Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Individualized Offers and the Measurement of Consumer Inertia: Evidence from Auto Insurance

Helena Perrone

h.perrone@tbs-education.fr

TBS Education

Fabricio Valiati

Agila Haustierversicherung AG

Research Article

Keywords:

Posted Date: November 9th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11206268/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Individualized Offers and the Measurement of Consumer Inertia: Evidence from Auto Insurance[‡]

September 2026

Abstract

Demand estimates can overstate consumer inertia when researchers do not observe the alternatives and prices available to each consumer. We study this issue using auto insurance renewal data that record both the chosen contract and all insurer quotes returned for the policyholder, including the premiums on alternative offers. We exploit these data by estimating the same demand model as we sequentially remove information on alternative prices and consumer-specific availability. With the observed offers, the premium-equivalent incumbency effect is about € 106, or 19% of the mean annual premium. Replacing alternative premiums with predicted prices while holding each consumer's observed offer set fixed raises this measure to about € 183. Replacing the consumer-specific offer set with a common market-level set raises it further to about € 232. The first step mainly reduces estimated price sensitivity, while the second raises the estimated incumbency effect. Missing offer-level information can therefore substantially affect measured price responsiveness and consumer inertia.

Keywords: Consumer Inertia, Offer Sets, Consideration, Auto Insurance, Demand Estimation, Switching

JEL Codes: D12, D83, G22, L11, M31

[*] Previously circulated as "Drivers of Switching in Autoinsurance: Evidence from observable and exogenous consideration sets"

[†]

1 Introduction

Persistent consumer choice is often interpreted as evidence of substantial switching costs. In many markets, however, researchers observe the product a consumer chooses but not the alternatives and prices available to that consumer at the time of choice. Demand estimation must then rely on model structure and assumptions about the consumer's choice environment. These modeling choices can matter because consumers may differ in the alternatives available to them, while price sensitivity is identified from price differences among those alternatives.

We study this measurement problem using insurance renewal data with a rare feature: for each policyholder, we observe both the chosen contract and the full set of premium quotes returned by insurers to the broker. We therefore observe information that is usually missing in demand applications. We first estimate demand using the returned offers and quoted premiums and then remove components of that information in a controlled way, asking how estimates of price sensitivity and incumbency change when the premiums on alternative offers, and then the consumer-specific offer set, are no longer observed.

Our empirical setting is the Brazilian auto insurance market, where annual renewals are intermediated by brokers. Insurers respond to renewal requests with individualized premium quotes, generating offer sets that vary across consumers in their size, composition, and prices. Insurers choose whether to return a quote, so these sets reflect supply-side selection rather than random assignment. We refer to the set of quotes returned by insurers to the broker as the observed offer set.[1]

The empirical analysis has three parts. We first document how switching varies with the offer environment and examine renewal pricing. We then estimate

[1] We use observed offer set rather than consideration set because quotes returned to the broker need not coincide with the alternatives ultimately considered by the consumer. Brokers may filter or recommend a subset of offers, and consumers may evaluate only some of the alternatives they see. We return to this distinction and its implications in Sections 3 and 6.

demand conditional on the observed offer set and quoted premiums. Finally, we remove two pieces of offer-level information in sequence. In the first exercise, we retain each consumer's actual observed offer set and chosen premium but replace the premiums on alternative offers with predicted prices. In the second, we also replace the consumer-specific observed offer set with a common state–month set. This sequential comparison lets us trace how the demand estimates change as alternative-price information and then consumer-specific availability are removed.

The reduced-form evidence shows a strong relationship between the offer environment and switching. Switching rates rise from 10% to 40% as the number of offers increases from 2 to 12, and the incumbent's relative price is strongly associated with switching. Specifications using only information from chosen contracts make price responsiveness look weak, whereas adding incumbent and alternative-offer information makes relative prices a strong predictor of switching. We also find that incumbents quote lower renewal premiums on average than competing insurers. Policyholders with a recent claim are also more likely to switch, consistent with prior evidence that claims experience affects insurer retention (Israel, 2005).

The demand estimates provide our main measurement result. Conditional on observed offer sets and quoted premiums, the premium-equivalent incumbency effect evaluated at sample-mean policyholder characteristics is approximately € 106, or about 19% of the mean annual premium. Because persistent consumer–insurer preferences and other forms of state dependence may also contribute to this coefficient, we interpret it as an incumbency effect rather than a pure switching cost. When we hold the actual offer sets fixed but replace the premiums on alternative offers with predicted prices constructed from observed policyholder and vehicle characteristics, insurer identities, location and time, and the chosen premium, estimated price sensitivity falls substantially and the premium-equivalent incumbency effect rises to about € 183. When we

also replace the consumer-specific offer set with the broader state–month set, price sensitivity changes little further but the estimated incumbency effect rises, producing a premium-equivalent value of about € 232. The two steps therefore show that missing prices and missing availability affect measured inertia through different margins.

