing consumer switching, search frictions, a limited consideration set, and multi-dimensional private information), are promising avenues for future research.

A Additional Tables and Figures

Table 8: Model fit: Simulated results using model estimates

Firm ID	Risk (€)		Premium (€)		Market Share
	Mean	Std. Dev.	Mean	Std. Dev.	
1	347.47	443.19	520.40	214.06	0.17
2	315.44	422.84	554.28	227.86	0.06
3	338.93	436.49	572.07	224.15	0.07
4	342.33	435.77	429.26	205.71	0.04
5	329.80	426.81	512.97	215.76	0.03
6	337.78	416.07	449.12	180.72	0.13
7	352.73	439.70	531.53	220.52	0.02
8	372.54	459.20	569.44	237.85	0.03
9	323.63	424.45	540.86	239.75	0.01
10	345.11	443.98	533.11	228.38	0.02
11	318.41	415.79	432.40	200.96	0.40

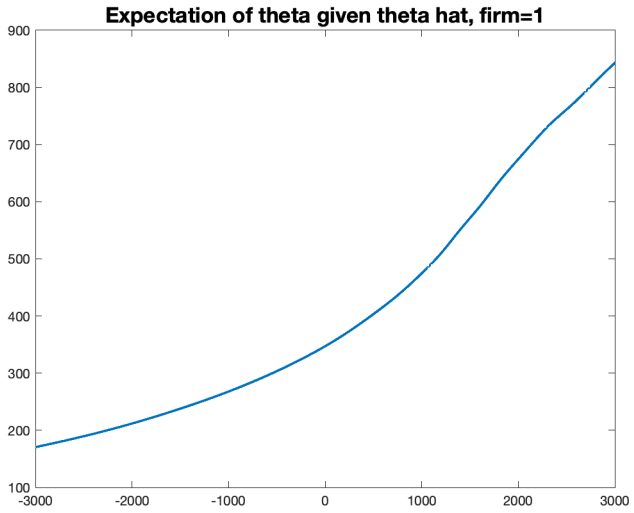


Figure 7: Risk rating versus signal for Firm 1: The y-axis represents the risk rating, and the x-axis denotes the signal.

B Identification of Signal Variance

We consider the identification of the variance of the signal distribution when the signals are drawn from a normal distribution. In this case, the density function of the signal distribution for firm j takes the following form:

$$f_j(\hat{\theta}_j|\theta) = \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(\hat{\theta}_j - \theta)^2}{2\sigma_j^2}\right), \tag{B.1}$$

where σ_j^2 is the variance of the signal distribution to be identified.

We first derive the posterior density of $\hat{\theta}_j$ for those who self-select into firm j. Let $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_J)$ denote the vector of signals received by all firms. Let $\hat{\boldsymbol{\theta}}_{-j}$ denote the vector of signals excluding firm j's signal. To simplify notation, we use $f(\hat{\boldsymbol{\theta}}|\theta)$ and $f(\hat{\boldsymbol{\theta}}_{-j}|\theta)$, respectively, to denote the densities of $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\theta}}_{-j}$ conditional on θ.

$$\begin{aligned} f(\hat{\theta}_j|\theta, j) &= \frac{\int_{\hat{\boldsymbol{\theta}}_{-j}} Pr(j|\hat{\boldsymbol{\theta}}) f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}}_{-j}}{\int_{\hat{\boldsymbol{\theta}}} Pr(j|\hat{\boldsymbol{\theta}}) f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}}} \\ &= \frac{\exp(-\gamma p_j(\hat{\theta}_j) + \xi_j) f_j(\hat{\theta}_j|\theta) \left[\int_{\hat{\boldsymbol{\theta}}_{-j}} \frac{f(\hat{\boldsymbol{\theta}}_{-j}|\theta) f_0(\theta)}{\sum_{j'=1}^{J} \exp(-\gamma p_{j'}(\hat{\theta}_{j'}) + \xi_{j'})} d\hat{\boldsymbol{\theta}}_{-j}\right]}{\int_{\hat{\theta}_j} \exp(-\gamma p_j(\hat{\theta}_j) + \xi_j) f_j(\hat{\theta}_j|\theta) \left[\int_{\hat{\boldsymbol{\theta}}_{-j}} \frac{f(\hat{\boldsymbol{\theta}}_{-j}|\theta) f_0(\theta)}{\sum_{j'=1}^{J} \exp(-\gamma p_{j'}(\hat{\theta}_{j'}) + \xi_{j'})} d\hat{\boldsymbol{\theta}}_{-j}\right] d\hat{\theta}_j} \end{aligned} \tag{B.2}$$

