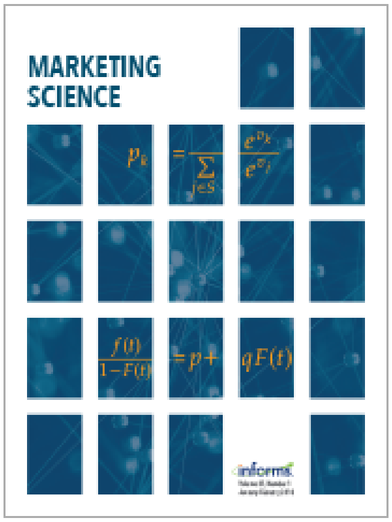


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Skimming from the Bottom: Empirical Evidence of Adverse Selection When Poaching Customers

Przemysław Jeziorski, Elena Krasnokutskaya, Olivia Ceccarini

To cite this article:
Przemysław Jeziorski, Elena Krasnokutskaya, Olivia Ceccarini (2019) Skimming from the Bottom: Empirical Evidence of Adverse Selection When Poaching Customers. Marketing Science 38(4):543-566. https://doi.org/10.1287/mksc.2018.1134

informs
http://pubsonline.informs.org/journal/mksc/

MARKETING SCIENCE
Vol. 38, No. 4, July–August 2019, pp. 543–566
ISSN 0732-2399 (print), ISSN 1526-548X (online)

Skimming from the Bottom: Empirical Evidence of Adverse Selection When Poaching Customers

Przemysław Jeziorski,[a] Elena Krasnokutskaya,[b] Olivia Ceccarini[c]

[a] University of California at Berkeley, Berkeley, California 94720; [b] Johns Hopkins University, Baltimore, Maryland 21218; [c] Porto Business School, University of Porto, 4460-312 Matosinhos, Portugal

Contact: przemekj@berkeley.edu, http://orcid.org/0000-0001-8351-0708 (PJ); ekrasno1@jhu.edu (EK); olivia.ceccarini@gmail.com (OC)

Received: January 13, 2017
Revised: October 18, 2017; January 23, 2018
Accepted: January 30, 2018
Published Online in Articles in Advance: July 17, 2019

https://doi.org/10.1287/mksc.2018.1134

Abstract. This paper studies implications of competitive customer poaching in markets with heterogeneous and privately known costs to serve. Using individual-level driving records from a large car insurer in Portugal, we show that poached customers generate a 21% higher cost to serve than observationally equivalent own customers. Screening on all available consumer characteristics and behavioral variables, with the exception of switching behavior, can alleviate only 50% of adverse selection. We develop and estimate an empirical framework based on a dynamic churn model that rationalizes this adverse selection. Our estimates imply that risky customers have more incentive to search and switch, and that the population of switchers is itself heterogeneous in riskiness. We propose a new consumer lifetime value measure that accounts for switchers' risk endogeneity. We apply this measure to study actuarial pricing and insurance contract design.

History: K. Sudhir served as the editor-in-chief and Harikesh Nair served as associate editor for this article.

Supplemental Material: The estimation code and online appendix are available at https://doi.org/10.1287/mksc.2018.1134.

Keywords: customer poaching • adverse selection • unobserved heterogeneity • cost to serve • behavior-based pricing • structural model

1. Introduction

In many industries, consumers are differentially expensive to serve. A canonical example is provided by insurance markets, where consumers differ by their inherent riskiness. However, differential cost to serve is present in many other industries, such as credit markets, services, and retail. The heterogeneity in cost to serve, when private information of the customer, may result in adverse selection and lead to inferior profitability or even market failures (see Akerlof 1970).

Prior marketing literature has examined the theoretical implications of heterogeneous cost to serve in the context of differential monopoly pricing between good and bad customers (see Shin et al. 2012). Building on these insights, this paper examines empirical implications of heterogeneous cost to serve in a competitive setting. Many industries have long recognized that new customers are different from old customers, and therefore offer differential pricing, such as higher interest rates for people with no credit history and higher insurance rates for new drivers. At the same time, firms frequently poach competitors' customers using targeted advertising and discounts;[1] examples include lower refinance rates for mortgages, lower interest rates for balance transfers, subsidies for switchers in the cell phone industry, and discretionary discounts for drivers insured with the competitor.[2] Such strategies are becoming more widespread with the advent of information technology and availability of big data including consumer purchase history. The literature provides some evidence that these strategies may be beneficial to firms when consumers have homogeneous cost to serve and heterogeneous willingness to pay (see Fong et al. 2015). However, this may no longer be true when cost to serve is heterogeneous and unobserved to the firm. Advantageous selection resulting in lower switchers' cost to serve reinforces the attractiveness of poaching strategies. Conversely, adverse selection resulting in higher switchers' cost may significantly impede the effectiveness of poaching or even lead to losses. Thus, firms should evaluate the extent of adverse selection when designing competitive promotions and pricing strategies.

This paper provides the empirical evidence for adverse selection when poaching competitors' customers.[3] We provide three sets of results. First, we empirically identify a gap in the cost to serve between consumers that switched from the competitor and the company's otherwise observationally equivalent own customers. Second, we propose and provide evidence in favor of strategic selective attrition as a mechanism behind the cost gap. Third, we use a structural model to develop a new measure of consumer lifetime value (LTV) that accounts for adverse selection and endogeneity of cost to serve. We use the model to evaluate several pricing and contract policies aimed at combating adverse selection.

The evidence is obtained by analyzing the Portuguese car insurance industry, an established, multibillion-dollar market.[4] Using individual-level data on insurance claims from a leading Portuguese car insurer, we show that the average switcher generates a 32% larger volume of liability claims than the average nonswitcher. Moreover, we find that commonly employed actuarial screening mechanisms can only partially alleviate this problem. In particular, screening based on observable characteristics and detailed driving history accounts for less than 50% of the adverse selection. Specifically, the average switcher is approximately 20% more risky than the observationally equivalent nonswitcher. This suggests that drivers exhibit a large degree of unobservable heterogeneity in riskiness, and that larger risk is correlated with switching. Current pricing does not reflect this risk gap. Switchers obtain 1.5 percentage point higher discretionary discounts; thus, they pay lower premiums than observationally equivalent nonswitchers.

To design a pricing policy that can address the riskiness gap, we need to describe what the nature of the gap across customer segments is. If the riskiness gap is the same across all consumers, a simple surcharge for switchers would correctly price in the extra risk. However, if the riskiness gap depends on customer characteristics, the optimal pricing would have to incorporate heterogeneous switcher surcharges.

First, we examine whether risk gap depends on observable characteristics. We find that the gap depends on tenure; that is, customers who are with the company for one to two years are significantly more risky than the otherwise equivalent customers with three or more years of tenure. The relationship of riskiness and tenure is flat beyond three years of tenure. To explain this pattern, we analyze the customers that cancel their contracts. We find that 20% of clients churn within one year; thus, some customers frequently switch insurance companies. Moreover, 35% of customers that incur a claim do not renew their contract. Such selective attrition can explain the relationship between tenure and riskiness. We also find that switchers with bad driving histories generate a 100% larger volume of claims than observationally equivalent nonswitchers with similar histories. In comparison, switchers with excellent driving histories generate a 38% larger volume of claims. These results suggest that the switcher surcharge decrease with tenure and that it should be lower for switchers with better driving histories.

Furthermore, we test for the existence of the unobserved variation in riskiness gap between switchers and the company's own clients. For this purpose, we demonstrate that two observationally equivalent switchers have heterogeneous riskiness. We use a test for unobserved heterogeneity developed by Puelz and Snow (1994) and Chiappori and Salanie (2000). Particularly, we show that switchers with extra collision insurance generate a 27% larger volume of liability claims than otherwise equivalent switchers with minimum coverage. Consequently, we find statistically significant unobserved heterogeneity in riskiness among switchers. The direct implication of this result is that even the most sophisticated switcher surcharges, those that vary across observable segments, may not be able to price in the entirety of the riskiness gap. We explore this point further using a structural model.[5]

To study the implications of the switcher–stayer risk gap for pricing and contract design, we develop a dynamic churn model that rationalizes this gap. We postulate a model in which heterogeneous consumers perform costly comparative shopping and switch to the competitor when offered a lower premium. Consumers can also decide to stop driving with an option of coming back to the market in the future. The model allows the consumers to be heterogeneous in their riskiness and search cost. This allows the framework to accommodate (i) markets in which switching is driven by risk heterogeneity, producing a large switcher–stayer gap, and (ii) markets in which switching is driven by nonrisk heterogeneity, producing a small or no switcher–stayer gap. The extent of these forces is driven by the data.

We identify the model in two steps. In the first step, we recover the distribution of risk in the population from the variation in the realized risk across drivers with different driving histories. In the second step, we separate incentives to switch to the competitor from the incentives to quit, controlling for the distribution of risk. To estimate the incentives to churn, we use variability in churn rates across consumers with the same demographics and riskiness that pay different premiums. The key identifying variation is provided by the institutional feature of the Portuguese car insurance market, that is, the prevalence of driving history discounts and discretionary discounts. Driving history discount is common across insurers, whereas the discretionary discount is insurer specific.[6] Drivers with a low driving history discount and high discretionary discount have little incentive to switch to the competitor, because they are unlikely to obtain a lower premium. Conversely, drivers with a high driving history discount and low discretionary discount have relatively more incentive to search than to quit. Exploring this mechanism, we identify incentives to quit using the variation in churn rates across driving history discounts for drivers with high discretionary discounts. Similarly, we identify incentives to search using the variation in churn rates across discretionary discounts for drivers with high driving history discounts.

We show that the model can explain the riskiness gap between switchers and nonswitchers, as well as the large size of the gap in realized risk between churners and nonchurners. Our estimates imply that 31% of observed churners switch to the competitor, whereas the remaining churners quit driving. The model implies that, in the estimation subsample, switchers and quitters are respectively 8% and 31% riskier than nonchurners. The gap obtains endogenously, because the riskiest drivers have the most incentives to quit, moderately risky drivers have the most incentives to search, and the safest drivers have little incentives to either search or quit. The greater search incentives of risky drivers explain why switchers obtain higher discretionary discounts, despite their higher riskiness. Importantly, this holds even when companies do not explicitly price discriminate on either switching or the unobserved portion of the riskiness.

We use the model to develop a new LTV measure. We define LTV as the discounted stream of profits from the customer under the customer's optimal behavior. The LTV measure builds on ideas from customer relationship management literature (see Reinartz et al. 2004, Venkatesan and Kumar 2004) and accounts for selective churn and return of customers. Because our measure takes into account that the customer behavior changes, as we alter pricing and contract structure, it enables us to conduct robust contract and pricing counterfactuals. We demonstrate that accounting for this endogeneity has consequences when evaluating actuarial pricing.

Excluding other operational costs besides claims, we show that the average LTV is €339 for an average own customer, and €302 for a switcher, respectively. The gap reflects (i) the riskiness gap, and (ii) differential premiums paid by switchers and stayers. The immediate prescription of actuarial pricing theory would be to raise the premiums on switchers to cover losses from extra underwritten risks. This prescription assumes that the riskiness of switchers is fixed. To the contrary, we show that a uniform price increase for switchers discourages better customers of the competitor from searching. This results in higher switchers' riskiness and overall deterioration in the company's risk pool. In particular, decreasing the average discretionary discount by approximately 50% increases the overall riskiness of the customer pool by approximately 4%. This results in a 2.3% decrease in the LTV of the company's own clients and a staggering 84% decrease in the LTV of switchers. This has immediate implications for firms, who must consider not only the impact of the change in premiums on their market share, but also on the risk pool of their customers and switchers.

We examine two ways to combat adverse selection and compare their relative effectiveness. First, we document that charging higher prices only to riskier switchers alleviates adverse selection and decreases the riskiness of the customer pool. Specifically, decreasing the offered discount by 50% (equivalent to an approximately 2.5 percentage point increase in transacted premiums) but only to those switchers that are riskier than average, leads to a 6% improvement in the risk pool and a 5.1% increase in the LTV of the company's own client pool. This result shows that even coarse information on the unobserved portion of switchers' riskiness can be pivotal for the effectiveness of actuarial pricing.

Furthermore, we allow for one firm to unilaterally deviate to a steeper incentive contract. A steeper contract should discourage riskier switchers from searching, which should have a similar effect as the differential price increase. We examine a range of counterfactual dynamic contracts and show that only large changes have economically significant effects; namely, a 100% increase of surcharges for drivers with bad driving histories has a negligible impact on the selection of drivers. An increase in incentives by 500% has an effect comparable to that of a selective 50% increase in the discretionary discount (2.5 percentage point increase in transacted premiums). Because such dramatic changes in the incentive structure may be hard to implement, we conclude that using a dynamic contract to substitute for extra information may have limited practical effectiveness.[7]

The closest paper to this work is that by Jeziorski et al. (2017). Whereas we concentrate our attention on switchers, Jeziorski et al. (2017) study the subset of the same data that includes only consumers that never switch insurance providers. Using this subsample, Jeziorski et al. (2017) provide a complementary set of results documenting the importance of moral hazard and adverse selection within the same firm, across contracts and risk classes. We use these results to conduct several robustness checks relaxing some of our assumptions (e.g., to account for moral hazard and heterogeneous risk aversion), which are presented in the online appendix.

This paper is organized as follows. Section 2 contains the description of Portuguese car insurance industry. The data description is contained in Section 3. The descriptive evidence of adverse selection is presented in Section 4. Section 5 contains a structural analysis and pricing counterfactuals. Section 6 concludes.

