Buying Data from Consumers

The Impact of Monitoring Programs in U.S. Auto Insurance

Yizhou Jin and Shoshana Vasserman*

September 2019

Abstract

We study the impact of a voluntary monitoring program by a major U.S. auto insurer, in which drivers accept short-term tracking in exchange for potential discounts on future premiums. We acquire a detailed proprietary dataset from the insurer and match it with competitor price menus. We first quantify the degree to which monitoring incentivizes safer driving and allows more accurate risk-based pricing. We then model the demand and supply forces that determine the amount of information revealed in equilibrium: structural demand estimates capture correlations among cost and demand for insurance and for monitoring; a dynamic pricing model links the firm's information on driver risk to prices. We find large profit and welfare gains from introducing monitoring. Safer drivers self-select into monitoring, with those who opt in becoming 30% safer when monitored. Given resource costs and price competition, a data-sharing mandate would have reduced short-term welfare.

*Jin: UC Berkeley, corresponding author, jyz@berkeley; Vasserman: Stanford, svass@stanford. Latest version: yjin.io/jmp. We thank our advisors Ariel Pakes, Nathan Hendren, Robin Lee, Dennis Yao, Leemore Dafny, and Elie Tamer; our data providers Quadrant Information Services and an unnamed auto insurer; Harvard and the Geneva Association for financial support; Alberto Abadie, Nikhil Agarwal, Isiah Andrews, Jie Bai, Alonso De Gortari, Liran Einav, Ashvin Gandhi, Nir Hak, Ben Handel, Oliver Hart, Kevin He, Panle Barwick, Ginger Jin, Myrto Kalouptsidi, Scott Kominers, Jonathan Kolstad, Jing Li, Alex MacKay, Jonathan Roth, James Savage, Steve Tadelis, Andrew Sweeting, Chad Syverson, John Wells, Thomas Wollmann, and various seminar participants for valuable comments. Ability to publish is not contingent on results (data usage agreement contact carolina_harvey@harvard.edu).

New technologies and data privacy regulations have led to a proliferation of *direct transactions of consumer data*. Firms directly incentivize consumers to voluntarily reveal information, while keeping the collected data as proprietary. How does this type of data collection influence firm profit and consumer welfare?

In this paper, we develop an empirical framework to answer this question and quantify the impact of an auto-insurance *monitoring program* ("pay-how-you-drive") in the U.S. New customers are invited to plug a simple device into their cars, which tracks and reports their driving behavior for up to six months (Figure A.1). In exchange, the insurer uses the data to better assess accident risk and adjust future premiums. Unlike most traditional pricing factors such as age or claim history, monitoring data is not shared with other firms. In 2017, insurers serving over 60% of the $267 billion U.S. auto insurance industry offered monitoring programs.[1] Similar programs have been introduced in other industries, such as life insurance and lending (Figure A.2).[2] Despite this growing relevance, empirical evidence on the economic impact of monitoring programs or other types of direct transactions of consumer data is sparse.

We construct a novel dataset by merging proprietary individual-level data from a major U.S. auto insurer (hereinafter referred to as "the Firm") with prices charged by its competitors. The resulting panel data details drivers' characteristics, the set of price menus that they face from top insurers, insurance contracts purchased, and realized insurance claims. Our research window covers the introduction of the firm's monitoring program. For each driver who opts in, we observe a monitoring score and the corresponding premium adjustments. Taken together, our analysis uses a panel dataset of over 1 million drivers and 50 million insurance quotes.

We take a two-step approach in our empirical analysis. First, we evaluate the monitoring technology by quantifying its ability to both incentivize safer driving and allow more accurate risk-based pricing. Second, we model the demand and supply forces that shape the amount of information revealed in equilibrium. Our demand model jointly captures self-selection into monitoring, into coverage plans, and into the Firm. On the supply side, proprietary data allow the Firm to raise markups, but it faces resource costs and price competition to "produce the data in the first place" (Posner 1978). We capture both factors with a two-period pricing model that make the Firms' information on driver risk dependent on prices. Our model allows us to jointly characterize market and information structures in counterfactuals. Us-

[1] 2017 annual report of the National Association of Insurance Commissioners.

[2] The Vitality program from life insurer John Hancock tracks and rewards exercise and health-related behaviors. Ant Financial incentivizes users to conduct more personal finance transactions in exchange for borrowing discounts.

ing this, we evaluate the impact of introducing monitoring, optimal pricing for the program, and a counterfactual regulation that eliminates proprietary data.[3]

We find three main results: (i) data collection changes consumer behavior. Drivers become 30% safer when monitored. (ii) Safer drivers are more likely to opt in, even holding financial risk and rewards fixed. (iii) But monitoring take-up remains low due to both demand frictions against monitoring and attractive outside options from other insurers. Overall, compared to a counterfactual with no monitoring, short-term consumer welfare and profit both increase. Forcing the firm to make monitoring data public would have done the opposite and reduced the amount of monitoring information revealed by consumers in equilibrium.

We start with a pair of reduced-form facts that characterize the relationship between consumers' accident risk and behavior under the monitoring technology. We first show that drivers become safer when monitored – an incentive effect. Monitoring is only done during the first semester of insurance for any new customer who opts in. We capture the corresponding within-driver across-period variation in claims with a difference-in-differences estimator. We find that the average opt-in driver *becomes* 30% safer when monitored. However, this incentive effect only explains 64% of the risk difference between monitored and unmonitored groups in the first period. Furthermore, monitoring scores remain highly predictive of risk in subsequent periods conditional on observables. These patterns suggest that the monitoring program captures previously unobserved risk differences across consumers, driving advantageous selection into monitoring.

The bulk of our analysis relies on a structural model of demand. In order to capture the correlation between cost, demand, and monitoring, we adopt a choice framework with three interrelated parts. First, a stochastic cost model maps claims into drivers' latent risk types. It also explains how risk covaries with observable characteristics and how it changes during monitoring. Second, a signal model formalizes how monitoring scores can further inform the Firm about driver risk. Third, a choice model connects consumers' information (monitoring opt-in) and product (insurer and coverage) choices by mapping both into a unified set of demand primitives.

Taken together, the model allows monitoring opt-in to depend on several forces. Drivers anticipate risk reduction during monitoring. Meanwhile, safer drivers ex-

[3] The General Data Protection Regulation (2016) in the EU aims to curb the accumulation of proprietary data by allowing consumers to rescind consent and take their data to other firms, and by requiring firms to be transparent about how consumer data is used in pricing (see EUGDPR (2018)). Similar regulatory proposals are being considered in the U.S. (see press release NTIA (2018)).

pect higher future discounts, but the monitoring signal noise raises reclassification risk. Lastly, drivers incur unobserved privacy, effort, or decision costs from being monitored. We model these jointly with a disutility term for monitoring.

To achieve this, our model augments the canonical insurance framework *a la* Cohen and Einav (2007) to feature inertia costs (path dependence in choices), as well as heterogeneous monitoring disutility and renewal price expectation across *unobserved* consumer risk types. Identification of demand parameters leverages rich time and geographic variation in prices and in coverage options, conditional on other observables used in firms' pricing rules. This includes variations in the eligibility and pricing of the monitoring program, which pin down monitoring disutility. Our estimates produce a close fit to the empirical distribution of monitoring scores and opt-in choices. It also makes good predictions out-of-sample, in which the mandatory minimum coverage changed in one (U.S.) state.

We find that the average driver suffers a $93 disutility from being monitored, contributing to the low opt-in rate in the data. But this disutility is lower for safer drivers, enhancing advantageous selection beyond what is implied by financial risk and rewards alone. Meanwhile, the average driver forgoes $284 in financial gain annually by not exploiting outside options from competitors. This suggests that the market may remain imperfectly competitive even with perfect information on driver risk. Further, drivers are only modestly risk-averse. Monitoring score's (signaling) precision therefore has little influence on monitoring demand.

To evaluate the impact of the monitoring program, we compare the current regime with a counterfactual one without monitoring, holding baseline prices fixed.[4] Introducing monitoring raises both firm profit and consumer welfare. Total annual surplus increases by $13.3 (1.7% of premium), 64% of which can be attributed to the risk reduction during monitoring. Without the incentive effect, overall profit drops in the market, highlighting that better information facilitates direct cream-skimming that push the market towards the first-best benchmark.[5]

Next, we propose a pricing model that endogenizes the production of monitoring data and therefore the firm's information set. This is used to study (i) the optimal pricing of the monitoring program given its observed marginal cost, and (ii) the equilibrium impact of a mandate forcing the firm to share its proprietary data with competitors. The two-period two-product model features an "invest-and-harvest"

[4] Appendix B shows that the firm did not raise prices for unmonitored drivers when introducing monitoring.

[5] This is in contrast to Rothschild and Stiglitz (1976), in which cream-skimming leads to unraveling when asymmetric information is present and fixed, while firms conduct competitive screening by offering lower coverage.

pricing dynamic.[6] Holding fixed competitor prices, the Firm reaches optimal pricing by reducing ex-post rent-sharing with monitored drivers while increasing ex-ante effort to produce monitoring data. The latter is achieved primarily with a large opt-in discount because price competition limits the Firm's ability to profitably surcharge unmonitored drivers. Moreover, a regulation that requires the firm to share monitoring data curbs ex-post markups but undermines ex-ante incentives for the Firm to produce monitoring data. Despite driver risk reduction during monitoring and high firm-switching inertia (imperfect competition), the Firm reduces the incentives it offers for monitoring opt-in. Compared to the equilibrium without data-sharing, this leads to a large drop in monitoring opt-in rate. Annual consumer welfare and firm profit both decrease.

Related Literature Our research contributes to several literatures. First, we extend the empirical literature on insurance and selection markets. We are among the first to investigate firms' strategy to acquire – and consumers' willingness to reveal – risk information, formalizing the linkage between (product) market and information structures.[7] Specifically, consumers self-select into monitoring, while the firm can unilaterally mitigate information asymmetry and enhance market power through monitoring. Our work thus extends the literature on competitive screening with predetermined asymmetric information on consumer risk (Rothschild and Stiglitz 1976; Hendren 2013; Jeziorski, Krasnokutskaya, and Ceccarini 2019) or on changes in public information in the market.[8]

Second, we are related to the literature on dynamic contracting and information revelation. Monitoring allows the Firm to learn about consumer risk over time (Hart 1983; Cohen 2012; Hendel 2017). We empirically show that this distorts consumer incentives and behavior.[9] Third, our study contributes to the economics

[6]This is common in markets with high switching costs, see Beggs and Klemperer (1992), Farrell and Klemperer (2007), and Dubé, Hitsch, and Rossi (2009).

[7]Screening is multi-dimensional in our setting (Cohen and Einav 2007; Fang, Keane, and Silverman 2008; Barseghyan, Molinari, O'Donoghue, and Teitelbaum 2013; Handel 2013; Handel, Kolstad, and Spinnewijn forthcoming). We also allow consumers to be forward-looking, related to studies on reclassification risk (Hendel and Lizzeri 2003; Handel, Hendel, and Whinston 2015; Aron Dine, Einav, Finkelstein, and Cullen 2015).

[8]Regulations such as community-rating mandates (limits to risk categorization) are most common (Finkelstein, Poterba, and Rothschild 2009; Einav, Finkelstein, and Schrimpf 2010; Einav, Levin, and Jenkins 2012; Agarwal, Chomsisengphet, Mahoney, and Stroebel 2015; Cox 2017; Nelson 2018). Lewis (2011) and Tadelis and Zettelmeyer (2015) examine disclosure rule change in online auctions. Mahoney and Weyl (2017) posit that market power further depresses quantity under adverse selection, which is contradicted empirically by Crawford, Pavanini, and Schivardi (2018)'s study in the Italian small-business lending market.

[9]A related theory literature focuses on price discrimination enabled by consumers' online purchase histories. See Rossi, McCulloch, and Allenby (1996), Acquisti and Varian (2005), Taylor (2004), Fudenberg and Villas-Boas (2006), and Bonatti

of privacy by characterizing the equilibrium (implicit) price and quantity of consumer information in a competitive market, as well as its social value. Specifically, we extend the literature by studying not only consumers' privacy choices,[10] but also how their choice environments are affected by product market competition and by data property rights (Posner 1978; Stigler 1980; Hermalin and Katz 2006).

The rest of the paper proceeds as follows. Section I describes our data and provides background information on auto insurance and the monitoring program we study. Section II conducts reduced-form tests that measure monitoring's ability to reduce risk and to mitigate information asymmetry. Section III presents our structural model, identification arguments, and estimation procedures to recover key demand and cost parameters. Section VI discusses estimation results and counterfactual simulation procedures for welfare analyses. Section V proposes a model of monitoring pricing and investigates equilibrium implications for optimal pricing and information sharing. Section VI concludes.

1 Background and Data

In this section, we provide background information on U.S. auto insurance and the monitoring program we study. We also describe our datasets.

and Cisternas (2018). Some empirical work have looked at monitoring among truck drivers and consumer lending (Hubbard 2000; Wei, Yildirim, Van den Bulte, and Dellarocas 2015). Soleymanian, Weinberg, and Zhu (2019) is closest to our setting. They analyze driving data, as opposed to claim outcomes, from a U.S. auto insurance monitoring program and find that monitoring reduces several dimensions of unsafe driving behaviors but not the amount driven. Another literature focus on usage-based pricing (Narayanan, Chintagunta, and Miravete 2007; Chung, Steenburgh, and Sudhir 2013; Lambrecht, Seim, and Skiera 2007; Liu, Montgomery, and Srinivasan 2014; Nevo, Turner, and Williams 2016). The main difference being that the temporary nature of monitoring and its dynamic price impact turn our problem from a standard moral hazard one into one with a signaling equilibrium.

[10]See Milgrom (1981), Jovanovic (1982), Jin and Leslie (2003), Dranove and Jin (2010), and Lewis (2011) about imperfect advantageous selection in information disclosure. See Goldfarb and Tucker (2011), Goldfarb and Tucker (2012), Tucker (2012), Acquisti, John, and Loewenstein (2012), Burtch, Ghose, and Wattal (2015), Acquisti, Taylor, and Wagman (2016), Kummer and Schulte (2019), and Lin (2019) for privacy preference.

1.1 Auto Insurance

Auto insurers in the U.S. collected \$267 billion dollars of premiums in 2017.[11] There are two main categories of insurance: liability and property. Property insurance covers damage to one's own car in an accident, regardless of fault. Liability insurance covers injury and property liability associated with an at-fault accident. In all states we study, liability insurance is mandatory, with the minimum required coverage ranging from \$25,000 to \$100,000.[12]

Insurance prices are heavily regulated. Major insurers collect large amount of consumer information in risk-rating, most of which is public or shared across firms. Firms are required to publish filings that detail their pricing algorithms. In most states, the insurance commissioner needs to approve such filings.[13] An important focus of the regulator is deterring excessive price discrimination based on demand elasticity.[14] In general, a pricing rule can be summarized by the following equation, where price p for a (single-driver-single-vehicle) policy choosing certain liability coverage is:[15]

$$\begin{aligned} p = &\text{ base rate} \times \text{driver factor} \times \text{vehicle factor} \times \text{location factor} \\ &\times \text{tier factor} \times \text{coverage factor} + \text{markups and fees} \end{aligned} \tag{1}$$

Within each firm, price variation is based on observable characteristics, time, and coverage choice. Base rates vary only by state and calendar time. Driver, vehicle, and location factors include age, vehicle model, and zipcode-level population density, etc. This information is verified and cross-referenced among various public and industry databases. Tier factors incorporate information from claim and credit databases, which include accident, traffic violation (DUI, speeding, etc.), or financial (delinquency, bankruptcy, etc.) records in the past.[16] Choosing a higher coverage scales prices by a positive factor. Lastly, firms charge a fee that includes markups and overhead for operational and marketing expenditures.[17]

As in Figure 1a, new customers to the firm must report observable characteristics at time $t = 0$. This facilitates risk rating, based on which the firm generates individualized price menu. Consumers make coverage choice or go to other firms. There is

[11] This is according to the National Association of Insurance Commissioners. This number is calculated as premiums from property annual statements plus state funds.

[12] All states that we study follow an "at-fault" tort system and mandate liability insurance. In reality, liability insurance is specified by three coverage limits. For example, 20/40/10 means that, in an accident, the insurer covers liability for bodily injuries up to \$40,000 overall, but no more than \$20,000 per victim; it also covers liability for property damage (cars or other infrastructure) for up to \$10,000. We quote the highest number here.

[13] Some states follow a "use-and-file" system, which means that insurers can seek pricing approval ex-post as long as

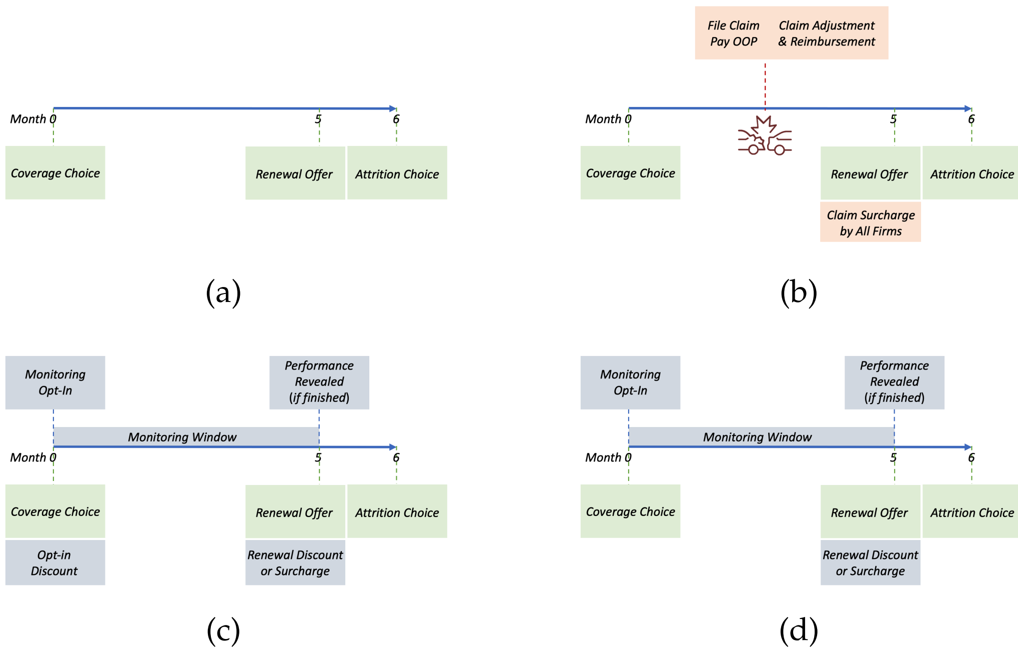


Figure 1: Timing Illustration of Auto Insurance and Monitoring Program

no long-term commitment in U.S. auto insurance. Each period lasts for six months, at the end of which consumers decide to stay or leave given the firm's renewal quotes provided at the end of month five. If an auto accident occurs (Figure 1b), the insured files a claim immediately and, given evaluation and adjustment by the insurer, gets reimbursed and pays out-of-pocket accordingly. Meanwhile, the claim is recorded in industry databases in real time. The consumer will likely face a claim surcharge at renewal or higher prices when switching to other firms.

Dataset 1 - Panel data from an auto insurer Our first dataset comes from a national auto insurer in the U.S. It is a panel that spans 2012 to 2016, and covers 22 states. For tractability, we focus only on *single-driver-single-vehicle* insurance policies sold online or via phone. Nonetheless, we observe more than 1 million drivers

any price changes are reflected in public filings.

[14]"Price optimization" on top of risk rating is typically not allowed by state insurance commissioners.

[15]See Appendix H, e.g. Figure H.1.

[16]See Appendix H, Figures H.7 and H.8

[17]The latter is often referred to as the loading factor in the literature.

for an average duration of 1.86 years (3.73 periods)[18]. The date range spans periods pre- and post-introduction of monitoring.

At the beginning of each period, we observe each driver's observable characteristics[19] as well as the price menu offered, which include all available options from the firm and their prices. We also see the driver's coverage choice. For simplicity, we limit our attention to *liability coverage* (limits). Not only is it the most expensive coverage for the average driver, its mandatory nature also strongly influences firms' competitive strategy and monitoring's allocative benefit. These cover auto accidents involving two or more parties, in which the policy holder is at least partially at-fault. As such, our focus also mitigates concerns about under-reporting.[20]

During renewals, those with a claim will experience a surcharge that ranges from 10% to 50% (Figure A.4).[21] Otherwise, the average driver experiences close to no price change in a typical renewal period. Overall, about 5% to 20% of drivers leave the firm after each period.[22]

Table 1(a) presents summary statistics of prices, coverage levels, and claims. In addition, the average driver is 33 years old, drives a 2006 vehicle, lives in a zipcode area with average annual income of $142,000, and has 0.3 recorded accidents in the past 5 years. Per six-month period, he pays $380 in liability premium and files 0.05 liability claims (1 in ten years). We also observe his assigned risk class, which is the premium calculated for him before coverage factor, markups, and fees.

Dataset 2 - Price menus of competitors based on price filings To understand competition, we need to account for drivers' outside options. Therefore, we complement our main dataset with the firm's competitor price menus. Our data include quotes from all liability coverage options offered by the firm's top five competitors in each state based on price filings, harnessed using Quadrant Information Services' proprietary software. We are able to achieve precise matches based on a rich set of consumer characteristics, including state and calendar day.[23] Table 1(b)

[18]The panel is right-censored, but the censoring is plausibly uninformative.

[19]Main observables include driver gender, age, martial status, education, out-of-state status, home-ownership, vehicle model, year, and financing, license and vehicle history, violation and accident records, credit history, prior insurance history, and zipcode population density. See Table A.3 for a list of observables used in our estimation procedure.

[20]In contrast, claim filing for single-car accidents is almost entirely discretionary.

[21]The surcharge varies only based on existing claims and traffic violation records.

[22]The first renewal sees some one-time discounts being removed, such as those for online processing.

[23]We match based on available observable characteristics including those in Table A.3, violation records, zipcode, vehicle make and model.

(a) Premium, Coverage and Claims (6-month Period)

Statistic	Mean	St. Dev.	Min	Median	Max
Total premium ($)	632	364	69	548	22,544
Liability premium ($)	380	208	32	336	10,177
Risk class ($)	255	172	50	212	9,724
Total claim ($)	323	2,822	0	0	544,814
Claim count	0.18	0.67	0	0	12
Liability claim ($)	164	2,209	0	0	513,311
Liability claim count	0.05	0.32	0	0	7
Liability coverage ($000)	126	119	25	60	500
Liability coverage (index)	2.10	1.15	1	2	8
Mandatory minimum ind.	0.36	0.48	0	0	1
Renewal count	1.76	2.01	0	1	9
Calendar year (index)	2.66	1.38	0	3	5

Notes: Risk class is the pre-markups-pre-fees premium for liability coverage. Coverage index ranks coverage options in ascending order and sets the mandatory minimum in each state as 1.

(b) By Coverage (a representative U.S. State)

Liability coverage ($000)	40	50	100	300	500
Quotes ($)	335.14	343.43	382.03	422.13	500.48
- Competitor 1	482.68	506.11	564.34	626.81	730.56
- Competitor 2	263.14	279.15	314.46	347.69	405.22
- Competitor 3	319.42	348.97	388.48	428.64	464.36
- Competitor 4	511.24	567.58	613.74	682.87	790.83
- Competitor 5	421.84	363.96	403.64	433.17	497.79
Share within firm (%)	19	39	20	19	3
Liability claim ($)	154.98	155.54	154.16	143.43	107.54
Liability claim count	0.05	0.05	0.04	0.03	0.03

Notes: This table reports the average quotes and claims of the Firm and its top 5 competitors by market share. We focus on one U.S. state to avoid pooling across states with different coverage options. In this state, the mandatory minimum and the most popular coverage changed from $40,000 to $50,000 during the research window.

Table 1: Summary Statistics

compares the quotes for the five most common liability coverage options across competitors in a representative U.S. state. Due to large menu size, we end up with millions of quotes per state. While our reduced-form analysis and our cost model estimation utilize the full dataset, our demand estimation relies only on three adjacent mid-western states, with 283,000 drivers and over 50 million quotes.

Looking ahead, observing competitor prices enables us to understand consumers' inertia to switch firms. In counterfactual analyses, we can then enumerate the full market (by simulating competitor quantities) and capture price competition under various information environments.

1.2 Monitoring Program

Our research focuses on the Firm's one-time and voluntary monitoring program for new customers. The monitoring process is summarized in Figures 1c and 1d. When customers first arrive, they choose whether to opt into monitoring immediately before seeing the coverage price menu. They are provided with information on the kinds of driving behavior that are tracked and rewarded, although the exact discount schedule is opaque. Specifically, high mileage driven, driving at night, high speed, and harsh braking are highlighted as monitored behaviors. The firm also spells out an opt-in discount applied on the first period premium as well as the mean and range of renewal discount that will be applied to all subsequent (renewal) periods.[24]

Opt-in drivers will receive a simple device via mailed within a week. They then have until the end of month five to accumulate around 100-150 days of monitored driving. If completed, the Firm evaluates their performance and includes an appropriate renewal discount when giving out renewal quotes.[25] If an accident occurs, monitoring data do not influence claim reporting, handling, or future premium adjustment. Monitoring continues after any disruptions from the accident.

During the monitoring period, monitored drivers receive real-time feedback on their performance. The Firm posts key statistics of recorded trips online. It also

[24]The average opt-in discount is 4.6% in our estimation dataset. We cannot disclose the renewal discount range exactly, but it centers around 7% and spans zero (-15% to 40%, for example).

[25]27% of drivers who start monitoring do not finish. Our main analysis ignores these drivers and focus on consumers' decision to *start and finish* monitoring. 97% of non-finishers drop out during a two-month grace period (no penalty) in which the firm sends out emails about projected renewal discounts. Afterwards, dropping out results in the maximum amount of renewal surcharge. Our analysis does not account for the costs and benefits associated with this learning process.

offers more active reminders, such as sending text messages, mobile app push notifications, or beeping from the monitoring device when punishable behaviors are records.

Nevertheless, monitoring data is *proprietary*. We verify this by confirming that the Firm's monitoring information appear nowhere in any of its competitors' price filings. In reality, other firms face many practical hurdles in getting and using monitoring information. First, verifying monitoring outcome with consumers alone is hard without heavy manual labor.[26] More importantly, firms may have very different preexisting risk assessment, underlying costs, and markups for serving the same type of consumers.[27] This greatly reduces how other firms can learn about consumer risk with the discount or price charged by our firm.

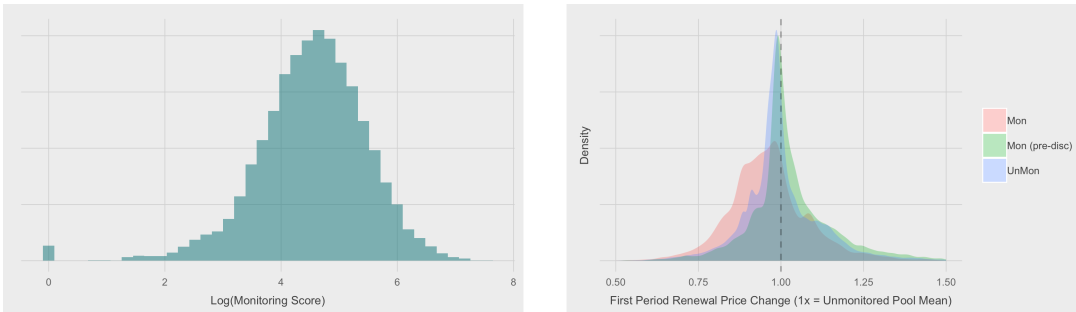
The proprietary nature of monitoring data also prevents us from observing details of competitive monitoring programs. Public filings contain very limited information on these programs; even the monitoring introduction dates often far lags behind the proposed dates in public filings. However, during our research window, monitoring takes up a small fraction of the market in general. In addition, until the second half of 2016, the firm is the only one offering monitoring in all three states in our estimation sample. We therefore do not consider this as a significant factor influencing our empirical results.