A large literature develops methods for demand estimation when search, consideration, or choice-set heterogeneity is unobserved (e.g., Sovinsky Goeree, 2008; Conlon and Mortimer, 2013; Barseghyan et al., 2021; Abaluck and Adams-Prassl, 2021; Crawford et al., 2021). In most applications, the researcher must recover or approximate the relevant alternatives. Our setting provides a complementary empirical benchmark: because we observe the insurer offers returned for each renewal and their quoted premiums, we can measure directly how the same demand estimates change when information on alternative prices and consumer-specific availability is removed.

The results also speak to work on switching costs and consumer inertia (e.g., Shum, 2004; Kim et al., 2003; Dubé et al., 2010; Ho et al., 2017). This literature emphasizes the roles of switching frictions, state dependence, and persistent preferences in generating observed persistence. Our analysis highlights an additional measurement margin: estimates of price responsiveness and the monetary value attached to incumbency can change substantially with the information available about the alternatives and prices consumers face.

The rest of this paper is organized as follows. The next section reviews the relevant literature. Section 3 describes the institutional setting, including the Brazilian auto insurance market, the renewal process, and the policy calculation platform behind our data. Section 4 presents the data and summary statistics. Section 5 reports reduced-form evidence on switching behavior, pricing, and their determinants. Section 6 introduces the demand model with choice inertia, details empirical implementation and identification, and presents estimates as offer-level information is removed. This section also reports sensitivity to smaller effective

offer sets. The final section concludes.

2 Relevant literature

This paper relates to a broad literature in marketing and economics that studies switching behavior, choice persistence, and demand estimation in markets with consumer frictions. A central empirical finding is that consumers exhibit substantial inertia, often interpreted as evidence of switching costs. Empirical studies document large switching costs in consumer packaged goods (Shum, 2004), banking (Kim et al., 2003), and insurance markets (Ho et al., 2017). Several mechanisms may explain this persistence, including transaction costs, psychological loyalty, search and attention frictions, and learning about firm-specific quality (Klemperer, 1995; Dubé et al., 2010). Within marketing, this literature is central to understanding customer retention, loyalty, and dynamic pricing.

A growing body of work emphasizes that observed persistence may reflect not only switching costs but also the fact that consumers often choose from restricted and heterogeneous consideration sets rather than from the full set of available alternatives. Sovinsky Goeree (2008) models the role of advertising in shaping consideration sets and shows how this affects estimated preference parameters. Conlon and Mortimer (2013) demonstrate that exogenous variation in product availability due to stockouts induces heterogeneous choice sets and can lead to biased demand estimates if ignored. Recent methodological advances have developed structural approaches to recover consideration sets when unobserved. Honka et al. (2017) study how advertising affects consumer awareness and choice in the U.S. banking market. Moraga-González et al. (2023) propose methods to identify consideration sets from choice data under different behavioral assumptions. Reviews by Honka et al. (2019) and Crawford et al. (2021) provide comprehensive overviews of both the theoretical foundations and empirical applications of consideration set models.

Our paper is most closely related to Honka (2014), who uses survey data to observe which alternative insurers policyholders considered at renewal in auto insurance markets. She imputes prices for considered insurers that were not selected and estimates a structural model that jointly identifies search costs, switching costs, and preferences, showing that search frictions are important for explaining retention. Using the same data, Honka and Chintagunta (2017) show that search technology (simultaneous versus sequential search) can be identified from consideration-set information. Our data capture a different object. We do not observe the consumer's consideration set; instead, we observe all insurer quotes returned to the broker, including the premiums quoted by alternative insurers. This allows us to ask a complementary measurement question: how much do estimates of price sensitivity and incumbency change when policyholder-specific offer sets and premiums on alternative offers are not observed and must be reconstructed? Because offer sets arise from insurer quote/no-quote decisions rather than consumer search, they are not generated by consumers' search choices, but insurer participation may still be selective. We therefore condition on realized offer sets rather than model their formation.

Within auto insurance markets, several papers document switching costs and learning dynamics. Israel (2005) shows that consumers learn about insurer quality through claims experience and finds evidence of switching costs in the U.S. market. Jõeveer and Kepp (2023) study switching in the Estonian auto insurance market using standardized insurance panels provided through leasing companies. They document an 11% annual switching rate and show that switching behavior depends on claim type: consumers are less likely to switch after self-induced claims but more likely to switch after third-party claims, suggesting behavioral mechanisms such as gratitude. In the same setting, Kepp and Männasoo (2025) find that attention frictions are the primary source of inertia.

That misspecified choice sets can distort demand estimates is well understood. Our contribution is to exploit data that observe both individualized

insurer offer sets and the premiums quoted by alternative insurers to examine how demand estimates change as exact alternative-premium information and consumer-specific offer availability are successively replaced. Unlike Jõeveer and Kepp (2023), where offers are standardized through leasing arrangements, our data contain individualized insurer quotes, providing direct information on the offer environment faced by policyholders at renewal.