2. Industry Description

According to 2011 Organisation for Economic Cooperation and Development (OECD) data, Portuguese motor vehicle ownership is comparable with that in other developed European economies. In particular, per 100 residents, there are 55 privately owned motor vehicles in Portugal, 55 in Germany, 52 in Great Britain, and 60 in France (OECD 2013). Such high density of car ownership, along with the fact that insurance is compulsory, results in a saturated motor insurance market. Including commercial vehicles, as of 2012, Portugal, with a population of 10.5 million, had approximately 7 million registered motor insurance policies. Total insured assets amounted to €120 billion, and total industry revenue to €2 billion. The revenues from liability premiums amounted to approximately €1 billion, whereas the total cost of liability claims amounted to €800 million. Thus, the industry profit margin before subtracting operating costs amounted to approximately 20% (APS 2013).

The consumers are offered a menu of insurance contracts that includes (i) a compulsory liability insurance contract fully covering damage to the counterparty's car in the case of an at-fault accident and (ii) a set of collision insurance contracts with varying deductibles covering damage to one's own car in the case of an at-fault accident or an accident with no counterparty.

The contracts are priced using a set of variables describing the driver, the car, and the location. In particular, the driver demographic characteristics used in pricing are gender, years since obtaining a driving license, and age. The car characteristics are age, make, horsepower, actuarial value, and weight. The location is defined by the postal code. In addition, insurers employ a dynamic contract with behavior-based pricing (BBP) using a risk-class rating scheme. Under this system, each driver is placed in one of the 18 risk classes depending on their claim history. New drivers start in class 10. Every year, the risk class is updated: if the policyholder did not have any claims in the previous year, then his risk class is reduced by one. For every claim that he had in the previous year he is moved three classes up. Policyholders in classes below the reference class are given a discount over the base premium. Policyholders in classes higher than the reference class pay a surcharge over the base premium. The risk-class transition depends exclusively on the policyholder's number of claims in previous year and not on driver characteristics, vehicle characteristics, or the volume of each claim. In addition, only claims in which the policyholder is at least partially at fault trigger upward transition.

Pricing of the basic and collision parts of the insurance contract are based on separate risk classes. Thus, drivers are assigned basic and collision risk classes that depend on the past claims on their basic and collision policies, respectively. For example, the accidents with no counterparty do not affect the basic risk class because no liability claim is necessary. Also, the drivers with no collision contract cannot submit collision claims by construction; thus, their latent collision risk class decreases by one every year.

Table 1 summarizes the slope of the premium function with respect to the risk class. Experience rating schemes and base premiums are freely set by the insurance company, but are subject to regulatory approval by the supervising authority (for a historical context of these regulations, see Barros 1996). The exact risk-class fee tables are part of physical insurance contracts; thus, they are known by all market participants. In the current equilibrium, the largest insurance carriers employ very similar risk-class rating schemes and similar risk-class pricing schedules.

Although the history of an individual's claims is not public knowledge, in practice, less than 0.1% of drivers do not bring their driving history when switching insurance providers. A policyholder who switches insurance companies and is not providing his new insurer with his driving record is penalized by a placement in class 16 and thus surcharged 250% over the baseline premium.[8] Effectively, the industry employs an information-sharing program, by which the risk classes are common knowledge.

The risk class is quite informative of the driving history of the insured; nevertheless, some information about past claims is lost during the aggregation. For example, an individual in risk class 1 did not have a claim in the last 3 years and had at most two claims in the last 10 years. An individual in risk class 9 either

Table 1. Scaling Coefficient for Various Risk Classes

Risk class	Liability insurance (%)	Collision insurance (%)
1	45	45
2	45	45
3	50	45
4	55	45
5	60	60
6	65	65
7	70	70
8	80	80
9	90	90
10	100	100
11	110	110
12	120	120
13	130	130
14	150	150
15	180	150
16	250	150
17	325	150
18	400	150

had no accidents last year but had a bad (or insufficient) driving history before that, or had one or more accidents last year but had a relatively good driving history before that. Thus, we expect that even after conditioning on demographic characteristics and risk class, there is significant residual private information regarding risk.

In Portugal, insurance contracts are mainly sold via sales-force agents. Agents can offer discretionary premium discounts to prospects. The amount of the discount is determined by bargaining between the insured and the agent. The personal interaction between the agent and the client can provide the former with additional risk-related information beyond what is known to the firm (such as wealth or behavioral cues). Thus, if the agent is able to allocate discounts based on this extra information, the delegation of pricing should improve screening. However, if the sales-force agent does not have superior information, allocates discounts based on nonrisk characteristics (such as the bargaining skills of a client), or has misaligned contractual incentives, the delegation of pricing would not improve or may even decrease the efficiency of screening. We provide evidence for the latter in Section 4.3.

The next section describes the data used in the analysis.

3. Data

We obtained the complete data set of insurance policies and claims from one of the largest auto insurers in Portugal. The company is part of a large international insurance conglomerate and operates across all geographic areas in Portugal. It offers a full range of vehicle insurance products such as auto, motorcycle, and boat insurance. The data are a panel of all insured individuals for the years 2007–2012 and contain all information about these insured individuals available to the firm. Beyond the individual characteristics, we observe the types of contracts selected, premiums paid, and the amounts of discretionary discounts. The observed premiums reflect the exact amounts paid and include all applicable taxes passed through to the customer. To standardize the analysis, we drop individuals for whom the length of the contract is less than one year and for whom any of the demographic variables are unobserved. We also drop commercial vehicles and motorcycles. We obtain a sample of 439,639 unique individuals with an average of 2.36 years of data per individual. Each observation is a person-year combination.

The data include the number and volume of insurance claims for each customer. We henceforth use these numbers to measure the cost to serve. This measure has some limitations. In particular, observed claim volumes account for first-time claim assessments only and do not include claims management costs and claim readjustments.[9] According to the company annual reports, the latter two items can amount to as much as 20% of the first-time claim assessments. Also, our measure does not include operating costs such as agent commissions and fixed costs, which can amount to as much as 50% of gross claims written. Thus, the observed claim costs and premiums alone do not determine net accounting profits per customer. Net profits can be obtained from the annual reports, which disclose negligible or negative probability of the motor insurance division. The observed claim costs can be used to rank customers according to cost to serve as long as the unobserved costs are not systematically different across consumers. In the case of the car insurance industry, this assertion is mild because the majority of the cost heterogeneity is likely to be captured by the heterogeneity in the observed frequency and volume of filed claims.

For the purpose of our analysis, we define switchers as observations in their first year with the firm that joined directly from competition or have a past driving history with a negligible insurance gap. These two categories are observationally equivalent to the firm. Switchers bring their driving history and are accordingly assigned to one of the 18 risk classes. All the other observations are labeled as being nonswitchers, which includes clients with tenure of more than one year and new clients without driving history.

We observe tenure; thus, we can identify drivers that are with the company for more than a year, which we label as nonswitchers. Observations in their first year with the firm can be of one of three categories: drivers with new licenses, drivers with old licenses but no driving history, and switchers. We can identify the first category of observations as those that join the company within a year of obtaining their license. The second category of observations are always assigned to risk class 10. The new clients in classes other than 10 have a recent driving history, so they are switchers. Drivers in risk class 10 are either nonswitchers or they are switchers that reached risk class 10 organically. The probability of reaching class 10 organically is less than 0.5%, whereas the proportion of drivers with old licenses in risk class 10 is more than 5%. Thus, the share of switchers among drivers with old licenses in risk class 10 is less than 10%. Without much loss of generality, we label all new clients in risk class 10 as nonswitchers. Mislabeling switchers in class 10 as nonswitchers lowers our estimate of the riskiness of switchers, as reaching class 10 organically is associated with bad or insufficient driving history. As a result, the estimates in this paper should be viewed as conservative.

4. Descriptive Analysis

In this section, we provide descriptive evidence of adverse selection during consumer switching across firms. First, we conduct an analysis that compares incoming switchers to the firm's own customers. We study the extent to which current behavior-based pricing alleviates adverse selection. Second, we conduct a corresponding analysis of outgoing churners. We demonstrate that filing a claim is related to subsequent churn. Third, we examine whether pricing delegation can alleviate adverse selection. The results about delegation facilitate identification of the structural model presented in the next section.

4.1. Switchers

Table 2 contains descriptive statistics of our sample divided into switchers and nonswitchers. The population is 19% switchers. Overall, the observable characteristics of both populations are comparable. Switchers are slightly younger than nonswitchers and have slightly less driving history. This difference is not driven by outliers and persists over the whole distribution; that is, switchers under 30 are younger than nonswitchers under 30, and switchers over 50 are younger than nonswitchers over 50. Nonswitchers are also driving newer and more expensive cars with more horsepower.

Switchers and nonswitchers differ by driving history. Nonswitchers tend to have better driving history and occupy an average risk class of 1.9 compared with an average risk class of 2.2 occupied by switchers. This difference indicates statistically significant observed heterogeneity of driving ability between switchers and nonswitchers. Conversely, switchers occupy lower collision risk classes than nonswitchers. We observe the collision risk class only if the individual buys a collision contract, so this comparison is polluted by selection. The selection is not present when working with liability contracts because such contracts are compulsory.

The 11th to 17th rows of Table 2 present the comparison of pricing between switchers and nonswitchers. The baseline price is the premium incorporating observable driver and car characteristics, without including discretionary discount and discount based on driving history. Nonswitchers are assigned slightly lower baseline liability premiums, but pay slightly higher collision premiums. Switchers obtain higher discretionary and driving history discounts. As a result, their final liability and collision premiums are lower than those of nonswitchers. Like collision risk classes, the collision premiums are hard to compare because of selection. Nevertheless, it is instructive to look at these premiums to exclude the possibility that competitive poaching is driven mostly by the profits from collision contracts.

The 18th to 21st rows of Table 2 compare the cost to serve for switchers and nonswitchers. An average nonswitcher has approximately 0.04 claims per year, which is about 20% lower than the number of claims for switchers. Similarly, on average, switchers generate about 13% more collision claims than nonswitchers (modulo selection into the collision contract). Average cost to serve is presented in the last two rows. Switchers are €19 per year more expensive to serve than nonswitchers when it comes to liability insurance, which is about 10% of the average premium. The difference is even greater for drivers with collision contracts and amounts to €50 per year, which is approximately 12% of the average collision premium. These differences are significant, because, as we noted earlier, the variable profit margin in the whole industry, not accounting for agent commissions, wages, and other fixed costs, amounts to only 20%. More importantly, the net profit of the motor division of the focal insurer, not including returns on capital investments, is usually negative.

The descriptive analysis yields that switchers pay lower premiums than nonswitchers and are more expensive to serve. However, it is useful to know how much the difference in cost to serve is observable to the firm, and thus could be potentially passed through to the consumer. Because the firm is allowed to price discriminate on observable characteristics, the efficiency and profitability of the market is tightly related to the fraction of "lemons" among switchers that can be weeded out by pricing on observables.

Next we compare cost metrics of switchers to observationally equivalent nonswitchers.[10] We regress the number of liability claims on the *switcher* dummy and other observable characteristics and present the results in Table 3. The *switcher* dummy represents the residual risk difference between switchers and nonswitchers. Column (1) contains a baseline comparison without using observables, but including year-month dummies to control for seasonal unobservables and dummies for new clients with new driving licenses and new clients with old driving licenses. We show that switchers generate 0.0074 more claims per year than nonswitchers. Not surprisingly, both new drivers and new clients that have not driven for a while are significantly riskier than average own customers.

In column (2), we include results with a full set of dummies for age, number of years since obtaining a driving license, and gender. We find that differences in demographics can explain only approximately 14% of the gap between switchers and nonswitchers. Whereas we can compare switchers to the company's own drivers with an equal number of years since obtaining a driving license, we cannot do the same for drivers with new licenses. To quantify the difference between new and seasoned drivers, we use the

Table 2. Descriptive Statistics

	Total			
	Mean	SD	Nonswitcher mean	Switcher mean
Switcher	0.19 (0.0003)	0.40	0	1
Age	48.0 (0.014)	12.6	48.0 (0.014)	44.4 (0.028)
Years since driving license	23.1 (0.011)	10.3	23.1 (0.011)	20.0 (0.023)
Car age	10.5 (0.006)	5.4	10.5 (0.006)	9.9 (0.012)
Car value (€)	6,991.5 (7.07)	6,511.9	6,991.5 (7.07)	7,045.0 (14.89)
Car horsepower	85.6 (0.03)	29.1	85.6 (0.03)	86.3 (0.07)
Car weight (kg)	1,377.2 (0.6)	565.0	1,377.2 (0.6)	1,359.4 (1.2)
Comprehensive contract	0.134 (0.0004)	0.339	0.134 (0.0004)	0.128 0.(0007)
Risk class, liability	1.95 (0.002)	2.00	1.95 (0.002)	2.29 (0.004)
Risk class, comprehensive	2.44 (0.007)	2.46	2.48 (0.008)	2.28 (0.011)
Baseline price, liability (€)	507.1 (0.14)	131.3	507.1 (0.14)	502.0 (0.30)
Baseline price, comprehensive (€)	995.6 (0.960)	356.4	982.7 (1.042)	1,051.5 (2.380)
Discretionary discount (%)	11.7 (0.008)	6.9	11.7 (0.008)	13.2 (0.014)
Driving discount, liability (%)	51.1 (0.012)	10.6	51.1 (0.012)	51.7 (0.018)
Driving discount, comprehensive (%)	50.0 (0.036)	13.3	49.5 (0.042)	52.5 (0.050)
Final price, liability (€)	217.6 (0.085)	76.7	217.6 (0.085)	209.5 (0.158)
Final price, comprehensive (€)	425.2 (0.51)	189.7	425.7 (0.58)	423.2 (1.04)
Claims, liability	0.039 (0.0002)	0.208	0.039 (0.0002)	0.048 (0.0005)
Claims, comprehensive	0.075 (0.001)	0.280	0.073 (0.001)	0.083 (0.002)
Claim volume, liability (€)	69.49 (1.47)	1,360.93	69.49 (1.47)	87.75 (3.15)
Claim volume, comprehensive (€)	179.7 (3.09)	1,146.0	170.6 (3.23)	219.4 (8.70)

Notes. Collision statistics are conditional on buying a collision contract. Standard errors are in parentheses. SD, standard deviation.

company's own drivers with 10 years since obtaining a driving license as a baseline comparison group. The resulting coefficient for drivers with new licenses is greater than the corresponding coefficient in column (1). This increase merely indicates that the gap between drivers with new licenses and drivers with 10 years since obtaining a driving license is greater than the corresponding gap between drivers with new licenses and drivers with an average number of years since obtaining a license.