Dataset 3 - Monitoring Our data on the firm's monitoring program includes its pricing schedule, drivers' opt-in choices, and realized monitoring scores and renewal discounts for monitored drivers. The firm's monitoring pricing is discussed in Section 5 as well as in Appendix B. Across calendar time and states, the average monitoring finish rates are around $10 - 20\%$ (Figure B.1).

Monitored drivers' performance is summarized by a score, the distribution of which is plotted in Figure 2(a). The more punishable behavior recorded for a given monitored driver, the *higher* her score. We treat this score as the output of the monitoring technology that provides additional information on drivers' *future* accident risk. To see this, Figure 3 plots the average claim count in period two based on monitoring choice and outcome in period one. Compared to unmonitored drivers, those who finished monitoring are 22% safer. Among finishers, the quintile of their monitoring score strongly predicts their second-period risk, which ranges from 60% better

[26]According to the privacy policy agreed upon when opting into monitoring, the firm cannot share personally identifiable data.

[27]Cost differences can come from competitive creaming or claim management.



(a) (b)

Figure 2: Monitoring Score and Renewal Discounts

Notes: (a) plots the density of the (natural) log of monitoring score for all monitoring finishers. The lower the score the better. Drivers that received zero score plugged in the device continuously for enough days but did not drive. We ignore these drivers in all subsequent tests. (b) plots the benchmarked (per firm request) distribution of renewal price change at the first renewal, by monitoring group. 1x represents the average renewal price change factor for the unmonitored group. The one-time monitoring opt-in discount is taken out in order to isolate the renewal discount for monitored drivers. "Mon" and "UnMon" are monitored and unmonitored groups, while "Mon (pre-disc)" is the renewal price change for monitored drivers without the monitoring discount.

Figure 3: Comparison of subsequent claim cost across monitoring groups

Notes: This is a binned-scatter plot comparing average claim count of the first renewal period ($t = 1$, after monitoring ends) across various monitoring groups. The benchmark is the unmonitored pool, which is the "opt-out" group. Group "opt-in" includes all monitored drivers that finished the program per definition in section 1.2. Groups "1" to "5" breaks down the "finish" group based on the quartile of the drivers' monitoring score. Lower monitoring score means better performance.

to 40% worse than the opt-out pool.

Monitoring finishers face the same renewal choices as other drivers, except that their renewal quotes include appropriate monitoring discounts or surcharge. Figure Figure 2(b) compares the distribution of first-renewal pricing change across monitoring groups. We benchmark the baseline price change to center around one. On average, monitored drivers received a 7% discount. Moreover, the monitoring discount is persistent after monitoring ends (Figure A.3). This is consistent with the firm's upfront communication with consumers during their opt-in decision.

2 Reduced-form Evidence

This section documents two reduced-form facts on the degree to which monitoring mitigates incentive and information problems. Drivers that opt into monitoring become safer when they are monitored. Despite this change in behavior, monitoring still reveals previously unobserved risk differences across drivers, which leads to advantageous selection into monitoring.

2.1 Risk Reduction and the Incentive Effect

If monitoring technology is effective, drivers may want to appear safer when monitored.[28] If this incentive effect is important and if drivers' risk is modifiable, then we should expect the *same* drivers to be riskier in unmonitored periods than in the monitored one.

Since monitoring is temporary, we can directly measure this effect by comparing claim outcome for the *same* monitored drivers before and after monitoring ends. This exercise requires us to balance our panel. We focus on the first three periods (18 months).[29] There may be spurious trends in claim rate across periods that are irrelevant to monitoring. We account for this effect with exhaustive observable controls and a difference-in-differences approach. Among monitored drivers, we

[28] This effect is studied in Fama (1980) and Holmström (1999). A similar setting is online tracking of consumers' purchase history (Taylor 2004; Fudenberg and Villas-Boas 2006). If consumers know that buying expensive items online may label them as inelastic shoppers and lead to higher prices in the future, they may refrain from purchasing those items online.

[29] In our robustness check, we show results with only two periods. Attrition is about $10 - 15\%$ per period and our data is right-censored, so balancing the panel eliminates 46% of our data.

take the first difference in claim counts[30] between post-monitoring and monitored periods. This difference is then benchmarked against its counterpart among unmonitored drivers (control group).

$$C_{it} = \alpha + \tau m_i + \omega \mathbf{1}_{post,t} + \theta_{mh} m_i \cdot \mathbf{1}_{post,t} + \mathbf{x}'_{it}\beta + \epsilon_{it} \tag{2}$$

Here, i, t index driver and period in our panel dataset. C denotes claim count, and m_i is a driver-specific indicator for whether i has finished monitoring. x is a rich set of observable characteristics that the firm uses in pricing.[31]

Our main specification includes only monitored drivers who finish monitoring in the first period. To test for parallel trends of the monitored and unmonitored groups, we conduct the same test in subsequent periods after monitoring. In reality, some monitored drivers do not finish monitoring until subsequent periods.[32] To make use of this plausibly exogenous variation in monitoring duration and timing across the first and subsequent periods, we introduce another specification, adding additional variation in relative monitoring duration in the pre-period, z_i. It is calculated as the fraction of days monitored in the first period minus the same fraction in post periods.[33]

Results are reported in Table 2. We find a large and robust incentive effect. Column (3) corresponds to the specification in Equation 2, with the addition of insurance coverage fixed effects.[34] It shows that monitored drivers' average claim count is 0.009 or 23% lower during the monitoring period, compared to after it. Adjusting for the average monitoring duration of first-period monitoring finishers (142 days), a fully-monitored period would be 29.5% less costly to insure for the same driver. Incorporating additional variations in monitoring duration generates similar results (Column (6)). We test for parallel trends between the monitored and unmonitored groups by repeating the baseline specification in subsequent (unmonitored)

---

[30]Throughout our reduced-form analyses, we use claim count as our cost proxy. This is because claim severity is extremely noisy and skewed. This is also common practice in the industry, where many risk-rating algorithms are set to predict risk occurrence only. We therefore present our estimates mostly in percentage comparison terms.

[31]See Table **??** for a list of main observable characteristics. We also include controls for trends and seasonality including third-order polynomials of the calendar year and the month when each driver i starts period t with the firm.

[32]Based on interviews with managers, among finishers, delays in finishing is predominantly caused by device malfunction or delayed start of monitoring due to mailing issues, etc.

[33]As discussed above, some drivers started monitoring but dropped out without finishing. This would bias our results if claims itself leads to non-finish. Out of more than 10,000 claims we observe among monitored drivers, only 13 occurs within 7 days before or after monitoring drop-out. In Table C.1, we further test the robustness of our results by repeating our main analyses on all drivers who started monitoring. This implies larger moral hazard effect adjusting for monitoring duration. However, if some monitored drivers drop out as they discover that they cannot change their risk, the incentive effect estimate would be contaminated by this selection effect.

[34]This soaks up any coverage adjustments between periods.

periods. As shown in Columns (7-10), no differential claim change across periods can be detected between the two groups.

We discuss two important caveats of our results. First, monitoring mitigates moral hazard because it signals drivers' future risk after monitoring as opposed to because it directly rewards effort (Fama 1980; Holmström 1999). The magnitude of risk reduction can be different in the latter setting.[35] On the flip side, our result provides evidence that at least some drivers are forward-looking and respond greatly to future incentives.

Second, our estimate measures a treatment-on-treated effect. If significant heterogeneity in the incentive effect exists across drivers and that it influences consumers' opt-in decision, the effect we find may be larger than the population average (or the average treatment effect) (Einav, Finkelstein, Ryan, Schrimpf, and Cullen 2013), raising external validity concerns in counterfactuals. [36] Our analysis therefore maintains the opt-in structure of the monitoring program and do not extrapolate to scenarios where the market monitoring rate is high.

2.2 Private Risk and the Selection Effect

Are drivers who choose monitoring safer than those who do not? Table 3 reports the results of regressing claim count in the first period ($t = 0$) on monitoring indicator, controlling for the same variables as in Column (3) of 2. The incentive effect only accounts for 64% of the risk differences across the two group. Had the monitored drivers not been monitored in the first semester, they would still be safer than the average unmonitored driver. It thus suggests that drivers possess private information on their own risk. Therefore, there may be strong advantageous selection into monitoring.

Selection into monitoring suggests that the technology is effective at capturing previously unobserved differences in drivers' risk types, further allowing the firm to

[35]We are also unable to disentangle the "Hawthrone effect" from drivers' responsiveness to financial incentives in our estimate. Since consumers must be aware of the data collection to be incentivized for it, we consider this effect as part of the incentive effect.

[36]In equilibrium, the firm assesses the signal monitored drivers send based on future claim records when drivers are no longer monitored, which corresponds to the renewal discount it gives. Therefore, risk reduction is compensated only to the extent to which it correlates with drivers' future risk type. If safer drivers' risk levels are also more responsive to incentives, as suggested by a pure effort cost model for example, selection on the incentive effect can be important. In particular, perfect revelation of a continuum of risk types is possible, as characterized in Mailath (1987), with a monotonicity condition similar to the single-crossing condition. However, consumers likely have multidimensional heterogeneity in reality, so drivers' performance during monitoring may not perfectly reveal their risk types (Frankel and Kartik 2016).

Table 2: **Estimates From Incentive Effect Regression**

explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)	Parallel Trend/Placebo		
	dependent variable: claim count (C)								
constant	0.045^{***}	0.002	0.003	0.046^{***}	0.003	0.004	0.001	0.002	0.002
	(0.000)	(0.005)	(0.005)	(0.000)	(0.005)	(0.005)	(0.006)	(0.006)	(0.008)
post monitoring indicator	-0.001^{*}	-0.003^{***}	-0.003^{***}	-0.001^{**}	-0.003^{***}	-0.003^{***}	0.001^{*}	-0.001	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.002)
monitoring indicator (m)	-0.013^{***}	-0.012^{***}	-0.012^{***}	0.008^{***}	0.005^{***}	0.005^{***}	-0.005^{***}	-0.005^{***}	-0.004^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
monitoring duration (z)				-0.026^{***}	-0.020^{***}	-0.020^{***}			
				(0.002)	(0.002)	(0.002)			
interaction ($\mathbf{1}_{post} \times m$)	0.008^{***}	0.009^{***}	0.009^{***}	-0.005^{**}	-0.005^{***}	-0.005^{***}	-0.001	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)
interaction ($\mathbf{1}_{post} \times z$)				0.015^{***}	0.016^{***}	0.016^{***}			
				(0.002)	(0.002)	(0.002)			
observables controls (x)	N	Y	Y	N	Y	Y	Y	Y	Y
coverage fixed effects	N	N	Y	N	N	Y	Y	Y	Y
implied risk reduction (%)	28.0	29.4	29.5	27.5	29.4	29.6			
pre- / post-periods - "1st diff"	0 / 1-2						1 / 2	2 / 3	3 / 4
treatment / control - "2nd diff"	$t = 0$ finisher / unmonitored			all finishers / unmonitored			$t = 0$ finisher / unmonitored		
number of drivers per period	755,614			809,784			755,614	539,296	397,642

Notes: This table reports results of equation (2). The estimate on the interaction term ($\mathbf{1}_{post} \times m$ or z) measures the "treatment effect" of monitoring ending on claim count across periods. We first balance our panel data to include all drivers who stay till the end of the third semester ($t = 3$). This gives us two renewal semesters ($t \in \{1, 2\}$) after the monitoring semester ($t = 0$). We control for a full set of observables, including driver and vehicle characteristics and tiers (past records of violations or claims). It also includes third-order polynomials of calendar year and month. Continuous observable characteristics are normalized. We report estimates with and without these controls.

Columns (3) and (6) are our main specification. Column (3) focuses on monitored drivers who finished within the first period, while Column (6) introduces additional variation in monitoring duration and timing and looks at all monitoring finishers. Columns (1,2,4,5) show robustness of our estimates to observable and coverage fixed-effect controls. The right-most columns are placebo tests for parallel trends among treatment/control groups after monitoring ends. We first try to detect a similar change from $t = 1$ to $t = 2$. We drop all observations from period 0, and roll the post-period cutoff one period forward, so that $1_{post,t} = 1 \iff t \geq 2$ (changed from $t \geq 1$). Naturally, we look at the future trends of monitored drivers who finished within the first semester and drop other monitored finishers. We find similar results by repeating this test in subsequent periods. As we need to balance panels, number of drivers drop in these tests.

Table 3: First Period Claim Comparison

	Dependent variable:
	Claim Count ($t = 0$)
constant	-0.004^{**}
	(0.009)
monitoring indicator	-0.014^{***}
	(0.001)
observable controls	Y

Notes: This table reports results of a regression where the dependent variable is first period claim count, and the independent variables are the monitoring indicator and observable controls. This is done within all first-period finishers of the monitoring program. This variable is consistent with the monitoring indicator in the incentive effect regression (2) (Table 2), so as to facilitate comparison and decomposition. *p<0.1; **p<0.05; ***p<0.01

dynamically select safer drivers. The following regression examines both factors. It shows how average costs in future (unmonitored) periods vary based on monitoring choice and score among all drivers.

$$C_{it} = \alpha_t + \theta_{m,t} m_i + \theta_{s,t} s_i + \mathbf{x}_{it}' \beta_{\mathbf{t}} + \epsilon_{it} \tag{3}$$

Again, $m = 1$ for monitored drivers who finished within the first period. s denotes monitoring score, which is normalized among monitored drivers and set to 0 for others. The estimates suggest that a monitored driver who scores one standard deviation above the mean has a 29% higher average claim count in the first renewal. Further, controlling for claims does not alter our estimate much. The sparsity of claims therefore greatly limits its informativeness on driver risk in the short run. Figures A.6 and A.7 report $\hat{\theta}_{m,t}$ and $\theta_{s,t}$ for renewal periods $t = 1$ to 5 (three years).

In order to further disentangle selection into monitoring and selective attrition, or to detect selection across various coverage options, structural assumptions are called for. This is because *unilateral* variation in the pricing of monitoring and coverage options is rare. As in Equation 1, any price revision triggers inter-dependent price movements that activate several demand margins at once. Therefore, in the next section, we propose a structural model to jointly account for firm, coverage, and monitoring choices.

3 Cost and Demand Models of Auto Insurance and Monitoring

This section develops a structural model for consumer risk and insurance demand. In the first period, consumers observe their types and make three choices: firm, insurance coverage, and monitoring opt-in. Following this, claims are realized; the monitoring scores for opt-in drivers are revealed to the firm. Consumers are then offered the corresponding renewal price for the second period.

We describe our model in two parts. First, we characterize choice utility conditional on the realization of claim and monitoring score ("realized choice utility"). It features risk aversion, path-dependence (choice inertia and disutility for monitoring), and expectation for future prices. We then describe the data generating processes for claims and for monitoring scores in a cost model that features risk heterogeneity, the incentive effect, and monitoring score's signaling precision. We can then unify cost and demand factors with an expected utility framework to capture selection. We also discuss estimation procedures and sources of identification for key parameters, before demonstrating model fit and validation out-of-sample.

Realized choice utility Besides consumers' risk type, our choice model highlights three factors. (i) Risk aversion governs both preference for insurance and distaste from price fluctuations. (ii) Demand frictions: firm-switching inertia leads to imperfect competition among insurers. Consumers' disutility from being monitored accounts for factors such as privacy or effort cost associated with monitoring. They also sustain partial pooling equilibrium, in which only a fraction of the population is monitored. (iii) Future prices contain most of the benefit of monitoring and depends on claims and monitoring score.

Denote consumers, periods and decision menu options ("plans") by i, t, and d, respectively.[37] Plans, $d = \{f, y, m\}$, consist of firm (f), coverage (y), and monitoring (m) choices. Consumer preferences are characterized by a standard von Neumann-Morgenstern utility function u_{idt} with absolute risk aversion, denoted by γ.

Each driver i starts period t with annual income w_{it} and evaluates insurance choices entirely based on their impacts on his utility through the consumption term h_{idt}, as summarized below.

[37]Monitoring takes place in the first period ($t = 0$).

$$u_{idt}(C,s) = u_\gamma\left(w_{it} + h_{idt}(C,s)\right) \tag{4}$$
$$h_{idt}(C,s) = -p_{idt} - \underbrace{\mathbf{1}_{d,t-1}\cdot\psi_{idt}}_{\text{friction}} - \underbrace{e(C,y_d)}_{\text{oop}} - \underbrace{p_{idt}\cdot R_{idt}(C,s)}_{\text{renewal price}} \tag{5}$$
$$\text{where } \psi_{idt} = \underbrace{\mathbf{1}_{d,t-1}\cdot\eta_0}_{\text{baseline inertia}} + \underbrace{\mathbf{1}_{f_d,t-1}\cdot\eta_{it}}_{\text{firm-switching inertia}} + \underbrace{\mathbf{1}_{m_d}\cdot\mathbf{1}_{t=0}\cdot\xi_{it}}_{\text{monitoring disutility}} \tag{6}$$

Consumption h spans a one-year horizon and consists of two types of components: upfront costs, p and ψ, and stochastic costs, $e(C,y)$ and $R(C,s)$.[38]

p_{idt} is the price for plan d at period t. The term ψ_{idt} captures the degree of path-dependence in consumer choice in monetary terms. This includes a cost of overcoming inertia: baseline η_0 that hinders any choice adjustment (indicated by $\mathbf{1}_{d,t-1} = 1$), and a firm-switching inertia η_{it} that deters consumers from exploiting financially lucrative outside options.[39] It also includes disutility from monitored, ξ_{it}, which may reflect unobserved factors such as hassle costs and privacy concerns.[40]

Out-of-pocket expenditures, e, and renewal prices charged for each plan, R_{idt}, depend on the realization of claims C and the monitoring score s. Consumer pay the portion of claims that exceed the plan's coverage limit out-of-pocket. Renewal prices are adjusted by multiplying two factors: a baseline factor $R_{0,idt}(s)$ that may be influenced by monitoring results, and a surcharge for claims, $R_{1,C}$. We model the baseline factor by with a Gamma distribution with shape parameter β_R and rate parameter $\alpha_{R,imt(s)}$ that depends on observables and monitoring opt-in.

Claim and monitoring score Claims arrive according to a Poisson distribution. The rate parameter, λ_{imt}, has a time-varying mean $\mu_{\lambda,imt}$ that depends on observables x and on monitoring choice m. It also contains an additive error $\epsilon_{\lambda,i}$ that is individual-specific, persistent over time, and log-normally distributed with spread σ_λ. This error captures unobserved risk differences across consumers. Further,

[38]We assume that consumers are myopic beyond a one-year (two-period) horizon. This is the simplest model that captures the different types of costs and benefits of monitoring programs. In particular, dynamic premium risk (reclassification) is incorporated, as higher uncertainty regarding renewal prices diminishes ex-ante utility. Our model can also be interpreted as approximating a two-period dynamic model with infinite adjustment costs. See Kim, Sudhir, and Uetake (2018) for a fully dynamic approach in estimating linear models with private information and effort provision.

[39]These terms capture imperfect competition that supports the observed attrition rate given price dispersion in the data (**??**). Inertia accounts for the search and switching costs as well as potential brand differentiation (Farrell and Klemperer 2007; Honka 2012; Handel 2013).

[40]Monitoring is a one-time offering and choice for new customers, so ξ can only incur at $t = 0$.

each claim has a stochastic cost ℓ, drawn from an independent Pareto distribution. The monitoring score s is an informative signal of the consumer's risk types. For opt-in drivers, a score is drawn once after the first semester, according to a log-normal distribution with an individual-specific mean $\mu_{s,i}$ and precision σ_s.

At each period t, consumer i chooses d from his feasible choice set D_{it} so as to maximize her expected utility, subject to a random coefficient ζ_{idt} on plans offered by the monitoring firm $f^\star$, and an independently drawn Type-1 extreme value error ε_{idt}.[41] We evaluate utility using a normalized second-order Taylor approximation of vNM utility around income w: [42]

$$d_{it} = \arg\max_{d \in D_{it}} \{v_{idt} + \varepsilon_{idt}\} \tag{7}$$

$$\text{where } v_{idt} = \mathbb{E}_{C,s}\left[u_{idt}(C,s)\right] \tag{8}$$

$$= \mathbb{E}\left[h_{idt}\right] - \frac{\gamma}{2}\mathbb{E}\left[h_{idt}^2\right]. \tag{9}$$

Econometric assumptions and heterogeneity We model individual heterogeneity across consumers in key drivers of choice. Heterogeneity is captured by a vector of driver attributes x_{it} and individual random effects.[43]

Our *demand parameters* Θ_d include risk aversion γ, the type I error variance σ, baseline inertia η_0, linear coefficients on driver attributes for firm-switching inertia, θ_η, and for monitoring disutility, θ_ξ, as well as parameters that characterize (expectation for) renewal pricing $\theta_{\mathbf{R}} = (\theta_{\mathbf{R},\mathbf{0}}, \theta_{\mathbf{R},\mathbf{1}})$:

$$\eta_{it} = (1, x_{it})' \, \theta_\eta$$
$$\xi_{it} = (1, x_{it}, \ln \lambda_{it})' \, \theta_\xi$$
$$\alpha_{R,imt}(s) = \begin{cases} \mathbf{x}_{it}^{R\prime}\theta_{\mathbf{R},\mathbf{0}} & m = 0 \\ \left(\mathbf{x}_{it}^{R}, s\right)'\theta_{\mathbf{R},\mathbf{1}} & m = 1 \end{cases}$$

In order to fully capture selection into monitoring, we allow monitoring disutility to vary based not only on observables but also on unobserved risk λ. Without it,

[41] The random coefficient ζ is drawn according to an independent normal distribution with mean zero and standard deviation σ_ζ.

[42] See Cohen and Einav 2007 and Barseghyan, Molinari, O'Donoghue, and Teitelbaum 2013 for further discussion of this approximation. The key underlying assumption is that third- or higher-order derivatives are negligible.

[43] For each type of parameter, we use a set of driver attributes that is consistent with those used in related actual firm pricing rules: x_{it}, x_{it}^R and x_i^s.

given an observable type of consumers, the propensity to opt-into monitoring is fully determined by the financial rewards (lower accident likelihood and potential monitoring discounts). To the extent that there is unobserved heterogeneity in privacy and hassle costs, or misperception of own risk or of the monitoring program, $\theta_{\xi,\lambda}$ can capture it.

Our *cost parameters* Θ_c include linear coefficients on driver attributes and monitoring status for claim arrival rate, $\theta_\lambda = (\theta_{\lambda,\mathbf{0}}, \theta_{\lambda,m})$, as well as for the monitoring score $\theta_{\mathbf{s}}$. In addition, they include the unobserved risk spread for new and old drivers, $\sigma_{\lambda,\text{new}}$ and $\sigma_{\lambda,\text{old}}$, and the monitoring score precision σ_s, as well as the rate and location parameters of the accident loss Pareto distribution, ℓ_0 and α_ℓ.

For tractability, we abstract away from the structure of effort provision underlying the incentive effect. We assume that the effect is homogeneous across drivers and that it enters risk in a mechanical and additively-separable fashion via $\theta_{\lambda,m}$ as in Equations 10 and 11.[44] Under our specification, the monitoring score is informative of driver risk conditional on observables when (i) $\theta_{s,\lambda} \neq 0$, (2) σ_s is finite, and (3) s is not co-linear with x_i^s.

$$\mu_{\lambda,imt} = (1, x_{it})' \, \theta_{\lambda,\mathbf{0}} + \theta_{\lambda,m} \cdot \mathbf{1}_{m=1} \cdot \mathbf{1}_{t=0} \tag{10}$$

$$\mu_{s,i} = (1, \ln \lambda_i, x_i^s)' \, \theta_{\mathbf{s}} \tag{11}$$

3.1 Estimation

We estimate our model of driver cost and insurance demand using a two-step simulated maximum likelihood procedure.[45] First, we estimate the cost parameters Θ_c using the full dataset of claims and monitoring scores. We then estimate the demand parameters Θ_d using menu options, plan choices, and prices, taking the point estimates of the first stage as data.[46]

The Type-1 extreme value distribution of ε_{idt} implies a mixed-logit structure on

[44] For more careful treatment of moral hazard and risk determination, see Jeziorski, Krasnokutskaya, and Ceccarini (2014).

[45] We adopt the two-step procedure due to computational constraints. This comes at the cost of efficiency in the estimator.

[46] Standard errors for the demand estimates are not currently adjusted for two-step estimation.

plan choice with choice probabilities:

$$\begin{aligned}\Pr(d_{it}|\Theta_i) &= \Pr(\epsilon_{idt} - \epsilon_{id't} > [v_{idt}(\Theta_i) - v_{id't}(\Theta_i)] \;\; \forall d' \neq d \\ &= \frac{\exp\left[v_{idt}(\Theta_i)/\sigma\right]}{\sum_{d'} \exp\left[v_{id't}(\Theta_i)/\sigma\right]}\end{aligned} \tag{12}$$

Our model includes random coefficients that enter utility nonlinearly. Private risk, in particular interacts with various observed monitoring and coverage characteristics (renewal price, out-of-pocket expenditure), as well as unobserved demand parameters (risk aversion and monitoring cost). To account for this, we simulate 50 independent draws of private risk (ϵ_λ) and the zero-mean firm dummy (ζ) for every proposal of Θ_d.[47] We then compute the likelihood for observed choices, claim count and severity, monitoring score, and renewal price change and average over the simulated draws.[48]

3.2 Identification

We now provide an informal discussion of the variation in our data that allows us to identify the parameters of our model.

For the cost parameters Θ_c, variation in average claim counts and monitoring scores across observable groups help identify the associated slope parameters θ_λ and $\theta_{\mathbf{s}}$. Variation in claims between monitored and unmonitored periods and drivers helps identify $\theta_{\lambda,m}$. Given the claim arrival rate of an observable group, the variance in claim counts may deviate from that implied by the Poisson structure and therefore identify the spread of risk across drivers σ_λ. The same quantities in the data, when conditioned on not only observables but also the monitoring score, help identify σ_s, the precision of the monitoring score signal. The rate parameter characterizing loss severity is identified by observed claim amounts.[49]

Identification of demand parameters Θ_d relies on price and contract space variation. Controlling for the attributes used in firms' pricing rules, the remaining

[47]We test the effect of increasing the number of draws in estimation on a 10,000 sub-sample. The effect of going from 50 to 200 draws is minimal.

[48]The Taylor approximation allows us to derive closed-form solutions for the first two moments of out-of-pocket expenditures and renewal prices. We therefore do not simulate claim losses or monitoring scores within each draw of random coefficients.