3 Auto Insurance in Brazil and the Policy Calculation Platform

The Brazilian auto insurance market offers two useful features for our analysis: contracts require an active annual renewal decision, limiting purely passive persistence from automatic renewal, and the platform records insurers' quote responses before the policyholder makes the renewal choice. This section describes the market's institutional structure, contract features, and renewal process.

3.1 Market and contract characteristics

Auto insurance represents a substantial share of non-life insurance markets worldwide. In Brazil, auto insurance accounted for 44% of total non-life written premiums in 2022 (SUSEP, 2022). The Brazilian insurance market is regulated by the Private Insurance Superintendence, which supervises insurers, brokers, and related entities and enforces compliance with legal and technical requirements.

Third-party motor vehicle liability insurance has historically been mandatory for all registered vehicles in Brazil. It is collected annually at the time of vehicle license renewal and provides limited coverage for third-party death, permanent disability, and medical expenses (SUSEP, 2006). All additional auto insurance coverage is voluntary and offered by private insurers through annual contracts. Insurers provide a range of coverage options, including comprehensive vehicle coverage, third-party liability, passenger accident coverage, and fully

comprehensive policies that bundle these components (SUSEP, 2022).

Premiums and deductibles are determined using characteristics of both the policyholder and the insured vehicle. Pricing typically combines actuarial risk assessment with cost loadings and strategic adjustments reflecting market conditions (Laas et al., 2016). For new contracts, insurers estimate claims probabilities using demographic and vehicle characteristics together with proprietary historical data. Their ability to assess and competitively price a particular risk can therefore vary with the composition of their own client databases, including geographic concentration. More generally, insurers differ in underwriting models, risk appetite, and target segments. At renewal, insurers can additionally use information about the existing relationship to predict retention and adjust renewal offers to meet internal retention and profitability targets.[2]

The Brazilian auto insurance market also features a Bonus–Malus system that links premiums to claims history. Policyholders begin with a bonus class of zero and gain one bonus point for each claim-free year, up to a maximum of ten. Submitting a claim results in a loss of bonus points, though the score cannot fall below zero. Because bonus class accumulates with claim-free experience, it is correlated with driver age and driving history and plays an important role in premium determination.

3.2 Renewal Process and the Policy Calculation Platform

Auto insurance contracts in Brazil are valid for one year, and renewal is not automatic. Each policyholder has an individual contract anniversary date, so renewals occur continuously throughout the year. Consumers cannot purchase insurance directly from insurers; all policies must be intermediated by licensed insurance brokers. Brokers are licensed and supervised by SUSEP and are com-

[2]This information is based on interviews with pricing managers in Brazilian auto insurance companies.

pensated through commissions paid by insurers, typically ranging from 10% to 20% of the premium value.

When a policy approaches expiration, the broker enters updated policyholder and vehicle information into a centralized calculation platform and submits a renewal request. The platform transmits the request to insurers, which then decide whether to return a quote and, if so, at what premium, based on their tariffs, risk assessments, and business strategies. An insurer may decline to quote, for example, because the applicant falls outside its risk appetite or target segment, or because it lacks sufficient information to price that risk competitively. Because insurers choose whether to quote, the resulting offer set reflects supply-side selection and need not be random. We therefore estimate demand conditional on the realized offer set rather than assuming exogenous offer availability.

The platform standardizes the information contained in the renewal request and the sequence through which insurers respond. Once the quotes are returned, the broker observes the full offer set and may present or recommend some or all of those offers to the policyholder, who then selects a contract. Importantly for our identification, the platform does not allow the broker to select case by case which eligible insurers receive the request or to privately transmit policyholder-specific preference information to insurers before they quote.

Our data include all price offers returned by insurers to brokers for each renewal request. We refer to this as the observed offer set. It is not necessarily the same as the set presented by the broker or actively considered by the policyholder. If brokers filter offers or consumers ignore some presented alternatives, the observed offer set is an upper bound on the effective choice set. Section 6.6 examines sensitivity to smaller effective sets.

4 Data and Summary Statistics

This section describes our dataset and presents summary statistics. Our data observe both consumer choices and the full set of renewal offers returned to the broker, enabling analysis of switching behavior with heterogeneous offer sets.

4.1 Dataset and sample construction

Our data come from a policy calculation and policy management platform that connects insurance brokers to insurers. The platform, which we refer to as Level-Up, provides centralized tools for managing renewals and other policy-related tasks and is widely used by brokers in the Brazilian auto insurance market. Level-Up handles approximately 14% of auto insurance calculation requests in Brazil.[3] Level-Up therefore covers an economically meaningful but not necessarily representative segment of the market. Insurer shares on the platform differ from national shares (Table 2), so our results are most directly informative about broker-mediated renewals processed through this platform. Table 2 compares insurer shares in the platform with industry-wide market shares.