Column (3) includes postal-code controls. The firm classifies the postal codes into four bins based on riskiness, and we include a dummy for each bin. Postal codes account for another 7% of the difference between switchers and nonswitchers. Interestingly, as reported in column (4), including flexible controls (quadratic functions and first order interactions of car value, weight, and horsepower) for car characteristics does not aid screening, but rather introduces noise. Therefore, we find no evidence that car preferences

Table 3. Efficiency of Switchers' Screening Assessed by Linear Count Model

	Claims							
	Number, LPM						Volume (€), tobit	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Switcher	0.0074***	0.0064***	0.0060***	0.0061***	0.0062***	0.0046***	22.1***	14.6***
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0006)	(1.64)	(1.84)
No driving history, new driving license	0.0485***	0.0590***	0.0592***	0.0626***	0.0626***	0.0362***	93.2***	57.0*
	(0.0089)	(0.0107)	(0.0107)	(0.0107)	(0.0107)	(0.0111)	(24.19)	(30.57)
No driving history, old driving license	0.0204***	0.0134***	0.0131***	0.0136***	0.0135***	–0.0082**	54.3***	–11.7
	(0.0020)	(0.0021)	(0.0021)	(0.0021)	(0.0021)	(0.0039)	(5.83)	(11.22)
Demographics	No	Yes	Yes	Yes	Yes	Yes	No	Yes
Location	No	No	Yes	Yes	Yes	Yes	No	Yes
Car	No	No	No	Yes	Yes	Yes	No	Yes
Extra insurance	No	No	No	No	Yes	Yes	No	Yes
Behavioral	No	No	No	No	No	Yes	No	Yes
N	1,039,403	1,039,403	1,039,403	1,039,403	1,039,403	1,039,403	1,039,403	1,039,403

Notes. Standard errors are in parentheses. LPM, linear probability model.
$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

reveal the riskiness of the driver after controlling for demographic characteristics. Column (5) includes a dummy equal to 1 if an individual purchased a collision contract. Naturally, such a decision is endogenous; thus, buying extra insurance should be informative about risk. Nevertheless, we find that this method of screening is not very effective. Cohen and Einav (2007) and Jeziorski et al. (2017) obtain similar results and explain them by heterogeneity in risk aversion and moral hazard, respectively.

Column (6) introduces behavior-based screening by including fixed effects for the current risk class. In this specification, we control for all available covariates, so a switcher is compared with a nonswitcher who is observationally equivalent for the firm. Thus, the size of the *switcher* dummy represents the limits of screening available to the firm. We observe that behavioral-based screening is relatively effective and eliminates another 18% of the gap between switchers and nonswitchers. However, we also find that nearly 50% of the initial gap is private information of the switchers.

Interestingly, the dummy for drivers with no driving history but with an old driving license changes sign after controlling for the risk class. This indicates that such drivers are riskier than the company's own drivers in an average risk class, but safer than the company's own drivers residing organically in risk class 10. This rationalizes placing such drivers in a high risk class, but indicates that their default placement in risk class 10 overstates their riskiness. Not surprisingly, controlling for the risk class does not eliminate the gap for drivers with a new driving license. In other words, drivers with a new license are significantly riskier than the company's own drivers in risk class 10 who obtained their license 10 years ago. This gap explains the large surcharge that new drivers pay for the first five years after obtaining a license, beyond obtaining no risk-class discount.[11]

Columns (7) and (8) of Table 3 contain the estimates of the average marginal effects obtained from the Tobit model predicting the claim volume. We find that switchers generate a €22 or 32% larger volume of liability claims than nonswitchers, as reported in Table 2. After accounting for all observable differences, this gap shrinks to €14, or 20% of nonswitchers' cost to serve. The residual cost difference is large considering that the focal company makes negligible profits from car insurance and suggests that poaching may not be profitable.

As yet, we established that, on average, switchers are more risky than observationally equivalent nonswitchers. This implies that currently employed actuarial price discrimination based on observables would not be sufficient to capture the entirety of risk heterogeneity. A natural solution based on actuarial pricing principles would be to include switching behavior as a pricing variable, which would generate a switcher surcharge. Next, we investigate whether such surcharge should depend on the segment of the switcher. In other words, we check whether the size of riskiness gap between switchers and nonswitchers varies across drivers with different observable characteristics.

First, we demonstrate the impact of driving history, a particularly good predictor of risk, on the size of the riskiness gap. We find that the riskiness gap persists for switchers with any driving history, however, its severity varies. The results are presented in columns (1) and (2) of Table 4. Switchers with excellent histories generate a €23 larger volume of claims than the company's own drivers with similar histories, which makes them comparable to the company's own drivers with fair histories. We find that switchers with bad driving histories have a four percentage point larger chance of having a claim within a year than the company's own drivers with similar histories, which results in a €151 gap in the volume of claims. Although accidents in general are somewhat random and are a noisy measure of future risk, accidents accompanied by switching indicate high risk with more certainty. Consequently, the combination of bad driving history and switching is a red flag.

Furthermore, we investigate how the riskiness gap depends on tenure. The results are presented in columns (3)–(5) of Table 4. In this analysis, we exclude the switcher dummy from the regression and include controls for tenure. As a result, we compare a switcher to an observationally equivalent nonswitcher with a varying tenure.[12] We start by employing the simplest specification in which the riskiness is linear in years of tenure, and we present the results in column (3). We find that, on average, one year of tenure decreases the number of liability claims by 0.5%—that is, loyal customers generate statistically significantly lower costs to serve—however, the magnitude of the effect seems economically small. To investigate this relationship further, we employ a more flexible specification in which tenure is expressed by a series of dummy variables that allow for less parametric relationship between tenure and riskiness. We present the results in columns (4) and (5). We identify a nonlinear

Table 4. Observed Heterogeneity of the Gap Between Switchers and Nonswitchers Depending on Driving History and Discount

	Claims						
	(1) Number	(2) Volume (€)	(3) Number	(4) Number	(5) Volume (€)	(6) Number	(7) Volume (€)
Switcher, excellent driving history, risk class 1	0.006*** (0.0009)	22.5*** (3.2)					
Switcher, good driving history, risk classes 2–4	0.003*** (0.0008)	7.5*** (2.8)					
Switcher, fair driving history, risk classes 5–10	0.009*** (0.002)	21.1*** (5.3)					
Switcher, bad driving history, risk classes 11–18	0.21*** (0.04)	151.3** (76.8)					
Nonswitcher			−0.004*** (0.0006)				
Nonswitcher × *tenure*			−0.0002*** (0.00005)				
Nonswitcher, 1–2 *years tenure*				−0.003*** (0.0006)	−10.1*** (2.5)		
Nonswitcher, 3–4 *year tenure*				−0.006*** (0.0007)	−23.1*** (3.2)		
Nonswitcher, 5–9 *year tenure*				−0.006*** (0.0007)	−21.3*** (3.3)		
Nonswitcher, 10 *year tenure*				−0.007*** (0.0007)	−23.6*** (3.3)		
Switcher, discount 2.5%						0.003** (0.001)	16.15*** (3.862)
Switcher, discount 7.5%						0.008*** (0.002)	19.48*** (5.218)
Switcher, discount 12.5%						0.007*** (0.001)	19.61*** (4.797)
Switcher, discount 17.5%						0.004*** (0.0008)	10.28*** (2.748)
Switcher, discount 22.5%						0.007*** (0.002)	18.73*** (5.879)
All controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,039,403	1,039,403	1,039,403	1,039,403	1,039,403	1,039,403	1,039,403

Note. Standard errors are in parentheses.
$^{**}p < 0.05$; $^{***}p < 0.01$.

relationship between riskiness and tenure. The company's own clients with one to two years of tenure and three to four years of tenure are, respectively, 8% and 16% less risky than switchers; however, the relationship between tenure and riskiness is flat beyond three to four years of tenure. This concave relationship manifests itself as a small coefficient in the linear model.

So far we have shown that riskiness gap depends on driving history and tenure, which suggests that optimal switcher surcharges should be larger for bad drivers and should decrease with tenure. The exercise can be repeated for other observables. Next, we show whether there is residual heterogeneity in the switcher–stayer riskiness gap after controlling for all observables. We note that if switchers within an observationally homogeneous segment are heterogeneous in riskiness, the riskiness gap between the switchers and nonswitchers within the same segment must vary as well. Thus, to test whether the heterogeneity in the riskiness gap can be explained by observables, one can simply test for unobserved heterogeneity in the riskiness among switchers.

We employ a modified version of a well-established test proposed by Puelz and Snow (1994) and extended by Chiappori and Salanie (2000).[13] We test whether, after controlling for all observables, the realized risk is different between switchers who had collision insurance and those who did not. Low-riskiness switchers have an incentive to buy less insurance; thus, showing realized risk gap between contracts provides evidence for unobserved heterogeneity among switchers.[14] As mentioned before, collision insurance is a distinct product with its own risk-class system. Therefore, buying collision insurance has no impact on the incentives to file liability claims. Our estimates show that switchers with collision contracts submit 0.008 (p-value less than 0.001) more liability claims than otherwise equivalent switchers with only liability contracts. Also, the switchers with more insurance generate a €23 (p-value less than 0.001) larger volume of liability claims. Using these estimates, we can reject the null hypothesis of homogeneous switchers and that the heterogeneity is entirely related to observable characteristics.

The above result has implications for pricing. The existence of the unobserved variation in the riskiness gap suggests that segment-based switcher surcharges will not price in the entirety of risk heterogeneity. Moreover, a price increase to switchers is likely to decrease the amount of switching. In the case where switching behavior is strategic, the new population of switchers would have a different unobserved risk profile than the old population. This would be particularly concerning for the firm, if only the riskiest switchers remained after increasing the price. Unfortunately, we do not observe switchers surcharges, so we cannot test this hypothesis directly. Instead, we use the structural model in the next section to investigate this possibility.

Beyond the implications for optimal pricing, the relationship between riskiness and tenure suggests two possible mechanisms generating riskiness gap: (i) the riskiness gap is a result of selective attrition, and the attrition of "bad" clients occurs within the first one to two years; (ii) clients start driving more safely after three years with the company (controlling for all other observables). The next section provides evidence for the former.

4.2. Churners

In this section, we analyze the realized risk gap between clients that churn and those that do not churn. Switchers described in the previous section are churners from the competitor. If the riskiness pool of clients is comparable across companies, the riskiness gap between switchers and nonswitchers implies a realized risk gap between churners and nonchurners. Thus, finding a gap between churners and nonchurners provides further evidence for the gap between switchers and nonswitchers.

For the purpose of our analysis, we define churners as client-year observations in which the client has canceled their contract. Churners either sign a contract with one of the competitors or quit the market.

We start by comparing descriptive statistics between churners and nonchurners (see Table 5). Customers churn with an average probability of 19%; however, the churn rate for customers who generate a claim amounts to a staggering 35%. From the opposite perspective, the customers that churned generated 135% more accidents and three times more cost to serve than customers that did not churn. The large differences between churners and nonchurners persist even after accounting for observable characteristics (see columns (1) and (2) of Table 6). The gap

Table 5. Descriptive Statistics of Churners vs. Nonchurners

	Nonchurner	Churner	Total
Number of claims	0.0324	0.0764	0.0409
	(0.000199)	(0.000641)	(0.000204)
Total claims (€)	51.15	163.8	73.04
	(1.003)	(5.462)	(1.335)
Age	47.87	44.93	47.30
	(0.0137)	(0.0278)	(0.0123)
Years since driving license	23.00	20.48	22.51
	(0.0112)	(0.0226)	(0.0101)
Discretionary discount (%)	12.10	11.56	12.00
	(0.00746)	(0.0164)	(0.00680)
Risk class	1.836	2.783	2.020
	(0.00191)	(0.00597)	(0.00196)

in realized risk does not automatically mean that churners have higher ex ante riskiness than customers that do not churn, because of reverse causality, that is, churning being a result of an exogenously random accident. Note that the customers have incentives to seek better prices (and churn) after an accident, everything else equal, because their current premium increases after losing the driving discount. Thus, even if all drivers are ex ante identical, we should observe more accidents in the year of churning. To somewhat alleviate these concerns, we regress lagged claims on churn dummies and obtain a large, albeit smaller than before, coefficient of 0.02. We admit that the reverse causality may still persist to some degree, because past claims are correlated with current claims, or because customers may churn because of the claims with a lag. We address the first issue by including current claims in the regression and showing that the estimate is numerically the same (unreported). We also address the latter issue using double lag and obtain qualitatively the same result; see column (4) of Table 6. It is noteworthy that all obtained coefficients on the *churn* dummy are larger than the corresponding coefficient on the *switcher* dummy in column (6) of Table 3. This suggests that churners to the outside option (quitters) are, on average, more risky than switchers and stayers.