[49]The claim amount is capped above by empirical coverage limits. The Pareto distribtion is sufficiently long-tailed so that loss events significantly larger than coverage limits still have non-degenerate support in consumer's expectation.

price variation depends on location and calendar time. Specifically, price changes associated with the Firm's and its competitors' rate revisions (back-end changes in pricing rules) as well as cross-zipcode variation that are plausibly exogenous from consumer demand.[50] Notably, the Firm altered monitoring opt-in discount over time, generating a useful source of variation in monitoring incentives.

We also observe variation in consumers' contract space. Specifically, monitoring eligibility differs based on state, time, specific vehicle models, and renewal period. For instance, drivers arriving before monitoring introduction in their states or with vehicles older than 1995 are not eligible. Monitoring is also only available to new customers. Meanwhile, mandatory minimum coverage changed in two states within our research window. We use one in our demand estimation and reserve the other for cross-validation (see Table 4).

Our primary concern is in identifying monitoring disutility (ξ) well. Given cost parameters and risk aversion, we can determine the relative attractiveness of the same coverage option with and without monitoring based on objective financial risk and rewards alone. On top of that, the monitoring disutility is pinned down by the actual monitoring share (under various pricing environments). The slope parameter on risk type ($\theta_{\xi,\lambda}$) controls the share of each risk type opting into monitoring. It therefore helps us fit both the share of monitoring and selection on risk.[51]

Another parameter of interest is risk aversion γ. For a given i, t, different γ values imply different gradient of Δv_{idt} across the multiple coverage options we observe in the data.[52] Therefore, conditional on risk parameters, risk aversion can be identified by how the empirical coverage share changes given contract space and pricing environment.[53] In our demand estimation, the Pareto severity parameters can also affect changes in coverage attractiveness. However, we restrict the Pareto distribution to approximate the actual (truncated) claim severity that we observe.

We also need to separately identify baseline inertia (η_0) and consumers' firm-switching

[50] To hone in on this variation, our model include each consumers' assigned risk class in the cost model, and include controls for yearly trends, seasonality, and zipcode characteristics like income and population density in our demand parameters.

[51] Simply raising baseline monitoring cost for all risk types (conditional on observables) enhances selection but also necessarily reduces monitoring share.

[52] This is conditional on the fixed effect for the mandatory minimum plan (ψ_1). The fixed effect adds an additional degree of freedom to more flexibly fit the gradient of willingness-to-pay across coverage options.

[53] Specifically, based on the company's pricing rule in Equation 1, the price gradient across coverage options only depends on the actuarial risk class assigned to each consumer and the coverage factor. The latter is heavily regulated. Each state offers an official guidance on the coverage options that auto insurers should offer and the corresponding coverage factors. Firms need to provide actuarial support to deviate from the guidance in order to avoid regulatory scrutiny. Empirically, coverage factor is rarely changed in our demand estimation states based on rate revision filings.

inertia (η). Conditional on observables, different levels of these parameters imply unique combinations of the share of drivers who adjust coverage versus leaving the firm at renewals. We also observe rich variation in competitive pricing environments conditional on observables. Under a given pricing environment, these parameters imply a corresponding threshold under which drivers would stay with the firm, and another one under which drivers would not adjust choices at all.

3.3 Fit and Validation

We demonstrate that our demand model is flexible enough to produce accurate fit for four critical moments of the data in Figure 4 and in Table A.5. As Table **??** shows, we match monitoring and coverage shares of the Firm well. Further, first-renewal attrition rates, the share of outside option, is also broadly consistent. More importantly, we also accurately fit the expected monitoring score. This demonstrates that the model is capable of capturing selection as well as the effectiveness of the monitoring score. Figure 4 confirms this graphically: we calculate the expected monitoring score for each driver over all random-coefficient draws. The red line plots the simulated score weighted by the corresponding monitoring choice probability in each draw. The orange line plots the full distribution of expected monitoring scores, had everyone in the data finished monitoring.

Using these estimates, we can calculate the expected unmonitored risk type (no incentive effect) of *monitored* drivers in the first period. Specifically, when we numerically integrate over private risk ϵ_λ, we simply weight it by the choice probability of monitoring. This gives us the expected (unmonitored) risk type in the monitored pool. Vice versa for the unmonitored pool. The selection effect is therefore a ratio between the two. The 21% ratio between the two pools is similar to the 17% back-of-the-envelope calculation we did in the reduced-form section.[54]

We also cross validate our demand estimates. In particular, one state in our dataset increased its mandatory minimum from $30,000 to $50,000. In our demand estimation, we draw from only the pre-change period for this state. The hold-out sample, however, contains all drivers in that state arriving in the post-period. As shown in Table A.6, our model performs well out of sample.

[54]In Tables A.5 and A.6, we compare our model fit and cross validation to a basic model specification that excludes the Firm random coefficients ζ and the private monitoring disutility $\theta_{\xi,\lambda}$.

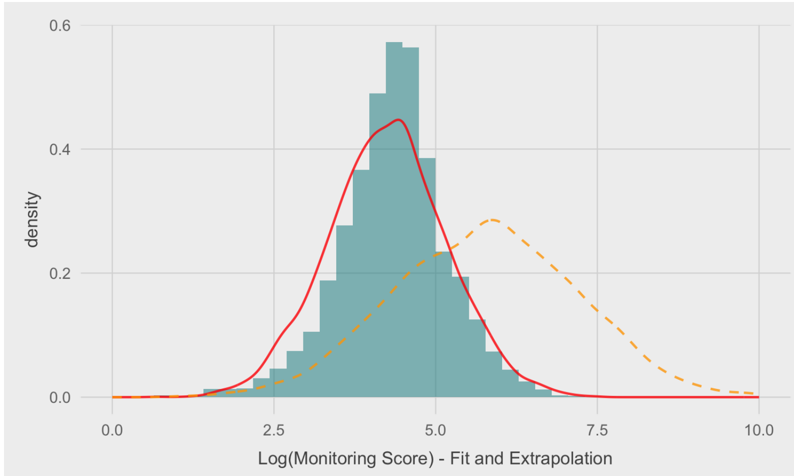


Figure 4: Monitoring Score - Fit and Extrapolation

Notes: The green histogram is the empirical distribution of monitoring score for monitoring finishers in our demand estimation data. The red line plots the fitted distribution as outlined above. The orange dotted line plots the density of the extrapolated distribution of monitoring scores had all drivers finished monitoring.

Table 4: Demand Model Fit and Cross Validation

	Model Fit		**Cross Validation**	
	Fit	**Data**	**Prediction**	**Hold-out Data**
Monitoring share (when eligible)	15.6%	15.3%	17.9%	17.6%
Expected score	4.25	4.30	3.97	4.17
Coverage share				
30K	12.5%	12.7%	-	-
40K	8.2%	8.5%	7.6%	7.2%
50K	49.8%	47.1%	60.5%	58.1%
100K	15.4%	17.0%	17.5%	19.6%
300K	11.9%	12.3%	10.9%	12.8%
500K	2.3%	2.4%	3.6%	2.4%
First renewal attrition	15.6%	15.2%	15.4%	14.7%

Notes: This table reports the fit of our demand model and cross validation results. Our demand estimation data pools across three states with different mandatory minimum. One state changed mandatory minimum from 30K to 50K; estimation data is drawn from only the pre-period of that state to capture monitoring introduction. First renewal attrition rate is benchmarked to data per the firm's request (reporting percent differences, not percentage point differences).

4 Estimation Results and Welfare Calculations

The raw estimates of our models are reported in Tables A.3 to A.2. In this section, we highlight some key results and provide intuition. In particular, we use a simulation exercise to demonstrate the relative importance of different demand factors. We also conduct welfare calculations. Importantly, all simulation exercises in this section hold observed prices as fixed.

The magnitude of private risk and the monitoring score's signal precision are presented in the left panel of Table A.2. Compared to Cohen and Einav (2007), we find significantly more unobserved heterogeneity in driving.[55] This can be attributed to our ability to capture information contained in an additional signal of private risk – the monitoring score. New drivers who do not have past claim records see particularly high spread of private risk. our estimates also capture the monitoring technology and the firm's renewal prices well. In particular, monitoring score rises with driver risk, as do renewal prices for monitored drivers (Table A.4).

We find that drivers are not risk averse in their auto insurance and monitoring choices. Our primary specification assumes homogeneous risk aversion, and the estimate of $\hat{\gamma} = 9.8 \times 10^{-5}$ is broadly consistent with the literature.[56]

Also consistent with prior literature, demand frictions are empirically important. This implies that many drivers who can benefit from monitoring do not participate. In Table 5, we show the empirical distribution of both firm-switching and monitoring costs in the population. The average driver foregoes $283 of gain by not choosing an outside option from other firms, which is 36% of annual premium (two periods). Monitoring cost is also large and is heterogeneous across drivers. In particular, the average driver needs to expect a gain of $93 to participate in monitoring.

Moreover, monitoring disutility increases with private risk.[57] This further accelerates advantageous selection into monitoring, while suggesting that observed renewal prices alone are not enough to explain the empirical selection pattern. At the same time, we see that older and more educated drivers tend to have lower

[55] Our private risk spread is 0.43 ($\exp(\ln \sigma_\lambda)$) for non-new drivers, compared to Cohen and Einav (2007)'s estimate of 0.15.

[56] Figure A.8 benchmarks our risk-aversion parameter against the literature. In the graph, risk aversion is interpreted as the indifference value between inaction and taking a 50-50 bet on gaining $1000 versus losing that value. Barseghyan, Molinari, O'Donoghue, and Teitelbaum (2013), in particular, differentiate between probability distortion (wrong belief about one's own risk) and risk aversion.

[57] Column (2) of table A.3 in the appendix reports the slope parameter for private risk.

monitoring costs, as well as those with newer cars, better prior insurance records and less traffic violation points.

Table 5: Latent Parameters

Statistic	Mean	Std. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
firm-switching inertia (η_{it}, \$)	284	35	158	265	286	307	407
(% annual premium)	36	5	20	34	37	39	52
monitoring disutility (ξ_{it}, \$)	93	19	10	80	93	105	187
(% annual premium)	12	2	1	10	12	14	24
claim risk (λ_{it})	0.05	0.05	0.001	0.02	0.03	0.06	1.48

Notes: This table reports the distribution of heterogeneous latent parameters in our dataset. We simulate a distribution of private risk and calculate these parameters based on our demand estimates.

Looking at the right panel of Table A.2, the fixed inertia cost that drivers need to overcome when adjusting coverages is $134. This adds to firm-switching and monitoring costs and further prevents safe drivers from being monitored. All else equal, the average driver only prefers the mandatory minimum coverage by $26, which seems low given that the plan commands almost 50% market share. This suggests that the rational amount of coverage for many drivers may be below the mandatory minimum, which restricts how monitoring can affect allocative changes across coverage. Appendix G calculate counterfactual demand pattern and Firm profit of removing the incentive effect, reclassification risk, firm-switching inertia, and monitoring disutility from consumer demand.

4.1 Fixed-price Counterfactuals and Welfare Calculations

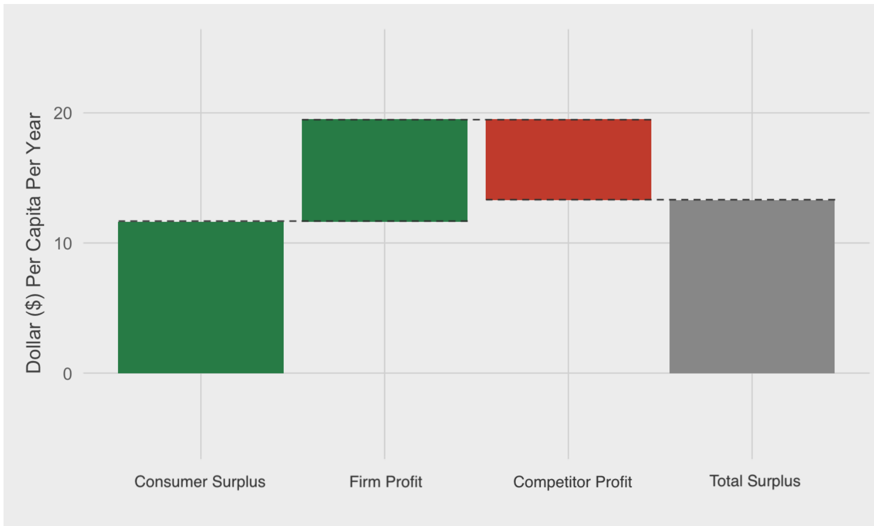
In this section, we simulate a counterfactual scenario in which monitoring was never introduced in order to calculate the profit and welfare impact of introducing monitoring. We observe the marginal cost of monitoring. Prices are held fixed here, and study equilibrium implication in the next section.

We detail our simulation methodology in appendix F. We first enumerate a market, maintaining a *no-brand-differentiation* assumption. This step makes use of our demand model and observed competitor prices. In counterfactual scenarios, we calculate consumers' first-period choice probability for the Firm, for insurance coverage, and for monitoring opt-in. In doing so, we obtain the annual (ex-ante) consumer welfare because the utility horizon is over two periods. We also get the Firm's first-period profit. Next, we simulate claim and monitoring score realizations, pinning down the Firm's second-period information set about consumers and the renewal prices charged. We then obtain second-period choice probabilities and therefore the annual profit of the Firm.

Welfare calculation We evaluate the welfare and total surplus of introducing monitoring by comparing the current monitoring regime to a simulated counterfactual where no monitoring is offered. As mentioned above, we take a certainty equivalent approach in calculating ex-ante welfare. Total surplus is the difference between welfare and total expected cost over two periods. Profits are given by observed prices (and renewal pricing parameters) minus the same expected cost. We also take into account the resource cost for the firm to administer monitoring. It is unobserved and is difficult to estimate since actual prices may be suboptimal. In our simulations, the resource cost is set at $35 per monitored period, based on interviews with the program manager and on industry estimates. This includes manufacturing, wireless data transmission, depreciation, inventory, and mailing costs as well as R&D, marketing, and other overheads.

Figure 5 plots the results in per-capita per-year terms. The average consumer gains $11.6 in certainty equivalent, or 1.5% of premium. Profit increases by $7.9 per capita, a 23.6% increase. Under our symmetric cost and no-brand-preference assumptions, competitors see a profit decline of $6.2. This isolates the impact of cream skimming by the monitoring firm because the firm can offer lower prices to some monitored drivers despite charging higher markups. The combined total surplus increases by $13.3 (1.7% of premium) over the no-monitoring scenario.

To disentangle the welfare consequence of the incentive effect (risk reduction) and allocative changes from mechanical monetary transfers across drivers, we first redo the welfare calculation without the incentive effect. Consumers' expected utility from monitoring and firms' expected cost for monitored drivers will both suffer, reducing the total surplus to $4.8 per capita. The top panel of Figure 6 plots this effect. This attributes almost 64% of total surplus gain to better driving, implying small allocative efficiency gains. To investigate this further, we look at changes in the quantity of insurance purchased, comparing the observed regime with the no-



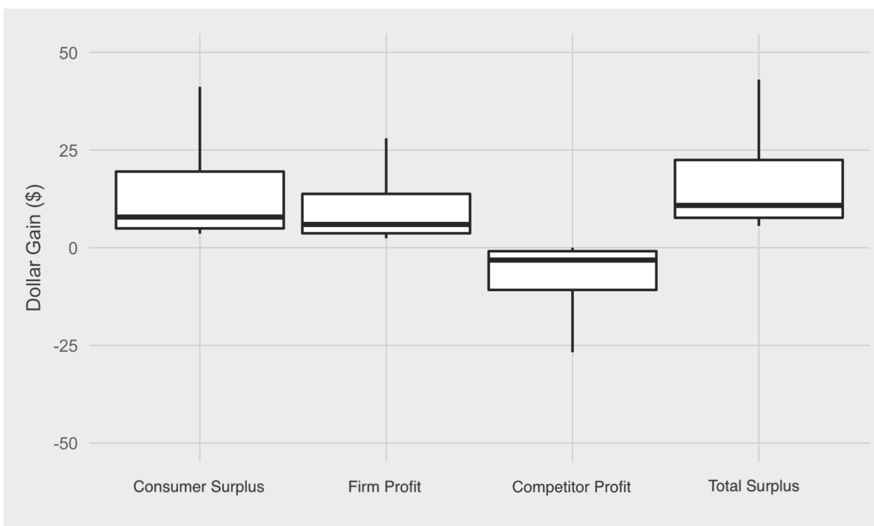


Figure 5: Welfare Calculations

Notes: These figures plot results from our welfare exercise outlined in Section 4.1. The amount denotes the change moving from a regime where no monitoring is offered to one we observe in the data. We plot the differences in ex-ante certainty equivalent, expected profit (across two-periods) for both the monitoring firm and its competitors, as well as total surplus (welfare minus expected cost). The top graph is a waterfall graph decomposing how the components of total surplus changes. The color green indicates an increase while red indicates a decrease. The box plot show 10/25/50/75/90 percentiles.

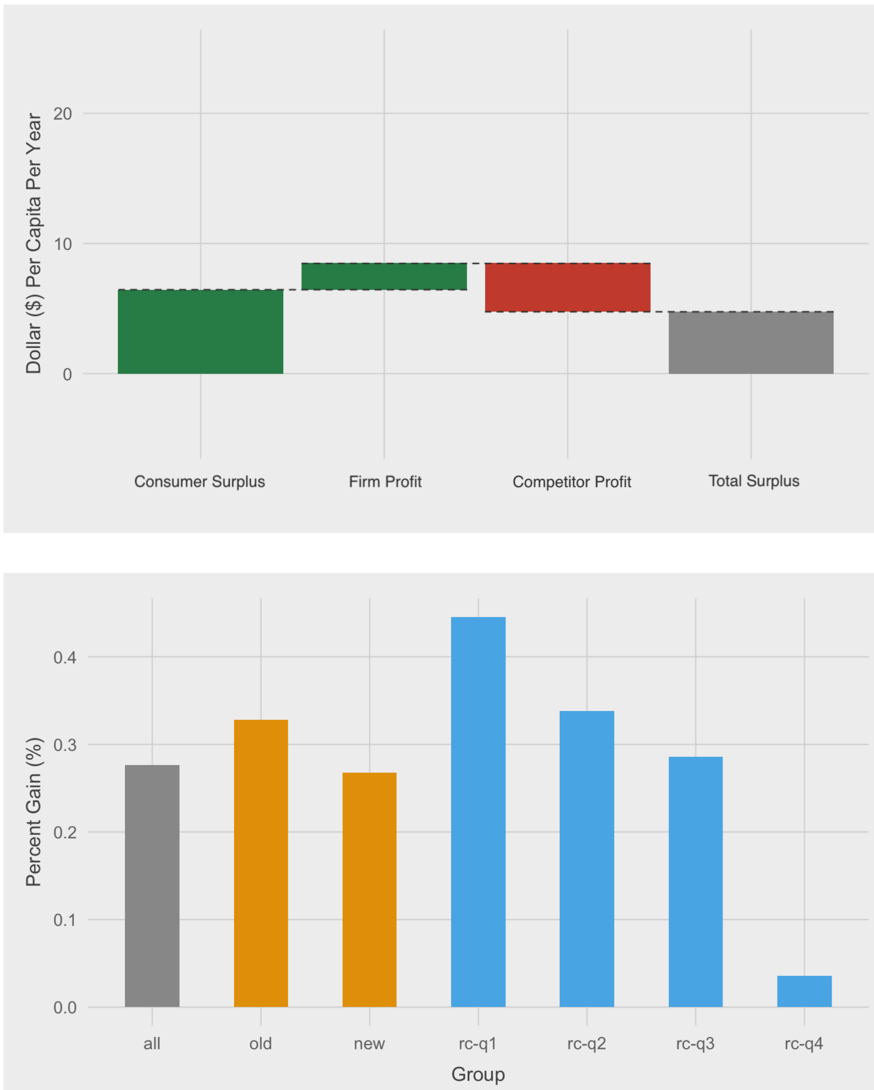


Figure 6: Incentive Effect and Coverage Reallocation

Notes: The top figure plots the same welfare calculation assuming away risk reduction during monitoring based on the incentive effect, per our discussion in the main text. The bottom figure plots average change in coverage amount in percentage across observable groups. "rc-q1" means risk class being in the first quartile at time of choice.

monitoring one. Because liability insurance is mandatory, the result we find here is entirely due to changes in coverage levels. Overall, insurance coverage increases, but only by 0.28%. Looking across various observable pools, the safer risk classes stand out despite the fact that they already pay lower premiums. Meanwhile, without risk reduction, overall profit in the industry falls as the monitoring firm offers lower prices to good monitored drivers at the expense of its competitors' profit.

Importantly, our simulation in this section do not consider how the introduction of monitoring may have changed baseline firm prices for unmonitored drivers. This is because, as shown in Appendix B and Figure B.2, the firm did not raise prices on the unmonitored pool during the introduction of monitoring. Therefore, any cream-skimming effect in our simulation would reduce profit in the unmonitored pool as opposed to reduce welfare of unmonitored drivers. In the next section, we propose a model for pricing where the firm can freely surcharge unmonitored drivers.

5 Monitoring Pricing and Equilibrium Implications

In this section, we propose a two-period two-product model of firm pricing that jointly considers the Firm's ex-ante decision to produce monitoring data and its ex-post decision to extract rent from the data. The model endogenizes the Firm's information set in counterfactual simulations. We use it to first determine optimal pricing of the monitoring program without constraints, given observed resource costs and holding competitor prices fixed. This highlights an "invest-and-harvest" pricing dynamic. Next, we allow competitor prices to respond and simulate an equilibrium in which the firm must disclose monitoring data to competitors. This strips the firm of its property rights over the monitoring data, corresponding to real-world regulations that aim to curb ex-post markups by restricting proprietary data.

5.1 Firm Pricing

In our data, the Firm uses two pricing levers for the monitoring program. First, it uses upfront discounts to encourage monitoring opt-in. Second, it uses non-uniform markups in the form of monitoring discounts.[58] These discounts are the levers by which the firm induces the selection and incentive effects discussed earlier in the paper.[59]

We now specify the Firm's action and profit (payoff) function. In order to highlight the pricing dynamics associated with the monitoring program, we focus on differential pricing based on monitoring opt-in choice ($p_{m,0}$), before monitoring takes place, and the ex-post monitoring score ($p_{m,1}$), after additional risk information is revealed for opt-in drivers. For simplicity, we assume that the Firm maximizes profit over a two-period horizon and chooses a vector of pricing adjustments $\vec{\kappa}$ to form $\mathbf{p_m} = \{p_{m,0}, p_{m,1}\}$.

$$\Pi(\vec{\kappa}) = \sum_i \int_\lambda \Bigg\{ \sum_m \underbrace{\Pr\left(f^\star, m \middle| \lambda, \mathbf{p_m}, \mathbf{p}_{-f^\star}; \Theta\right)}_{\text{demand share}} \cdot \Bigg[\underbrace{\left(p_{m,0}(\vec{\kappa}|\mathbb{I}_{0,i}) - c(\lambda, m) - m \cdot c_m\right)}_{\text{markup}} \\ + \delta \cdot \mathbb{E}_{C,s|\lambda} \Bigg[\underbrace{\Pr\left(f^\star \middle| \lambda, p_{m,1}, \mathbf{p}_{-f^\star}; \Theta\right)}_{\text{retention rate}} \cdot \underbrace{\left(p_{m,1}(\vec{\kappa}|\mathbb{I}_{1,m,i}(C,s)) - c(\lambda, 0)\right)}_{\text{retention markup}} \Bigg] \Bigg] \Bigg\} g(\lambda) d\lambda$$

In the first period, the Firm observes characteristics x and monitoring choice m for each potential customer ($\mathbb{I}_{0,i} = \{x_i, m\}$), while its competitors see only x ($\mathbb{I}_{-f^\star,0,i} = \{x_i\}$). Given these variables and the resulting conditional distribution of latent risk types λ, the Firm forms expectations over the consumer's demand for its insurance products, as well the costs to insure and monitor her: $c(\lambda, m)$ and $c_m = 35$, respectively. We allow the Firm to choose any discount κ_1 and any surcharge κ_0 for those that do and do not opt into monitoring, respectively. These are applied on top of a baseline price schedule $p(x)$:[60]

$$p_{m,0}(\kappa_0, \kappa_1 | x, m) = p(x) \cdot \begin{cases} \kappa_0 & m = 0 \\ \kappa_1 & m = 1 \end{cases}$$

[58]We conduct an event study around the introduction of monitoring to show that the firm did not raise prices for the unmonitored pool. Meanwhile, we show that the retention elasticity drops as the firm gives more discounts. See Appendix B for more details.

[59]The prices ultimately charged to monitored drivers in our data may not fully reflect profit maximization by the firm, largely because prices are heavily regulated in the insurance industry.

[60]As the firm's complete pricing rule is very complex with price filings that often span thousands of pages, we start from the Firm's existing price rule $p(x)$ for each set of observables x observed in the data, and parameterize price discrimination between the monitored and unmonitored pools through multiplicative adjustments κ_m to this price.

κ_0 and κ_1 influence profits in three ways, *ceteris paribus*. First, they scale premiums directly in the first period. Second, they change the competitive share of all plans offered by the Firm. Finally, they change the relative attractiveness of monitoring among the Firm's plans, thereby nudging drivers to opt-in and improving the Firm's information set in the second period. In this way, $\{\kappa_0, \kappa_1\}$ constitute an "investment" in the production of information for the Firm.

In the second period, the Firm gains an informational advantage over its competitors for all monitored consumers: $\mathbb{I}_{1,m,i}(C,s) = \{x_i, C, m \cdot s\}$, $\mathbb{I}_{-f^\star,1,i}(C) = \{x_i, C\}$. The monitoring score s affords more precise estimate of the cost to insure each driver. Thus, for a monitored driver who is, say, 30% safer than previously expected, the Firm may be able to offer a discount that is much smaller than 30% and still be confident that the driver would not leave for a competitor.

As discussed in Section 3 (Equation 5), the renewal price offered to a driver with observables x is given by first period price multiplied by factor $R(C,s) = R^C_{1,C} \cdot R_{0,idt}(s)$, where $R(C,0)$ represent the factor for unmonitored consumers. The wedge between $R(C,s)$ and $R(C,0)$ constitutes the amount of *rent-sharing* between the Firm and the monitored driver that is observed in the data. This is extent to which the Firm "harvests" the value of the monitoring data that is collected. We model the optimal level of rent-sharing by the choice of a parameter κ_s that adjusts the existing rent-sharing schedule linearly.

$$p_{m,1}(\vec{\kappa}|x_i, C, s) = p(x_i) \cdot R^C_{1,C} \cdot \begin{cases} 1 & m = 0 \\ \left[1 - \kappa_s \cdot (1 - R_{0,idt(s)})\right] & m = 1 \end{cases}$$

If $\kappa_s = 0$, then the Firm keeps all the rent: performance in monitoring has no bearing on renewal pricing. On the other hand, if $\kappa_s > 1$, then the Firm shares more rent with consumers than it does in the current regime.

5.2 Equilibrium Implication

Optimal pricing We find the optimal pricing rule $\{\kappa_0, \kappa_1, \kappa_s\}$ that maximizes the Firm's two-period profit under the demand and cost estimates in Section 4. Details on the procedures to compute profit under each pricing regime are outlined in Appendix F. Our results show that in the first period, the Firm should optimally surcharge the unmonitored pool by 2.7%, while offering a 22.1% upfront discount for opting into monitoring.[61]

[61] Consistent with our model, this discount is given for all drivers that *finish* monitoring.