We observe insurance policy renewals and the full set of premium offers returned in each renewal query. Our analysis focuses on fully comprehensive auto insurance contracts, which account for 82% of renewed contracts in the data. To ensure meaningful comparisons across offers within a query, we restrict attention to fully comprehensive premium offers. Restricting to fully comprehensive contracts ensures comparability across offers within a renewal query, since consumers rarely alter coverage tier at renewal.[4] In addition, changing core contract features may imply losing accumulated bonus classes and being reassigned to

[3] We use a fictitious name for the platform and label insurers as I1, I2, I3, etc., to comply with the non-disclosure agreement signed with the data provider. Insurer labels are ordered by platform policy count.

[4] This assumption is consistent with a two-stage decision process in which consumers first choose coverage and then select the contract with the highest indirect utility, as in Honka (2014).

class 0, which entails substantial premium increases and therefore creates strong incentives to keep coverage stable.

The sample consists of renewal queries for policies expiring between September 2021 and August 2022. For each query, we observe policyholder and vehicle characteristics, including age, gender, marital status, region, bonus class (both previous and current), claims history, vehicle age and value, and annual mileage. We also observe the set of offers returned in the query, including premiums, insurer identities, an indicator for the incumbent insurer, and key contract characteristics such as deductibles and coverage. All queries include an offer from the previous insurer.[5] This allows us to observe the incumbent's quoted premium for all policyholders.

We exclude from the sample queries in which the policyholder received only one offer, as switching is not feasible in these cases. These single-offer queries represent roughly 1% of overall queries.

For the demand analysis, we focus on the metropolitan regions of Rio de Janeiro and São Paulo, the two densest markets in the platform. Concentrating on these markets provides substantially greater consumer and insurer support for demand estimation. After restricting the sample to these markets and excluding renewal queries with only one remaining offer, the demand-estimation sample contains 42,069 renewal queries and 252,648 insurer offers.

4.2 Summary Statistics

Table 1 reports summary statistics separately for renewal queries that resulted in switching and those that did not. The upper panel summarizes characteristics of the offers within a query, while the lower panel reports policyholder

[5]This reflects institutional practice rather than universal retention effort. Industry interviews reveal that insurers prefer offering uncompetitive (high) premiums to customers they wish to shed rather than sending explicit non-renewal letters, which damage relationships with brokers and policyholders. Some incumbent offers thus represent "polite exits" rather than competitive bids.

characteristics. Premium PI denotes the premium offered by the previous insurer (incumbent). Premiums are denominated in Brazilian reais in the raw data. For ease of interpretation, we report monetary values in euros using a fixed conversion rate of R$5.40 per euro.

Table 1: Summary Statistics: Offers and Policyholders

	No Switch		Switch	
	mean	sd	mean	sd
Offers per query				
Premium PI is the lowest	0.72	0.45	0.05	0.22
Mean Premium (€)	570.42	378.46	549.74	324.77
% Premium PI – Mean Premium	-31.06	46.40	-7.34	28.47
Number of Offers	6.36	2.56	7.28	2.31
Policyholder				
Female	0.41	0.49	0.43	0.50
Age	52.25	13.65	50.62	13.32
Married	0.78	0.42	0.78	0.41
Current Bonus	6.67	3.43	6.07	3.57
Claimant	0.04	0.20	0.07	0.25
Observations	57,118		24,271	

Notes: Premium PI is the lowest = 1: previous insurer's premium is the lowest in the offer set. % Premium PI – Mean Premium = 100 × (Premium PI – Mean Premium)/Mean Premium. Current Bonus = policyholder's bonus class (0–10). Claimant = 1 if filed a claim in the past year. Number of Offers is the count of insurers returning a quote in a given query, excludes single-offer queries. sd = standard deviation

The two subsamples differ along several observable dimensions. The mean premium among non-switchers is € 570.42, compared with € 549.74 among switchers, with substantial dispersion in both groups. In 72% of non-switcher queries, the incumbent insurer offers the lowest premium in the observed offer set; among switchers, this occurs in only 5% of cases. Consistent with this pattern, the relative premium difference between the incumbent insurer and the query mean premium is substantially higher for non-switchers than for switchers.

Observed offer sets also differ systematically across groups. Non-switchers receive an average of 6.4 offers per query, compared with 7.3 offers among switchers, suggesting that larger observed offer sets are associated with higher switching

rates. The lower panel of Table 1 shows that switchers are slightly younger and have lower average bonus classes than non-switchers. Correspondingly, the incidence of recent claims is higher among switchers. Gender and marital status distributions are nearly identical across the two groups.