We find that 20% of switchers leave the company within a year, which suggests that there is a segment of clients that frequently switch insurance providers. According to the estimates presented in columns (3) and (4) of Table 6, the segment of switcher-churners is the most expensive to serve. They generate €89.1 more cost than the company's own clients that did not churn, and €14.2 more cost than the company's own clients that did churn. Because we do not have the data on the riskiness after churning, we can only speculate that the existence of switcher-churners segment contributes to the previously detected unobserved heterogeneity of riskiness among switchers. Thus, identifying this segment of drivers at the point of signing the contract could help price in the cost of riskier switchers.

We have shown previously that the tenure of the customer does not indicate increased riskiness beyond three years with the company. Results in this section indicate that this is at least partly due to selective attrition of riskier customers. Such selective attrition involves temporarily serving the expensive segment of switcher-churners. Consequently, the company should take into account potentially high cost of selective attrition when poaching competitors' customers.

4.3. Delegation

We investigate one remaining possibility of screening bad switchers currently available to the firm. Basu et al. (1985), Dolan and Simon (1996), and Lal (1986) argue that the delegation of pricing decisions to the sales force may help to improve pricing under asymmetric information. Such delegation is effective when the sales force has more information than the sales manager or, in our case, the centralized actuarial pricing analyst. The delegation is usually analyzed in the context of price discrimination among consumers with heterogeneous willingness to pay. In contrast, we analyze whether pricing by the sales force improves screening of the unobserved cost to serve. Columns (6) and (7) of Table 4 shows that the riskiness gap is not systematically related to the discretionary discount. In other words, more risky drivers do not obtain lower discretionary discounts. The results

Table 6. Marginal Effect of Churning on the Number and Volume of Claims per Year

	Claims		Lagged claims	Double-lagged claims	Claims	
	Number	Volume (€)	Number	Number	Number	Volume (€)
	(1) LPM	(2) Tobit	(3) LPM	(4) LPM	(5) LPM	(6) Tobit
Churner	0.0362***	76.74***	0.021***	0.005***		
	(0.0005)	(1.197)	(0.0007)	(0.0008)		
Switcher, nonchurner					0.00388***	13.83***
					(0.000623)	(1.776)
Nonswitcher, churner					0.0348***	74.89***
					(0.000600)	(1.304)
Switcher, churner					0.0457***	89.05***
					(0.00109)	(2.411)
All controls	Yes	Yes	Yes	Yes	Yes	Yes
N	1,039,403	1,039,403	599,762	336,483	1,039,403	1,039,403

Notes. Standard errors are in parentheses. LPM, linear probability model.
***$p < 0.01$.

suggest that the involvement of sales force does not improve screening (for similar results, see Stephenson et al. 1979). There are several theories explaining this phenomenon. The sales force may have inferior or biased information about the riskiness of the client. Also, the sales force may be risk averse and give away discounts to more insistent but riskier switchers (see Berger 1972, Weinberg 1975). Additionally, we have anecdotal evidence that the current compensation scheme incentivizes the sales force to increase the volume of policies and does not promote screening on future profitability of the customer.

The negative result on the efficiency of delegation has one more practical implication; namely, when estimating the structural model presented in the next section, we can assume stationarity of the price distribution (a standard assumption in search literature). Particularly, we will assume that the distribution of the available discretionary discounts is unrelated to the unobserved portion of the riskiness of the driver.

5. Structural Analysis

In this section, we develop and estimate a stylized structural model of churn. There are three reasons for this exercise. First, the model allows us to postulate and test the mechanism driving the results from the previous section. In particular, we demonstrate that a churn model, embedding the framework of Cohen and Einav (2007), explains various facts on switching and churning demonstrated earlier.

Second, the model allows us to provide more insight into the nature of switching and churning. In particular, it allows to assess the riskiness of an incoming switcher before he signs the contract with the company, as well as assess the riskiness of the churner before he quits the company. Because it is possible to identify the model from the risk pool of a single company, the model provides a feasible tool for the practitioners, who do not have access to the consumer pool of the competitor. This restriction is binding in most markets with asymmetric information.

Third, the model allows us to develop a robust LTV measure and apply it to analyze the impact of counterfactual pricing policies and contracts. We consider (i) charging higher premiums to switchers and (ii) applying a steeper incentive contract.

There are several necessary features that we need to include in the model to flexibly capture the adverse selection observed in the raw data:

1. The model needs to allow for the unobserved heterogeneity in risk. If all observationally equivalent drivers had homogeneous risk, observationally equivalent switchers and own clients would have the same risk as well.

2. The model needs to incorporate a demand friction, which prevents the dynamics of the market to unravel; that is, in the world of stationary prices without demand friction, all consumers would obtain the lowest price quote upon entering the market, leaving no room for switching. Following the previous literature on car insurance, we achieve demand friction by postulating that customers have imperfect information about insurance premiums and experience search cost.[15]

3. Because we would like to use the model to rationalize the size of the switcher–stayer riskiness gap, the model must be capable of generating a large gap, as well as no riskiness gap, depending on the primitives. For this reason, the model needs a second dimension of customer heterogeneity that is capable of driving switching behavior and that is unrelated to risk. The most natural choice in our context is heterogeneity in search cost; however, we also investigate other viable options, such as heterogeneity in preferences (risk aversion). If heterogeneity in risk is large (small) relative to search cost heterogeneity, the model will generate a large (small) switcher–stayer risk gap.

5.1. Model

Consider a market with N active insurance firms and I consumers. Consumers make decisions in discrete time and live T periods.[16] Firms offer two insurance products: a liability-only policy, Y^L, and a comprehensive policy, Y^C. Liability-only policy is compulsory and covers damages to the counterparty's car. The products are priced accordingly to a three-tier pricing formula, consistent with the description in Section 2. Companies set a baseline premium $P(X_{it}, Y)$, which depends on observable characteristics X_{it} of consumer i, such as car covariates, location, age, and driving experience. Each driver obtains a multiplicative driving history discount $H_{it} = H(M_{it})$, where M_{it} is the current risk class, and a discretionary discount D_{it}. The final premium[17] is given by

$$\bar{P}_{it}(Y_{it}) = H_{it} D_{it} P(X_{it}, Y).$$

We assume that each company uses the same pricing function P and the same driving history discount schedule H. Companies compete by offering different discretionary discounts. We further discuss this assumption when we describe the identification of the model.

As mentioned above, each consumer is characterized by a set of observable characteristics X_{it}, driving history M_{it}, and grandfathered level of discretionary discount D_{it-1}. In addition, consumers possess a set of unobserved (both to the firms and to the econometrician) characteristics denoted by Z_{it}.

We assume that $Z_{it} = (\lambda_{it}, \gamma)$, where λ_{it} is the ex ante riskiness[18] of the driver, and γ is his risk aversion.[19] We assume that λ_{it} is distributed in the population as $F_\lambda(X_{it})$. Let σ_λ be the variability parameter of this distribution. We denote the vector of all consumer characteristics, except for the discretionary discount, as $s_{it} = (X_{it}, Z_{it}, M_{it})$.

At the beginning of each period, the consumer is either active, $\omega_{it} = 1$, or inactive, $\omega_{it} = 0$. Active consumers have a contract with one of the providers at time t. In contrast, inactive consumers do not have a contract with any of the providers at time t. Each consumer starts inactive, that is, $\omega_{it_0 - 1} = 0$.

Each consumer enters the period t with the information about s_{it}, D_{it-1}, and ω_{it-1}. Previously active consumers, that is, when $\omega_{it-1} = 1$, can make three choices:

1. The customer can stay with the current provider, in which case the customer grandfathers the current discount, that is, $D_{it} = D_{it-1}$, and receives a choice-specific continuation payoff given by

$$\bar{V}_t^{STAY} = U(s_{it}, D_{it}) + \beta E V_{t+1}(s_{it+1}, D_{it}, \omega_{it} = 1) + \sigma_\epsilon \epsilon_{it}^{STAY},$$

where $U(s_{it}, D_{it})$ represents per-period monetary payoffs from driving, and V is the continuation value. The term ϵ_{it}^{STAY} captures nonpecuniary benefits from deciding not to search in period t. The expectation before the continuation value represents integration over s_{it+1} conditional on s_{it}. The details of this integration are discussed later. The term β is the discount factor.

2. The customer can search for deals, in which case he pays search cost $C - \sigma_\epsilon \epsilon_{it}^{SEARCH}$. He receives a random draw $\tilde{D}$ from the equilibrium distribution of discretionary discounts F_D.[20] He switches to a new provider only if he receives a better discount, that is, $D_{it} = \min\{D_{it-1}, \tilde{D}\}$. His per-period utility from searching is given by

$$\bar{V}_t^{SEARCH} = E\Big[U(s_{it}, D_{it}) + \beta E V_{t+1}(s_{it+1}, D_{it}, \omega_{it} = 1)\Big] - C + \sigma_\epsilon \epsilon_{it}^{SEARCH},$$

where the expectation is taken over D_{it}.

3. The customer can become inactive, in which case he receives utility of not driving U^0 and loses a grandfathered discount D_{it}. He may reenter the market in the next period, and his payoff is given by

$$\bar{V}_t^{LEAVE} = U^0 + \beta E V_{t+1}(s_{it+1}, D_{it} = 1, \omega_{it} = 0) + \sigma_\epsilon \epsilon_{it}^{LEAVE},$$

where ϵ_{it}^{LEAVE} represents random transitory events that cause an individual to stop driving, such as transitory health shocks.

Each previously inactive consumer has two choices:

1. The consumer can stay inactive, in which case he receives the utility

$$\bar{V}_t^{INACTIVE} = U^0 + \beta E V_{t+1}(s_{it+1}, D_{it} = 1, \omega_{it} = 0) + \sigma_\epsilon \epsilon_{it}^{INACTIVE}.$$

2. The consumer can activate, in which case he receives the utility

$$\bar{V}_t^{ACTIVE} = E\Big[U(s_{it}, D_{it}) + \beta E V_{t+1}(s_{it+1}, D_{it}, \omega_{it} = 1)\Big] - C + \sigma_\epsilon \epsilon_{it}^{ACTIVE}.$$

where the expectation is taken with respect to $D_{it} = \tilde{D}$. Note that, compared with searching action, there is no minimization operator, because inactive consumers do not have grandfathered discounts. Also, we do not allow the consumers to stay inactive if they receive low discounts. It reflects the reality of the insurance industry, because one needs to register the car before obtaining a binding quote on D_{it}.

The consumers follow a Markovian policy function, which specifies optimal search, switching, and quitting actions. Under the assumption that payoff shocks ϵ_{it} are distributed as normalized type 1 extreme distributions (mean-centered at zero),[21], [22] the Bellman equations for the consumers are given by

$$V_t(s_{it}, D_{it-1}, \omega_{it-1} = 1) = \sigma_\epsilon \log\Big[\exp(\bar{V}_t^{STAY}\sigma_\epsilon^{-1}) + \exp(\bar{V}_t^{SEARCH}\sigma_\epsilon^{-1}) + \exp(\bar{V}_t^{LEAVE}\sigma_\epsilon^{-1})\Big], \quad (1)$$

$$V_t(s_{it}, D_{it-1}, \omega_{it-1} = 0) = \sigma_\epsilon \log\Big[\exp(\bar{V}_t^{INACTIVE}\sigma_\epsilon^{-1}) + \exp(\bar{V}_t^{ACTIVE}\sigma_\epsilon^{-1})\Big]. \quad (2)$$

Next, we discuss the structure of the utility function from driving $U(\cdot)$ and transition process for s_{it}. The function $U(\cdot)$ embeds the optimal contract choice, insurance premiums, and potential losses resulting from uninsured accidents in a way that follows the model of Cohen and Einav (2007). The consumer chooses one of the contracts, $Y_{it} \in \{Y^L, Y^C\}$, and starts driving after paying the corresponding premium, $\bar{P}_{it}(Y_{it})$. While driving, during the contract year, the consumer incurs R_{it} accidents that generate damage to his own car, denoted by L_{it}. (We do not model losses to the counterparty, because they are always covered.) Losses to his own car are covered only if consumer has purchased comprehensive contract. We assume that R_{it} is distributed as Poisson random variable with parameter λ_{it} that is truncated from above at three (we never observe more than three claims in one contract year). Moreover, following other papers in the car insurance literature (see Cohen and Einav 2007, Abbring et al. 2008, Dionne et al. 2013), we assume that $L_{it} = \sum_r^{R_{it}} L_{itr}$, where L_{itr} is independent and identically distributed conditional on X_{it}. This implies that L_{itr} is independent of λ_{it} conditional on X_{it}, so that λ_{it} affects L_{it} only through the number of accidents R_{it}.

Consumer obtains a per-period expected utility

$$U(s_{it}, D_{it}, \omega_{it} = 1) = U^D + \max_{Y_{it} \in \{Y^L, Y^C\}} \cdot E\left[u\left(-L_{it}\mathbf{1}_{Y_{it}=Y^L} - \bar{P}_{it}; \gamma\right) \middle| \lambda_{it}\right],$$

where U^D is the utility of driving. We note that, without loss of generality, we can normalize $U^D = 0$ and reparametrize the model by replacing U^0 with $U^0 - U^D$. In this new formulation U^0 is the disutility of using alternative means of transportation. We assume constant absolute risk aversion utility, $u(x; \gamma) = -\exp(-\gamma x)$.

After the period ends, the state $(s_{it}, D_{it-1}, \omega_{it-1})$ is updated. Updating of D_{it-1} and ω_{it-1} has been explained earlier. The risk class of the consumer is updated by the number of accidents $M_{it+1} = F_M(M_{it}, R_{it})$ according to the plus three, minus one rule.[23] The variables X_{it} and Z_{it} are updated according to the deterministic functions $F_X(\cdot|X_{it}, Z_{it})$ and $F_Z(\cdot|X_{it}, Z_{it})$.