$$\approx \frac{\exp(-\gamma p_j(\hat{\theta}_j)) f_j(\hat{\theta}_j|\theta)}{\int_{\hat{\theta}_j} \exp(-\gamma p_j(\hat{\theta}_j)) f_j(\hat{\theta}_j|\theta) d\hat{\theta}_j}. \tag{B.3}$$

Denote the term in square brackets in Equation (B.2) by $\Delta(\hat{\theta}_j, \theta)$. Since it is inside of the integral in the denominator, $\Delta(\hat{\theta}_j, \theta)$ cannot be cancelled out. However, if firm j has a small market share, the effect of $\hat{\theta}_j$ on $\Delta(\hat{\theta}_j, \theta)$ is small, and therefore we can treat this term as if it does not depend on $\hat{\theta}_j$. When firm j is large, then as long as $\hat{\theta}_j$ has a relatively small variance, the overall impact of $\Delta(\hat{\theta}_j, \theta)$ varying with $\hat{\theta}_j$ is small in the denominator. With this approximation, we obtain Equation (B.3).

If the risk-rating term $E(\theta|\hat{\theta}_j, j)$ can be well approximated by a linear function of the signal $\hat{\theta}_j$,[29] i.e., $E(\theta|\hat{\theta}_j, j) \approx a_j + b_j\hat{\theta}_j$, we can further simplify the posterior distribution

[29] We empirically verify this assumption in Figure 7, where we plot the relationship between risk rating $E(\theta|\hat{\theta}_j, j)$ and signal $\hat{\theta}_j$ after we estimate the model. This figure suggests that the assumption that risk rating can be approximated by a linear function of the signal is reasonable.

of $\hat{\theta}_j$ to be

$$f(\hat{\theta}_j|\theta, j) \approx \frac{\exp(-\gamma\beta_j b_j \hat{\theta}_j) f_j(\hat{\theta}_j|\theta)}{\int_{\hat{\theta}_j} \exp(-\gamma\beta_j b_j \hat{\theta}_j) f_j(\hat{\theta}_j|\theta) d\hat{\theta}_j}$$
$$= \frac{1}{\sqrt{2\pi\sigma_j^2}} \exp\left(-\frac{(\hat{\theta}_j - (\theta - \gamma\beta_j b_j \sigma_j^2))^2}{2\sigma_j^2}\right). \tag{B.4}$$

Equation (B.4) implies that after selection, the signal still approximately follows a normal distribution with the same variance (σ_j^2) as before. However, the mean of the posterior normal distribution shifts to $\theta - \gamma\beta_j b_j \sigma_j^2$, which is lower than the original mean θ. This is intuitive as consumers who receive lower prices from j are more likely to self-select into the firm.

The following lemma provides the key identification argument for the signal variance. We focus on the joint distribution of premiums and risk types of consumers selected into firm j. To simplify notation, let p_j and θ_j be the premiums and risk types of consumers selected into firm j. The random variable θ_j admits distribution m_j, which can be directly estimated from the data.