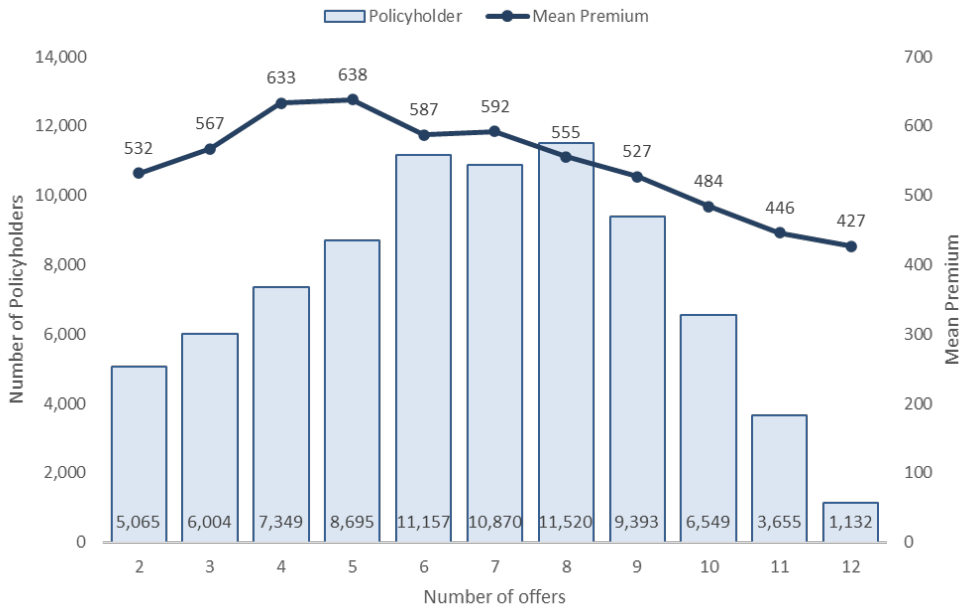
Table 2 reports insurer-level statistics, including market shares, switching rates, conversion rates, and the frequency with which each insurer returns a quote in renewal queries ("Quoted"). Market shares are measured both using industry-wide gross written premiums and within-platform policy counts. There is substantial heterogeneity across insurers along all dimensions. The largest insurer in the platform is I1, with approximately 23% of policies, followed by I2 (20%) and I3 (19%). Conversion rates vary widely: insurers I9 and I8 have conversion rates of 60% and 55%, respectively, while other insurers exhibit much lower conversion rates. Switching rates from incumbent insurers range from approximately 20% to 40%, indicating meaningful differences in retention performance across insurers. Insurers also differ in how frequently they appear in consumers' observed offer sets, with quote frequencies ranging from 16% to 80%.

Table 2: Insurer Market Shares and Switching Outcomes

Insurer	Industry Share	Platform Share	Switching	Conversion	Quoted
I1	0.13	0.23	0.25	0.20	0.80
I2	0.12	0.20	0.21	0.23	0.79
I3	0.19	0.19	0.37	0.09	0.80
I4	0.07	0.15	0.32	0.20	0.68
I5	0.07	0.08	0.37	0.36	0.60
I6	0.01	0.04	0.34	0.16	0.16
I7	0.03	0.03	0.36	0.28	0.50
I8	0.06	0.03	0.40	0.55	0.49
I9	0.02	0.03	0.35	0.60	0.32
I10	0.01	0.02	0.21	0.40	0.23
Others	0.28	–			

Notes: Industry Share is based on 2022 Gross Written Premium from SUSEP (`https://www2.susep.gov.br/menuestatistica/SES/principal.aspx`). Platform Share is based on policies due for renewal in the platform. Switching is the share of incumbent policyholders who switch away from the insurer. Conversion is the share of accepted quotes originating from policyholders who were not previously insured by the insurer. Quoted is the share of renewal queries in which the insurer returned a quote. "Others" aggregates smaller insurance firms.

Figure 1: Mean Premium conditional on number of offers



Notes: The bars represent the number of policyholders that received a given number of offers, whereas the line represents the mean premium.

Figure 1 plots the mean premium within a query against the number of offers received. The figure reveals a negative relationship between average premiums and the size of the observed offer set: queries with more returned offers tend to be associated with lower mean premiums.

5 Reduced-form evidence on switching, pricing, and offer sets

This section documents empirical patterns in switching and renewal pricing that motivate the demand analysis, with particular attention to the offer environment and the incumbent's relative price.

5.1 Switching patterns at renewal

Table 3 shows that approximately 30% of policyholders switch insurers at renewal. In 52% of queries, the previous insurer offers the lowest premium, underscoring that incumbent insurers frequently maintain competitive pricing at renewal.