One of the goals of this paper is to show that a simple search model with competition can explain the stayer–switcher riskiness gap. Because the gap will be estimated from the data, it is important that the model is capable to generate a large riskiness gap, as well as no riskiness gap, depending on the primitives. The crucial primitives are variability in risk σ_λ and variability in search cost σ_ϵ. If σ_λ is large compared with σ_ϵ, then the variation in switching behavior is driven by the variation in riskiness. In such a case, the riskiness gap would be large. Conversely, if σ_ϵ is large compared with σ_λ, the variation in switching behavior is driven by the variation in search cost, generating a small riskiness gap. In the next subsection, we explain how we identify these and the remaining primitives.

5.2. Estimation and Identification

The model is estimated from the panel data described in Section 3. The parameters to be estimated, θ, involve the distribution of risk in the population F_λ, risk aversion γ, search cost C, disutility of nondriving U^0, equilibrium discount distribution F_D, and the variance of payoff shocks σ_ϵ. We calibrate the discount factor to $\beta = 0.95$ and T to 90 years old.

The identification of the model relies on two assumptions: (i) that the insurer cannot price discriminate on unobserved riskiness and (ii) that companies play symmetric pricing equilibrium. As already described is Section 4.3, we find no evidence that sales force has extra information; thus, assumption (i) is likely to be true. As mentioned in Section 2, insurance market in Portugal is highly competitive, and the contract structure is regulated; thus, symmetric pricing equilibrium provides a good approximation.[24] Nevertheless, we show in the online appendix that our main findings are robust to relaxing this assumption.

The literature on car insurance, such as Puelz and Snow (1994) and Chiappori and Salanie (2000), demonstrates that it is important to properly control for observed heterogeneity when estimating the distribution of the unobserved heterogeneity in the population, F_λ. If the specification linking risk to observables is not sufficiently flexible, one can overestimate unobserved heterogeneity in risk. For example, one can misinterpret variation in risk generated by higher order terms that are missing from the model as truly unobserved heterogeneity. Estimating the degree of unobserved heterogeneity in risk is pivotal for accurately describing adverse selection. Fortunately, our data allow us to take a conservative semiparametric approach. We subsample our data and keep only one specific car make and model that is modal in our sample, namely, a particularly popular Renault car.[25] This allows for comparing the moments while keeping car characteristics fixed. Furthermore, we specify F_λ as a truncated normal distribution with parameters $\mu_\lambda(X_{it})$ and σ_λ. The dependence of μ on X_{it} signifies that we allow for postal-code fixed effects. We control for age and experience by allowing young age (under 25) and low experience (less than four years of driving) riskiness multipliers. Such specification mimics the dependence of the actuarially set premium function $P(\cdot)$ on location, age, and experience; thus, it captures relevant observed heterogeneity.

We estimate the model with a method of simulated moments (MSM), as in McFadden (1989) and Pakes and Pollard (1989). Let m_j for $j \in 1, \ldots, J$ be a set of moments used in the estimation. The particular choice of moments is discussed below together with the identification. The caveat when estimating the model is that the data contain a selected sample of customers that are active in a particular period. This selection is important, because the decision to be active is endogenous and is a function of λ_i. Thus, during the estimation, we have to use $E[m_{jit}|\omega_{it} = 1]$ instead of the unconditional moments. This has two implications for sampling the moments. First, conditional on λ_i, we have to compute a conditional moment $E[m_{jit}|\omega_{it} = 1, \lambda_i]$. Second, we need to use a conditional distribution of λ_i, that is $F(\lambda_i|\omega_{it} = 1)$, to integrate the conditional moments. The estimation procedure is as follows:

1. Fix θ.
2. For each consumer i, draw R parameters λ_i^r from the unconditional distribution $F_\lambda(\cdot; \theta)$.
3. Use the importance sampling procedure to obtain conditional moments $E[m_{jit}|\omega_{it} = 1]$:

a. Compute moments $E[m_{jit}|\omega_{it} = 1, \lambda_i^r]$ and probability of being active $\text{Prob}(\omega_{it} = 1|\lambda_i^r)$. It is possible to obtain these values analytically.

b. Reweight the moments according to the importance sampling formula with the instrumental density $f(\lambda_i; \theta)$,

$$E[m_{jit}|\omega_{it} = 1] = \int E[m_{jit}|\omega_{it} = 1, \lambda_i] dF(\lambda_i|\omega_{it} = 1)$$
$$= \text{Prob}(\omega_{it} = 1)^{-1} \int E[m_{jit}|\omega_{it} = 1, \lambda_i] \cdot \text{Prob}(\omega_{it} = 1|\lambda_i) dF(\lambda_i),$$

with sample analogues

$$\widehat{E}[m_{jit}|\omega_{it} = 1] = \widehat{\text{Prob}}(\omega_{it} = 1)^{-1} \sum_{r=1}^{R} E[m_{jit}|\omega_{it} = 1, \lambda_i^r] \cdot \text{Prob}(\omega_{it} = 1|\lambda_i^r)$$
$$\widehat{\text{Prob}}(\omega_{it} = 1) = \sum_{r=1}^{R} \text{Prob}(\omega_{it} = 1|\lambda_i^r).$$

4. Aggregate $E[m_{jit}|\omega_{it} = 1]$ across i and t to obtain population moments. Compute the MSM objective function.

All parameters are estimated jointly; however, each parameter of the model corresponds to a set of identifying moments. We start by discussing the identification of the distribution of λ_i. Suppose that we know other parameters of the model, besides the parameters of $F(\lambda_i)$. Because the model provides selection equation for active drivers, we can identify $F(\lambda_i|X_{it})$ from the distribution of realized risk across risk classes conditional on X_{it}. We have 18 risk classes; thus, theoretically, we could identify 18 parameter family of conditional distributions $F_\lambda(\cdot|X_{it})$. In practice, we use the average number of accidents conditional on three groups, risk class 1, risk classes 2–9, and risk classes 10–18. This overidentifies the two-parameter conditional Gaussian of λ_i. The location, age, and experience parameters are identified from the corresponding variation across conditional moments. Knowing the distribution of risk, the risk aversion parameter, γ, is identified from the share of the comprehensive contract in the population.

Joint identification of the disutility of driving, U^0, search cost, C, and equilibrium distribution of discounts, F_D, is more complicated. We observe only transacted discounts, and we do not see offers that were rejected. Also, we do not know whether the consumers churn to an inactive state or to the competitor. The identification of the incentives to churn relies on the variation in churn rates at different level of premiums, controlling for the riskiness. Customers paying larger premiums have more incentives to churn, both to the competitor and to an inactive state. If the utility from churning is high, then we should observe that the churn rate increases steeply in the current premium.

Suppose that we know the distribution F_D and the parameter σ_ϵ. To distinguish churning to the competitor from churning to an inactive state, we assume that people churn to the competitor only if they receive a greater discretionary discount (see assumption (ii)). This means that the insurance products are homogeneous if we fix premium levels, observable contract characteristics, all consumer observed characteristics, and some consumer unobserved characteristics, such as risk aversion and riskiness. In the current specification, we do not allow other dimensions of heterogeneity, such as service quality, because those would be hard to identify separately from the heterogeneity in the search cost.

Consider two churning customers with the same riskiness and final premiums, but with different compositions of risk-class and discretionary discounts.[26] The churner with the higher discretionary discount but lower risk-class discount has likely churned to an inactive state. Conversely, the churner with an opposite discount composition has likely churned to the competitor. Thus, we identify the utility from churning to an inactive state, U^0, from the variation in churn rates across risk classes among people with the same discretionary discount. In the extreme, U^0 is directly identified from the variation in churn rates across risk classes for the people with the highest possible discretionary discount. Once we know U^0, we can identify the search cost, C, from the variation in the churn rates within a fixed risk class, across discretionary discounts. In the extreme, customers with the lowest risk class are the most likely to churn to the competitor. In practice, matching churn rates at different levels of discretionary discount and risk class would separate U^0 from C. We match churn rates conditional on risk classes 1, 2–9, and 10–18; discretionary discounts 2.5%, 7,5%, 12.5%, and 17.5%; and several interactions between risk classes and discretionary discounts. Once we know U^0 and C, we can identify F_D by matching the distribution of transacted discounts.

The last parameter to identify is the variability, σ_ϵ, of the private shock ϵ. The parameter σ_ϵ embodies, among other things, the idiosyncratic variability of search cost, which introduces extra flexibility in the distribution of churn-related statistics. We note that the risk-class moments and discretionary discount moments overidentify U^0 and C; thus, σ_ϵ is chosen to match the residual variation. In addition, to obtain a more nonparametric identification, we match the second moment of the number of accidents conditional on churning.

5.3. Results

The estimated structural parameters are presented in Table 7. The first three rows present measures of the central tendency of the distribution F_λ. In particular, they represent means of the underlying normal distribution, before truncation, across postal codes. Aggregating across age and experience of drivers, implied mean accident counts are 0.038, 0.035, and 0.048 for postal codes 1, 2, and 3, respectively. The difference between the first two numbers is not significant. These results suggest that postal code 3 contains a more risky population of drivers than the other two postal codes. On average, the population of drivers is estimated to have 0.039 accidents per contract period, aggregating across postal codes, age, and experience of drivers.

The fourth and fifth rows contain risk multipliers for young drivers (less than 25 years old) and inexperienced drivers (three years or fewer years since obtaining a driving license). We show that both young drivers and new drivers generate approximately twice as many claims than older and experienced drivers, respectively. Also, drivers that are both young and new, generate approximately four times more claims than those that are both older and experienced. This variation proves important in explaining higher churn rates among young and inexperienced drivers.

Table 7. Structural Parameters

Parameter	Baseline
Risk: central tendency, μ_λ	0.023*** (0.002)
Risk: central tendency, postal code 2 *fixed effect*	–0.004 (0.003)
Risk: central tendency, postal code 3 *fixed effect*	0.017*** (0.003)
Risk: variability, σ_λ	0.034*** (0.002)
Risk: young multiplier, λ^{YOUNG}	1.845*** (0.503)
Risk: inexperienced multiplier, λ^{INEX}	2.112*** (0.713)
Risk aversion, γ	0.129*** (0.028)
Search cost (utils)	13.0* (7.1)
Disutility of not driving (€)	429.6*** (13.0)
2.5% *discount probability*	0.686*** (0.021)
7.5% *discount probability*	0.102*** (0.005)
12.5% *discount probability*	0.075*** (0.005)
17.5% *discount probability*	0.123*** (0.010)
Variability of private shock (utils), σ_ϵ	0.033*** (0.006)

*$p<0.1$; ***$p<0.01$.

The sixth row contains the standard deviation of the Gaussian underlying the distribution F_λ. This implies 0.031 standard deviation of the average number of accidents across people, std$[E[R_{it}|i]]$, resulting in a large (0.75) coefficient of variation.

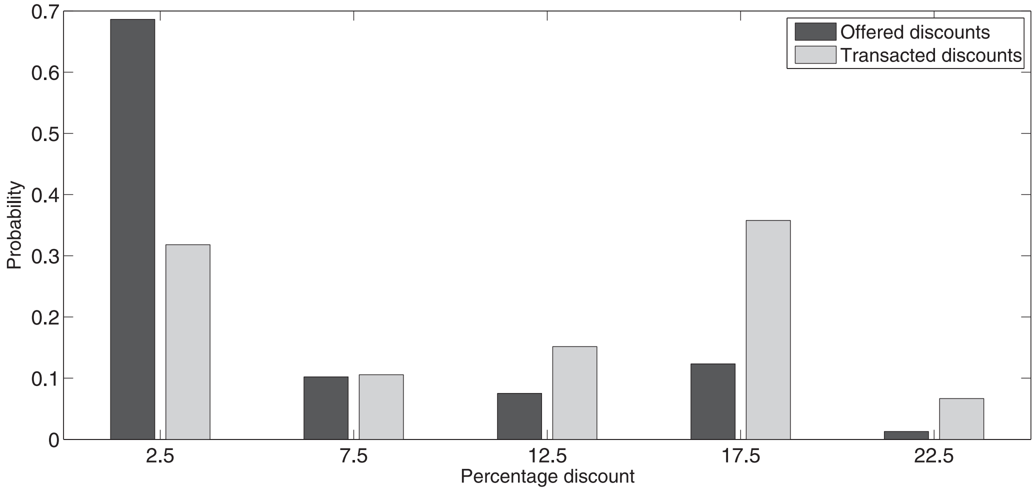
The seventh row contains the estimate of risk aversion. The utility parameters are hard to interpret alone. Thus, we use a modal population of older and experienced drivers to derive monetary interpretation. The estimated risk aversion parameter implies that a modal driver would be willing to pay €99 to avoid damages to his own car, which amounts to €70 in expectation. We contend that drivers exhibit moderate risk aversion, which corresponds to a low market share of comprehensive contracts in the data.

As reported in eighth row, search cost amounts to 0.013 utils. To better understand this number, we computed the compensating variation resulting in the drop of 0.013 utils for the modal driver. We found that an average drop of €102 in the payoff of the driving individual was equivalent to the estimated value of the search cost (or between €96 and €143 after relaxing the symmetric pricing assumption; see the online appendix). This contrasts with the $42 search cost estimate of Honka (2014), who studies user behavior in the U.S. car insurance market and allows the user to perform multiple searches in one period. Honka (2014) finds that in the U.S. market, users obtain 2.96 quotes at a time. Since our estimate of search costs is approximately three times larger than Honka's, per-period volume of incurred search costs is strikingly similar in both papers. Dahlby and West (1986) found search cost between $131 and $570 in Canada; however, their analysis was conducted in an earlier time period, when telecommunication and digital marketing were not as prevalent, and they considered only rural locations.