Lemma 1. *Given m_j, the correlation $corr(p_j, \theta_j)$ is monotonically decreasing in σ_j.*

Proof. Since p_j and θ_j are positively correlated (empirically verifiable),

$$corr(p_j, \theta_j) = \frac{cov(p_j, \theta_j)}{\sqrt{var(\theta_j) var(p_j)}} = \sqrt{\frac{cov^2(p_j, \theta_j)}{var(p_j) var(\theta_j)}} = \sqrt{\frac{var(E(\theta_j|\hat{\theta}_j))}{var(\theta_j)}}. \tag{B.5}$$

The last equality in Equation (B.5) holds because, by Equation, (4.12)

$$var(E(\theta_j|\hat{\theta}_j)) = \frac{var(p_j)}{\beta_j^2} = \frac{var(p_j)}{[var(p_j)/cov(p_j, \theta_j)]^2} = \frac{cov^2(p_j, \theta_j)}{var(p_j)}.$$

Thus, for a fixed type distribution m_j within firm j, it is equivalent to show $var(E(\theta_j|\hat{\theta}_j))$ is monotonically decreasing in σ_j.

We first show that $var(E(\theta_j|\hat{\theta}_j)) = var(E(\theta_j|\tilde{\theta}_j))$, where $\tilde{\theta}_j \equiv \hat{\theta}_j + \gamma\beta_j b_j \sigma_j^2$. Let ϕ

denote the density of $\hat{\theta}_j$. We get that

$$\begin{aligned} var(E(\theta_j|\hat{\theta}_j)) &= \int (E(\theta_j|\hat{\theta}_j = x))^2 \phi(x)dx - \mu_j^2 \\ &= \int (E(\theta_j|\tilde{\theta}_j = x + \gamma\beta_j b_j \sigma_j^2))^2 \phi(x)dx - \mu_j^2 \\ &= \int (E(\theta_j|\tilde{\theta}_j = y))^2 \phi(y - \gamma\beta_j b_j \sigma_j^2)dy - \mu_j^2 \\ &= var(E(\theta_j|\tilde{\theta}_j)). \end{aligned}$$

Next, we show that $var(E(\theta_j|\tilde{\theta}_j))$ is monotonically decreasing in σ_j. Note that by Equation (B.4), $\tilde{\theta}_j \sim \mathcal{N}(\theta, \sigma_j^2)$. Consider a new normal random variable $\tilde{\theta}'_j \sim \mathcal{N}(\theta, \sigma_j'^2)$ with a larger variance $\sigma_j'^2 > \sigma_j^2$. Since both $\tilde{\theta}_j$ and $\tilde{\theta}'_j$ follow a normal distribution, we get that

$$\tilde{\theta}'_j = \tilde{\theta}_j + \eta,$$

where η is an independent and normally distributed random variable. By independence, the distribution of $E(\theta_j|\tilde{\theta}_j)$ is the same as the distribution of $E(\theta_j|\tilde{\theta}'_j, \eta)$, so $var(E(\theta_j|\tilde{\theta}_j)) = var(E(\theta_j|\tilde{\theta}'_j, \eta))$. Moreover,

$$var(E(\theta_j|\tilde{\theta}'_j, \eta)) > var(E(\theta_j|\tilde{\theta}'_j)),$$

because on the left-hand side of the inequality we project θ_j onto a larger space. We therefore obtain the desired result that $var(E(\theta_j|\tilde{\theta}_j)) > var(E(\theta_j|\tilde{\theta}'_j))$. □

For a fixed distribution of risk types within a firm, higher variance in the firm's signal distribution leads to a lower correlation between the premiums and risk types. Intuitively, when the signal distribution is very informative, consumers' risk types are better reflected in their premiums, and vice versa.[30] Since the type distribution m_j and $corr(p_j, \theta_j)$ can be easily estimated from the data, the one-to-one mapping between $corr(p_j, \theta_j)$ and σ_j^2 in Lemma 1 uniquely pins down the signal variance.