Table 3: Switching Behavior at Renewal

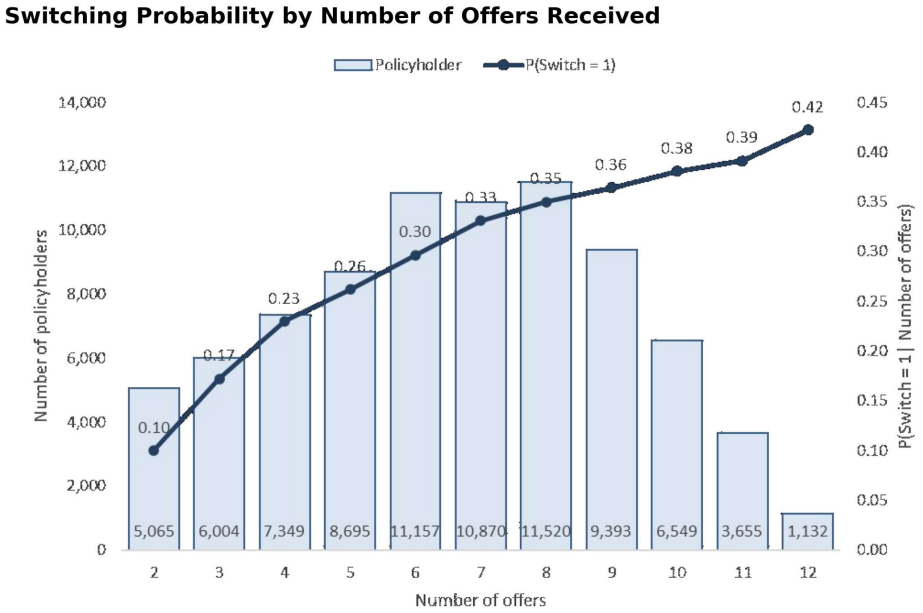
Variable	n (%)
Switched insurer	24,271 (29.8%)
Previous insurer offered lowest premium	42,353 (52.0%)

Notes: Sample includes 81,389 renewal queries from the Level-Up platform in 2021–2022. Percentages are calculated over the total sample.

Figure 2 shows that switching rates increase sharply with the size of the observed offer set: the probability of switching rises from 10% when two offers are received to over 40% when twelve offers are received. This pattern shows that switching is strongly associated with the number of returned offers, underscoring the importance of the observed offer environment for interpreting persistence.

Figure 3 documents two additional patterns. Panel A shows switching probabilities as a function of the premium distance between the incumbent's offer and the mean premium in the query. Switching rates are close to zero when the incumbent's offer is substantially cheaper than alternatives and increase monotonically as the incumbent's relative price worsens. The monotonic relationship is consistent with substantial responsiveness to the incumbent's relative price and underscores the importance of modeling relative prices within the observed offer set rather than focusing on accepted premiums alone. Panel B reports switching probabilities by previous insurer. Switching rates vary substantially across incumbents, ranging from roughly 21% to 40%. These differences in retention across insurers motivate the inclusion of insurer fixed effects in subsequent analysis.

Figure 2: Switching probability and number of offers



Notes: Bars represent the number of policyholders receiving each number of offers; the line shows the corresponding probability of switching.

5.2 Monetary gains from switching and incumbent-relative premium gaps

When policyholders switch insurers, they often realize substantial monetary savings. We define the saving rate for switchers as the percentage difference between the premium offered by the previous insurer (P_i^{PI}) and the accepted premium (P_i^A):

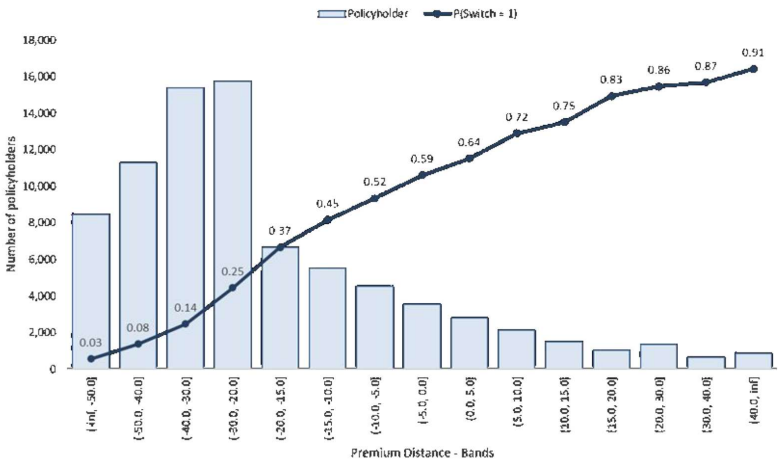
$$\text{Saving}_i = \frac{P_i^{PI} - P_i^A}{P_i^A}. \tag{1}$$

Figure 4 Panel A displays the distribution of saving rates conditional on switching. The distribution is wide and right-skewed, with a mean of 36%, indicating that many switchers achieve substantial savings relative to the incumbent offer.