Without competition, our model contains no outside option for consumers other than the Firm's plans. This corresponds to the mandatory nature of auto liability insurance. It also means that the Firm can arbitrarily surcharge unmonitored drivers (κ_0) to force them into monitoring. By contrast, the optimal pricing only includes a modest κ_0 surcharge of 2.7%. Price competition in the (insurance) product market therefore significantly limits the Firm's ability to coerce drivers into monitoring and to extract excessive rent. Instead, a large monitoring opt-in discount suggests that the Firm can benefit from more "investment" to elicit monitoring data, which not only enhances the Firm's ex-post competitive advantage, but also directly reduces the cost to insure drivers in the first period.

In the renewal period, we find that optimal pricing implies 19.6% less rent-sharing than observed in the data, offering a smaller discount for good drivers and a smaller surcharge for bad ones. This implies that the Firm should engage in more aggressive price discrimination conditional on risk. Within the monitored driver pool, safe ones receive a discount only from the Firm and are therefore less prone to attrition. Surcharged drivers can avoid the surcharge by switching to a competitor and are therefore more price-sensitive. This pattern is documented descriptively in Appendix B.

Overall, the monitoring opt-in rate increases to 4.4% (unconditional on coming to the Firm). Consumer welfare and market surplus both increase. Intuitively, although the Firm is taking a larger share of the surplus, it also creates more surplus in the first place by eliciting more monitoring data.

Information sharing Building on the optimal pricing regime, we now endogenize competitor prices to study the impact of a regulation that would have required the Firm to share its monitoring data with competitors. This is especially relevant for information and data regulations due to the *non-rival* nature of monitoring data and close-to-zero marginal cost of replication and distribution.

In a data-sharing regime in which only the Firm controls the monitoring technology, monitoring data becomes a public good. Competitors can poach monitored drivers of the Firm with more attractive rent-sharing schedules. This diminishes the extent to which the Firm can "harvest" the information produced by its (costly) monitoring program. However, it does not fully eradicate the returns from monitoring because (1) significant firm-switching inertia may form an effective barrier against competitive poaching of monitored drivers, and (2) the Firm directly benefits from the risk reduction during monitoring. In this section, we analyze the

Table 6: Counterfactual Equilibrium Simulations

	Current Regime	**Optimal Pricing**	**Data Sharing**
Firm Profit	46.5	61.2	49.3
Competitor Profit	149.2	138.2	147.1
Consumer Welfare (CE)	-	+4.7	+2.2
Total Surplus	-	+8.4	+2.9
Monitoring Market Share	3.0%	4.4%	3.4%
Invest			
Unmonitored surcharge	0.0%	2.7%	1.6%
Opt-in discount	4.6%	22.1%	8.3%
Harvest			
Rent-sharing (κ_s)	1	0.80	1.14
Competitor rent-sharing ($\kappa_{s,-f}$)	-	-	1.81

Notes: This table reports results from our counterfactual equilibrium simulations in Section 5. The simulation procedure to calculate welfare, profits, and total surplus is outlined in Section 4.1. These quantities are reported in dollar per driver per year terms as we translate utility with a certainty equivalent approach. We further enumerate our sample of new customers to the full market by calculating driver weight as in Appendix Section F. The time frame we report is one year (two-period). The level of consumer welfare and total surplus is not identified, so we report only the change in those values in counterfactual regimes compared to the current regime. "Optimal Pricing" represents our equilibrium simulation in Section 5.2. "Data Sharing" represents the equilibrium simulation in Section 5.2, where the monitoring firms is required to share monitoring data to competitors. The "Current Regime" uses monitoring pricing we observe in the data. The rent-sharing parameter (κ_s) is indexed against the one observed in the "Current Regime". Empirically, it is a scalar on top of the firm's existing monitoring renewal schedule. $\kappa_s = 0$ means no rent sharing with consumers (flat pricing schedule regardless of monitoring outcome). $\kappa_s > 1$ means a steeper monitoring discount schedule than observed. This represents more rent-sharing with the consumers. *p<0.1; **p<0.05; ***p<0.01

equilibrium effects of the information-sharing regulation, when competitors cannot offer their own monitoring program, but can adjust their prices in response to the monitoring Firm's pricing and information production.

We make two additional assumptions to facilitate this exercise. First, information sharing is complete and credible. Therefore, firms have symmetric knowledge about the expected cost of monitored drivers, given observables and monitoring scores. Second, competitors have symmetric profit functions and their action space only consists of setting a single competing rent-sharing schedule $\kappa_{-f^{\star},s}$ for monitored drivers. This eliminates baseline price adjustments in order to highlight the competitive poaching motive to "free-ride" on the monitoring information revealed.[62] Similarly, we do not allow for the competitive adoption of monitoring.[63]

Results under competitive equilibrium are presented in Table 6. We find that competitors offer an 81% "steeper" rent-sharing schedule than what the monitoring Firm offers in the current regime. Consequently, the Firm is forced to share more rent with monitored drivers: 14% more compared to the current regime and 43% more compared to the optimal pricing regime. This diminishes the value of "investment" in the monitoring program for the Firm, and so it offers only an 8.3% opt-in discount for monitoring uptake and reduces the surcharge to the unmonitored pool to 0.8%. Overall, as profit is reallocated across firms, consumer welfare and total surplus decrease slightly compared to the equilibrium without the information sharing mandate (the optimal pricing regime). The positive impact of information sharing on curbing ex-post markups is outweighed by the Firm's adjustments to its "investment" in monitoring, lowering participation and social surplus generation. This suggests that the existing levels of price competition and consumer demand frictions already significantly limit the firm's pricing power.

Limitations There are several important limitations to our equilibrium simulations. First, our simplistic pricing framework may not fully capture the firm's action space. The latter can vary nonlinearly and interact with baseline prices in complex ways. We also restrict our simulation to two periods, as we find that the value of monitoring data diminishes dynamically (Figure A.7). Moreover, we

[62] Note that re-optimizing competitors' baseline prices largely captures the effect of our symmetric cost assumption as opposed to competitive response to the pricing of the Firm's monitoring program.

[63] In our setting, competitive adoption can mitigate the benefit of introducing monitoring if competing programs cream skim a large portion of the market. But monitoring is voluntary and monitoring rates are very low in our simulations and empirically during our research window. Therefore, we believe that our results will be robust to competitive adoption of similar monitoring programs.

maintain our assumption of symmetric cost and zero-brand-differentiation across firms due to data constraints. In reality, competitors have different preexisting beliefs about driver risk given observables. Similarly, we hold competitors' baseline prices fixed and do not allow them to adjust in response to the data-sharing regulation. Further, our myopia assumptions hold that different regimes influence consumers' ex-ante expected utility only by altering accident risk and by changing the prices (including renewal prices) that they face at the monitoring Firm. This is because they do not anticipate potential adjustments after the first period in our model. Lastly, firms' profit function do not take into account loading factors (overhead and administrative expenses unrelated to monitoring) on top of claim costs because we cannot separate loading factors from markups charged in our micro data. We may therefore exaggerate the firm's marginal profit.

6 Conclusion

Firms are increasingly collecting consumer data in direct transactions. This influences social surplus and its division in complex ways. Beyond testing for the presence of various economic forces, it is important to quantify the underlying primitives and incentives to understand the formation of information structure and its interaction with prices and with market structure.

In this paper, we acquire novel datasets that give us direct visibility into how valuable proprietary data are collected and used by firms. We also develop an empirical framework that links together the information market in which data transactions occur with the underlying product market. We conclude by revisiting our main results and discussing their real-world implications.

First, data collection changes consumer behavior. Drivers become 30% safer when monitored. This is the primary reason why the monitoring program boosts total surplus in the short run. In general, firms learning about consumers can change consumer incentives and behavior, depending on how consumers perceive their information will be used.[64]

[64] In other settings, consumer behavior may be distorted in a way that harms social surplus. For example, if consumers know that buying expensive items may label them as inelastic shoppers and lead to higher prices in the future, they may delay or refrain from purchasing those items.

Without the incentive effect, the overall allocative efficiency gain from monitoring repricing is small, mostly due to path dependence, the popularity of the mandatory minimum plan, and a lack of unmonitored surcharge. Nonetheless, overall profit decreases with significant profit shift from competitors to the Firm, pushing the market towards the first-best benchmark.

Further, despite strong advantageous selection into monitoring, most drivers who would expect a monitoring discount do not. One reason is large demand friction against monitored, implying inelastic supply firms face when buying consumer data.[65] Another reason is price competition in the product market. Attractive outside options from other insurers limits the firm's ability to coerce drivers into monitoring by raising prices on unmonitored drivers. This highlight the intuition that firms' product market power can spillover to increase their buyer power when eliciting consumer data.[66]

Lastly, our counterfactual simulation demonstrates that, despite the non-rival nature of consumer information, the government should protect the Firm's ownership to the monitoring data in the short run in order to preserve its incentives to produce the data. In the long run, however, markup implications may dominate, in which case the optimal regulation for proprietary data may resemble a patent mechanism.

References

Acquisti, Alessandro, Leslie K John, and George Loewenstein (2012). "The impact of relative standards on the propensity to disclose". In: *Journal of Marketing Research* 49.2, pp. 160–174.

[65] Our data does not allow us to identify the micro foundation of this disutility term. It may include "real" costs like privacy and effort costs. It can also incorporate the effect of salience issues (that goes away in a monitoring mandate) or systematic misconceptions of monitoring's benefit. See Lin (2019) for more careful treatment.

[66] In many online settings, large firms hold significant market power and can afford to make their service contingent upon data collection without losing too many customers. For instance, after the EU's sweeping privacy regulation GDPR went into effect in 2018, the *Wall Street Journal* reports that large firms such as Google and Facebook achieved far higher consent rate for targeted ads than most competing online-ad services (Kostov and Schechner 2018). This further reinforces large firms' competitive advantage. In light of our results, the reason for their high opt-in rates is perhaps not only the value of their services but also the poor outside options consumers face. See Schechner (2018) for the opt-in process used by large multinational firms following the implementation of the GDPR.

Acquisti, Alessandro, Curtis Taylor, and Liad Wagman (2016). “The Economics of Privacy”. In: *Journal of Economic Literature* 54.2, pp. 442–92.
Acquisti, Alessandro and Hal R. Varian (2005). “Conditioning Prices on Purchase History”. In: *Marketing Science* 24.3, pp. 367–381.
Agarwal, Sumit, Souphala Chomsisengphet, Neale Mahoney, and Johannes Stroebel (2015). “Regulating Consumer Financial Products: Evidence from Credit Cards”. In: *The Quarterly Journal of Economics*, pp. 111–164.
Aron Dine, Aviva, Liran Einav, Amy Finkelstein, and Mark Cullen (2015). “Technology Diffusion and Productivity Growth in Health Care”. In: *Review of Economics and Statistics* 97.4, pp. 725–741.
Barseghyan, Levon, Francesca Molinari, Ted O’Donoghue, and Joshua C. Teitelbaum (2013). “The nature of risk preferences: Evidence from insurance choices”. In: *American Economic Review* 103.6, pp. 2499–2529.
Beggs, Alan and Paul Klemperer (1992). “Multi-period competition with switching costs”. In: *Econometrica: Journal of the Econometric Society*, pp. 651–666.
Bonatti, Alessandro and Gonzalo Cisternas (2018). “Consumer scores and price discrimination”. In:
Burtch, Gordon, Anindya Ghose, and Sunil Wattal (2015). “The hidden cost of accommodating crowdfunder privacy preferences”. In: *Management Science* 61.5, pp. 949–962.
Chung, Doug J, Thomas Steenburgh, and K Sudhir (2013). “Do bonuses enhance sales productivity? A dynamic structural analysis of bonus-based compensation plans”. In: *Marketing Science* 33.2, pp. 165–187.
Cohen, Alma (2012). “Asymmetric learning in repeated contracting: An empirical study”. In: *Review of Economics and Statistics* 94.2, pp. 419–432.
Cohen, Alma and Liran Einav (2007). “Estimating Risk Preference from Deductible Choice”. In: *American Economic Review* 97.1994, pp. 745–788.
Cox, Natalie (2017). “The Impact of Risk-Based Pricing in the Student Loan Market: Evidence from Borrower Repayment Decisions”. In: *Working Paper, Princeton University*.
Crawford, Gregory S, Nicola Pavanini, and Fabiano Schivardi (2018). “Asymmetric information and imperfect competition in lending markets”. In: *American Economic Review* 108.7, pp. 1659–1701.
Dranove, David and Ginger Zhe Jin (2010). “Quality disclosure and certification: Theory and practice”. In: *Journal of Economic Literature* 48.4, pp. 935–63.
Dubé, Jean-Pierre, Günter J Hitsch, and Peter E Rossi (2009). “Do switching costs make markets less competitive?” In: *Journal of Marketing research* 46.4, pp. 435–445.

Einav, Liran, Amy Finkelstein, Stephen P. Ryan, Paul Schrimpf, and Mark R. Cullen (2013). "Selection on moral hazard in health insurance". In: *American Economic Review* 103.1, pp. 178–219.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf (2010). "Optimal Mandates and The Welfare Cost of Asymmetric Information". In: *Econometrica* 78.3, pp. 1031–1092.

Einav, Liran, Jonathan Levin, and Mark Jenkins (2012). "Contract Pricing in Consumer Credit Markets". In: *Econometrica* 80.4, pp. 1387–1432.

EUGDPR (2018). "GDPR Key Changes". In: *https://eugdpr.org/the-regulation/*.

Fama, Eugene F (1980). "Agency Problems and the Theory of the Firm". In: *Journal of Political Economy* 88.2, pp. 288–307.

Fang, Hanming, Michael Keane, and Dan Silverman (2008). "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market". In: *Journal of Political Economy* 116.2, pp. 303–350.

Farrell, Joseph and Paul Klemperer (2007). "Chapter 31 Coordination and Lock-In: Competition with Switching Costs and Network Effects". In: *Handbook of Industrial Organization* 3.06, pp. 1967–2072.

Finkelstein, Amy, James Poterba, and Casey Rothschild (2009). "Redistribution by insurance market regulation: Analyzing a ban on gender-based retirement annuities". In: *Journal of Financial Economics* 91.1, pp. 38–58.

Frankel, Alex and Navin Kartik (2016). "Muddled information". In: *Working Paper, University of Chicago, Booth School of Business and Columbia University*.

Fudenberg, Drew and J Miguel Villas-Boas (2006). "Behavior-based price discrimination and customer recognition". In: *Handbook on Economics and Information Systems* 1, pp. 377–436.

Goldfarb, Avi and Catherine Tucker (2011). "Privacy Regulation and Online Advertising". In: *Management Science* 57.1, pp. 57–71.

– (2012). "Shifts in privacy concerns". In: *American Economic Review* 102.3, pp. 349–53.

Handel, Ben, Igal Hendel, and Michael D. Whinston (2015). "Equilibria in Health Exchanges: Adverse Selection versus Reclassification Risk". In: *Econometrica* 83.4, pp. 1261–1313.

Handel, Benjamin R. (2013). "Adverse Selection and Switching Costs in Health Insurance Markets: When Nudging Hurts". In: *American Economic Review* No. 17459, pp. 1–48.

Handel, Benjamin R, Jonathan T Kolstad, and Johannes Spinnewijn (forthcoming). "Information frictions and adverse selection: Policy interventions in health insurance markets". In: *The Review of Economics and Statistics*.

Hart, Oliver (1983). “Optimal labour contracts under asymmetric information: an introduction”. In: *The Review of Economic Studies* 50.1, pp. 3–35.
Hendel, Igal (2017). “Dynamic Selection and Reclassi … cation Risk : Theory and Empirics”. In: *Advances in Economics and Econometrics: Eleventh World Congress*. Vol. 1, p. 99.
Hendel, Igal and Alessandro Lizzeri (2003). “The Role of Commitment in Dynamic Contracts : Evidence from Life Insurance”. In: *The Quarterly Journal of Economics* 118.1, pp. 299–327.
Hendren, Nathaniel (2013). “Private Information and Insurance Rejections”. In: *Econometrica* 81.5, pp. 1713–1762.
Hermalin, Benjamin E and Michael L Katz (2006). “Privacy, Property Rights and Efficiency: The Economics of Privacy as Secrecy”. In: *Quantitative Marketing and Economics* 4, pp. 209–239.
Holmström, Bengt (1999). “Managerial Incentive Problems : A Dynamic Perspective”. In: *Review of Economic Studies* 66.1, pp. 169–182.
Honka, Elisabeth (2012). “Quantifying Search and Switching Costs in the U.S. Auto Insurance Industry”. In: *The RAND Journal of Economics* 45.4, pp. 847–884.
Hubbard, Thomas (2000). “The Demand for Monitoring Technologies: The Case of Trucking”. In: *The Quarterly Journal of Economics* 115.2, pp. 533–560.
Jeziorski, Przemyslaw, Elena Krasnokutskaya, and Olivia Ceccarini (2014). *Adverse Selection and Moral Hazard in the Dynamic Model of Auto Insurance*. Tech. rep. Mimeo.
Jeziorski, Przemysław, Elena Krasnokutskaya, and Olivia Ceccarini (2019). “Skimming from the bottom: Empirical evidence of adverse selection when poaching customers”. In: *Marketing Science*.
Jin, Ginger Zhe and Phillip Leslie (2003). “The effect of information on product quality: Evidence from restaurant hygiene grade cards”. In: *The Quarterly Journal of Economics* 118.2, pp. 409–451.
Jovanovic, Boyan (1982). “Truthful disclosure of information”. In: *The Bell Journal of Economics*, pp. 36–44.
Kim, Minkyung, K Sudhir, and Kosuke Uetake (2018). “A Structural Model of a Multi-tasking Salesforce with Private Information”. In:
Kostov, Nick and Sam Schechner (2018). “Google Emerges as Early Winner From Europe’s New Data Privacy Law”. In: *Wall Street Journals*.
Kummer, Michael and Patrick Schulte (2019). “When private information settles the bill: Money and privacy in Google’s market for smartphone applications”. In: *Management Science*.

Lambrecht, Anja, Katja Seim, and Bernd Skiera (2007). “Does uncertainty matter? Consumer behavior under three-part tariffs”. In: *Marketing Science* 26.5, pp. 698–710.

Lewis, Gregory (2011). “Asymmetric information, adverse selection and online disclosure: The case of eBay motors”. In: *American Economic Review* 101.4, pp. 1535–46.

Lin, Tesary (2019). “Valuing Intrinsic and Instrumental Preferences for Privacy”. In:

Liu, Xiao, Alan Montgomery, and Kannan Srinivasan (2014). “Overhaul overdraft fees: Creating pricing and product design strategies with big data”. In: *Carnegie Mellon University Working paper*.

Mahoney, Neale and E Glen Weyl (2017). “Imperfect Competition in Seleciton Markets”. In: *Review of Economics and Statistics* 99.4, pp. 637–651.

Mailath, George J (1987). “Incentive compatibility in signaling games with a continuum of types”. In: *Econometrica*, pp. 1349–1365.

Milgrom, Paul R (1981). “Good news and bad news: Representation theorems and applications”. In: *The Bell Journal of Economics*, pp. 380–391.

Narayanan, Sridhar, Pradeep K Chintagunta, and Eugenio J Miravete (2007). “The role of self selection, usage uncertainty and learning in the demand for local telephone service”. In: *Quantitative Marketing and economics* 5.1, pp. 1–34.

Nelson, Scott T (2018). “Private Information and Price Regulation in the US Credit Card Market”. In: *Working Paper, University of Chicago, Booth School of Business*, pp. 1–79.

Nevo, Aviv, John L. Turner, and Jonathan W. Williams (2016). “Usage-Based Pricing and Demand for Residential Broadband”. In: *Econometrica* 84.2, pp. 411–443.

NTIA (2018). *NTIA Seeks Comment on New Approach to Consumer Data Privacy*. Tech. rep.

Posner, Richard A. (1978). “The Right of Privacy”. In: *Georgia Law Review* 12.3, p. 393.

Rossi, Peter E, Robert E McCulloch, and Greg M Allenby (1996). “The value of purchase history data in target marketing”. In: *Marketing Science* 15.4, pp. 321–340.

Rothschild, Michael and Joseph Stiglitz (1976). “Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information”. In: *The Quarterly Journal of Economics* 90.4, pp. 629–649.

Schechner, Sam (2018). “Agree to Facebook’s Terms or Don’t Use It”. In: *Wall Street Journals*.

Soleymanian, Miremad, Charles B Weinberg, and Ting Zhu (2019). "Sensor Data and Behavioral Tracking: Does Usage-Based Auto Insurance Benefit Drivers?" In: *Marketing Science* 38.1, pp. 21–43.
Stigler, George J . (1980). "An Introduction to Privacy in Economics and Politics". In: *The Journal of Legal Studies* 9.4, pp. 623–644.
Tadelis, Steven and Florian Zettelmeyer (2015). "Information disclosure as a matching mechanism: Theory and evidence from a field experiment". In: *American Economic Review* 105.2, pp. 886–905.
Taylor, Curtis R. (2004). "Consumer Privacy and the Market for Customer Information". In: *The RAND Journal of Economics* 35.4, p. 631.
Train, Kenneth E. (2009). *Discrete Choice Methods with Simulation*. Vol. 2, pp. 1–370.
Tucker, Catherine E (2012). "The economics of advertising and privacy". In: *International journal of Industrial organization* 30.3, pp. 326–329.
Wei, Yanhao, Pinar Yildirim, Christophe Van den Bulte, and Chrysanthos Dellarocas (2015). "Credit scoring with social network data". In: *Marketing Science* 35.2, pp. 234–258.

A Additional Figures and Tables

Figure A.1: Examples of Telematics Devices in U.S. Auto Insurance

Notes: These are some examples of the in-vehicle telecommunication devices (or "telematics") technology used in monitoring programs in U.S. auto insurance. These devices can be easily installed by plugging them into the on-board diagnostics (OBD) port. The OBD-II specification that these monitoring devices rely on has been mandatory for all cars (passenger cars and light trucks) manufactured or to be sold in the U.S. since 1996.

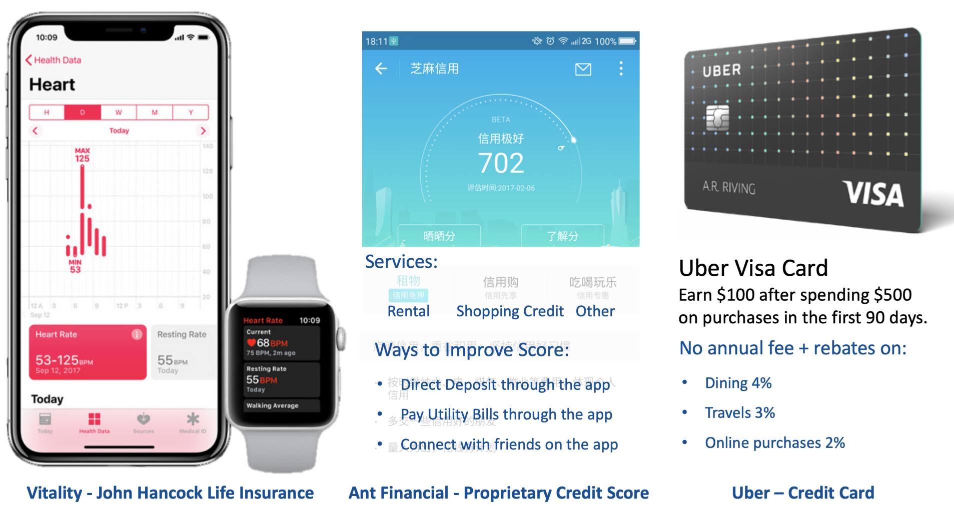


Figure A.2: Other Examples of Direct Transactions of Consumer Data

Notes: Examples of direct transactions of consumer data in other settings. The Vitality program from life insurer John Hancock tracks and rewards exercise and health-related behaviors in exchange for discounts on life insurance premiums. Ant Financial incentivizes users to conduct more personal finance transactions through the platform, such as setting up direct deposit or paying utility bills, in exchange for discounts on various borrowing and rental services. The Uber credit card offers much larger incentives for consumers to use it intensively than the transaction fees charged. One of the plausible business rationales is that the transaction data can be linked back to improve Uber's main businesses in ride sharing and in food delivery.

Table A.1: Summary Statistics on Select Observable Characteristics

Statistic	Mean	St. Dev.	Min	Median	Max
Number of Drivers	1	0	1	1	1
Number of Vehicles	1	0	1	1	1
Calendar month	6.25	3.43	1	6	12
Female Ind.	0.49	0.50	0	0	1
Driver Age	33.42	11.68	15	30	103
Adult Ind.	0.96	0.19	0	1	1
Age <25 Ind.	0.22	0.41	0	0	1
Age <60 Ind.	0.04	0.20	0	0	1
Years of Education	14.46	2.05	9	14	18
College Ind.	0.73	0.44	0	1	1
Post Graduate Ind.	0.41	0.49	0	0	1
Years of License	2.44	1.14	0	3	3
Driver Credit Tier	106	26	0	101	239
Credit Available Ind.	0.96	0.19	0	1	1
Credit Report Ind.	0.83	0.38	0	1	1
Homeowner Ind.	0.17	0.38	0	0	1
Garage Verification Ind.	0.84	0.37	0	1	1
Out-of-state Ind.	0.11	0.32	0	0	1
Population Density Percentile	51	21	0	54	99
Vehicle Model Year	2006	6.05	1928	2007	2018
Vehicle on Lease Ind.	0.51	0.50	0	1	1
Length of Ownership	0.42	0.92	0	0	4
Class C Vehicle indicator	0.89	0.31	0	1	1
ABS Ind.	0.13	0.34	0	0	1
Safe Device Ind.	0.35	0.48	0	0	1
Accident Point	1.53	2.80	0	0	82
At-Fault Accident Count	0.33	0.65	0	0	11
DUI Count	0.05	0.23	0	0	8
Clean Record Ind.	0.64	0.48	0	1	1
Prior Insurance - Some	0.08	0.27	0	0	1
Prior Insurance - Yes	0.57	0.49	0	1	1
Length of Prior Insurance	1.59	1.45	0	2	4
Zipcode AGI ('\$000)	142	162	1	114	100,508

Notes: Our data only consist of single-driver-single-vehicle insurance policies. Years of license data is capped at 3 in compliance with regulations that limit risk rating. Zipcode AGI is merged into the dataset by researchers based on zipcode.

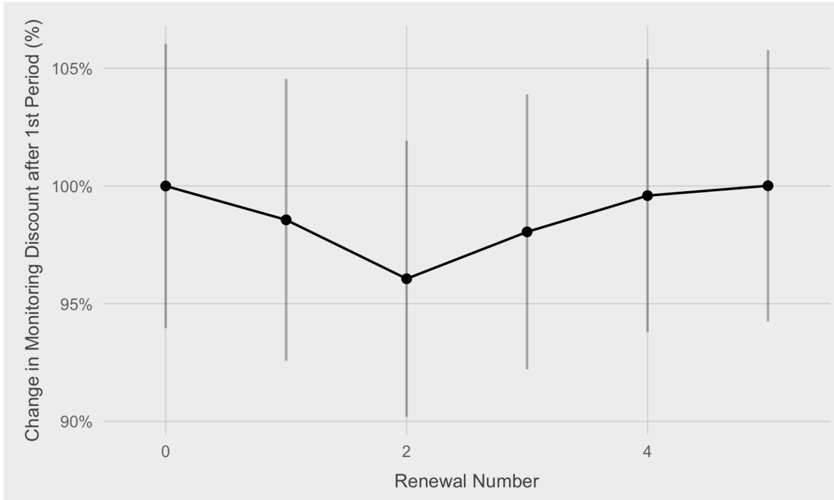


Figure A.3: Persistence of Monitoring Discount

Notes: This graph plots the empirical progression of monitoring discount for all monitoring finishers in one state that stayed with the firm till at least the end of the 5th periods (so we observe monitoring discount in the renewal quote for the 6th period). The benchmark is monitoring discount in the first renewal quote ($t = 0$). Fluctuations and noises are due to ex-post adjustments. Firm may change their discount schedule slightly. Monitored drivers can also report mistakes in their records and have their discount adjusted.