[30]To see this from another angle, the proof of Lemma 1 shows that $var(E(\theta_j|\hat{\theta}_j))$ decreases with σ_j^2. When σ_j^2 is small, the signals received by the firm are precise, and therefore the posterior mean of θ is very sensitive to the signals. As a result, the variance of the posterior mean is large. By contrast, when σ_j^2 is large, the signals received by the firm are not informative, and therefore the posterior mean is similar across different signals. As a result, the variance of $E(\theta|\hat{\theta}_j)$ is small.

C Solving the Equilibrium for Counterfacual Analysis

To evaluate counterfactual policies, we need to solve the market equilibrium for any given set of model parameters. We propose an iterative procedure to solve for all firms' pricing coefficients $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, the equilibrium expectation $E_j(\theta|\hat{\theta}_j, j)$, and the selection probabilities $Pr(j|\hat{\theta}_j, \theta)$ for all j. The iterative algorithm for solving the equilibrium works as follows:

1. In the outer loop, we solve for firms' pricing coefficients $(\boldsymbol{\alpha}, \boldsymbol{\beta})$. Denote the pricing coefficients at the r-th iteration by $(\boldsymbol{\alpha}^r, \boldsymbol{\beta}^r)$.

2. Given the model primitives $(\gamma, \boldsymbol{\sigma}, \boldsymbol{\xi}, \boldsymbol{c}, \boldsymbol{k})$ and pricing coefficients $(\boldsymbol{\alpha}^r, \boldsymbol{\beta}^r)$, we solve the r-th equilibrium sorting pattern in an inner loop:

   (1) Set the initial value $E_j^{r,0}(\theta|\hat{\theta}_j, j) = E_j^{r-1}(\theta|\hat{\theta}_j, j)$.

   (2) Compute the premium offered by each firm using the following equation:

   $$p_j^{r,0}(\hat{\theta}_j) = \alpha_j^r + \beta_j^r E_j^{r,0}(\theta|\hat{\theta}_j, j).$$

   (3) Compute the choice probabilities:

   $$Pr_j^{r,0}(j|\hat{\theta}_j, \theta) = \int_{\hat{\boldsymbol{\theta}}_{-j}} \frac{\exp(-\gamma p_j^{r,0}(\hat{\theta}_j) + \xi_j)}{\sum_{j'=1}^{J} \exp(-\gamma p_{j'}^{r,0}(\hat{\theta}_{j'}) + \xi_{j'})} f(\hat{\boldsymbol{\theta}}_{-j}|\theta; \boldsymbol{\sigma}_{-j}) d\hat{\boldsymbol{\theta}}_{-j}.$$

   (4) Update the expectation by

   $$E_j^{r,1}(\theta|\hat{\theta}_j, j) = \frac{\int_\theta \theta Pr^{r,0}(j|\hat{\theta}_j, \theta) f_j(\hat{\theta}_j|\theta; \sigma_j) f_0(\theta) d\theta}{\int_\theta Pr^{r,0}(j|\hat{\theta}_j, \theta) f_j(\hat{\theta}_j|\theta; \sigma_j) f_0(\theta) d\theta}.$$

   (5) Iterate this process until the expectation converges. Denote the limits by $E_j^r(\theta|\hat{\theta}_j, j)$ and $Pr_j^r(j|\hat{\theta}_j, \theta)$. These represent the r-th equilibrium sorting pattern.

3. Having obtained $E_j^r(\theta|\hat{\theta}_j, j)$ and $Pr_j^r(j|\hat{\theta}_j, \theta)$ from the inner loop, we now compute the first-order derivatives of the profit function with respect to pricing coefficients