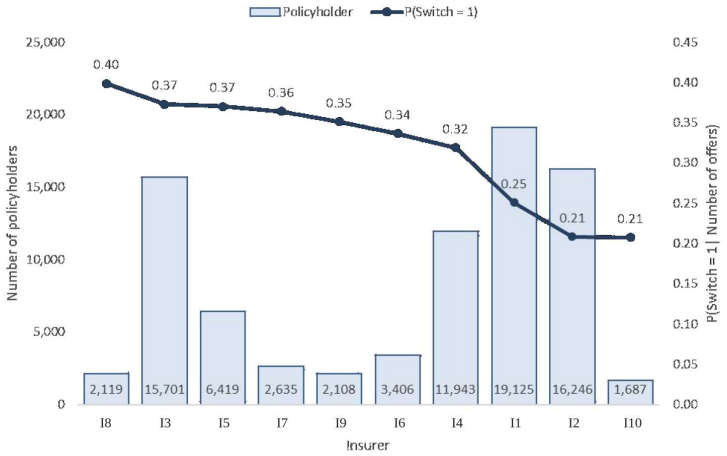
At the same time, a substantial fraction of non-switchers forgo cheaper alternatives. Among policyholders who do not switch, 25% receive at least one offer

Figure 3: Switching probability by price and insurer

Panel A: Switching Probability by Previous Insurer's Relative Price



Panel B: Switching Probability by Previous Insurer



Notes: Panel A shows switching probability as a function of the previous insurer's relative price, measured as the percentage difference between the previous insurer's premium and the mean premium in the observed offer set. Bars represent the number of policyholders in each premium distance band; the line shows the probability of switching. Panel B shows switching probability by previous insurer. Insurers are ordered by switching rate. Bars represent the number of policyholders insured by each firm; the line shows the probability of switching.

with a premium lower than that of the previous insurer. For these consumers, we define the incumbent-relative premium gap as the percentage difference between the lowest available premium and the previous insurer's premium:

$$\text{Gap}_i = \frac{P_i^{\min} - P_i^{PI}}{P_i^{PI}}. \tag{2}$$

Figure 4 Panel B shows the distribution of the incumbent-relative premium gap among non-switchers who had cheaper alternatives available. The mean gap is -13%. Equivalently, among nonswitchers with a cheaper alternative, mean foregone savings are 13% of the incumbent premium. The presence of sizable foregone savings is consistent with the existence of switching frictions beyond simple price comparisons.

5.3 Reduced-form determinants of switching

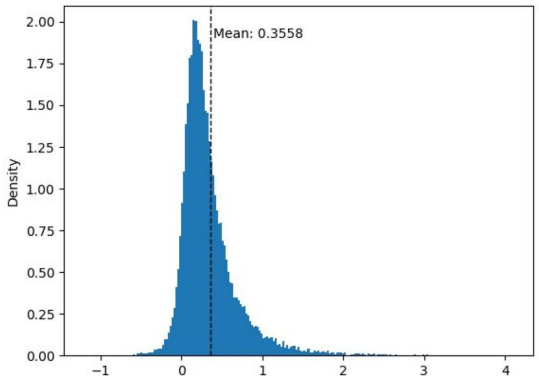
Table 4 reports results from logit regressions that estimate how observable characteristics relate to the probability of switching. Column (1) includes only variables that are typically available in datasets containing only accepted contracts, not alternative offers. In this specification, the accepted-premium coefficient is economically very small, a pattern that by itself could be read as weak price responsiveness.

Columns (2) and (3) progressively incorporate variables measuring incumbent prices and the characteristics of the observed offer set. Once the incumbent premium and within-query relative price measures are included, relative prices become economically important predictors of switching. Whether the incumbent offers the lowest premium is a particularly strong predictor of retention. These results show why accepted prices alone provide an incomplete description of price incentives at renewal.

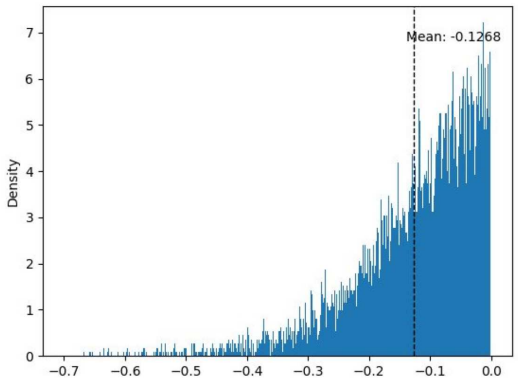
Across all specifications, having filed a claim in the previous year is a strong predictor of switching, even in specifications that include prices, bonus class, and

Figure 4: Monetary gains from switching and incumbent-relative premium gaps

Panel A: Savings Rate Distribution (Switchers Only)



Panel B: Cost Rate Distribution (Non-Switchers with Cheaper Alternatives)



Notes: Panel A shows the distribution of saving rates for policyholders who switched insurers, defined as $(P_i^{PI} - P_i^A)/P_i^A$ where P_i^{PI} is the premium offered by the previous insurer and P_i^A is the accepted premium. Panel B shows the distribution of incumbent-relative premium gaps for policyholders who did not switch despite having cheaper alternatives available, defined as $(P_i^{\min} - P_i^{PI})/P_i^{PI}$ where $P_i^{\min}$ is the lowest premium in the observed offer set.

insurer fixed effects (Column 3). This pattern is consistent with prior evidence that claims experience affects insurer retention (Israel, 2005).