The ninth row reports the intrinsic benefit of driving over nondriving, which amounts to €429 per year (or between €391 and €457 after relaxing the symmetric pricing assumption; see the online appendix). This number accounts for auxiliary costs, such as gasoline, car maintenance, and public transit, as well as nonpecuniary costs and benefits. The number excludes insurance related monetary costs, such as insurance premium and potential losses from uninsured accidents. This estimate should be viewed in relation to the average car value in the subsample, which amounts to only €2,455.

The 10th to 13th rows contain the estimates of the equilibrium price distribution F_D. (The probability of obtaining the largest discount is 1 minus the sum of

Figure 1. Distribution of Offered and Transacted Discounts



reported probabilities.) As described in Figure 1, the distribution of the offered discounts is shifted to the left relative to the distribution of the discounts in the data. This shift is generated by the fact that consumers are grandfathered into the discounts from the current insurance provider, and switch to a new provider only if offered a better deal.

Table 8 provides measures of goodness of fit of the model. The model is able to correctly predict the increase in the number of claims for drivers in higher risk classes. The model slightly overpredicts riskiness in the highest risk classes, which may be a result of small sample of people in these risk classes (approximately 2% of the observations). The model is able to accurately fit the variation of risk across location as well as age and experience of drivers. Importantly, the model correctly fits the churn rate as a function of the risk class and discretionary discount, which is a key variation used in our identification strategy. We are also able to explain larger churn rates among young and new drivers, without explicit parameters capturing this variation. However, we note that predicted churn rates for these groups are large, but still smaller than those in the data. This gap suggests that young and new drivers may have larger outside options or smaller search cost. These drivers constitute less than 2% of our sample; thus, we are unable to accurately estimate interactions between risk class, discretionary discounts, and churn rates for this subsample; such interactions are necessary to identify search cost from the outside option. For the same reason, this gap in churn has a minimal impact on our final conclusions.

5.4. Insights on Selective Churn

After estimating the primitives, we can use the model to draw insights into the selective churn process. The average ex ante riskiness of the driver that does not churn amounts to 0.038 accidents a year, which compares to 0.047 for the average churner. The model allows us to decompose churn into switching to a different insurance provider and quitting driving. The average driver switches to the competitor with 5.6% probability and quits driving with 12.6% probability per year. Thus, about 31% of the observed churn in the data is to the competitor. Switchers have an average riskiness of 0.041, which is 8% more than that of nonchurners, resulting in adverse selection. The mechanism is that riskier drivers have a higher incentive to search in anticipation of potential claims, because their bonus-malus penalty is smaller, if their discretionary discount is greater. This switcher–stayer gap is close to the switcher–stayer gap established in Section 4.1 using only the population of incoming drivers. Because we estimate the model using the variation in the outgoing churn rates and

Table 8. Goodness of Fit

Statistic	Model	Data
Claims: risk class 1	0.037	0.032
Claims: risk classes 2–9	0.050	0.050
Claims: risk classes 10–18	0.097	0.078
Claims: postal code 1	0.038	0.039
Claims: postal code 2	0.035	0.035
Claims: postal code 3	0.048	0.047
Claims: young	0.084	0.081
Claims: new	0.098	0.104
Churn: risk class 1 (%)	18	16
Churn: risk classes 2–9 (%)	21	25
Churn: risk classes 10–18 (%)	61	61
Churn: 2.5% *discount* (%)	27	22
Churn: 7.5% *discount* (%)	21	19
Churn: 12.5% *discount* (%)	17	18
Churn: 17.5% *discount* (%)	14	16
Churn: young (%)	35	43
Churn: new (%)	43	55
Collision contract (%)	1.2	1.3

we do not use the information about the incoming drivers, the correspondence of both results serves as validation of the model "out of sample."

The average ex ante riskiness of quitters amounts to 0.050, which is about 31% more than that of stayers, and 22% more than that of switchers. Most drivers do not have comprehensive policies; thus, they are exposed to damages to their own cars. The risky drivers thus face higher expected costs of driving overall. This generates an ex ante riskiness gap between quitters and stayers.

In Section 4.2, we documented a large realized risk gap between stayers and churners. Indeed, the model predicts that the average number of realized accidents in the period prior to churning amounts to 0.054, which is 15% larger than the ex ante riskiness of churners. This discrepancy is the direct consequence of the impact of accident occurrence on risk-class transition and the inherent randomness of accidents. One implication of this discrepancy is that companies should not be using realized risk to assess the ex ante riskiness of churners when designing their churn management programs. Another implication is that some good drivers may be priced out of the market and churn if they incur a claim randomly. Specifically, we find that drivers with lower than average riskiness have two percentage point greater churn rates in periods with an accident than in periods without an accident. This is a consequence of the inefficiency in the incentive contract based on realized risk instead of ex ante riskiness, the former being only a noisy signal of the latter. This inefficiency has implications for both consumers and firms. Some unlucky, but otherwise good, drivers are being priced out of the market, which lowers their consumer surplus—a classic mechanism of adverse selection at work. On the other side, firms forgo profits by not serving an attractive consumer segment.

Furthermore, we investigate how heterogeneous the pool of switchers and quitters is. Assessing it is important, because we are interested in the implications of the riskiness gap for pricing. If we find that switchers experience large unobserved variation in riskiness, a uniform price hike for switchers should be less effective, compared with a case in which switchers are homogeneous. Using the reduced form test in Section 4.1, we have already rejected the null that all switchers have the same ex ante riskiness. Beyond this binary test, the model provides a way to quantify the exact degree of this variability. In particular, the standard deviation of the riskiness of the switcher amounts to 0.027, which results in sizable coefficient of variation of 0.66. In the next section, we investigate implications of these results for pricing.

5.5. Implications for Pricing and Contract Design

In this section, we use the model to conduct pricing counterfactuals. Knowing that switchers are more risky than the company's own clients, the company may be tempted to increase the price to switchers. In this section, we analyze the consequences of such a decision. To conduct this counterfactual, we consider a duopoly in which Firm B keeps the prices fixed, and Firm A raises the prices to switchers.[27] We integrate the statistics using the empirical distribution of covariates.

The model implies a natural measure of the LTV, which is defined as discounted stream of profits of the customer, conditional on being a customer of the focal firm. There are three main components of each churn model: (i) revenue per period, (ii) cost to serve per period, and (iii) churn rate. Our LTV measure is based on the model in which the current insurance premium and churn rates are endogenous. As a result, the LTV measure takes into account that customer revenue and churn rates vary as a function of the driving history, pricing strategy, and contract structure. Effectively, we will recompute revenues and churn rates for every customer using the model and update the LTV.

There are two effects of the price increase:

1. Price increase is likely to result in more switching and quitting from Firm A. We know that the marginal nonswitcher (nonquitter) is riskier than the inframarginal nonswitcher (nonquitter). Thus, a price increase would improve the risk pool in Firm A.
2. Price increase is likely to result in less switching from Firm B to Firm A. Because the inframarginal switcher is more risky than marginal switcher, the shrinking incoming population of switchers to Firm A becomes more risky.

The overall impact of the price increase on the risk pool is theoretically ambiguous. We evaluate this impact empirically.

The above argument relies on the observation that switchers and quitters are heterogeneous. For example, if the inframarginal switcher is as risky as the marginal switcher, an increase in the number of switchers would not result in a different risk composition of switchers. For this reason, a sizable coefficient of variation of risk among switchers, which we documented in the previous section, is a preliminary indication that the described mechanisms are important.

We implement the pricing counterfactual by conducting a first order stochastic dominance shift in the discretionary discount distribution to switchers. Such change should be easier to implement for the firm, because it does not involve changes to baseline premiums, which are set by complicated actuarial formulas and therefore take time to calibrate. In particular, we

Table 9. Counterfactual Pricing

	Average discount offered		Average discount transacted	
Mass shift, firm A	Firm A (%)	Firm B (%)	Firm A (%)	Firm B (%)
0.0	5.9	5.9	10.8	10.8
0.1	5.5	5.9	10.7	10.7
0.2	5.2	5.9	10.5	10.5
0.3	4.9	5.9	10.3	10.4
0.4	4.5	5.9	10.1	10.2
0.5	4.2	5.9	9.9	10.0
0.6	3.8	5.9	9.6	9.9
0.7	3.5	5.9	9.3	9.7
0.8	3.2	5.9	8.9	9.5
0.9	2.8	5.9	8.5	9.4

proportionally shave x mass off the probability distribution of discounts greater than 2.5% and increase the mass of 2.5% discount accordingly. The impact of this change on the average offered discount is presented in the second and third columns of Table 9. The fourth and fifth columns present the impact of these changes on the transaction prices at both firms. The transaction price in Firm B increases; however, it does so less than the transaction price in Firm A.

Table 10 presents the results of the pricing counterfactual. First part of the table shows the results of the uniform price increase on switchers. Not surprisingly, the market share of Firm A decreases. The change in market shares is coupled with the selection of the consumers on risk. As price of Firm A increases, the risk pool of Firm A deteriorates, and the risk pool of Firm B improves. As a result, the LTV of own clients of Firm A decreases. This change is a result of amplified adverse selection when switching away from Firm B, illustrated by sharply decreasing LTV of switchers (from B to A).

Large unobserved heterogeneity in risk among the population of switchers suggests that the firm may benefit from further screening of switchers. Indeed, many insurance companies implement so-called usage-based or telematic car insurance. In the past, telematic data were usually self-reported and hence unreliable. However, more recently, companies have started using devices that pull data directly from the car's computer and send it to the insurance company via a mobile telecom network in real time. One form of usage-based insurance is monitoring mileage with a GPS tracker. High mileage is likely to result in higher number of accidents. Thus, mileage may be useful to screen riskier switchers. Another form of such insurance is pay as you drive, in which the driver is priced using statistics about speed, time of day, and braking patterns.

We consider an illustrative example of an insurance policy that contains pricing on currently unobserved characteristics. We presume that the company can use a separate attribution model to make an assessment, whether the drivers unobserved riskiness is above or below the population average. Such assessment could potentially be made using additional data on mileage and driving patterns. After making the assessment, Firm A offers smaller discount to switchers, whose unobserved riskiness is above average. Second part of Table 10 contains the results. Firm A, loses market share; however, it does so slower than in the previous pricing policy. This is not surprising, because Firm A increases the price only on a subset of the

Table 10. Counterfactual Pricing: Price Increase on Switchers

	Discount PMF shift	Market share	Riskiness	Stayer LTV (€)	Switcher LTV (€)
Baseline		0.50	0.039	339	320
Uniform price increase	0.1	0.49 (−2.3%)	0.039 (+0.2%)	339 (−0.1%)	302 (−5.7%)
	0.3	0.46 (−7.3%)	0.040 (+0.7%)	338 (−0.3%)	259 (−19.0%)
	0.5	0.43 (−13.2%)	0.040 (+1.5%)	337 (−0.7%)	206 (−35.7%)
	0.7	0.40 (−20.0%)	0.040 (+2.7%)	335 (−1.3%)	138 (−56.9%)
	0.9	0.36 (−27.8%)	0.041 (+4.6%)	332 (−2.3%)	52 (−83.9%)
	0.1	0.50 (−0.9%)	0.039 (−0.1%)	340 (+0.2%)	314 (−1.7%)
Selective price increase	0.3	0.49 (−2.9%)	0.039 (−1.2%)	343 (+1.1%)	303 (−5.2%)
	0.5	0.47 (−5.2%)	0.038 (−2.5%)	347 (+2.2%)	290 (−9.4%)
	0.7	0.46 (−7.7%)	0.038 (−4.0%)	351 (+3.5%)	274 (−14.3%)
	0.9	0.45 (−10.5%)	0.037 (−5.8%)	356 (+5.1%)	256 (−20.1%)

Note. PMF, probability mass function.

population. Moreover, the risk pool in Firm B improves, which is reflected by increased LTV of the firm's own clients. The LTV of switchers decreases, because the risk pool of Firm B (potential switchers) deteriorates.

A selective poaching policy could potentially be achieved through dynamic contract, without collecting additional information on switchers. Particularly, the company could increase the slope of the bonus-malus discount scheme to produce larger premiums in cases where claims are incurred. This should, in theory, discourage riskier clients of the competitor from searching, as well as riskier quitters from coming back to the market. We conduct a counterfactual that entails a unilateral change in the bonus-malus penalty by Firm B. Specifically, we keep the discount in risk class 1 unchanged, and change the difference in premiums between risk class 1 and subsequent risk classes.

Table 11 presents the results of the contract counterfactual. We consider a wide range of contracts, from a 1.3 times steeper contract to an extreme, 6 times steeper contract. We show that even significant changes in the contract structure do not result in improvement in the company's own riskiness pool. Only a dramatic change to a 6 times steeper contract delivers effects comparable to those of a medium selective price increase on switchers. Such change is, however, impractical because the regulator would be unlikely to allow it. This shows that dynamic contract design based on driving history has its limits and, in practice, cannot substitute for collecting more information about riskiness.[28]

The above results have implications for both firms and regulators. To set an optimal price, the firm has to take into account the change in risk composition of their customers, in addition to the overall price elasticity of demand.[29] The firm should also consider investing in screening technologies that enable screening riskier switchers. From a regulatory standpoint, the deterioration of risk pool as the price increases echoes the classical results in the literature on adverse selection. The decrease of the price offered in the market for "lemons" leads to increase in the proportion of lemons in the market, because nonlemons decide not to enter [or, as noted by Ackelof (1970, p. 489), "The 'bad' cars tend to drive out the good"]. In our case, the increase in premiums discourages nonlemons to search. In extreme cases, this dynamics may lead to market failure, in which no profitable contract can be written that would be acceptable for switchers [for an example of a similar mechanism in the monopoly market, see Rothschild and Stiglitz (1976)].