(α_j^r, β_j^r) for all firms. Specifically,

$$\begin{aligned}
\frac{\partial \pi_j^r(\boldsymbol{\alpha}^r, \boldsymbol{\beta}^r)}{\partial \alpha_j^r} &= \int_\theta \int_{\hat{\boldsymbol{\theta}}} Pr^r(j|\hat{\boldsymbol{\theta}}) f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}} d\theta \\
&+ \int_\theta \int_{\hat{\boldsymbol{\theta}}} (\alpha_j^r + \beta_j^r \theta + c_j - k_j \theta) \frac{\partial Pr^r(j|\hat{\boldsymbol{\theta}})}{\partial \alpha_j^r} f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}} d\theta, \\
\frac{\partial \pi_j^r(\boldsymbol{\alpha}^r, \boldsymbol{\beta}^r)}{\partial \beta_j^r} &= \int_\theta \int_{\hat{\boldsymbol{\theta}}} \theta Pr^r(j|\hat{\boldsymbol{\theta}}) f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}} d\theta \\
&+ \int_\theta \int_{\hat{\boldsymbol{\theta}}} (\alpha_j^r + \beta_j^r \theta + c_j - k_j \theta) \frac{\partial Pr^r(j|\hat{\boldsymbol{\theta}})}{\partial \beta_j^r} f(\hat{\boldsymbol{\theta}}|\theta) f_0(\theta) d\hat{\boldsymbol{\theta}} d\theta.
\end{aligned}$$

4. Update the pricing coefficients as follows:

$$\begin{aligned}
\alpha_j^{r+1} &= \alpha_j^r + \Delta\alpha_j^r \mathbf{1}\left\{ \frac{\partial \pi_j^r(\boldsymbol{\alpha}^r, \boldsymbol{\beta}^r)}{\partial \alpha_j^r} \geq 0 \right\} - \Delta\alpha_j^r \mathbf{1}\left\{ \frac{\partial \pi_j^r(\boldsymbol{\alpha}^r, \boldsymbol{\beta}^r)}{\partial \alpha_j^r} < 0 \right\}, \\
\beta_j^{r+1} &= \beta_j^r + \Delta\beta_j^r \mathbf{1}\left\{ \frac{\partial \pi_j^r(\boldsymbol{\alpha}^r, \boldsymbol{\beta}^r)}{\partial \beta_j^r} \geq 0 \right\} - \Delta\beta_j^r \mathbf{1}\left\{ \frac{\partial \pi_j^r(\boldsymbol{\alpha}^r, \boldsymbol{\beta}^r)}{\partial \beta_j^r} < 0 \right\},
\end{aligned}$$

   where $\Delta\alpha_j^r$ and $\Delta\beta_j^r$ are the r-th incremental changes in the pricing coefficients.

5. Finally, we reduce the size of incremental changes $(\Delta\alpha_j^r, \Delta\beta_j^r)$ if sign switching in the first-order conditions is observed in two consecutive iterations.

$$\begin{aligned}
\Delta\alpha_j^{r+1} &= \begin{cases} \frac{\Delta\alpha_j^r}{2} & \text{if} \quad \frac{\partial \pi_j^r(\boldsymbol{\alpha}^r, \boldsymbol{\beta}^r)}{\partial \alpha_j^r} \frac{\partial \pi_j^{r-1}(\boldsymbol{\alpha}^{r-1}, \boldsymbol{\beta}^{r-1})}{\partial \alpha_j^{r-1}} < 0 \\ \Delta\alpha_j^r & \text{otherwise,} \end{cases} \\
\Delta\beta_j^{r+1} &= \begin{cases} \frac{\Delta\beta_j^r}{2} & \text{if} \quad \frac{\partial \pi_j^r(\boldsymbol{\alpha}^r, \boldsymbol{\beta}^r)}{\partial \beta_j^r} \frac{\partial \pi_j^{r-1}(\boldsymbol{\alpha}^{r-1}, \boldsymbol{\beta}^{r-1})}{\partial \beta_j^{r-1}} < 0 \\ \Delta\beta_j^r & \text{otherwise,} \end{cases}
\end{aligned}$$

6. Iterate this process until $(\Delta\alpha_j^r, \Delta\beta_j^r)$ converge to 0 for all j.

Finally, given that the current solution of the pricing coefficient might be a local maximizer, we check whether it is globally optimal. That is, for each firm j, fixing other firms' pricing strategy at the current solution, we search for the value of (α_j, β_j) that maximizes profit π_j. If the global maximizer differs from the current pricing coefficient, we update firm j's pricing coefficient to that global maximizer and repeat steps 1–6. Otherwise, we have found the Nash equilibrium.