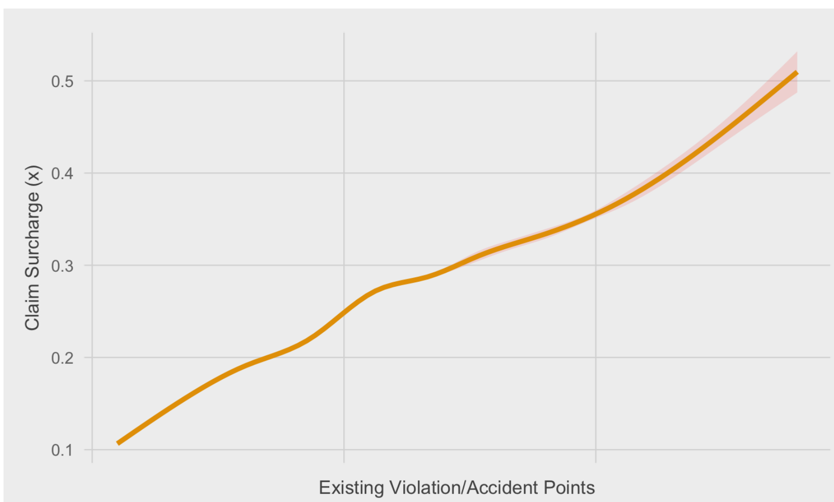


Figure A.4: Renewal Price Claim Surcharge

Notes: This graph plots the empirical claim surcharge function for at-fault accidents. Claim surcharge varies with existing violation points and calendar time. 0.1 means 10% surcharge. This differs from the filed factors because the latter is applied on the base rate only, while this function represents the surcharge percentage on top of overall premium. This is done by regressing renewal price change on violation point last period and current period at-fault claim, controlling for all other observables.

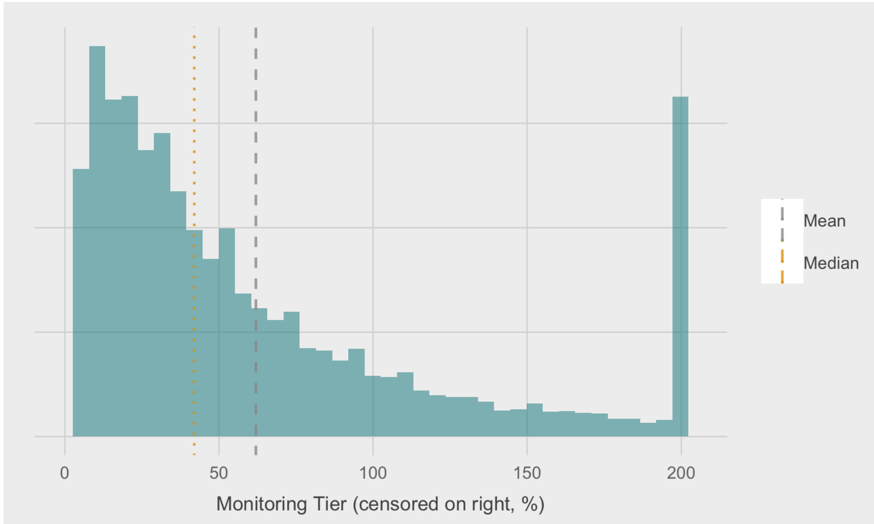


Figure A.5: Distribution of monitoring tier

Notes: This figure plots the empirical density of monitoring tier for all monitored drivers who finished monitoring. It is calculated as the quotient of realized monitoring score over ex-ante expected monitoring score. For monitored driver i, the expected score is derived based on the average driver in i's observable (x_i) group. It does not take into account the fact that i has selected into monitoring. The graph has a long right tail and is truncated at 200%.

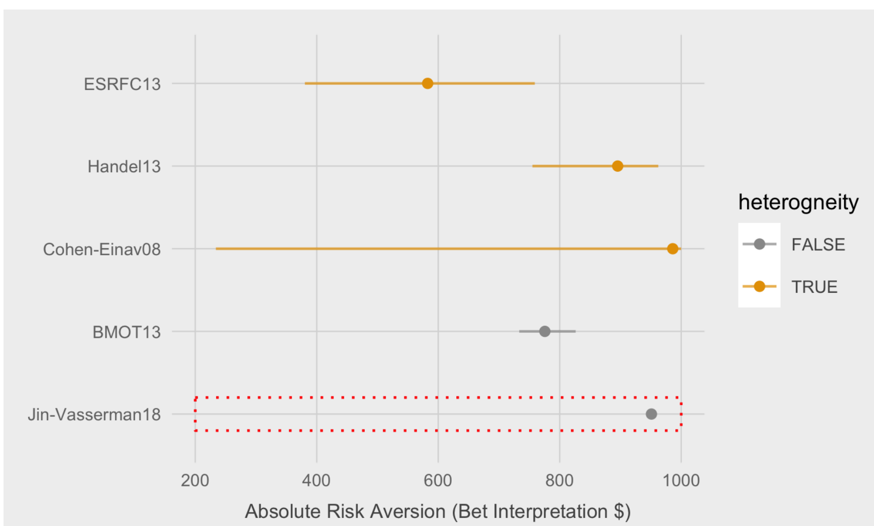


Figure A.8: Risk Aversion Parameter Estimates - Benchmark

Notes: This figure benchmarks our risk aversion parameter estimate to the literature. Heterogeneity indicator means that the author allows risk aversion to vary across people, in which case we plot the range of risk aversion paramters in the population. Otherwise we plot the 95% confidence interval of the homogeneous risk aversion parameter.

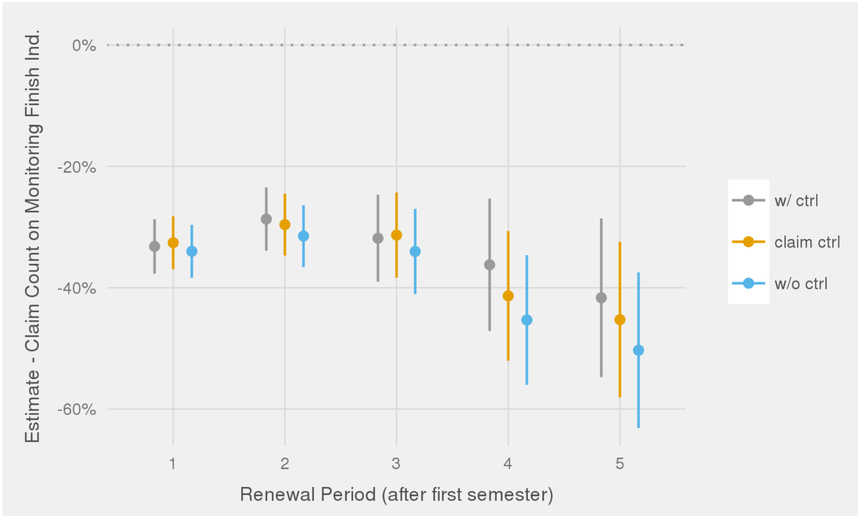


Figure A.6: Estimates - dynamic informativeness of monitoring participation

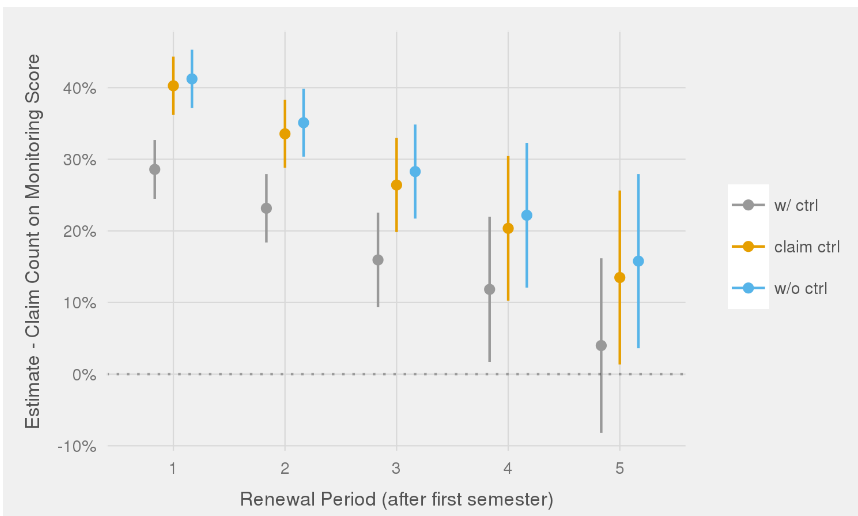


Figure A.7: Estimates - dynamic informativeness of monitoring score

Notes: Figures A.6 and A.7 report the estimate for θ_t and γ_t from regression (3) in percent increase terms. Monitoring participation is an indicator for finishing monitoring. For each $t > 0$, we take all drivers who stayed with the firm till at least the end of period t. θ_t is the coefficient of claim count of driver i in period t on monitoring score of i, and γ_t is that on monitoring finish indicator of i. Monitoring score is normalized, and defaulted as 0 for unmonitored drivers. So θ_t measures the effect of getting a score one standard deviation above the mean during the monitoring period ($t = 0$). γ_t compares unmonitored drivers with the average monitoring finisher. To further translate these effects into percent increase terms, we divide the estimate of θ_t and γ_t by the average claim count in period $t$of all *monitored* drivers. The horizontal axis represents different regressions for different renewal period $t > 0$. Different colors within each t value represent different specifications of control variables (x_{it}). The grey (left-most) series represents estimates from regressions with the full set of x_{it}; the orange (middle) one includes only claim records revealed since $t = 0$; the blue (right) series includes no control.

Table A.2: Estimates: Homogeneous Parameters

Cost		Score & Pricing		Demand	
$\ln \sigma_{\lambda,\text{new driver}}$	-0.266^{***}	$\ln \sigma_s$	-0.081^{***}	$\ln \gamma$	-9.235^{***}
	(0.060)		(0.007)		(0.089)
$\ln \sigma_{\lambda,\text{old driver}}$	-0.840^{***}	$\beta_{R,\text{new}}$	66.953^{***}	η_0	134.262^{***}
	(0.070)		(0.403)		(2.228)
$\ln \alpha_\ell$	-1.480^{***}	$\beta_{R,\text{monitoring}}$	59.680^{***}	σ_ζ	98.989^{***}
	(0.063)		(0.902)		(2.303)
		$\beta_{R,\text{renw}}$	78.571^{***}	σ	39.213^{***}
			(0.315)		(0.632)

Note: This table reports estimates for homogeneous parameters of our structural model. *Cost*: spread of private risk $\sigma_{\lambda,\text{new driver}}$ and $\sigma_{\lambda,\text{old driver}}$ (new drivers are defined as those licensed in the past three years), claim severity Pareto distribution parameters ℓ_0 and α_ℓ (ℓ_0 is set at $3,000 per discussion in the text). *Score and Pricing*: monitoring score's signal precision σ_s, rate parameters for the renewal price change (R_0) Gamma distribution β_R's. *Demand*: absolute risk aversion coefficient γ, baseline inertia η_0 in dollar term, variance of own firm random coefficient σ_ζ, scale of the logit error σ. *p<0.1; **p<0.05; ***p<0.01 *p<0.1; **p<0.05; ***p<0.01

Table A.3: Estimates: Heterogeneous Latent Parameters

	Log Claim Rate (μ_λ)	Monitoring Disutility (ξ/\$)	Firm-switching Inertia (η/\$)
Intercept	-3.294^{***}	96.773^{***}	228.559^{***}
	(0.080)	(2.813)	(6.213)
Private Risk		25.238^{***}	
		(1.657)	
Monitoring Ind.	0.404^{***}		
	(0.063)		
Monitoring Duration	-0.796^{***}		
	(0.081)		
Driver			
Driver Age	-0.240^{***}	-1.049^{**}	4.526^{***}
	(0.053)	(0.437)	(1.641)
– Square	0.156^{***}	-1.047^{***}	3.816^{**}
	(0.055)	(0.309)	(0.742)
Age < 25	0.081^{**}	0.326	-0.500
	(0.032)	(0.339)	(0.922)
Age > 21	-0.064	-0.059	3.195^{***}
	(0.053)	(0.403)	(0.449)
Age > 60	-0.046	-0.139	-0.275
	(0.068)	(1.689)	(0.340)
Year of Education	0.001	-2.452^{***}	-7.526^{***}
	(0.025)	(0.331)	(0.915)
College Ind.	-0.00001	-0.952^{***}	0.234
	(0.038)	(0.339)	(0.237)
Post Grad Ind.	0.005	-0.728	-1.547
	(0.039)	(1.644)	(1.686)
Female Ind.	0.099^{***}	-0.261	1.007
	(0.021)	(1.643)	(1.686)
Driver License Year	-0.018	-0.016	16.776^{***}
	(0.019)	(0.905)	(0.338)
Home Ownership	-0.020	-0.039	0.058
	(0.038)	(0.447)	(1.653)
Out-of-State License	-0.104^{***}	-0.380	-0.406
	(0.030)	(0.339)	(0.922)
Location			
Garage Verified Ind.	-0.069^{*}	0.008	1.847^{**}
	(0.036)	(0.521)	(0.922)
Population Density	0.076^{***}	0.359	-4.902^{***}
	(0.015)	(0.419)	(0.445)
Zipcode Income	-0.058^{***}	0.610	-2.936^{*}
	(0.017)	(1.615)	(1.677)

	μ_λ	$\xi/\$$	$\eta/\$$
Log Zipcode Income	0.031***	0.284	−0.808
	(0.008)	(2.949)	(1.850)
Vehicle			
Length of Ownership	0.017	−0.918	−0.084
	(0.012)	(0.887)	(0.338)
Vehicle on Lease Ind.	0.092***	−1.058	4.789***
	(0.024)	(1.677)	(0.343)
Model Year	−0.026*	−1.621***	3.211***
	(0.014)	(0.421)	(0.445)
ABS Ind.	−0.058*	0.034	−1.626***
	(0.035)	(0.741)	(0.422)
Airbag Ind.	0.014	0.199	1.225
	(0.021)	(1.644)	(1.686)
Class C Ind.	0.023	0.079	3.843**
	(0.053)	(0.448)	(1.655)
Tier			
Credit Report Ind.	0.044	0.414	1.832***
	(0.035)	(0.429)	(0.448)
Delinq. Score*	−0.016	2.114***	10.959***
	(0.014)	(0.331)	(0.917)
Prior Ins. Length	−0.038**	−2.293	−3.993***
	(0.017)	(1.648)	(0.338)
Has Prior Ins.	−0.067*	−1.183***	−0.759*
	(0.035)	(0.427)	(0.448)
– w/ Lapse	−0.050	0.204	0.001
	(0.043)	(1.686)	(0.620)
Violation Points	−0.032	1.084***	4.333***
	(0.030)	(0.337)	(0.429)
Clean Record Ind.	−0.097***	−0.909	−1.392***
	(0.035)	(0.916)	(0.342)
Total Accident Count	0.115***	0.470	−0.139
	(0.029)	(1.638)	(1.690)
Total DUI Count	−0.233***	0.031	0.326
	(0.065)	(0.922)	(0.536)
Log Risk Class	0.275***		
	(0.046)		
Risk Class	0.042		
	(0.074)		
– Square	−0.124*		
	(0.073)		
– Cube	0.0002		
	(0.046)		
Seasonality	0.026**	−0.764**	−1.585***
	(0.011)	(0.331)	(0.427)

	μ_λ	$\xi/\$$	$\eta/\$$
– Square	0.063	-0.364	-0.519
	(0.046)	(0.340)	(0.430)
Trend Year	0.083*	-1.570	7.417***
	(0.043)	(1.660)	(0.338)
– Square	-0.102***	-1.413	6.199***
	(0.039)	(1.830)	(1.674)

Note: This table reports intercept and slope estimates for heterogeneous latent parameters. Continuous covariates are normalized (except λ and monitoring duration). Discrete variables with more than two values are normalized so that the minimum is zero. Deliq. (delinquency) Score is based on records from a credit bureau. Higher scores mean worse records. *p<0.1; **p<0.05; ***p<0.01

Table A.4: Estimates: Renewal Pricing and Monitoring Score

	$\mathbb{E}[R_{0,m=0,t=0}]$	μ_s	$\mathbb{E}[R_{0,m=0,t=1}]$
Intercept	-0.362***	11.367***	-1.131***
	(0.001)	(0.506)	(0.132)
Log Risk Class	-0.413***	-0.384**	-0.080***
	(0.018)	(0.155)	(0.018)
Risk Class	0.367***	-0.077	0.063
	(0.051)	(0.304)	(0.034)
– Square	-0.290***	0.245	-0.155***
	(0.054)	(0.308)	(0.036)
– Cube	-0.229***	-0.039	0.031
	(0.022)	(0.140)	(0.019)
$\ln\lambda$		1.859***	
		(0.094)	
log(Monitoring Score)			0.150***
			(0.005)

Notes: This table reports estimates for the renewal pricing and monitoring score model. Instead of modeling the Gamma shape parameters (α), we use a change-of-variables technique to directly estimate the expected renewal rate. It is modeled with a Sigmoid function between 0.5 (50% cheaper) and 2 (twice as expensive). That is, $\mathbb{E}[R_0] = \sigma(\mathbf{x}'\theta_R) \times 1.5 + 0.5$. We include the appropriate Jacobian adjustments in estimation, and winsorize away extremely large or small renewal price change. *p<0.1; **p<0.05; ***p<0.01

Table A.5: Demand Model Fit

	Basic Specification	**Primary Specification**	**Data**
Monitoring share (when eligible)	17.7%	15.6%	15.3%
Expected score	5.46	4.25	4.30
Selection effect (risk)	6.7%	21.2%	-
Coverage share			
30K	13.7%	12.5%	12.7%
40K	9.1%	8.2%	8.5%
50K	53.2%	49.8%	47.1%
100K	13.0%	15.4%	17.0%
300K	9.3%	11.9%	12.3%
500K	1.8%	2.3%	2.4%
First renewal attrition (indexed)	133.0%	102.9%	100.0%

Notes: This table reports the fit of our demand model as described above. The primary specification is outlined in our econometric model section. Monitoring share is conditional on eligibility. For coverage shares, our demand estimation data pools across three states with different mandatory minimum. One state changed mandatory minimum from 30K to 50K; estimation data is drawn from only the pre-period of that state to capture monitoring introduction. First renewal attrition rate is benchmarked to data per the firm's request (reporting percent differences, not percentage point differences).

Table A.6: Cross Validation

	Basic Specification	**Primary Specification**	**Hold-Out Data**
Monitoring share (when eligible)	21.2%	17.9%	17.6%
Expected score	5.23	3.97	4.17
Selection effect (risk)	5.2%	23.7%	-
Coverage share			
30K	-	-	-
40K	9.4%	7.6%	7.2%
50K	66.3%	60.5%	58.1%
100K	13.4%	17.5%	19.6%
300K	9.7%	10.9%	12.8%
500K	1.3%	3.6%	2.4%
First renewal attrition	132.2%	104.2%	100.0%

Notes: This table reports our cross-validation result. All measures are calculated analogously as Table A.5. For the state that changed mandatory minimum, the hold-out data include all post-period data. For the other two states, the hold-out data include all observations that are not in our demand estimation data.

B Analysis of Actual Firm Pricing

Cream skimming effect Advantageous selection into monitoring may cream skim from the firm's unmonitored pool. As a result, firms may choose to raise prices in the unmonitored pool. In addition, they may also want to surcharge the unmonitored pool to indirectly encourage monitoring participation. To test the effect of monitoring introduction on the unmonitored pool more formally, we take advantage of the staggered introduction of monitoring across states. This gives rise to a regression discontinuity strategy that evaluates how prices and average cost changed in the *unmonitored* pool. We focus on a year before and after monitoring introduction; our observable characteristics also include state fixed effects and flexible controls for trends and seasonality. We only focus on the first semester ($t = 0$) to avoid contamination from attrition[67]. We therefore drop the t subscript, and run the following regression

$$dep.\ var._i = \alpha + \gamma Qtr_i + \kappa \mathbf{1}_{post,i} + \theta \cdot Qtr_i \times \mathbf{1}_{post,i} + \mathbf{x}_i'\beta + \xi_{y,i} + \epsilon_i \tag{13}$$

We use price p_i and claim count C_i as our dependent variable. Qtr is the running variable, which denotes the calendar quarter when driver i arrived at the firm[68]. $\mathbf{1}_{post}$ is an indicator for whether i arrived at the firm after the introduction of monitoring. $\mathbf{x}$ and a coverage fixed effect ξ_y soak up compositional changes in observable risk class and coverage plans. The coefficient θ reveals treatment effect of monitoring introduction on prices and claims in the unmonitored pool.

Estimates for $\hat{\theta}$ across various specifications are reported in figure B.2. The firm did not raise prices around monitoring introduction. We also find no evidence that the average cost of the unmonitored pool deteriorated by more than 2%.

In reality, monitoring is only a small fraction of the market. As our demand estimates will reveal in the next section, even when monitored drivers are significantly better, its influence on the unmonitored pool is significantly limited by its small size. Further, the firm does not make follow-up offers to customers who initially opted out monitoring, which is necessary for unraveling to occur empirically. Lastly, monitoring programs are subject to approval by state commissioners. And a new program that affects baseline pricing may be subject to more regulatory scrutiny. On the flip side, this suggests that the current monitoring regime is largely welfare-neutral for unmonitored drivers.

[67] This regression does not include monitored drivers, so there is no contamination from moral hazard.

[68] It is normalized so that the quarter immediately after monitoring introduction is indexed as 0.

Dynamic and non-uniform pricing Monitored drivers have 35% higher profitability overall, controlling for observables. On top of the risk reduction (during monitoring) and better risk rating, this can also be a result of higher profit margin and retention rate when information is revealed. We provide descriptive evidence on pricing and dynamic retention in this section.

First, the Firm faces a dynamic pricing problem as information is revealed at the end of the first period. It offers a opt-in discount to encourage all drivers to participate in monitoring. This averages to around 5% across states and time.

When monitoring information is revealed, the firm can use it to set non-uniform prices. Here, the firm's pricing schedule is based on a monitoring tier that measures how "surprising" a given driver's monitoring score is to the firm. In figure A.5, we plot the empirical distribution of monitoring tier, which is realized monitoring score divided by firm's expected score given observables[69]. Consistent with our findings above, the average monitored driver performed much better than expected[70].

Figure B.3 presents the discount schedule the firm uses given the percentile of monitoring tier as defined above. Surprisingly good drivers are on the left, who are offered the highest renewal discount, while around 25% of drivers that performed poorly (compared to firm's expectation) received a surcharge.

Figure B.4 plots the corresponding retention rate. It is clear that as discounts approach zero or negative, retention rate drops significantly. In fact, we can regress renewal choice (binary) on prices with monitoring discount, controlling for observables and price level without the discount. θ then measures the slope of the residual (retention) demand.

$$\mathbf{1}_{renew,i} = \alpha + \delta p_i + \theta disc_i + \mathbf{x}_i' \beta + \epsilon_i \tag{14}$$

The estimates for $\hat{\theta}$ are reported in figure B.5. Without monitoring discount, a $\$1$ increase in price (decrease in discount given) causes the retention rate to drop by 0.07 percentage points (7 basis points). When firms give discounts, however, the slope of the demand decreases, and by 56% when the discount given is larger than 10%. This suggests that

[69]For monitored driver i, the expected score is derived based on the average driver in i's observable (x_i) group. It also does not take into account the fact that i has selected into monitoring. The graph has a long right tail and is truncated at 200%.

[70]It is important to note that a driver with a monitoring tier of 30% is not necessarily 70% safer than the average person in her pool, especially in renewal period. This is because monitoring score does not capture risk perfectly, and it is also stochastic. Our structural model quantifies these effects more formally.

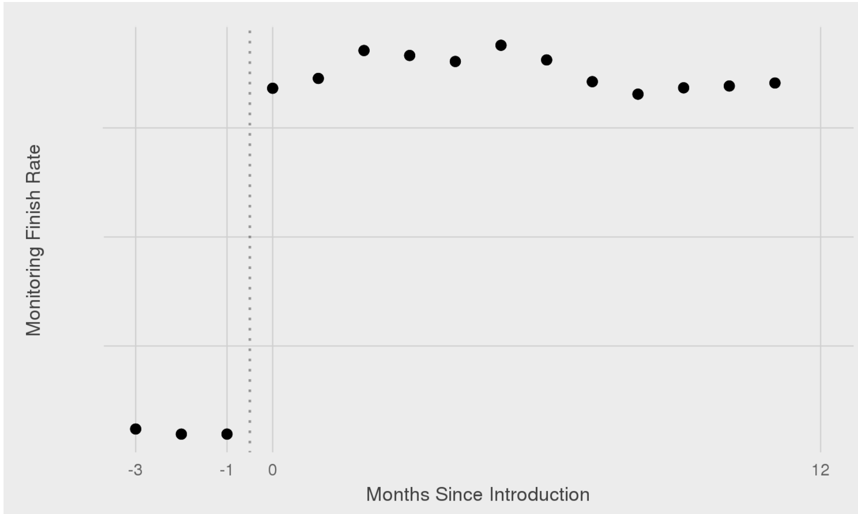


Figure B.1: Monthly monitoring finish rate around monitoring introduction

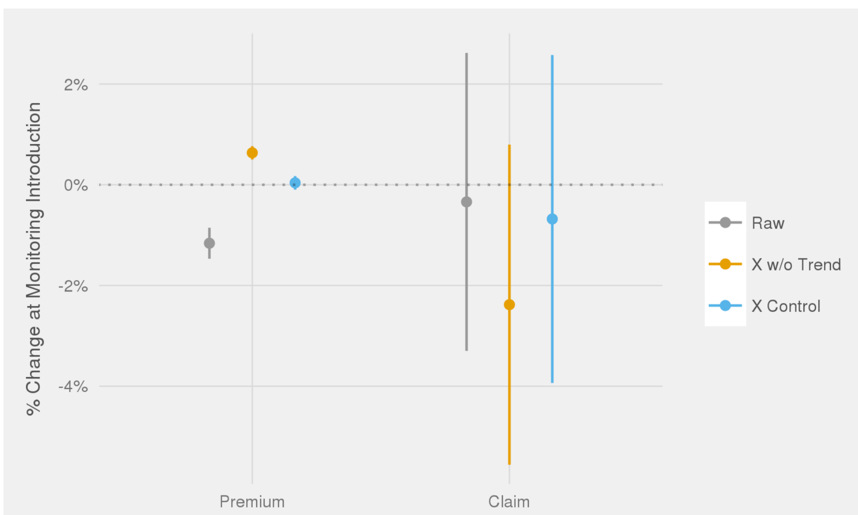


Figure B.2: Event Study: treatment effect of monitoring introduction on the unmonitored pool

Notes: figure B.1 the progression of monthly monitoring finish rate around the introduction of monitoring. The monthly finish rate are below 0.1% in all months before monitoring introduction. The reason why it is not exactly zero before monitoring introduction is due to small-scale trial and experimentation. We throw out states that introduced monitoring in the first three months or the last 12 months of our research window. This ensures that the trend we see do not pick up changes in state composition.
figure B.2 reports regression-discontinuity estimate θ of equation (13), where the horizontal axis distinguishes dependent variable used. These effects are translated in percentage terms by dividing the average of the dependent variable in the period immediately before monitoring introduction. We look at only first period outcomes, and include all *unmonitored* drivers arriving at the firm a year before or after the firm. States that introduced monitoring within a year after the beginning or a year before the end of our research window are excluded. The running variable is quarter since monitoring introduction. Different colors and positions represent different specifications of control variables (x_{it}). The grey (left-most) series represents estimates from regressions with the full set of x_{it}; the orange (middle) one includes a full set of observables, including flexible controls for trend and seasonality.