5.6. Moral Hazard

The analysis in Section 5 abstracts from both ex ante and ex post moral hazard. Ex ante moral hazard occurs when drivers are able to modify their riskiness before accidents are realized (see Abbring et al. 2003). Ex post moral hazard (see Einav et al. 2013) occurs when drivers settle the damages without filing a claim with the insurance company. Jeziorski et al. (2017) find no evidence of the latter in our data; however, they do find evidence for the former. In the remainder of this section, we discuss the implications of ex post moral hazard for our results.

Moral hazard may impact our conclusions in two ways. First, as demonstrated by Jeziorski et al. (2017), omitting risk adjustment leads to underestimation of the degree of unobserved heterogeneity in risk.[30] If the degree of risk heterogeneity is underestimated, the size of the risk gap between switchers and nonswitchers is likely to be underestimated as well. This is because the variance of λ (compared with the variable of ϵ) is directly driving the size of the risk gap. To address this issue, we reestimate the model with calibrated within-individual risk adjustments across risk classes (borrowing the numbers from Jeziorski et al. 2017). We find nearly 50% larger variability in unobserved risk, compared with the case with no moral hazard. This, however, translates to only 10% bias in the size of the switcher–stayer gap. This is possibly because the riskiest drivers are quitting driving altogether, which leads to truncation of the risk distribution. Note that the risk gap between quitters and stayers increases by 20%. This truncation attenuates the impact of tails of F_λ on the switcher–stayer gap.

Second, moral hazard may affect our counterfactuals. Changing the slope of the risk-class surcharges impacts the incentives to drive well. Naturally, the steeper the penalty structure, the less claims should

Table 11. Counterfactual Contracts: Steeper Incentive Contract

Contract slope	Market share	Riskiness	Stayer LTV	Switcher LTV
Baseline	0.50	0.039	339	320
×1.3	0.50 (−0.2%)	0.040 (+1.6%)	337 (−0.7%)	318 (−0.7%)
×1.5	0.50 (−0.4%)	0.040 (+1.6%)	338 (−0.5%)	319 (−0.4%)
×2	0.50 (−0.8%)	0.040 (+1.5%)	339 (+0.0%)	320 (+0.2%)
×6	0.48 (−3.0%)	0.039 (−0.8%)	351 (+3.6%)	332 (+3.9%)

occur. If moral hazard is indeed large, our findings should be interpreted as quantification of the impact of adverse selection on optimal pricing and contract design, keeping the individual risk fixed. Given that 81% of customers do not switch, the impact of moral hazard of our counterfactuals should be similar to the one estimated by Jeziorski et al. (2017), who analyze only customers who never switch. Thus, the full picture of the trade-off between contracts and extra information can be obtained by aggregating the results of both papers. This would require developing a model of joint churn and risk production, which we leave for future research.

6. Conclusion

We study switching in a market with heterogeneous cost to serve by analyzing the data from a leading car insurance provider in Portugal. We find evidence for adverse selection when poaching customers from competitors. New customers that switch from the competitor are significantly more costly to serve than the company's observationally equivalent own customers. In particular, switchers generate a 23% higher number and 20% larger volume of claims than observationally equivalent own customers with the same driving history. Furthermore, after controlling for all observable characteristics, including the number of years with a driving license, risk is related to tenure with the current provider. Specifically, customers with one to two years of tenure are more risky than customers with three or more years of tenure. The relationship between the number of years with the company and riskiness becomes flat beyond the third year of tenure. We also demonstrate that commonly used measures to mitigate imperfect information about riskiness, such as demographic characteristics and driving history, can account for only less than 50% of the riskiness gap between switchers and nonswitchers. Thus, pricing only on currently used covariates does not allow the firm to close the gap between switchers and nonswitchers.

We show that the population of switchers is heterogeneous in risk and that some of this heterogeneity can be screened using observables. For example, switchers with bad driving histories are exceptionally risky. However, a statistically significant part of the heterogeneity is unobserved. Specifically, frequent switchers are particularly costly to serve, and are possibly hard to detect at the point of contract signing. We conduct similar a analysis for churners and show a qualitatively comparable but quantitatively larger realized risk gap between churners and nonchurners. Specifically, churners generate three times more cost to serve than nonchurners. We find that filing a claim is related to 15 percentage point higher likelihood of churning.

We postulate a simple churn model, in which consumers vary with the their inherent riskiness. In the model, the consumer can stay with the current provider, search for better prices (and possibly switch), or quit driving. The least risky clients have the most incentive to stay, the moderately risky clients have the most incentive to search, and the riskiest clients have the most incentive to quit. These incentives rationalize the riskiness gap between switchers and nonswitchers, as well as the large gap in realized risk between churners and nonchurners. We show that increasing the price to switchers may lead to deterioration of the company's own risk pool of drivers, because higher prices discourage low-risk customers of the competitor to search. We show the limitations of contract design based on the current level of information on the riskiness of switchers, as well as benefits of obtaining additional information. The latter may explain why many major American car insurers are eager to introduce devices monitoring driving behavior combined with pay-as-you-drive pricing schemes. However, thus far, we have no evidence on the effectiveness of these practices. Another way to obtain additional information is to better incentivize sales forces by tying their compensation to the realized risk of the clients they acquire. The efficient design of such incentive contracts remains an open empirical question.

Acknowledgments

This paper benefited from discussions with Miguel Villas-Boas and presentation at the 2016 INFORMS Marketing Science Conference.

Endnotes

[1] We define poaching as any instance in which the customer receives a lower price offer from the competitor (see Fudenberg and Villas-Boas 2006). To receive that price offer, the customer may have been directly approached by a firm or may have actively searched for deals.

[2] Villas-Boas (1995) finds empirical evidence for price discrimination across "switchers" and "loyals" in coffee and saltine cracker categories. Rossi et al. (1996) elaborates on the value of pricing based on customer purchase history.

[3] Although there is little empirical literature on the effectiveness of poaching in markets with variation in cost to serve, there are several empirical papers that examine markets with heterogeneous willingness to pay. Neslin (1990) studies the effect of competitive couponing using retail scanner data and shows that this strategy may not always be profitable. In a related work, Deighton et al. (1994) examine brand switching in the grocery retail industry and show that advertising induces brand switching, but not repeated purchasing.

[4] Auto insurance proves to be a particularly well-suited laboratory to study the unobserved variation in riskiness. The contracts in this industry are heavily regulated and relatively standardized, observable characteristics are easy to define, and ex post riskiness is relatively straightforward to identify.

[5] Our results contribute to the vast literature on unobserved riskiness in insurance markets. This literature was spurred by the seminal work of Rothschild and Stiglitz (1976), who showed that heterogeneous

consumer-level risk that is unobservable to the insurer can lead to underprovision of insurance, deterioration of firms profits, and market failure. However, the empirical literature on auto insurance offers conflicting evidence on the mere existence of private information on riskiness (for relevant work outside of auto insurance, see Petersen and Rajan 1994, Padilla and Pagano 1997, Ausubel 1999, De Meza and Webb 2001, Jappelli and Pagano 2002, Finkelstein and Poterba 2004, Finkelstein and McGarry 2006, Brown et al. 2009, Karlan and Zinman 2009, Polyakova 2016). Puelz and Snow (1994) find unobserved heterogeneity, whereas Chiappori and Salanie (2000) argue that controlling for observables nonparametrically produces the opposite result. Cohen (2005) demonstrates informational asymmetries; however, a paper by Cohen and Einav (2007) argues that observationally equivalent drivers are essentially homogeneous in riskiness. More recent papers that incorporate moral hazard into the analysis (see Marcel Boyer 1989; Abbring et al. 2003, 2008; Ceccarini 2008; Dionne et al. 2013; Jeziorski et al. 2017) again document the existence of private information.

[6] Although it is plausible that the sales force may be able to screen for some of the unobserved variation in riskiness [Misra and Nair (2011) and Chung et al. (2013) describe the relationship between sales force incentive contracts and overall sales force performance; in a more related study, Kim et al. (2019) analyze the interaction between sale force incentive contracts and the level of customers' adverse selection], we show that partial delegation of pricing downstream fails to provide further screening of switchers. Particularly, we observe a nonmonotonic relationship between the allocation of discretionary discounts and ex post realized risk of switchers.

[7] The paper is related to the theoretical literature on BBP and targeted promotions (see Fudenberg and Tirole 2000, Villas-Boas 2004, Fudenberg and Villas-Boas 2006, Pazgal and Soberman 2008, Chen and Pearcy 2010, Esteves 2010, Zhang 2011, Caillaud and De Nijs 2014). This literature describes the implications of BBP in markets with heterogeneous consumer preferences and homogeneous cost to serve [a notable exception is Matsumura and Matsushima (2015), who consider heterogeneous marginal cost]. A more related theoretical literature studies implications of private information on borrowers' risk in credit markets (see Stiglitz and Weiss 1981, Bester 1985, Rajan 1992, Pagano and Jappelli 1993, Padilla and Pagano 1997, Hellmann and Stiglitz 2000). Finally, a related theoretical study by Shin and Sudhir (2010) describes conditions for the profitability of BBP in markets with heterogeneity in the consumer value to the firm (generated by varying purchase quantity) and time variability of consumer preferences.

[8] Drivers that refuse to disclose their driving history are not the same as drivers with no history, who are placed in risk class 10. Importantly, it is illegal for drivers with a bad history to pretend to have no history.

[9] The number of observed claims is not the same as the number of accidents, because some accidents may be unreported. However, from the perspective of cost to serve, we should measure the reported claims only. Hence, for the purpose of our analysis, the word "riskiness" should be understood as the propensity to generate cost to serve, not the as the propensity to generate social harm. We acknowledge that part of the riskiness variation documented in this paper may be due to underreporting or ex post moral hazard (see Abbring et al. 2008). Importantly, this possibility does not impact our main conclusions.

[10] We consider only liability policies, which are not contaminated by self-selection. The observable gap in cost to serve and premiums in collision contracts between switchers and nonswitchers is larger than the observable gap in liability contracts. Thus, the estimates of the corresponding differences between switchers and nonswitchers based on collision contracts should be larger than those based on liability contracts.

[11] The linear specification may provide inaccurate marginal effects for individuals with near-zero propensity toward accidents. To investigate this possibility, we replicated all linear regressions with the Poisson count model and obtained numerically identical results.

[12] Note that in the tenure regressions, we change the baseline riskiness group from nonswitchers to switchers; that is, we exclude the switcher dummy instead of excluding the nonswitcher dummy. The change is purely expositional and allows for easier comparison between nonswitchers with varying tenure and switchers.

[13] This test cannot distinguish between unobserved driving ability and moral hazard. This distinction is less relevant for the discussion in this paper; nevertheless, Abbring et al. (2003) and Jeziorski et al. (2017) present evidence that both factors are important.

[14] A gap in the realized risk across contracts is also the evidence for a separating equilibrium. Importantly, we do not need to assume a separating equilibrium for our test to be valid. See Puelz and Snow (1994) for a discussion of these issues.

[15] The insurance literature has developed two ways to introduce demand friction: nonzero search cost and nonzero switching cost. Historically, the car insurance literature was analyzed using the search cost paradigm; see Honka (2014) and Honka and Chintagunta (2016). This approach matches the industry well, because the price menu is obfuscated by discretionary discounts, and consumers are unlikely to know what discretionary discount a competitor can offer without incurring the cost of obtaining a quote. Apart from premiums, car insurance is composed of nearly homogeneous products, and sale-force agents are trained to facilitate nearly costless switching. Conversely, the health insurance literature tends to use a switching cost paradigm, because comparing across different health insurance options is usually quite difficult; see Handel (2013) and Handel and Kolstad (2015). That being said, we conducted simulations showing that the version of our model with switching instead of search cost generates similar switching patterns; thus, this choice is unlikely to be consequential for our results.

[16] We use a nonstationary model for two reasons: (i) because there is natural driving limit, the finite horizon setup better reflects the reality, and (ii) the nonstationary model is easier to solve numerically. This is because some payoff-relevant variables, such as age and driving experience, are deterministic functions of t, thus, they do not need to explicitly enter the state in a nonstationary model.

[17] The three-tier pricing schedule is common across insurance industries in many other countries, including the United States. The baseline premium and driving history discount schedule is usually set by a risk-management team composed of actuarial professionals. Discretionary discounts are usually under the control of the marketing team, including a regional manager and sales force.

[18] We allow λ_{it} to vary over time in an exogenous fashion; however, following Cohen and Einav (2007), we do not allow drivers to choose λ_{it}. This rules out moral hazard. We discuss the implication of this assumption for our results in Section 5.6.

[19] As demonstrated by Cohen and Einav (2007), heterogeneity in risk aversion is identified from our data; however, because our goal is to introduce the simplest model explaining switching behavior, we apply Occam's razor and refrain from introducing heterogeneity in risk aversion in our specification. Instead of estimating the degree of risk aversion heterogeneity, we leverage on the results by Jeziorski et al. (2017), who estimate the distribution of risk aversion in the same market. We use their estimates to calibrate and reestimate our model. We show that our results become stronger (see the online appendix).

[20] We assume that consumer can obtain only one quote at a time, but may obtain multiple quotes across time periods. This approach is similar to that used by Seiler (2013). The Portuguese market has a

substantially lower number of insurance firms than the United States; thus, it is unlikely that consumers obtain many quotes simultaneously. If they indeed do, our results are valid as long as the riskiness does not impact the number of obtained quotes in a meaningful way.