References

ACQUISTI, A., C. TAYLOR, AND L. WAGMAN (2016): "The economics of privacy," *Journal of Economic Literature*, 54, 442–492.

AKERLOF, G. A. (1970): "The market for 'lemons': Quality uncertainty and the market mechanism," *Quarterly Journal of Economics*, 84, 488–500.

ALCOBENDAS, M., S. KOBAYASHI, K. SHI, AND M. SHUM (2023): "The impact of privacy protection on online advertising Markets," in *Proceedings of the 24th ACM Conference on Economics and Computation*, 62–62.

ALLEN, J., R. CLARK, AND J.-F. HOUDE (2014): "Price dispersion in mortgage markets," *Journal of Industrial Economics*, 62, 377–416.

——— (2019): "Search frictions and market power in negotiated-price markets," *Journal of Political Economy*, 127, 1550–1598.

BARWICK, P. J., H.-S. KWON, AND S. LI (2023): "Attribute-based subsidies and market power: an application to electric vehicles," *Working paper*.

BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): "Automobile prices in market equilibrium," *Econometrica*, 63, 841–890.

BERRY, S. T. (1994): "Estimating discrete-choice models of product differentiation," *RAND Journal of Economics*, 242–262.

BIGLAISER, G., F. LI, C. MURRY, AND Y. ZHOU (2020): "Intermediaries and product quality in used car markets," *RAND Journal of Economics*, 51, 905–933.

BROWN, Z. Y. (2019): "Equilibrium effects of health care price information," *Review of Economics and Statistics*, 101, 699–712.

BYRNE, D. P., N. DE ROOS, M. S. LEWIS, L. M. MARX, AND X. A. WU (2023): "Asymmetric information sharing in oligopoly: A natural experiment in retail gasoline," *Available at SSRN 4513435*.

CABRAL, M., M. GERUSO, AND N. MAHONEY (2018): "Do larger health insurance subsidies benefit patients or producers? Evidence from Medicare Advantage," *American Economic Review*, 108, 2048–2087.

CAMPBELL, J., A. GOLDFARB, AND C. TUCKER (2015): "Privacy regulation and market structure," *Journal of Economics & Management Strategy*, 24, 47–73.

CHEN, N. AND H.-T. TSAI (2023): "Price competition dnder information (dis)advantage," *Available at SSRN 4420175*.

CHIAPPORI, P.-A., B. JULLIEN, B. SALANIÉ, AND F. SALANIE (2006): "Asymmetric information in insurance: General testable implications," *RAND Journal of Economics*, 37, 783–798.

CHIAPPORI, P.-A. AND B. SALANIE (2000): "Testing for asymmetric information in insurance markets," *Journal of Political Economy*, 108, 56–78.

CICALA, S. (2015): "When does regulation distort costs? Lessons from fuel procurement in us electricity generation," *American Economic Review*, 105, 411–444.

COHEN, A. (2005): "Asymmetric information and learning: Evidence from the automobile insurance market," *Review of Economics and Statistics*, 87, 197–207.

COHEN, A. AND L. EINAV (2007): "Estimating risk preferences from deductible choice," *American Economic Review*, 97, 745–788.

COOPER, Z., S. V. CRAIG, M. GAYNOR, AND J. VAN REENEN (2019): "The price ain't right? Hospital prices and health spending on the privately insured," *Quarterly Journal of Economics*, 134, 51–107.

COSCONATI, M. (2023): "Market-wide moral hazard and price walking: Evidence from automobile insurance market," *Working paper*.

CRAWFORD, G. S., N. PAVANINI, AND F. SCHIVARDI (2018): "Asymmetric information and imperfect competition in lending markets," *American Economic Review*, 108, 1659–1701.