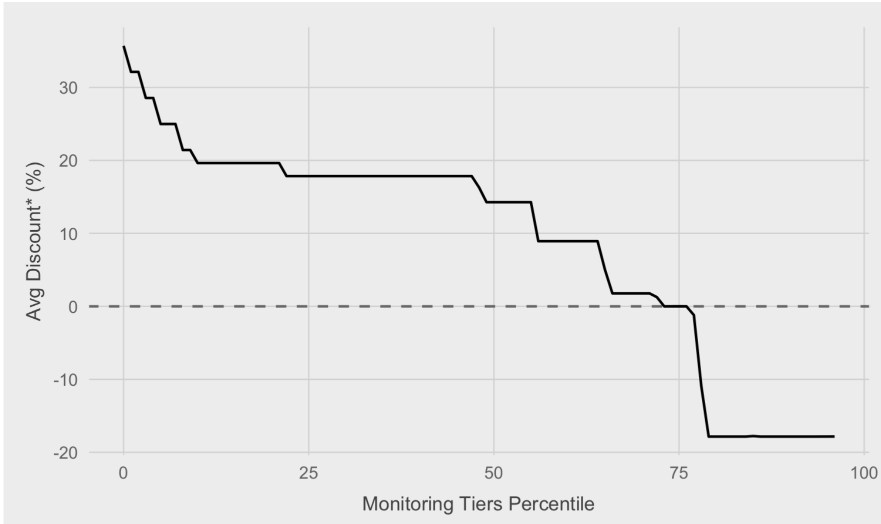


Figure B.3: Monitoring Discount Schedule

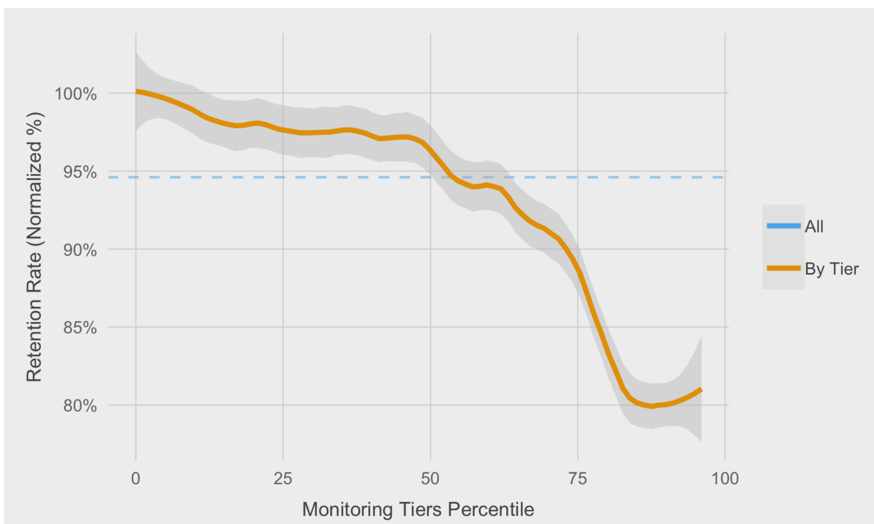


Figure B.4: Indexed Retention Rate

Notes: figure B.3 plots the firm's pricing schedule for giving monitoring discount. On the horizontal axis, we plot the percentile of monitoring tier, which is monitoring score divided by that expected by the firm given observables. 74% of people received a discount. The vertical axis is scaled by a factor between 0.5 and 1.5. This is to protect the firm's identity while demonstrating the scale and shape of the pricing algorithm. The firm went through two pricing schedules. This graph plots the second pricing schedule. The first one is similar, except that no surcharge was given.
figure B.4 uses the same horizontal axis, and non-parametrically plots the retention rate for the semester immediately after drivers finish monitoring (and thus when they first got monitoring discounts). Bandwidth is set as 5, and all numbers are benchmarked/normalized against the mean retention rate of the lowest 5 monitoring tiers. For 93% of monitored drivers, this is the first renewal period.

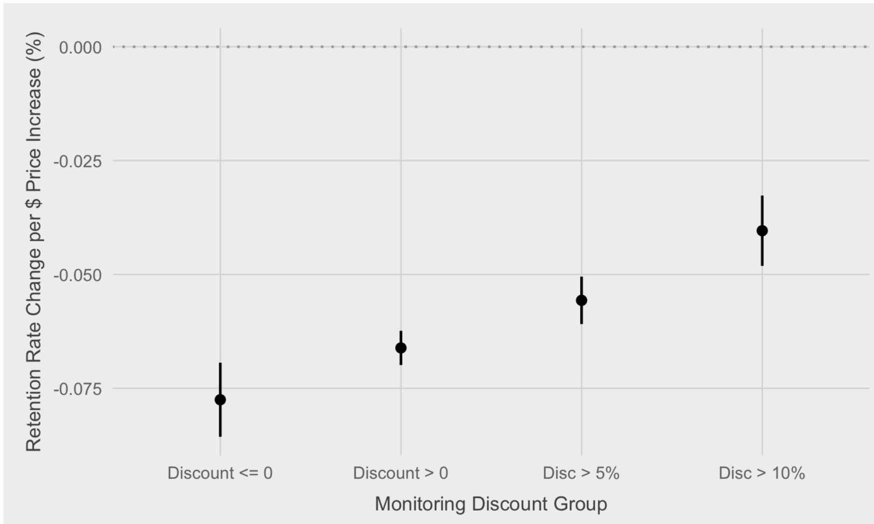


Figure B.5: Comparison of subsequent claim cost across monitoring groups

Notes: This figure plots the estimate of θ from equation (14) in various subsamples. These subsamples are represented on the horizontal axis. Notice that although we segment the data using discount percentage, we use the actual discount amount in the regression to measure demand elasticity. The results are scaled to percentage point terms. Therefore, -0.05 means that the slope of retention demand is such that a one dollar increase in price would lead to a 0.05 percentage point drop in retention rate.

C Additional Robustness Checks

Table C.1: **ESTIMATES FROM MORAL HAZARD REGRESSION**

	dependent variable: claim count (C)					
explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)
constant	0.046^{***}	0.003	0.004	0.046^{***}	0.003	0.004
	(0.000)	(0.005)	(0.005)	(0.000)	(0.005)	(0.005)
post monitoring indicator	-0.001^{**}	-0.003^{***}	-0.003^{***}	-0.001^{**}	-0.003^{***}	-0.003^{***}
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
monitoring start indicator (m_{start})	-0.009^{***}	-0.008^{***}	-0.008^{***}	0.025^{***}	0.020^{***}	0.020^{***}
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
monitoring intensity (M)				-0.050^{***}	-0.042^{***}	-0.042^{***}
				(0.004)	(0.004)	(0.004)
interaction ($\mathbf{1}_{post} \times m$)	0.006^{***}	0.006^{***}	0.006^{***}	-0.014^{***}	-0.014^{***}	-0.014^{***}
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
interaction ($\mathbf{1}_{post} \times M$)				0.028^{***}	0.029^{***}	0.029^{***}
				(0.005)	(0.005)	(0.005)
observables controls (x)	N	Y	Y	N	Y	Y
coverage fixed effects	N	N	Y	N	N	Y
implied risk reduction of a full period of monitoring (%)	28.0	29.4	29.5	27.5	29.4	29.6
pre- / post-periods for "first difference"	0 / 1-2					
treatment / control groups "second difference"	all starters / unmonitored					
number of drivers in balanced panel	812,924					

Notes: This table reports results of equation (2), but instead look at all monitored drivers regardless of whether they finish or not. Again, the estimate on the interaction term ($\mathbf{1}_{post} \times m_{start}$ or z) measures the treatment effect of monitoring ending on claim count. We first balance our panel data to include all drivers who stay till the end of the third semester ($t = 3$). This gives us two renewal semesters ($t \in \{1, 2\}$) after the monitoring semester ($t = 0$). We control for a full set of observables, including driver and vehicle characteristics and tiers (past records of violations or claims). It also includes third-order polynomials of calendar year and month. Continuous observable characteristics are normalized. We report estimates with and without these controls.

D Estimation Details

Intercept and slope parameters We parameterize heterogeneous latent parameters linearly. Broadly consistent with actual firm pricing rules, x_{it}^R and x_i^s only include a polynomial and the log of risk class, which represents firm's risk assessment without monitoring information.

Nest structure Incorporating additional alternative-level random effects can further enrich our model. In our primary specification, we add a random coefficient, ζ, on all choices within $f^{\star}$. This allows us to capture correlations between choices within the firm. Here, we assume ζ is an independently normally distributed with mean zero and standard deviation σ_ζ (Train 2009). This allows us to escape the Independence of Irrelevant Alternatives property of a simple logit model. The model can therefore achieve better fit on attrition rate differences across consumers facing different contract spaces across states or when mandatory minimum changes.

Taylor approximation approach for nonlinear utility Next, following the literature on auto insurance choices (Cohen and Einav 2007; Barseghyan, Molinari, O'Donoghue, and Teitelbaum 2013), we start with an approximation approach to model the utility function . Assuming that third- or higher-order derivatives are negligible, the utility function can be expressed by a second-order Taylor approximation of the utility function around income w. Normalizing by marginal utility evaluated at w, we get the following expression, in which γ is the absolute-risk-aversion term:

$$v_{idt}(\lambda, \zeta) = \mathbb{E}\left[h_{idt} | \lambda, \zeta\right] - \frac{\gamma}{2}\mathbb{E}\left[h_{idt}^2 | \lambda, \zeta\right] \tag{15}$$

This further simplifies product differentiation into consumption bundles with different mean and variance profiles. It also allows us to interpret v in monetary values, as the second term of Equation 15 is exactly the risk premium, while the first is expected consumption. We are currently running robustness checks for alternative utility assumptions such as CARA and CRRA, as well as to allow for richer heterogeneity in risk preference.

Estimation Our model includes random coefficients that enter utility nonlinearly. Private risk, in particular interacts with various observed monitoring and coverage characteristics (renewal price, out-of-pocket expenditure), as well as unobserved demand parameters (risk aversion and monitoring cost). Therefore, we use a simulated maximum likelihood approach (Train 2002; Handel 2013). In particular, the mix logit structure implies that the choice probability is numerically integrated as follows:

$$\begin{aligned}\Pr(d_{it}|\lambda) &= \Pr(\epsilon_{idt} - \epsilon_{id't} > [v_{idt}(\lambda) - v_{id't}(\lambda)] \ \forall d' \neq d \\ &= \frac{\exp\left[v_{idt}(\lambda)/\sigma\right]}{\sum_{d'} \exp\left[v_{id't}(\lambda)/\sigma\right]} \qquad (16)\\ \Pr(d_{it}) &= \int \Pr(d_{it}|\lambda) f_\lambda(\lambda) d\lambda \qquad (17)\end{aligned}$$

In general, for each parameter proposal Θ_d, we simulate 50 independent draws of private risk (ϵ_λ) and the zero-mean firm dummy (ζ).[71] Then, we compute the likelihood for observed choices, claim count and severity, monitoring score, and renewal price change. These are averaged over to get the simulated log likelihood. The estimator $\theta^\star$ maximizes the log likelihood. Notice that the Taylor approximation allows us to derive closed-form solutions for the first two moments of out-of-pocket expenditures and renewal prices.[72] We therefore do not simulate claim losses or monitoring scores within each draw of random coefficients.

As discussed above, our cost model is easier to estimate but requires a large amount of data to estimate precisely. Our demand model faces the opposite challenge, being computationally demanding but also making use of rich variations in choice environment and outcome. Therefore, we adopt a two-step estimation procedure. First, risk and monitoring score parameters ($\theta_\lambda, \sigma_\lambda, \theta_s, \sigma_s$) are estimated in the full dataset (except the loss severity parameter, per the discussion above). We then feed the estimates into the demand models as truth.[73] We lose precision by doing so, but both models are identified standalone.

Our model includes unobserved state variables (random coefficients) that enter utility non-linearly. Therefore, we use a random coefficient simulated maximum

[71] We test the effect of increasing the number of draws in estimation on a 10,000 sub-sample. The effect of going from 50 to 200 draws is minimal.

[72] Further, we restrict α_ℓ to be larger than 2 so that the mean and variance of the distribution are both finite, as both moments enter consumers' utility. The mean of the Pareto distribution is thus no more than $2\ell_0$. Therefore, to fit the average cost to the firm well, we set $\ell_0 = 3000$, roughly half the empirical mean of the claim distribution. This parameter is selected in cross-validation, om which we compare model performance in a hold-out dataset by directly calculating the likelihood. In a robustness check, we are also fitting a Gamma model for calculating the firm's cost only.

[73] Standard errors for the demand estimates are current not adjusted for two-step estimation. In a robustness check, we are correcting those standard errors and implementing a joint estimation.

likelihood approach (Train 2009; Handel 2013) to estimate the model.

For each parameter proposal θ, we simulate the model 50 times using Halton draws and compute the likelihood for all observations in the data. We then average over these to get the "simulated log likelihood", denoted as $\hat{\mathcal{L}}_{sim}(\theta)$. The estimator $\theta^\star$ maximizes the log likelihood. Simulated maximum likelihood suffer from simulation bias

Likelihood Function The log likelihood are sample analogs of four types of data likelihoods (denoted as $\mathcal{L}$) - claims, monitoring score, choices (of firm, coverage and monitoring participation), as well as renewal price. Utilities are history-dependent in our model. Therefore, we need to simulate choice sequence for each driver i. For notational simplicity, we suppress firm-dummy random effect ζ as in our baseline specification. The log likelihood function can then be expressed as follows.

$$\mathcal{L}_i \equiv \sum_{t \leq T_i} \int_\lambda \mathcal{L}(\underbrace{R_{it}, s_i, C_{it}, d_{it}|\lambda, \psi, x_{it}, \mathbf{p}_{it}, D_{it}, d_{i,t-1}; \mathbf{\Theta}}_{(A)\text{:obs. stoc outcome}}) \cdot \underbrace{g_\lambda(\lambda|x_{it}; \theta_\lambda, \sigma_\lambda)}_{(B)\text{:latent var.}} d\lambda$$

The simulation procedure allows us to numerically integrate over λ given parameter proposals θ_λ and σ_λ. We follow the timing of the model to decompose the likelihood component A as follows.

$$\begin{aligned}(A) = &\ln \Pr(d_{it}|\lambda, \mathbf{x}_{it}, \mathbf{p}_{it}, D_{it}, d_{i,t-1}; a, \psi_0, \psi_1, \theta_\eta, \theta_\xi, \alpha, \theta_\beta) + \\ &+ \ln \Pr(C_{it}|\lambda, \mathbf{x}_{it}) + \ln g(\ell_{it}|d_{it}, \mathbf{x}_{it}; \alpha, \theta_\beta) \\ &+ \ln g_s(s_i|\lambda, \mathbf{x}_{it}; \theta_s, \sigma_s) + \ln g_R(R_{idt}|C_{it}, s_i, \lambda, \mathbf{x}_{it}, \mathbf{p}_{it}; \theta_{\mathbf{R}}, \theta_{\mathbf{R},\mathbf{m}}, \sigma_R)\end{aligned}$$

Each component of (A) is modeled in the main text and given distributional assumptions.

Choice probability Our choice probability requires integration over all possible C, ℓ, R_0 and s. In our model, we assume away uncertainty in s, and our Poisson-Gamma model gives analytical solutions for expectation over C and ℓ.

For simplicity, in people's expectation, we only consider the possibility of one claim occurrence per term (Cohen and Einav 2007; Barseghyan et al. 2013). We can then capitalize on the attractive analytical property of gamma distributions and avoid numerical integration over C, ℓ, R_0 and s.

E Simulation Analysis of the Informativeness of Monitoring Signal

We can conduct a simple simulation exercise to quantify the spread of private risk and monitoring's effectiveness. To do so, we first simulate a large risk pool by taking the mean of all observable characteristics and simulating each driver's private risk. Figure E.1 plots the density of simulated true risk.[74] Next, Figure E.2 plots the firm's prior mean for all drivers in the risk pool. The firm has a flat prior for all drivers in the first period, which is far from the perfect belief (represented by the dotted and zoomed in 45-degree line). In Figure E.3, we calculate the evolution of firm belief (posterior mean) in subsequent periods as the firm observes potential claim realization. The firm's belief evolves towards the truth as claim is a direct measure of risk. However, the sparsity of claims, especially among safe drivers, dramatically slows down the firm's belief updating.

Monitoring score provides an immediate signal for driver risk after the first period. In Figure E.4, we plot, in orange, how the firm's belief updates after observing a one-time monitoring score. It is clear that monitoring is far more informative than observing a period of potential claim realization (dark grey line). Monitoring is especially useful in distinguishing the large mass of safe drivers, in which claims are even rarer. To quantify this measure, we can calculate the absolute deviation of firm belief from the true risk in our simulated risk pool. Overall, observing the monitoring score gets the firm 12.3% closer to the perfect belief (45-degree line).

[74] Our figures use private risk spread among new drivers for illustrative clarity.

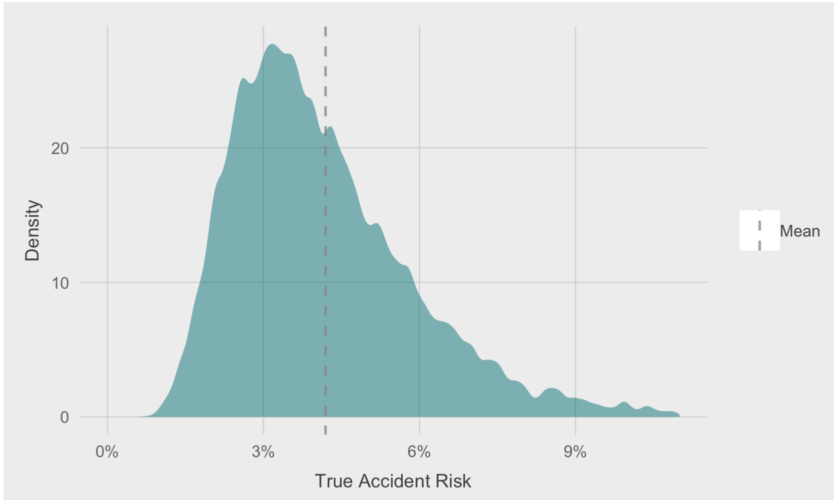


Figure E.1: A simulated mean risk pool given our cost estimate

Notes: This figure plots the distribution of a simulated mean risk pool given our cost estimates.

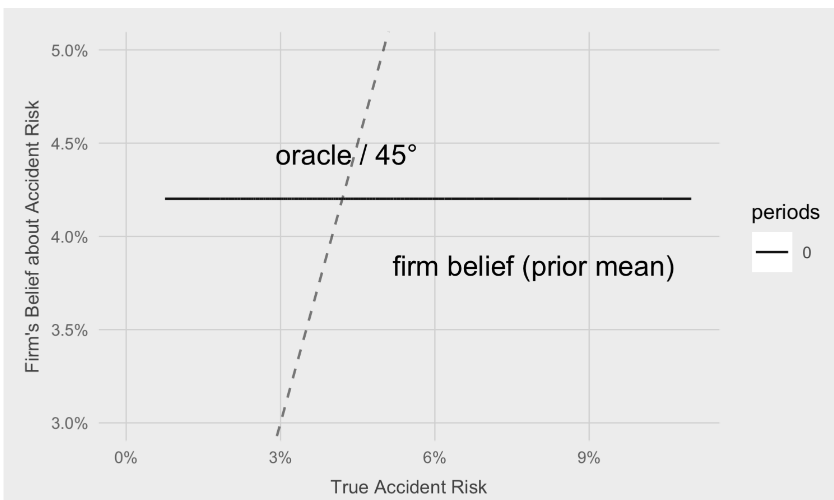


Figure E.2: Firm's prior on simulated risk pool

Notes: This figure plots firm's belief (prior mean / risk rating) for drivers in our simulated pool. In the first period, they are by definition pooled together. Therefore, firm has a flat prior for all drivers in the pool. The dotted line is the 45 degree line, which represents perfect belief.

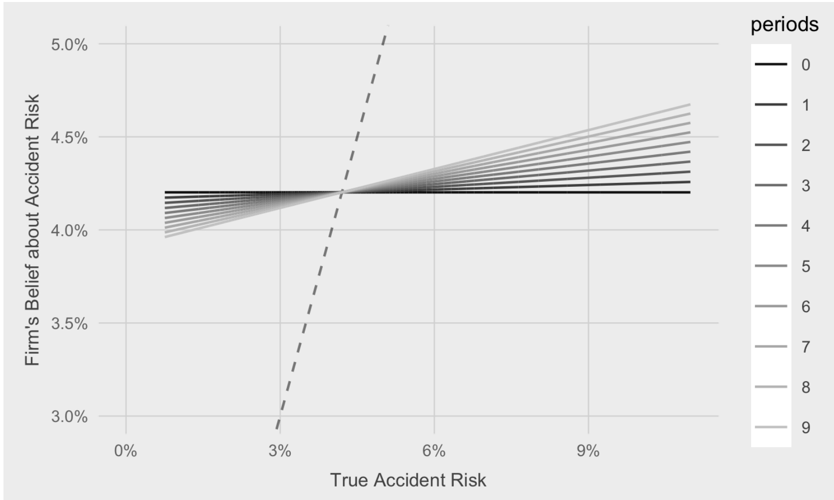


Figure E.3: Firm's posterior updating based on claims

Notes: This figure plots the evolution of firm belief (posterior mean) for drivers in our simulated pool based on liability claims alone. To make the updating analytically feasible, we first fit a gamma distribution on our risk pool by matching the mean and variance. Since gamma distribution is a conjugate prior for poisson updating, we are able to analytically derive the posterior mean.

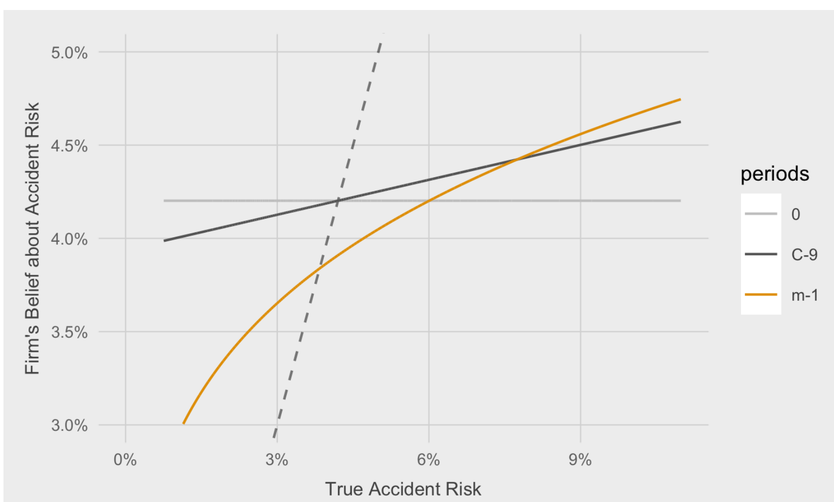


Figure E.4: Firm's posterior updating based on monitoring vs. claims

Notes: This figure plots the evolution of firm belief (posterior mean) for drivers in our simulated pool based on claims versus monitoring. Since lognormal distribution is a conjugate prior for lognormal updating, we are able to analytically derive the posterior mean.

F Counterfactual Simulation Methodology

Consistent with our demand model, we take a one-year horizon. The following procedure is used to calculate ex-ante and expected realized (ex-post) quantities.

1. For each driver i, simulate random coefficients (private risk and firm dummy) $L \in \mathbb{N}^+$ times.
2. For each draw $l \in \{1, ..., L\}$, calculate ex-ante utility directly and the corresponding certainty equivalent.[75] First-period choice probabilities are also calculated, which gives us the monitoring share. Expected cost of the first semester can be calculated directly. But we also need to form an expectation of the second-period cost (and prices) in order to calculate total surplus (and profit):
3. Simulate $K \in \mathbb{N}^+$ draws of first-period claim occurrence and monitoring score based on private risk.[76] Each draw pins down the renewal price change that driver i would face in the second period. All other prices remain constant. For each first-period choice d, we can then calculate the second period choice probability and the corresponding expected cost.

Sample enumeration Since we observe new customers' origins, as well as the competitive prices they face when coming to the firm, we can use our model to enumerate a full sample of potential new customers (Train 2009). To do so, we first calculate the probability of each new customer arriving at the firm. We then follow the same procedure as outlined above, but weight each driver by the inverse of the calculated probability. The simulation is carried out assuming that monitoring is available for all new customers.[77] Overall, our simulated dataset is expanded by a factor of 4.03, which gives us a market share (among the top six firms for which we have data) close to the reality in the states we study.[78] This also allows us to derive a realistic proxy for competitor profit under a symmetric cost assumption; that is, the distribution of risk that we estimate in our dataset is valid when extrapolated to the simulated market.

[75] Due to our Taylor approximation, this should be the negative root of the polynomial.

[76] For simplicity, we assume that R_0 is deterministic conditional on C and s. In reality, the spread of baseline R_0 without claims and monitoring may have subtle nonlinear effects on consumer choice, which we assume away.

[77] Part of the estimation data is pre-monitoring introduction. We use the average opt-in discount for these drivers.

[78] We winzorize the re-weighting scaling factor to be between 1 and 20 to deal with outliers.

In order to enumerate the market, we need to extrapolate the estimated attrition elasticity the firm faces to understand how the firm competes with other firms in the first period. To do so, we make a *no-brand-differentiation assumption*: liability insurance contracts offered by different firms only differ financially. This means that our firm-switching inertia estimate consists only of search and switching costs that are state-dependent (on consumers' preexisting firm choice) and that consumers have no unobserved preference for our firm, which is not state-dependent. In the context of our counterfactual simulations, this assumption essentially maintains that the price elasticity the firm's competitors face when the firm tries to poach customers away from them (in the first period) is the same as the price elasticity the firm faces when trying to retain existing customers.

This assumption follows naturally from our data limitation: we do not observe comprehensive micro-level choice or quantity data for the firm's competitors. But it is also supported by empirical evidence. Honka (2012) uses a survey dataset that includes individual consumer choices across auto insurers. She is then able to tease out switching cost from firm-specific preferences. She finds that the mean firm preferences are not significantly different from 0 for all companies.[79]

[79] Her estimate of search and switching cost is lower than our estimate. However, for the firm from which our administrative dataset comes from, the reported attrition rate in her dataset is more than three times as large as what we observe. Her estimate is therefore likely biased downwards.