[21] This assumption implies that the search cost is orthogonal to λ_i. In our setting, direct correlation of search cost and riskiness is not necessary to generate selective switching. In particular, the multiplicative insurance premium structure already has selective switching incentives. We illustrate it using a simple numerical example. Consider two risk-neutral customers, A and B, living two periods, without discounting. Customer A has 0% probability of an accident, and customer B has 50% probability. Both customers start with a \$100 baseline premium and 0% discretionary discount. If they search, they can obtain a 20% discretionary discount with 50% probability. The penalty in the case of an accident is 50%. The payoff of customer A without search is −\$200, and that with search is $-\$190 - C$. Consumer A searches iff $C < \$10$. The payoff of customer B without search is −\$225. Payoff with search is $-\$212.5 - C$. Consumer B searches iff $C < \$12.5$. A more risky consumer B would search more, and thus switch more often.

[22] Although we do not need the correlation to generate adverse selection in our market, estimating such correlation could help generalizable our results to other markets, especially if we want to generalize to markets that do not have the multiplicative discount structure. In the online appendix, we investigate the possibility of correlated riskiness and search cost. We find that riskiness and search costs are indeed correlated, and the model with correlation generates a 40% larger riskiness gap. Thus, little more than 60% of the adverse selection is related to the contract structure; the remaining adverse selection is a result of the correlation in the primitives. Results from our main specification should be regarded as conservative.

[23] In the Portuguese market, each driver is characterized by two risk classes; however, because 95% of the time they are equal in the data, we assume that the comprehensive risk class moves together with the liability risk class. This simplification results in computational improvements that make the estimation feasible. This can be violated when the accident results in damage to only the counterparty's car or only to the insured's own car. The former is extremely rare. The latter can be more common, because it occurs when no counterparty is involved in the accident. Although we assume away the differential effects of accidents with no counterparty on the risk-class transition, we allow for the probability mass on the zero value in the distribution of L_{itr}.

[24] See APS (2013). If necessary, it is straightforward to extend the model to contain asymmetric pricing; however, the identification of an extended model would require data on the prices of competitors. The data on the customer pool of the competitor are not required.

[25] We recognize that it is infeasible to use the full data and make no parametric assumptions because of the long tail of car characteristics. Thus, aiming for maximum internal validity, we restrict to the sample in which no parametric restrictions about car characteristics are necessary. However, recognizing that we trade off external validity for internal validity, we repeated the exercise using 95% of the sample containing 15 modal car makes. We report the results in the online appendix. We find that our baseline estimates generate a qualitatively and quantitatively similar risk gap; thus, our main results generalize to the full population. The reduced form results in the previous section use the full population.

[26] Two consumers with the same ex ante riskiness may occupy different risk classes because accidents are subject to some degree of randomness. Similarly, because of assumption (i), we are likely to observe two otherwise equivalent customers that possess different discretionary discounts.

[27] The pricing and contract counterfactuals study the short-run effects of unilateral deviations. Such analysis does not capture longer-run equilibrium effects. In particular, it does not account for the possibility of the competitors altering their pricing and contract menus. Full equilibrium analysis would require a dynamic model of competition, which is beyond the scope of this paper.

[28] This exercise is aimed at isolating the impact of the incentive structure on the riskiness pool through the selection of customers. If moral hazard is important, changes in contract additionally affect the riskiness of the firm's *own* clients (see Jeziorski et al. 2017).

[29] We refrain from suggesting an optimal price for the firm for two reasons: (i) as we discuss in Section 3, we do not observe the full extent of marginal cost, and (ii) we do not attempt to compute the Nash equilibrium of the model.

[30] The identification of the heterogeneity in risk relies on the variation in realized risk across risk classes. If moral hazard is present, this variation is usually endogenously compressed, leading to underestimation of the variance of λ. This is because inherently riskier drivers residing in higher risk classes face steeper penalties for accidents and thus put more effort toward reducing their riskiness.

References

Abbring JH, Chiappori PA, Pinquet J (2003) Moral hazard and dynamic insurance data. *J. Eur. Econom. Assoc.* 1(4):767–820.

Abbring JH, Chiappori PA, Zavadil T (2008) Better safe than sorry? Ex ante and ex post moral hazard in dynamic insurance data. Tinbergen Institute Discussion Paper No. 08-075/3, CentER Discussion Paper No. 2008-77, Tinbergen Institute, Amsterdam and Tilburg University, Tilburg, Netherlands.

Akerlof GA (1970) The market for "lemons": Quality uncertainty and the market mechanism. *Quart. J. Econom.* 84(3):488–500.

APS (2013) Insurance in Portugal: Insurance market overview 12/13. Associação Portuguesa de Seguradores, Lisboa. Accessed January 1, 2015, http://www.apseguradores.ptt.

Ausubel LM (1999) Adverse selection in the credit card market. Working paper, University of Maryland, College Park.

Barros PP (1996) Competition effects of price liberalization in insurance. *J. Indust. Econom.* 44(3):267–287.

Basu AK, Lal R, Srinivasan V, Staelin R (1985) Salesforce compensation plans: An agency theoretic perspective. *Marketing Sci.* 4(4): 267–291.

Berger PD (1972) On setting optimal sales commissions. *Oper. Res. Quart.* 23(2):213–215.

Bester H (1985) Screening vs. rationing in credit markets with imperfect information. *Amer. Econom. Rev.* 75(4):850–855.

Brown M, Jappelli T, Pagano M (2009) Information sharing and credit: Firm-level evidence from transition countries. *J. Financial Intermediation* 18(2):151–172.

Caillaud B, De Nijs R (2014) Strategic loyalty reward in dynamic price discrimination. *Marketing Sci.* 33(5):725–742.

Ceccarini O (2008) Does experience rating matter in reducing accident probabilities? A test for moral hazard. Unpublished doctoral dissertation, University of Pennsylvania, Philadelphia.

Chen Y, Pearcy J (2010) Dynamic pricing: When to entice brand switching and when to reward consumer loyalty. *RAND J. Econom.* 41(4):674–685.

Chiappori P-A, Salanie B (2000) Testing for asymmetric information in insurance markets. *J. Political Econom.* 108(1):56–78.

Chung DJ, Steenburgh T, Sudhir K (2013) Do bonuses enhance sales productivity? A dynamic structural analysis of bonus-based compensation plans. *Marketing Sci.* 33(2):165–187.

Cohen A (2005) Asymmetric information and learning: Evidence from the automobile insurance market. *Rev. Econom. Statist.* 87(2):197–207.

Cohen A, Einav L (2007) Estimating risk preferences from deductible choice. *Amer. Econom. Rev.* 97(3):745–788.

Dahlby B, West D (1986) Price dispersion in an automobile insurance market. *J. Political Econom.* 94(2):418–438.

De Meza D, Webb DC (2001) Advantageous selection in insurance markets. *RAND J. Econom.* 32(2):249–262.

Deighton J, Henderson CM, Neslin SA (1994) The effects of advertising on brand switching and repeat purchasing. *J. Marketing Res.* 31(1):28–43.

Dionne G, Michaud P-C, Dahchour M (2013) Separating moral hazard from adverse selection and learning in automobile insurance: Longitudinal evidence from France. *J. Eur. Econom. Assoc.* 11(4):897–917.

Dolan RJ, Simon H (1996) *Power Pricing: How Managing Price Transforms the Bottom Line* (Free Press, New York).

Einav L, Finkelstein A, Ryan SP, Schrimpf P, Cullen MR (2013) Selection on moral hazard in health insurance. *Amer. Econom. Rev.* 103(1):178–219.

Esteves R-B (2010) Pricing with customer recognition. *Internat. J. Indust. Organ.* 28(6):669–681.

Finkelstein A, McGarry K (2006) Multiple dimensions of private information: Evidence from the long-term care insurance market. *Amer. Econom. Rev.* 96(4):938–958.

Finkelstein A, Poterba J (2004) Adverse selection in insurance markets: Policyholder evidence from the UK annuity market. *J. Political Econom.* 112(1):183–208.

Fong NM, Fang Z, Luo X (2015) Geo-Conquesting: Competitive locational targeting of mobile promotions. *J. Marketing Res.* 52(5): 726–735.

Fudenberg D, Tirole J (2000) Customer poaching and brand switching. *RAND J. Econom.* 31(4):634–657.

Fudenberg D, Villas-Boas JM (2006) Behavior-based price discrimination and customer recognition. *Economics and Information Systems*, Handbooks in Information Systems, vol. 1 (Elsevier, Amsterdam), 377–436.

Handel BR (2013) Adverse selection and inertia in health insurance markets: When nudging hurts. *Amer. Econom. Rev.* 103(7): 2643–2682.

Handel BR, Kolstad JT (2015) Health insurance for "humans": Information frictions, plan choice, and consumer welfare. *Amer. Econom. Rev.* 105(8):2449–2500.

Hellmann T, Stiglitz J (2000) Credit and equity rationing in markets with adverse selection. *Eur. Econom. Rev.* 44(2):281–304.

Honka E (2014) Quantifying search and switching costs in the US auto insurance industry. *RAND J. Econom.* 45(4):847–884.

Honka E, Chintagunta P (2016) Simultaneous or sequential? Search strategies in the us auto insurance industry. *Marketing Sci.* 36(1): 21–42.

Jappelli T, Pagano M (2002) Information sharing, lending and defaults: Cross-country evidence. *J. Banking Finance* 26(10):2017–2045.

Jeziorski P, Krasnokutskaya E, Ceccarini O (2017) Adverse selection and moral hazard in a dynamic model of auto insurance. Technical report, University of California, Berkeley, Berkeley.

Karlan D, Zinman J (2009) Observing unobservables: Identifying information asymmetries with a consumer credit field experiment. *Econometrica* 77(6):1993–2008.

Kim M, Sudhir K, Uetake K, Canales R (2019) Multidimensional sales incentives in CRM settings: Customer adverse selection and moral hazard. *J. Marketing Res.* Forthcoming.

Lal R (1986) Technical note-delegating pricing responsibility to the salesforce. *Marketing Sci.* 5(2):159–168.

Marcel Boyer GD (1989) An empirical analysis of moral hazard and experience rating. *Rev. Econom. Statist.* 71(1):128–134.

Matsumura T, Matsushima N (2015) Should firms employ personalized pricing? *J. Econom. Management Strategy* 24(4):887–903.

McFadden D (1989) A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica* 57(5):995–1026.

Misra S, Nair HS (2011) A structural model of sales-force compensation dynamics: Estimation and field implementation. *Quant. Marketing Econom.* 9(3):211–257.

Neslin SA (1990) A market response model for coupon promotions. *Marketing Sci.* 9(2):125–145.

OECD (2013) *Environment at a Glance 2013: OECD Indicators* (OECD Publishing, Paris).

Padilla AJ, Pagano M (1997) Endogenous communication among lenders and entrepreneurial incentives. *Rev. Financial Stud.* 10(1): 205–236.

Pagano M, Jappelli T (1993) Information sharing in credit markets. *J. Finance* 48(5):1693–1718.

Pakes A, Pollard D (1989) Simulation and the asymptotics of optimization estimators. *Econometrica* 57(5):1027–1057.

Pazgal A, Soberman D (2008) Behavior-based discrimination: Is it a winning play, and if so, when? *Marketing Sci.* 27(6): 977–994.

Petersen MA, Rajan RG (1994) The benefits of lending relationships: Evidence from small business data. *J. Finance* 49(1):3–37.

Polyakova M (2016) Regulation of insurance with adverse selection and switching costs: Evidence from Medicare Part D. *Amer. Econom. J.: Appl. Econom.* 8(3):165–195.

Puelz R, Snow A (1994) Evidence on adverse selection: Equilibrium signaling and cross-subsidization in the insurance market. *J. Political Econom.* 102(2):236–257.

Rajan RG (1992) Insiders and outsiders: The choice between informed and arm's-length debt. *J. Finance* 47(4):1367–1400.

Reinartz W, Krafft M, Hoyer WD (2004) The customer relationship management process: Its measurement and impact on performance. *J. Marketing Res.* 41(3) 293–305.

Rossi PE, McCulloch RE, Allenby GM (1996) The value of purchase history data in target marketing. *Marketing Sci.* 15(4) 321–340.

Rothschild M, Stiglitz J (1976) Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quart. J. Econom.* 90(4):629–649.

Seiler S (2013) The impact of search costs on consumer behavior: A dynamic approach. *Quant. Marketing Econom.* 11(2):155–203.

Shin J, Sudhir K (2010) A customer management dilemma: When is it profitable to reward one's own customers? *Marketing Sci.* 29(4): 671–689.

Shin J, Sudhir K, Yoon D-H (2012) When to "fire" customers: Customer cost-based pricing. *Management Sci.* 58(5):932–947.

Stephenson PR, Cron WL, Frazier GL (1979) Delegating pricing authority to the sales force: The effects on sales and profit performance. *J. Marketing* 43(2):21–28.

Stiglitz JE, Weiss A (1981) Credit rationing in markets with imperfect information. *Amer. Econom. Rev.* 71(3):393–410.

Venkatesan R, Kumar V (2004) A customer lifetime value framework for customer selection and resource allocation strategy. *J. Marketing* 68(4):106–125.

Villas-Boas JM (1995) Models of competitive price promotions: Some empirical evidence from the coffee and saltine crackers markets. *J. Econom. Management Strategy* 4(1):85–107.

Villas-Boas JM (2004) Price cycles in markets with customer recognition. *RAND J. Econom.* 35(3):486–501.

Weinberg CB (1975) An optimal commission plan for salesmen's control over price. *Management Sci.* 21(8):937–943.

Zhang J (2011) The perils of behavior-based personalization. *Marketing Sci.* 30(1):170–186.