CRAWFORD, G. S., O. SHCHERBAKOV, AND M. SHUM (2019): "Quality overprovision in cable television markets," *American Economic Review*, 109, 956–995.

CUESTA, J. I. AND A. SEPÚLVEDA (2021): "Price regulation in credit markets: A trade-off between consumer protection and credit access," *Available at SSRN 3282910*.

EDELBERG, W. (2006): "Risk-based pricing of interest rates for consumer loans," *Journal of Monetary Economics*, 53, 2283–2298.

EINAV, L. AND A. FINKELSTEIN (2011): "Selection in insurance markets: Theory and empirics in pictures," *Journal of Economic Perspectives*, 25, 115–138.

EINAV, L., A. FINKELSTEIN, AND M. R. CULLEN (2010a): "Estimating welfare in insurance markets using variation in prices," *Quarterly Journal of Economics*, 125, 877–921.

EINAV, L., A. FINKELSTEIN, AND N. MAHONEY (2021): "The IO of selection markets," in *Handbook of Industrial Organization*, Elsevier, vol. 5, 389–426.

EINAV, L., A. FINKELSTEIN, AND P. SCHRIMPF (2010b): "Optimal mandates and the welfare cost of asymmetric information: Evidence from the UK annuity market," *Econometrica*, 78, 1031–1092.

EINAV, L., M. JENKINS, AND J. LEVIN (2012): "Contract pricing in consumer credit markets," *Econometrica*, 80, 1387–1432.

——— (2013): "The impact of credit scoring on consumer lending," *RAND Journal of Economics*, 44, 249–274.

ENGELBRECHT-WIGGANS, R., P. R. MILGROM, AND R. J. WEBER (1983): "Competitive bidding and proprietary information," *Journal of Mathematical Economics*, 11, 161–169.

ERICSON, K. M. M. (2014): "Consumer inertia and firm pricing in the Medicare Part D prescription drug insurance exchange," *American Economic Journal: Economic Policy*, 6, 38–64.

FAN, Y. AND C. YANG (2020): "Competition, product proliferation, and welfare: A study of the US smartphone market," *American Economic Journal: Microeconomics*, 12, 99–134.

FARRELL, J. AND P. KLEMPERER (2007): "Coordination and lock-in: Competition with switching costs and network effects," *Handbook of Industrial Organization*, 3, 1967–2072.

FRIEDRICH, B. U., M. B. HACKMANN, A. KAPOR, S. MORONI, AND A. B. NANDRUP (2023): "Asymmetric information in matching markets: Evidence from medical school programs in Denmark," *Working paper*.

FUDENBERG, D. AND J. TIROLE (1991): *Game Theory*, MIT Press.

GOLDBERG, P. K. (1996): "Dealer price discrimination in new car purchases: Evidence from the consumer expenditure survey," *Journal of Political Economy*, 104, 622–654.

GOLDBURD, M., A. KHARE, D. TEVET, AND D. GULLER (2016): "Generalized linear models for insurance rating," *Casualty Actuarial Society, CAS Monographs Series*, 5.

GOLDFARB, A. AND V. F. QUE (2023): "The economics of digital privacy," *Annual Review of Economics*, 15.

GRODZICKI, D. (2023): "The evolution of competition in the credit card market," *Available at SSRN 4493211*.

GUERRE, E., I. PERRIGNE, AND Q. VUONG (2000): "Optimal nonparametric estimation of first-price auctions," *Econometrica*, 68, 525–574.

HANDEL, B. R. (2013): "Adverse selection and inertia in health insurance markets: When nudging hurts," *American Economic Review*, 103, 2643–2682.

HENDRICKS, K. AND R. H. PORTER (1988): "An empirical study of an auction with asymmetric information," *American Economic Review*, 865–883.

HENDRICKS, K., R. H. PORTER, AND B. BOUDREAU (1987): "Information, returns, and bidding behavior in OCS auctions: 1954–1969," *Journal of Industrial Economics*, 517–542.