G Counterfactual Demand Models

In this section, we show simulation results of removing key components of the demand model, as an illustration of their relative importance in determining monitoring share and the firm's profitability.

Second, the "Perfect Sig." model assumes that the monitoring signal is perfect in consumers' expectation by setting σ_s to zero. The market share, unconditional and conditional monitoring shares increase by 0.4pp, 0.6pp, and 2.6pp, respectively. In reality, our specification is consistent with a dynamic framework in which firm-switching is infinitely costly within a year. This will likely overstate the effect of reclassification risk. Nevertheless, the impact of a perfect signal on demand is small compared to that of other forces.[80]

Demand frictions are the most important deterrent against monitoring participation. The third model removes firm-switching inertia, which dramatically lowers the barrier for drivers with good private risk to participate in monitoring. However, It also clears the way for drivers to explore attractive outside options. We find that the firm is able to gain market share by 12.6pp, while increasing its monitoring share by 12.1pp so that 5.9% of drivers in the market has monitoring. Lastly, we remove monitoring cost. This generates the biggest impact on monitoring by far. In particular, any driver with good private risk would prefer monitoring with any coverage within the firm. The monitoring share rises to 61.3%, with 16.2% of the market opting in the firm's monitoring program.

Firm profit is influenced not only by its market share, but also by risk selection. To directly visualize this, we isolate the risk selection effect from the overall profit impact in Figure G.6. It plots the expected private risk parameter ($\epsilon_{\lambda,i}$, mean 0) for the firm's customers, both monitored and unmonitored. This clarifies the changes in the private risk of the marginal customers that come to the firm as we relax demand factors, which is crucial in understanding competition in selection markets. As the firm cream-skims better drivers in its monitored pool, the unmonitored pool in and outside of the firm deteriorates. These pool may therefore eventually unravel as firms adjust prices.

[80] A caveat is that we assume rational expectation in our model. This means that the effect of a systematic over- or under-estimation of the monitoring signal's noise would show up in drivers' monitoring cost instead of be attributed to reclassification risk.

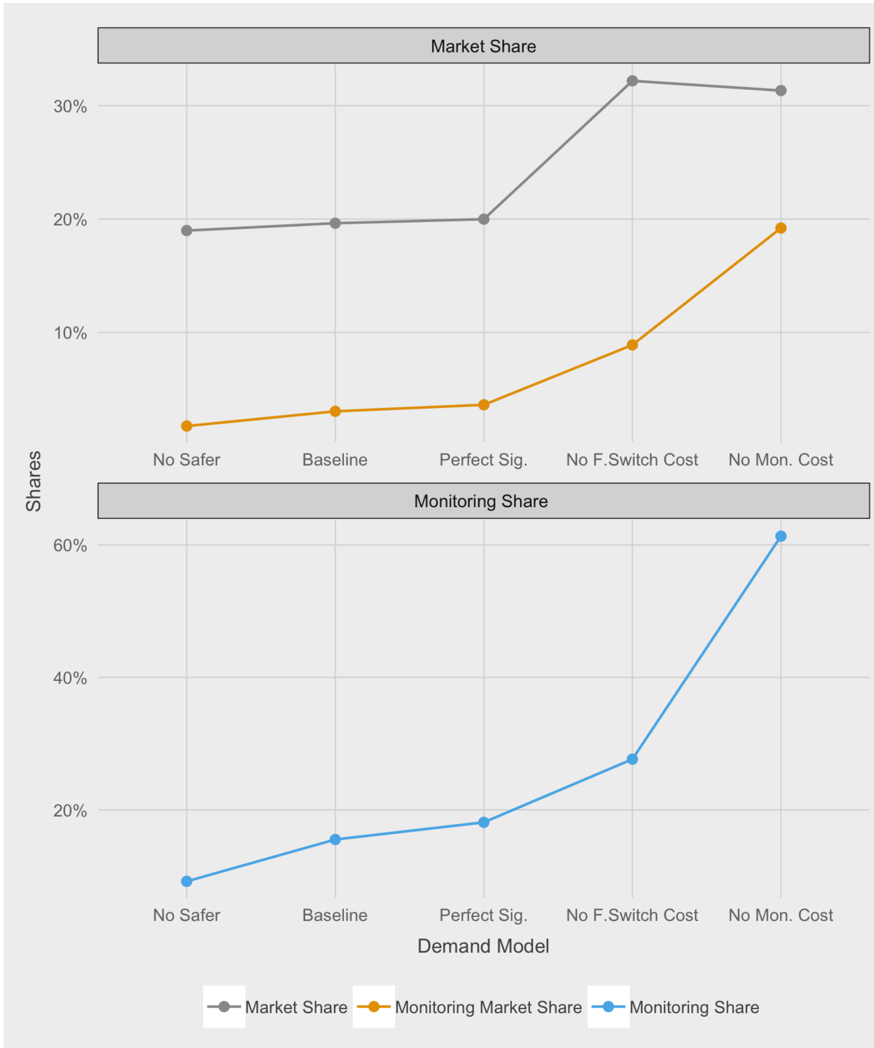


Figure G.5: Demand Share Simulation Across Demand Model Assumptions

Notes: These figures correspond to our analyses in **??**. The top graph plots the counterfactual market share of the firm, as well as the unconditional share of monitored drivers in the market, when prices are fixed but the demand model changes. The bottom graph plots the conditional monitoring share within the firm. See main text for definitions of each model - importantly, changes in model features are *not* cumulative from left to right. We also enumerate our sample of new customers to the full market with model-predicted likelihood of each new customer being in our dataset.

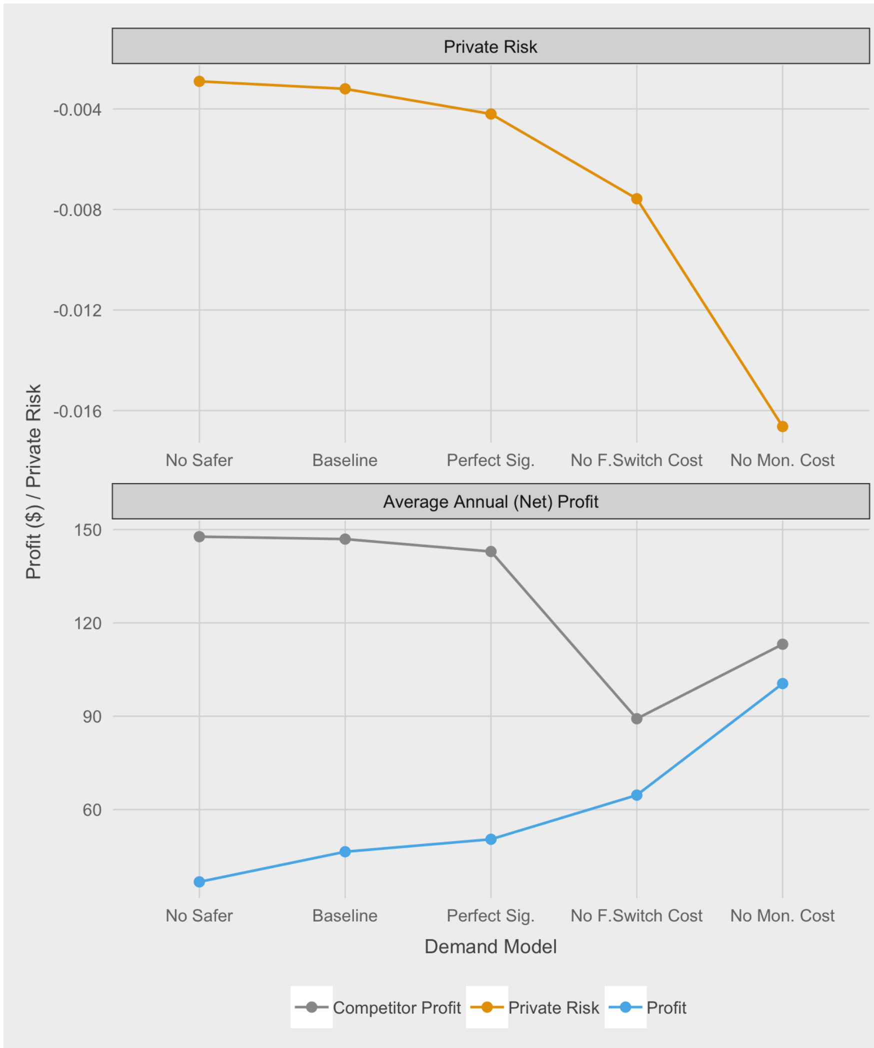


Figure G.6: Simulation - Profit Under Different Demand Model Assumptions

Notes: Corresponding to the figure above, these graphs plot firm profit and competitor profit, holding prices fixed. The top graph plots the expected private risk among the firm's customers. Notice that private risk has mean zero in the population. It is numerically integrated over in the counterfactual simulations. With each draw, we weight each person's private risk with her probability of arriving at the firm to get the number shown above. It therefore represents both the monitored and the unmonitored pools of the firm.

H Regulatory Filing Examples

OHIO
VOLUNTARY PRIVATE PASSENGER AUTO
PREMIUM CALCULATION

ROUND AFTER EACH CALCULATION TO THE NEAREST PENNY

STEP #		AA	BB	CC	DD	HH	DNC*^	HNC*^
1	TERRITORIAL BASE RATE (RP-1BR)							
2	RATE ADJUSTMENT FACTOR (PENNY ROUND)	x 1.598	x 1.594	x 1.410	x 1.121	x 1.111	x 1.121	x 1.111
3	INCREASED LIMIT FACTOR/ADDEND (RP-3A)	x	+	x				
4	POLICY GROUP FACTOR (RP-4A-1 through RP-4A-2)	x	x	x	x	x	x	x
5	RATING TIER FACTOR (RP-5A)	x	x	x	x	x	x	x
6	ALLSTATE® YOUR CHOICE AUTO INSURANCE OPTION PACKAGE FACTOR (RP-15A)	x	x	x	x	x	x	x
7	POLICY CLASS FACTOR (RP-7A-1 through RP-7A-4)	x	x	x	x	x	x	x
8	HOUSEHOLD COMPOSITION FACTOR (RP-8A-1 and RP-8A-2)	x	x	x	x	x	x	x
9	SMART STUDENT DISCOUNT FACTOR (RP 10A and RP-11A)	x	x	x	x	x	x	x
10	DEFENSIVE DRIVER DISCOUNT FACTOR (RP-10A and RP-12A)	x	x	x	x	x	x	x
11	MULTIPLE POLICY DISCOUNT FACTOR (RP-15A)	x	x	x	x	x	x	x
12	HOMEOWNER DISCOUNT FACTOR (RP-15A)	x	x	x	x	x	x	x
13	THE GOOD HANDS PEOPLE® DISCOUNT FACTOR (RP-15A)	x	x	x	x	x	x	x
14	RESPONSIBLE PAYER DISCOUNT FACTOR (RP-15A)	x	x	x	x	x	x	x
15	FULLPAY DISCOUNT (RP-15A)	x	x	x	x	x	x	x
16	ALLSTATE EASY PAY PLAN DISCOUNT (RP-15A)	x	x	x	x	x	x	x
17	EARLY SIGNING DISCOUNT (RP-15A)	x	x	x	x	x	x	x
18	ALLSTATE AUTO/LIFE DISCOUNT℠ (RP-15A)	x	x	x	x	x	x	x
19	ALLSTATE eSMART℠ DISCOUNT (RP-15A)	x	x	x	x	x	x	x
20	SAFE DRIVING CLUB (RP-10A and RP-13A through RP-14A)	x	x	x	x	x	x	x
21	PRIOR NON-STANDARD CARRIER SURCHARGE (RP-16A)	x	x	x	x	x	x	x
22	ACCIDENT SURCHARGE FACTOR (RP-17A)	x	x	x	x		x	
23	MAJOR VIOLATION SURCHARGE FACTOR (RP-18A)	x	x		x		x	
24	MINOR VIOLATION SURCHARGE FACTOR (RP-19A)	x	x		x		x	
25	MODEL YEAR FACTOR (RP-20A)				x	x	x	x
26	DEDUCTIBLE BY PGS FACTOR (RP-20A)				x	x	x	x
27	EXPERIENCE GROUP RATING FACTOR (EGR PAGES and RP-21A-24A)	x	x	x	x	x	x	x
28	ALLSTATE DRIVE WISE® ENROLLMENT DISCOUNT (RP-26A)	x	x	x	x	x	x	x
29	ALLSTATE DRIVE WISE® PERFORMANCE RATING (RP-26A)	x	x	x	x	x	x	x
30	ANNUAL VEHICLE MILEAGE FACTOR (RP-16A)	x	x	x	x	x	x	x
31	VEHICLE USAGE FACTOR (RP-16A)	x	x	x	x	x	x	x
32	FARM DISCOUNT FACTOR (RP-16A)	x	x		x		x	
33	ELECTRONIC STABILITY CONTROL DISCOUNT FACTOR (RP-16A)	x	x	x	x		x	
34	PASSIVE RESTRAINT DISCOUNT (RP-16A)			x				
35	ANTILOCK BRAKE DISCOUNT (RP-16A)	x	x		x		x	
36	NEW CAR DISCOUNT FACTOR (RP-16A)	x	x	x	x	x	x	x
37	CERTIFIED RISK SURCHARGE FACTOR (RP-16A)	x	x					
38	CAMPER UNIT ADDITIONAL PREMIUM (RP-25A)				+	+		
39	NEW CAR EXPANDED PROTECTION FACTOR (RP-25A)						x	x
40	RATE TRANSITION FACTOR (Rule 72)	x	x	x	x	x		
41	COMPLEMENTARY GROUP RATING (CGR) FACTOR (RP-9A-1 through RP-9A-13)	x	x	x	x	x	x	x
42	FIXED EXPENSE PREMIUM ^^ (RP-16A)	+						
43	SUB-TOTAL VEHICLE PREMIUM	=	=	=	=	=	=	=

	RENTAL REIMBURSEMENT (UU)	
	RENTAL REIMBURSEMENT BASE RATE (RP-52BR)	
	RENTAL REIMBURSEMENT INCREASED LIMIT FACTOR (RP-3A)	x
44	TOTAL RENTAL REIMBURSEMENT COVERAGE PREMIUM	=

	TOWING & LABOR COSTS (JJ) (RP-25A)	
	SOUND SYSTEMS (ZA) (RP-25A)	+
	TAPE (ZZ) (RP-25A)	+
45	TOTAL MISCELLANEOUS COVERAGES	=

	PER AUTO UM/UIM - PROPERTY DAMAGE COVERAGE (SSP)	
46	UM - PROPERTY DAMAGE PREMIUM RATE (RP-3A)	

	POLICY UM/UIM - BODILY INJURY COVERAGE (SS)	
	TERRITORIAL BASE RATE (RP-1BR)	
	RATE ADJUSTMENT FACTOR (PENNY ROUND)	x 0.872
	INCREASED LIMIT FACTOR/ADDEND (RP-3A)	x
	POLICY GROUP FACTOR (RP-4A-1 through RP-4A-2)	x
	RATING TIER FACTOR (RP-5A)	x
	POLICY CLASS FACTOR (RP-7A-1 through RP-7A-4)	x
	HOUSEHOLD COMPOSITION FACTOR (RP-8A-1 through RP-8A-2)	x
	SMART STUDENT DISCOUNT FACTOR (RP 10A and RP-11A)	x
	DEFENSIVE DRIVER DISCOUNT FACTOR (RP-10A and RP-12A)	x
	HOMEOWNER DISCOUNT FACTOR (RP-15A)	x
	RESPONSIBLE PAYER DISCOUNT FACTOR (RP-15A)	x
	FULLPAY DISCOUNT (RP-15A)	x
	SAFE DRIVING CLUB (RP-10A and RP-13A through RP-14A)	x
	ACCIDENT SURCHARGE FACTOR (RP-17A)	x
	MAJOR VIOLATION SURCHARGE FACTOR (RP-18A)	x
	MINOR VIOLATION SURCHARGE FACTOR (RP-19A)	x
	RATE TRANSITION FACTOR (Rule 72)	x
	COMPLEMENTARY GROUP RATING (CGR) FACTOR (RP-9A-1 through RP-9A-13)	x
47	TOTAL UM/UIM - BODILY INJURY COVERAGE	=

48	**TOTAL SEMI-ANNUAL VEHICLE 1 PREMIUM = 43 + 44 + 45 + 46 +47**	+
49	**TOTAL SEMI-ANNUAL VEHICLE 2 PREMIUM = 43 + 44 + 45 + 46 +47**	+
50	**TOTAL SEMI-ANNUAL VEHICLE 3 PREMIUM = 43 + 44 + 45 + 46 +47**	+
51	**TOTAL SEMI-ANNUAL VEHICLE 4 PREMIUM = 43 + 44 + 45 + 46 +47**	+
52	**TOTAL SEMI-ANNUAL POLICY PREMIUM = 48 + 49 + 50 +51**	=

A ALLSTATE FIRE AND CASUALTY INSURANCE COMPANY 04-

Figure H.1: Pricing Algorithm - Insurer 1 OH

Notes: This page is taken from an insurer's Ohio rate filing, which demonstrates their pricing algorithm.

Progressive Direct Insurance Company (DI)
Progressive Specialty Insurance Company (AG)
Ohio Private Passenger Automobile Program
NB Effective Date: 9/05/2014

Exhibit: 11

RATE ORDER OF CALCULATION

The first step of the rate calculation formula is to determine the Household Risk Factor. The Household Risk Factor is the average of the Developed Driver Risk Factors for all eligible to be rated drivers up to the number of vehicles (or at least one in the case of a named operator policy). For policies where there are more drivers than vehicles, the Household Risk Factor is the average of the highest ranked drivers, up to the number of vehicles. The rank is determined by the Developed Driver Risk Factor for BI (higher factor = higher rank). The Developed Driver Risk Factor is determined as follows:

Driver Risk Factor Items	BI	PD	COMP	COLL	LOAN	MED	RENT	ROADSIDE	UMPD
Driver Classification Factor									
Years Licensed Factor	x	x	x	x	x	x	x	x	x
Driving Record Points Factor	+	+	+	+	+	+	+	+	+
Violation Leniency Factor [1]	-	-	-	-	-	-	-	-	-
Subtraction of One	-1	-1	-1	-1	-1	-1	-1	-1	-1
(1 - Distant Student Discount)	x	x	x	x	x	x	x	x	x
(1 - Minor Child Discount)	x	x	x	x	x	x	x	x	x
(1 - Good Student Discount)	x	x	x	x	x	x	x	x	x
(1 - Senior Citizen Discount)	x	x	x	x		x			
Household Member Factor	x	x	x	x	x	x	x	x	x
Driver Age Point Factor	x	x	x	x	x	x	x	x	x
Financial Responsibility by Clean Factor	x	x	x	x	x	x	x	x	x
Developed Driver Risk Factor									

The second step of the rate calculation formula uses the Household Risk Factor and follows

	BI	PD	COMP	COLL	LOAN	MED	RENT	ROADSIDE	UMPD
Household Risk Factor									
Base Rate	x	x	x	x	x	x	x	x	x
Financial Responsibility Factor	x	x	x	x	x	x	x	x	x
Financial Responsibility by Number of Drivers Factor	x	x	x	x	x	x			x
Deductible Savings Bank Factor	x	x	x	x		x			
Occupation/Education Rating Factor	x	x	x	x	x	x			x
Full Coverage Factor	x	x				x			
Household Structure Factor	x	x	x	x	x	x			x
Residency Rewards Factor	x	x	x	x	x	x	x		x
Luxury Vehicle Factor	x	x	x	x	x	x			x
Tier Factor	x	x	x	x	x	x	x	x	x
Policy Term Factor	x	x	x	x	x	x	x	x	x
Vehicle Age Factor [3]	x	x	x	x	x	x	x	x	x
Excess Vehicle Factor	x	x	x	x	x	x			x
Limit Factor	x	x			x	x	x	x	x
Deductible Factor			x	x					
Vehicle Age by Deductible Factor			x	x					
Vehicle Symbol Factor	x	x	x	x	x	x			x
Value Class Factor (for Vehicle symbols 67 & 68)			x	x	x				x
Vehicle Garaging Location Factor	x	x	x	x	x	x	x	x	x
(1 - Homeowner/Mobile Home/Multi-car Discount)	x	x	x	x		x	x	x	x
(1 - Advance Quote [1]/Three-year Safe Driving/Five-year Accident Free Discount)	x	x	x	x		x			x
(1 - Three-year Safe Driving Bonus) [1]	x	x	x	x		x			x
(1 - Agent Discount) [1]	x	x	x	x		x			x
(1 - Electronic Funds Transfer Discount)	x	x	x	x		x			x
(1 - Paid In Full Discount)	x	x	x	x		x			x
(1 - Online Quote Discount) [2]	x	x	x	x		x			x
(1 - Loyal Customer Discount) [2]	x	x	x	x		x			x
(1 - Paperless Discount)	x	x	x	x	x	x	x	x	x
(1 - Continuous Insurance Discount)	x	x	x	x	x	x	x	x	x
(1 - Multi-policy Discount)	x	x	x	x		x	x	x	x
(1 + Business Use Surcharge)	x	x	x	x		x			x
(1 + Financial Responsibility Filing Surcharge)	x	x	x	x		x			x
Bad Debt Factor	x	x	x	x		x			x
Apply Rate Capping Rule P23 [4]	x	x	x	x	x	x	x	x	x
Usage-based Insurance Factor	x	x	x	x	x	x	x	x	x
(1 - E-signature discount) [2]	x	x	x	x	x	x	x	x	x
Round to the Whole Dollar									
Operations Expense [5]	+		+						
Acquisition Expense [2,6]	+		+						
Developed Premium [8]									

[1] Applies to Progressive Specialty Insurance Company (AG) Only
[2] Applies to Progressive Direct Insurance Company Only (DI)
[3] If coverage is BI, PD, UM/UIM, MED, RENT, or ROADSIDE and Vehicle Symbol = 66, then Vehicle Age Factor = 1.0.
If coverage is COMP, COLL, LOAN, or UMPD and Vehicle Symbol = 66, 67, 68, or 69, then Vehicle Age Factor = 1.0.
[4] Policy level rate changes are capped at +/- 10% as described in Rule P23. The Snapshop Usage Based Insurance Program (UBI) is not taken into consideration when applying the Rate Capping Rule
[5] Operations expense is added to BI if BI is selected; if BI is not selected, then Operations Expense is added to COMP.
[6] Acquisition expense is added to BI if BI is selected; if BI is not selected, then Acquisition Expense is added to COMP.
[7] Average factors are determined by taking the average of Location, Symbol, Vehicle Age factors, and Business Use Surchage for each vehicle, respectively
[8] There is a minimum premium of $5 for each coverage selected for each vehicle.
[9] The trailer coverages will receive the factors associated with COMP and COLL, unless otherwise noted.

NOTES
x means factor is to be used multiplicatively
/ means factor is to be used as a divisor
+ means factor is to be added
- means factor or amount is to be subtracted

Figure H.2: Pricing Algorithm - Insurer 2 OH 1/2

Notes: These pages are taken from a an insurer's rate filing in Ohio, which demonstrate their pricing algorithm.

Progressive Direct Insurance Company (DI)
Progressive Specialty Insurance Company (AG)
Ohio Private Passenger Automobile Program
NB Effective Date: 9/05/2014

Exhibit: 11

	UM/UIM
Base Rate	
Financial Responsibility Factor	x
Financial Responsibility by Number of Drivers Factor	x
Deductible Savings Bank Factor	x
Occupation/Education Rating Factor	x
Full Coverage Factor	x
Household Structure Factor	x
Residency Rewards Factor	x
Driver Count Factor	x
Luxury Vehicle Factor	x
Tier Factor	x
Policy Term Factor	x
Avg. Vehicle Age Factor [3, 7]	x
Excess Vehicle Factor	x
Limit Factor	x
Avg. Vehicle Symbol Factor [7]	x
Avg. Vehicle Garaging Location Factor [7]	x
(1 - Homeowner/Mobile Home/Multi-car Discount)	x
(1 - Advance Quote [1]/Three-year Safe Driving/Five-year Accident Free Discount)	x
(1 - Three-year Safe Driving Bonus) [1]	x
(1 - Agent Discount) [1]	x
(1 - Electronic Funds Transfer Discount)	x
(1 - Paid In Full Discount)	x
(1 - Online Quote Discount) [2]	x
(1 - Loyal Customer Discount) [2]	x
(1 - Paperless Discount)	x
(1 - Continuous Insurance Discount)	x
(1 - Multi-policy Discount)	x
(1 + Avg. Business Use Surcharge [7])	x
(1 + Financial Responsibility Filing Surcharge)	x
Bad Debt Factor	x
Apply Rate Capping Rule P23 [4]	x
(1 - E-signature discount) [2]	x
Round to the Whole Dollar	
Developed Premium [8]	

	ACPE	COMP-TRLR [1,9]	COLL-TRLR [1,9]	CONTENTS [1]	OPERATIONS EXPENSE [5]	ACQUISITION EXPENSE [2,6]
Base Rate				0.015 * Value		
Financial Responsibility Factor		x	x	x		
Deductible Savings Bank Factor		x	x			
Residency Rewards Factor		x	x			
Tier Factor		x	x	x		
Policy Term Factor	x	x	x	x	x	x
Limit Factor	x					
Deductible Factor		x	x			
Vehicle Symbol Factor		x	x			
Value Class Trailer Factor [1]		x	x			
Vehicle Garaging Location Factor	x	x	x			
(1 - Paperless Discount)	x	x	x	x	x	
(1 - Continuous Insurance Discount)		x	x			
(1 - Multi-policy Discount)						x
Operations Expense Factor 1					x	
Operations Expense Factor 2					x	
Operations Expense Factor 3					x	
Acquisition Expense Full Coverage Factor [2]						x
Acquisition Expense Homeowner Factor [2]						x
Acquisition Expense Online Quote Factor [2]						x
Acquisition Expense Prior Insurance Factor [2]						x
Acquisition Expense Vehicle Count Factor [2]						x
Number of Vehicles						/
Apply Rate Capping Rule P23 [4]	x	x	x	x	x	x
Bad Debt Factor		x	x			
Usage-based Insurance Factor	x				x	x
(1 - E-signature discount) [2]	x				x	x
Round to the Whole Dollar						
Developed Premium [8]						

Total Policy Premium = Sum of Developed Premiums

[1] Applies to Progressive Specialty Insurance Company (AG) Only
[2] Applies to Progressive Direct Insurance Company Only (DI)
[3] If coverage is BI, PD, UM/UIM, MED, RENT, or ROADSIDE and Vehicle Symbol = 66, then Vehicle Age Factor = 1.0.
If coverage is COMP, COLL, LOAN, or UMPD and Vehicle Symbol = 66, 67, 68, or 69, then Vehicle Age Factor = 1.0.
[4] Policy level rate changes are capped at +/- 10% as described in Rule P23. The Snapshop Usage Based Insurance Program (UBI) is not taken into consideration when applying the Rate Capping Rule
[5] Operations expense is added to BI if BI is selected; if BI is not selected, then Operations Expense is added to COMP.
[6] Acquisition expense is added to BI if BI is selected; if BI is not selected, then Acquisition Expense is added to COMP.
[7] Average factors are determined by taking the average of Location, Symbol, Vehicle Age factors, and Business Use Surchage for each vehicle, respectively
[8] There is a minimum premium of $5 for each coverage selected for each vehicle.
[9] The trailer coverages will receive the factors associated with COMP and COLL, unless otherwise noted.