HENDRICKS, K., R. H. PORTER, AND C. A. WILSON (1994): "Auctions for oil and gas leases with an informed bidder and a random reservation price," *Econometrica: Journal of the Econometric Society*, 1415–1444.

HONKA, E. (2014): "Quantifying search and switching costs in the US auto insurance industry," *RAND Journal of Economics*, 45, 847–884.

JEZIORSKI, P., E. KRASNOKUTSKAYA, AND O. CECCARINI (2017): "Adverse selection and moral hazard in the dynamic model of auto insurance," *UC Berkeley, Haas School of Business working paper*.

JIN, G. Z. AND L. WAGMAN (2021): "Big data at the crossroads of antitrust and consumer protection," *Information Economics and Policy*, 54, 100865.

JIN, Y. AND S. VASSERMAN (2021): "Buying data from consumers: The impact of monitoring programs in US auto insurance," Tech. rep., National Bureau of Economic Research.

JOHNSON, G. (2022): "Economic research on privacy regulation: Lessons from the GDPR and beyond," *NBER Working paper*.

JOHNSON, G. A., S. K. SHRIVER, AND S. G. GOLDBERG (2023): "Privacy and market concentration: Intended and unintended consequences of the GDPR," *Management Science*, 69, 5695–5721.

MAHONEY, N. AND E. G. WEYL (2017): "Imperfect competition in selection markets," *Review of Economics and Statistics*, 99, 637–651.

MASKIN, E. AND J. RILEY (2000): "Equilibrium in sealed high bid auctions," *Review of Economic Studies*, 67, 439–454.

MATCHAM, W. (2023): "Risk-based quantity limits in credit card markets," Tech. rep.

MILGROM, P. R. (1979): "A convergence theorem for competitive bidding with differential information," *Econometrica: Journal of the Econometric Society*, 679–688.

——— (1981): "Rational expectations, information acquisition, and competitive bidding," *Econometrica: Journal of the Econometric Society*, 921–943.

NELSON, S. (2018): "Private information and price regulation in the US credit card market," *Working paper*.

PANETTA, F., F. SCHIVARDI, AND M. SHUM (2009): "Do mergers improve information? Evidence from the loan market," *Journal of Money, Credit and Banking*, 41, 673–709.

PORTER, R. H. (1995): "The role of information in US offshore oil and gas lease auction," *Econometrica: Journal of the Econometric Society*, 1–27.

ROTHSCHILD, M. AND J. STIGLITZ (1976): "Equilibrium in competitive insurance markets: An essay on the economics of imperfect information," *Quarterly Journal of Economics*, 90, 629–649.

ROY, A. D. (1951): "Some thoughts on the distribution of earnings," *Oxford Economic Papers*, 3, 135–146.

SAGL, S. (2023): "Dispersion, discrimination, and the price of your pickup," *Working paper*.

SALANIÉ, B. (2017): "Equilibrium in insurance markets: An empiricist's view," *Geneva Risk and Insurance Review*, 42, 1–14.

SALZ, T. (2022): "Intermediation and competition in search markets: An empirical case study," *Journal of Political Economy*, 130, 310–345.

SERNA, N. (2023): "Non-price competition, risk selection, and heterogeneous costs in hospital networks," *Working paper*.

SWEETING, A. (2013): "Dynamic product positioning in differentiated product markets: The effect of fees for musical performance rights on the commercial radio industry," *Econometrica*, 81, 1763–1803.

TAYLOR, C. R. (2004): "Consumer privacy and the market for customer information," *RAND Journal of Economics*, 631–650.

WILSON, R. (1977): "A bidding model of perfect competition," *Review of Economic Studies*, 44, 511–518.

WILSON, R. B. (1967): "Competitive bidding with asymmetric information," *Management Science*, 13, 816–820.

WU, F. AND Y. XIN (in process): "Estimating selection models: A functional contraction approach," Tech. rep., California Institute of Technology.