NOTES
x means factor is to be used multiplicatively
/ means factor is to be used as a divisor
\+ means factor is to be added
\- means factor or amount is to be subtracted

Figure H.3: Pricing Algorithm - Insurer 2 OH 2/2

Notes: These pages are taken from an insurer's rate filing in Ohio, which demonstrate their pricing algorithm.

GEICO Casualty Company - Voluntary Private Passenger Automobile Insurance
Ohio Rate Pages Effective: New Business 10/2/2009 Renewals 10/2/2009 Rate Gen 01
Rate Order of Calculation: Private Passenger

Machine Rated, Exception: Licensed/Registered Dune Buggies rated as PPV are **Manually Rated**

Oper	Step	BI	PD	MED	UM-UND	UMBI	UMPD	COMP	COLL	ERS	RR	MBI
Base Rate												
	Base Rate	X	X	X	X	X	X	X	X	X	X	X
*	Limit Factor	X	X	X	X	X	X			X	X	
*	Deductible Factor							X	X			
*	Term Factor	X	X	X	X	X	X	X	X	X	X	X
*	Upgraded Accident Forgiveness Factor	X	X	X	X	X	X	X	X			
Driver Level Rating Steps- Composite Relativities												
*	Driver Class Factor (Composite Relativity)	X	X	X	X	X	X		X			
	Accident Factor	X	X	X	X	X	X		X			
	* Minor Violation Factor	X	X	X	X	X	X		X			
	* Major Violation Factor	X	X	X	X	X	X		X			
	* Speeding Violation Factor	X	X	X	X	X	X		X			
	* DUI Violation Factor	X	X	X	X	X	X		X			
	* Unverifiable Driving Record Factor	X	X	X	X	X	X		X			
	= Merit Factor	X	X	X	X	X	X		X			
*	Merit Factor (Composite Relativity)	X	X	X	X	X	X		X			
Driver Level Discounts: Composite Relativities												
*	*Good Driver Discount (Composite Relativity)*	X	X	X	X	X	X		X			
*	*Student Away at School Discount (Composite Relativity)*	X	X	X	X	X	X		X			
*	*Driving Experience Discount (Composite Relativity)*	X	X	X	X	X	X		X			
*	*Good Student Discount (Composite Relativity)*	X	X	X	X	X	X		X			
*	*Defensive Driver Discount (Composite Relativity)*	X	X	X	X	X	X		X			
*	*Deployed Driver Discount (Composite Relativity)*	X	X	X	X	X	X		X			
Vehicle Level Rating Steps												
*	Vehicle Type Factor											
*	Annual Mileage/ Vehicle Use Factor	X	X	X	X	X	X	X	X			
*	Vehicle Classification Factor	X	X	X	X	X	X	X	X			
*	Vehicle Cost Factor	X	X	X	X	X	X	X	X			
*	Model Year Factor	X	X	X	X	X	X	X	X			
*	Vehicle Age Factor	X	X	X	X	X	X	X	X	X		
*	MBI Model Year Factor											X
*	MBI Coverage Age											X
Vehicle Level Discounts												
*	*Anti-Theft Discount*							X				
*	*New Vehicle Discount*	X	X	X	X	X	X	X	X			
*	*Extra Vehicle Discount*	X	X	X	X	X	X	X	X			
*	*Anti-Lock Brake Discount*	X	X	X	X	X	X	X	X			
*	*Restraint Discount*			X	X	X						
Policy Level Rating Steps												
*	Household Composite Factor	X	X	X	X	X	X	X	X			
*	Maximum Named Insured Age Factor	X	X	X	X	X	X	X	X			
*	Policy Occurrence Factor	X	X	X	X	X	X	X	X			
*	Risk Tier Factor	X	X	X	X	X	X	X	X	X	X	X
Policy Level Discounts												
*	*Financial Responsibility Discount*	X	X	X	X	X	X	X	X	X	X	X
*	*Seat Belt Discount*			X	X	X						
*	*Multi-Vehicle Discount*	X	X	X	X	X	X	X	X	X	X	X
*	*Continuous Insurance Discount*	X	X	X	X	X	X	X	X			
*	*Military Discount*	X	X	X	X	X	X	X	X			
*	*Multi-Line Discount*	X	X	X	X	X	X	X	X	X	X	X
*	*CDL Discount*											
Policy Level Discounts 2												
*	*Sponsored Marketing Discount*	X	X	X	X	X	X	X	X	X	X	X
*	*Associate Discount*	X	X	X	X	X	X	X	X	X	X	X
*	*E-Banking Discount*	X	X	X	X	X	X	X	X	X	X	X
Expense Constants												
+	Vehicle Expense Load	X	X									
+	Policy Expense Load	X	X									

Figure H.4: Pricing Algorithm - Insurer 3 OH

Notes: These pages are taken from an insurer's rate filing in Ohio, which demonstrate their pricing algorithm.

OHIO
VOLUNTARY PRIVATE PASSENGER AUTO
POLICY CLASS FACTOR

POLICY CLASS FACTOR CALCULATION

Complete the steps below for all applicable coverages. Round to 4 decimals after each step.

Step 1: Based on number of male and female operators on the policy, obtain a value from Value Table 1a and find the corresponding factor from Factor Table 1a.

Step 2: Based on number of single operators, married operators, operators aged <25, and operators aged 25+ on the policy, obtain a value from Value Table 1b and find the corresponding factor from Factor Table 1b.

Step 3: Based on age and gender, obtain a value from Value Table 1c for **each** operator and calculate the geometric average of the corresponding factors from Factor Table 1c to obtain a factor for step 3.

Step 4: Multiply steps 1, 2, and 3 together to obtain the policy class factor to be applied to all vehicles on the policy.

See **RP-6A** for instructions on how to calculate the geometric average.

VALUE TABLE 1a - GENDER

Continuous Prior Insurance	# of Males	# of Females 0	1	2	3+
Yes	0	-	1	2	3
Yes	1	10	11	12	13
Yes	2	20	21	22	23
Yes	3+	30	31	32	33
No	0	-	4	5	6
No	1	14	15	16	17
No	2	24	25	26	27
No	3+	34	35	36	37

VALUE TABLE 1b - MARITAL STATUS AND AGE

Continuous Prior Insurance (Y/N)			Marital Status														
		# Married	0	0	0	1	1	1	1	2	2	2	2	3+	3+	3+	3+
		# Single	1	2	3+	0	1	2	3+	0	1	2	3+	0	1	2	3+
	# age >= 25	# age < 25															
Yes	0	1	1			16											
Yes	0	2		3			18			33							
Yes	0	3+			6			21	25		36	40	45	52	62	70	77
Yes	1	0	2			17											
Yes	1	1		4			19			34							
Yes	1	2			7			22			37			53			
Yes	1	3+			8				26			41	46	54	63	71	78
Yes	2	0		5			20			35							
Yes	2	1			9			23			38			55			
Yes	2	2			10				27			42		56	64		
Yes	2	3+			11				28				47	57	65	72	79
Yes	3+	0			12			24	29		39	43	48	58	66	73	80
Yes	3+	1			13				30			44	49	59	67	74	81
Yes	3+	2			14				31				50	60	68	75	82
Yes	3+	3+			15				32				51	61	69	76	83
No	0	1	84			99											
No	0	2		86			101			116							
No	0	3+			89			104	108		119	123	128	135	145	153	160
No	1	0	85			100											
No	1	1		87			102			117							
No	1	2			90			105			120			136			
No	1	3+			91				109			124	129	137	146	154	161
No	2	0		88			103			118							
No	2	1			92			106			121			138			
No	2	2			93				110			125		139	147		
No	2	3+			94				111				130	140	148	155	162
No	3+	0			95			107	112		122	126	131	141	149	156	163
No	3+	1			96				113			127	132	142	150	157	164
No	3+	2			97				114				133	143	151	158	165
No	3+	3+			98				115				134	144	152	159	166

VALUE TABLE 1c - AGE AND GENDER

Age	Male	Female
16 and Under	1	20
17	2	21
18	3	22
19	4	23
20	5	24
21	6	25
22	7	26
23	8	27
24	9	28
25-29	10	29
30-49	11	30
50-54	12	31
55-59	13	32
60-64	14	33
65-70	15	34
71-75	16	35
76-80	17	36
81-84	18	37
85+	19	38

RP-7A-1 ALLSTATE FIRE AND CASUALTY INSURANCE COMPANY 04-01-2014

Figure H.5: Variable Definition and Interactions

Notes: This is an excerpt from an insurer's rate filing on how observable information is used and interacted.

GEICO Casualty Company - Voluntary Private Passenger Automobile Insurance
Ohio Rate Pages Effective: New Business 06/07/2013 Renewals 07/22/2013 Rate Gen 12
Driver Class Factors

** Risk Group: B = B10, B20, and B30; C = C10, C20, C30; D = D10, D20, D30
** Z in Risk Tier represents all Risk Tiers
** Driver Age 999 = 80 and older
** RV Factor = 1.0

Risk Group	Risk Tier	Rated Vehicle Type	Coverage	Named Insured Indicator	Gender	Marital Status	Driver Age	Factor
B	Z	PP	BI	N	F	S	24	1.1660
B	Z	PP	BI	Y	M	M	24	0.9460
B	Z	PP	BI	N	M	M	24	1.1976
B	Z	PP	BI	Y	M	S	24	0.9361
B	Z	PP	BI	N	M	S	24	1.1387
B	Z	PP	BI	Y	F	M	25	0.7939
B	Z	PP	BI	N	F	M	25	0.8392
B	Z	PP	BI	Y	F	S	25	0.9649
B	Z	PP	BI	N	F	S	25	1.1458
B	Z	PP	BI	Y	M	M	25	0.9460
B	Z	PP	BI	N	M	M	25	1.1633
B	Z	PP	BI	Y	M	S	25	0.9361
B	Z	PP	BI	N	M	S	25	1.1178
B	Z	PP	BI	Y	F	M	26	0.8060
B	Z	PP	BI	N	F	M	26	0.8520
B	Z	PP	BI	Y	F	S	26	0.9649
B	Z	PP	BI	N	F	S	26	1.0819
B	Z	PP	BI	Y	M	M	26	0.9460
B	Z	PP	BI	N	M	M	26	1.1360
B	Z	PP	BI	Y	M	S	26	0.9361
B	Z	PP	BI	N	M	S	26	1.0359
B	Z	PP	BI	Y	F	M	27	0.8060
B	Z	PP	BI	N	F	M	27	0.8520
B	Z	PP	BI	Y	F	S	27	0.9649
B	Z	PP	BI	N	F	S	27	1.0525
B	Z	PP	BI	Y	M	M	27	0.9460
B	Z	PP	BI	N	M	M	27	1.0460
B	Z	PP	BI	Y	M	S	27	0.9361
B	Z	PP	BI	N	M	S	27	1.0251
B	Z	PP	BI	Y	F	M	28	0.8060
B	Z	PP	BI	N	F	M	28	0.8520
B	Z	PP	BI	Y	F	S	28	0.9649
B	Z	PP	BI	N	F	S	28	1.0398
B	Z	PP	BI	Y	M	M	28	0.9460
B	Z	PP	BI	N	M	M	28	1.0260
B	Z	PP	BI	Y	M	S	28	0.9361
B	Z	PP	BI	N	M	S	28	1.0172
B	Z	PP	BI	Y	F	M	29	0.8060
B	Z	PP	BI	N	F	M	29	0.8530
B	Z	PP	BI	Y	F	S	29	0.9649
B	Z	PP	BI	N	F	S	29	1.0118
B	Z	PP	BI	Y	M	M	29	0.9460
B	Z	PP	BI	N	M	M	29	1.0110
B	Z	PP	BI	Y	M	S	29	0.9361
B	Z	PP	BI	N	M	S	29	0.9821
B	Z	PP	BI	Y	F	M	30	0.8060
B	Z	PP	BI	N	F	M	30	0.8440
B	Z	PP	BI	Y	F	S	30	0.9649
B	Z	PP	BI	N	F	S	30	1.0100
B	Z	PP	BI	Y	M	M	30	0.9460
B	Z	PP	BI	N	M	M	30	0.9900
B	Z	PP	BI	Y	M	S	30	0.9361
B	Z	PP	BI	N	M	S	30	0.9800
B	Z	PP	BI	Y	F	M	31	0.8060
B	Z	PP	BI	N	F	M	31	0.8360
B	Z	PP	BI	Y	F	S	31	0.9648
B	Z	PP	BI	N	F	S	31	1.0010
B	Z	PP	BI	Y	M	M	31	0.9415
B	Z	PP	BI	N	M	M	31	0.9760
B	Z	PP	BI	Y	M	S	31	0.9360
B	Z	PP	BI	N	M	S	31	0.9710
B	Z	PP	BI	Y	F	M	32	0.8060
B	Z	PP	BI	N	F	M	32	0.8270
B	Z	PP	BI	Y	F	S	32	0.9648
B	Z	PP	BI	N	F	S	32	0.9900
B	Z	PP	BI	Y	M	M	32	0.9421
B	Z	PP	BI	N	M	M	32	0.9670

Figure H.6: Rating Factors based on Observables

Notes: This is an excerpt from an insurer's rate filing on how observable information is translated into pricing factors.

Progressive Direct Insurance Company
State of Ohio
New Business Effective: January 23, 2015
Renewals Effective: February 20, 2015

D06-Driving Violation Descriptions

The following chart lists the violation codes and their associated descriptions:

Violation Code	Violation Description
AAF	At Fault Accident
AFM	Accident found on MVR only at renewal - Not Chargeable
ANC	Waived Claim – Closed
ANO	Waived Claim – Open
ASW	Accident Surcharge Waived
CML	Commercial Vehicle Violation
CMP	Comprehensive Claim
CMU	Comprehensive Claim Less Than $1000
CRD	Careless or Improper Operation
DEV	Traffic Device/Sign
DR	Drag Racing
DWI	Drive Under Influence
FDL	Foreign Drivers Lic
FEL	Auto Theft/Felony Motor Vehicle
FFR	Failure to File Required Report
FLE	Fleeing from Police
FTC	Following Too Close
FTY	Failure to Yield
HOM	Vehicular Homicide
IP	Improper Passing
IT	Improper Turn
LDL	Operating Without Owner's Consent
LIC	License/Credentials Violation
LTS	Leaving the Scene
MAJ	Other Serious Violation
MMV	Minor Moving Violation
NAF	Not At Fault Accident
NFX	Waived Not At Fault Accident
PUA	Permissive Use At Fault Accident
PUN	Permissive Use Not At Fault Accident
RKD	Reckless Driving
SLV	Serious License Violations
SPD	Speeding
SUS	Driving Under Suspension
TMP	Dispute - At Fault Accident
UDR	Unverifiable Record
WSR	Wrong Way on a One Way Street

Figure H.7: Violation Captured in OH

Notes: This is an excerpt from an insurer's rate filing on the kinds of violations recorded in tier rating in Ohio.

GEICO Casualty Company - Voluntary Private Passenger Automobile Insurance
Ohio Rate Pages Effective: New Business 06/07/2013 Renewals 07/22/2013 Rate Gen 12
Accident Factors

** Risk Group: B = B10, B20, and B30; C = C10, C20, C30; D = D10, D20, D30
** Z in Risk Tier represents all Risk Tiers
** For Coverages BI,PD, COLL, COLL PP, and COLL TL Driver Age 18 = 18 and younger; 999 = 80 and older. All other Coverages Driver Age 18 = 18 and yo

Risk Group	Risk Tier	Rated Vehicle Type	Coverage	Driver Age	Number of Chargeable Occurrences	Months Since First Occurrence	Months Since Second Occurrence	Factor
B	Z	PP	BI	31	4	23	35	3.3112
B	Z	PP	BI	31	4	35	35	3.0748
B	Z	PP	BI	31	99	11	11	4.9426
B	Z	PP	BI	31	99	11	23	4.5307
B	Z	PP	BI	31	99	11	35	4.3248
B	Z	PP	BI	31	99	23	23	3.9644
B	Z	PP	BI	31	99	23	35	3.7842
B	Z	PP	BI	31	99	35	35	3.5140
B	Z	PP	BI	32	0	0	0	1.0000
B	Z	PP	BI	32	1	11	0	1.6375
B	Z	PP	BI	32	1	23	0	1.3267
B	Z	PP	BI	32	1	35	0	1.2320
B	Z	PP	BI	32	2	11	11	2.2925
B	Z	PP	BI	32	2	11	23	2.1014
B	Z	PP	BI	32	2	11	35	2.0059
B	Z	PP	BI	32	2	23	23	1.6550
B	Z	PP	BI	32	2	23	35	1.5797
B	Z	PP	BI	32	2	35	35	1.4669
B	Z	PP	BI	32	3	11	11	3.5525
B	Z	PP	BI	32	3	11	23	3.2565
B	Z	PP	BI	32	3	11	35	3.1083
B	Z	PP	BI	32	3	23	23	2.8493
B	Z	PP	BI	32	3	23	35	2.7199
B	Z	PP	BI	32	3	35	35	2.5256
B	Z	PP	BI	32	4	11	11	4.3248
B	Z	PP	BI	32	4	11	23	3.9644
B	Z	PP	BI	32	4	11	35	3.7842
B	Z	PP	BI	32	4	23	23	3.4689
B	Z	PP	BI	32	4	23	35	3.3112
B	Z	PP	BI	32	4	35	35	3.0748
B	Z	PP	BI	32	99	11	11	4.9426
B	Z	PP	BI	32	99	11	23	4.5307
B	Z	PP	BI	32	99	11	35	4.3248
B	Z	PP	BI	32	99	23	23	3.9644
B	Z	PP	BI	32	99	23	35	3.7842
B	Z	PP	BI	32	99	35	35	3.5140
B	Z	PP	BI	33	0	0	0	1.0000
B	Z	PP	BI	33	1	11	0	1.6375
B	Z	PP	BI	33	1	23	0	1.3267
B	Z	PP	BI	33	1	35	0	1.2320
B	Z	PP	BI	33	2	11	11	2.2925
B	Z	PP	BI	33	2	11	23	2.1014
B	Z	PP	BI	33	2	11	35	2.0059
B	Z	PP	BI	33	2	23	23	1.6550
B	Z	PP	BI	33	2	23	35	1.5797
B	Z	PP	BI	33	2	35	35	1.4669
B	Z	PP	BI	33	3	11	11	3.5525
B	Z	PP	BI	33	3	11	23	3.2565
B	Z	PP	BI	33	3	11	35	3.1083
B	Z	PP	BI	33	3	23	23	2.8493
B	Z	PP	BI	33	3	23	35	2.7199
B	Z	PP	BI	33	3	35	35	2.5256
B	Z	PP	BI	33	4	11	11	4.3248
B	Z	PP	BI	33	4	11	23	3.9644
B	Z	PP	BI	33	4	11	35	3.7842

Figure H.8: Tier Factors

Notes: This is an excerpt from an insurer's rate filing on how tier information is rated.

Progressive Direct Insurance Company (DI)
Progressive Specialty Insurance Company (AG)
Ohio Private Passenger Automobile Program
Effective Date: January 23, 2015

Usage-based Insurance Factor Table - Initial Discount (DI Experience) Exhibit: 9C

UBI SCORE	BI/PD	COLL	COMP	LOAN	MED	RENT	ROADSIDE	UMPD	ACPE	OPERATIONS EXPENSE	ACQUISITION EXPENSE
0	0.56	0.56	0.96	0.96	0.56	0.56	0.96	0.56	0.96	1.00	1.00
1	0.61	0.61	0.96	0.96	0.61	0.61	0.96	0.61	0.96	1.00	1.00
2	0.65	0.65	0.97	0.97	0.65	0.65	0.97	0.65	0.97	1.00	1.00
3	0.75	0.74	0.97	0.97	0.75	0.74	0.97	0.75	0.97	1.00	1.00
4	0.79	0.79	0.97	0.97	0.79	0.79	0.97	0.79	0.97	1.00	1.00
5	0.83	0.83	0.97	0.97	0.83	0.83	0.97	0.83	0.97	1.00	1.00
6	0.86	0.87	0.97	0.97	0.86	0.87	0.97	0.86	0.97	1.00	1.00
7	0.89	0.89	0.97	0.97	0.89	0.89	0.97	0.89	0.97	1.00	1.00
8	0.89	0.90	0.97	0.97	0.89	0.90	0.97	0.89	0.97	1.00	1.00
9	0.89	0.91	0.97	0.97	0.89	0.91	0.97	0.89	0.97	1.00	1.00
10	0.90	0.90	0.97	0.97	0.90	0.90	0.97	0.90	0.97	1.00	1.00
11	0.90	0.90	0.97	0.97	0.90	0.90	0.97	0.90	0.97	1.00	1.00
12	0.90	0.90	0.98	0.98	0.90	0.90	0.98	0.90	0.98	1.00	1.00
13	0.91	0.89	0.98	0.98	0.91	0.89	0.98	0.91	0.98	1.00	1.00
14	0.91	0.88	0.98	0.98	0.91	0.88	0.98	0.91	0.98	1.00	1.00
15	0.91	0.90	0.98	0.98	0.91	0.90	0.98	0.91	0.98	1.00	1.00
16	0.92	0.90	0.98	0.98	0.92	0.90	0.98	0.92	0.98	1.00	1.00
17	0.92	0.91	0.98	0.98	0.92	0.91	0.98	0.92	0.98	1.00	1.00
18	0.92	0.91	0.98	0.98	0.92	0.91	0.98	0.92	0.98	1.00	1.00
19	0.92	0.92	0.98	0.98	0.92	0.92	0.98	0.92	0.98	1.00	1.00
20	0.92	0.92	0.98	0.98	0.92	0.92	0.98	0.92	0.98	1.00	1.00
21	0.92	0.92	0.98	0.98	0.92	0.92	0.98	0.92	0.98	1.00	1.00
22	0.92	0.92	0.98	0.98	0.92	0.92	0.98	0.92	0.98	1.00	1.00
23	0.92	0.92	0.98	0.98	0.92	0.92	0.98	0.92	0.98	1.00	1.00
24	0.93	0.93	0.98	0.98	0.93	0.93	0.98	0.93	0.98	1.00	1.00
25	0.93	0.93	0.98	0.98	0.93	0.93	0.98	0.93	0.98	1.00	1.00
26	0.93	0.93	0.98	0.98	0.93	0.93	0.98	0.93	0.98	1.00	1.00
27	0.93	0.93	0.99	0.99	0.93	0.93	0.99	0.93	0.99	1.00	1.00
28	0.93	0.94	0.99	0.99	0.93	0.94	0.99	0.93	0.99	1.00	1.00
29	0.93	0.94	0.99	0.99	0.93	0.94	0.99	0.93	0.99	1.00	1.00
30	0.94	0.94	0.99	0.99	0.94	0.94	0.99	0.94	0.99	1.00	1.00
31	0.94	0.94	0.99	0.99	0.94	0.94	0.99	0.94	0.99	1.00	1.00
32	0.94	0.94	0.99	0.99	0.94	0.94	0.99	0.94	0.99	1.00	1.00
33	0.94	0.94	0.99	0.99	0.94	0.94	0.99	0.94	0.99	1.00	1.00
34	0.95	0.95	0.99	0.99	0.95	0.95	0.99	0.95	0.99	1.00	1.00
35	0.95	0.95	0.99	0.99	0.95	0.95	0.99	0.95	0.99	1.00	1.00
36	0.95	0.95	0.99	0.99	0.95	0.95	0.99	0.95	0.99	1.00	1.00
37	0.95	0.95	0.99	0.99	0.95	0.95	0.99	0.95	0.99	1.00	1.00
38	0.95	0.95	0.99	0.99	0.95	0.95	0.99	0.95	0.99	1.00	1.00
39	0.95	0.96	0.99	0.99	0.95	0.96	0.99	0.95	0.99	1.00	1.00

Note:
-The premium-weighted average factor for the vehicle is calculated and applied to all coverages for the vehicle as indicated in the Rate Order of Calculation. This factor cannot be lower than 0.70 or greater than 1.0.

-If a vehicle does not participate in the Usage-based Insurance program it is assigned a 1.0 factor.

Figure H.9: Violation Captured in OH

Notes: This is an excerpt from an insurer's rate filing on how monitoring pricing is filed.

Progressive Direct Insurance Company (DI) & Progressive Specialty Insurance Company (AG)
Private Passenger Automobile Program
Supporting Exhibits for the State of Ohio
Effective Date: September 5, 2014
Coverage: BI **Exhibit 10Y**

Limit Factor

Experience	Has Prior Insurance	Limit	Incurred Loss Capped	Indicated Factor	Proposed Factor	Current Factor	Percent Change
AG	N	$25,000/$50,000	243,943,611	1.00	1.00	1.00	0.0%
AG	N	$50,000/$100,000	102,950,757	1.16	1.08	1.08	0.0%
AG	N	$100,000 CSL	1,444,950	1.24	1.11	1.11	0.0%
AG	N	$100,000/$300,000	70,326,408	1.54	1.29	1.29	0.0%
AG	N	$300,000 CSL	3,758,408	2.04	1.50	1.50	0.0%
AG	N	$250,000/$500,000	9,874,286	2.15	1.68	1.68	0.0%
AG	N	$500,000 CSL	5,350,267	2.25	1.80	1.80	0.0%
AG	Y	$25,000/$50,000	302,253,249	1.00	1.00	1.00	0.0%
AG	Y	$50,000/$100,000	256,452,902	1.21	1.13	1.12	0.9%
AG	Y	$100,000 CSL	7,102,129	1.26	1.19	1.16	2.6%
AG	Y	$100,000/$300,000	388,729,047	1.53	1.37	1.33	3.0%
AG	Y	$300,000 CSL	25,394,374	1.85	1.45	1.46	-0.7%
AG	Y	$250,000/$500,000	85,216,412	2.10	1.69	1.80	-6.1%
AG	Y	$500,000 CSL	45,591,859	2.15	1.93	1.95	-1.0%
DI	N	$25,000/$50,000	94,310,074	1.00	0.95	0.95	0.0%
DI	N	$50,000/$100,000	71,807,198	1.16	1.00	1.00	0.0%
DI	N	$100,000 CSL	81,354	1.27	1.11	1.11	0.0%
DI	N	$100,000/$300,000	45,810,439	1.54	1.28	1.28	0.0%
DI	N	$300,000 CSL	254,864	1.56	1.41	1.41	0.0%
DI	N	$250,000/$500,000	10,296,001	2.00	1.49	1.49	0.0%
DI	N	$500,000 CSL	440,458	2.16	1.59	1.59	0.0%
DI	Y	$25,000/$50,000	182,880,315	1.00	1.00	1.00	0.0%
DI	Y	$50,000/$100,000	199,882,577	1.15	1.05	1.05	0.0%
DI	Y	$100,000 CSL	1,287,766	1.22	1.17	1.17	0.0%
DI	Y	$100,000/$300,000	286,763,971	1.40	1.33	1.33	0.0%
DI	Y	$300,000 CSL	4,867,338	1.74	1.39	1.39	0.0%
DI	Y	$250,000/$500,000	53,447,656	1.82	1.47	1.47	0.0%
DI	Y	$500,000 CSL	5,998,809	2.13	1.60	1.60	0.0%

Figure H.10: Tier Factors
Notes: This is an excerpt from an insurer's rate filing on how limit choices influence pricing.