Table 4: Determinants of switching

	(1)	(2)	(3)
Accepted Premium (€100s)	0.064***		
	(0.005)		
% Premium - Min Premium		0.002*	-0.116***
		(0.001)	(0.003)
% Premium PI - Mean Premium			0.037***
			(0.001)
Premium PI is the lowest=1			-3.505***
			(0.047)
Claimant=1	0.299***	0.293***	0.481***
	(0.043)	(0.041)	(0.059)
Current Bonus		-0.035***	-0.068***
		(0.003)	(0.004)
Female	0.016	-0.004	-0.136***
	(0.019)	(0.018)	(0.026)
Month and Region	Yes	Yes	Yes
Insurer and Previous Insurer FE	Yes	Yes	Yes
Marital status	Yes	Yes	Yes
Number of Offers	No	No	Yes
N	75,405	75,405	75,405

Notes: Logit regressions; the dependent variable equals 1 if the policyholder switches insurers at renewal. Logit coefficients are reported. Standard errors in parentheses. Accepted Premium is measured in hundreds of euros. % Premium - Min Premium = 100 × (Accepted Premium - Min Premium in Set) / Accepted Premium. % Premium PI - Mean Premium = 100 × (Premium PI - Mean Premium in Set) / Premium PI. Premium PI is the lowest = 1 if the incumbent's premium is the lowest in the observed offer set, 0 otherwise. Number of Offers = count of insurer quotes in the observed offer set. All columns include the same demographic and insurer, region, and month controls. Column (1) uses the accepted premium; column (2) introduces the accepted premium relative to the lowest available premium; column (3) additionally includes incumbent price and offer-set characteristics. The logit estimation drops 5,984 observations because two previous-insurer indicators perfectly predict switching. ***p <0.01, **p <0.05,*p <0.1.

5.4 Pricing at renewal

We next examine how renewal prices differ between incumbent and competing insurers. We regress log premiums on policyholder characteristics, firm and time fixed effects, and an indicator for whether the firm making the offer is the policyholder's previous insurer.

Table 5 reports the results. Incumbents' quoted premiums are approximately 27–30% lower than those of competitors ($\exp(-0.32) - 1 \approx -0.27$), and the difference remains similar with policyholder fixed effects.

Thus, the price pattern in our data does not exhibit the higher incumbent prices often associated with "bargain-then-ripoff" dynamics.[6] Lower incumbent prices are consistent with strong competitive pressure at renewal, although they may also reflect incumbents' informational advantage about existing customers.

Table 5: Determinants of pricing

	Log of premium offered	
	(1)	(2)
Previous Insurer	-0.317***	-0.349***
	(0.001)	(0.001)
Current Bonus	-0.040***	
	(0.000)	
Age	-0.008***	
	(0.000)	
Age Squared	0.000***	
	(0.000)	
Male	0.035***	
	(0.001)	
Claimant=1	0.165***	
	(0.002)	
Policyholder FE	No	Yes
Month and Region FE	Yes	No
Marital status	Yes	No
Car characteristics	Yes	No
Offer characteristics	Yes	Yes
N	906,160	940,295
R^2	0.542	0.175

Notes: OLS regressions with log premium as dependent variable. Column (1) includes observable policyholder and vehicle characteristics; Column (2) uses policyholder fixed effects. Standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

[6] In canonical models of competition with switching costs, firms face a trade-off between pricing aggressively to attract customers, who may become valuable repeat buyers, and subsequently charging higher prices to customers who are costly to switch. This mechanism can generate low introductory prices followed by higher prices for locked-in consumers (Klemperer, 1995).

6 Demand with Observed and Reconstructed Offer Sets

This section develops a discrete-choice model of insurer choice at renewal. We first estimate demand using each policyholder's observed offer set and quoted premiums. We then hold the utility specification fixed while sequentially removing observed offer-level information: first replacing non-chosen quoted premiums with predicted prices, while preserving each policyholder's observed offer set, and then replacing consumer-specific offer availability with a common set of insurers for policyholders renewing in the same state and month, defined as the insurers chosen by at least one policyholder in that state–month.

6.1 Demand Model

Policyholder i chooses among the insurer offers in the observed offer set S_i. Indirect utility from insurer $j \in S_i$ is

$$u_{ij} = \alpha_i p_{ij} + \beta_i I_{ij,t-1} + W_i' \delta_j + \gamma_j + \varepsilon_{ij}, \tag{3}$$

where p_{ij} is the premium quoted by insurer j and $I_{ij,t-1}$ is an indicator equal to one if insurer j is policyholder i's previous insurer. The vector W_i contains observed policyholder and renewal characteristics, and δ_j captures systematic heterogeneity in preferences for insurer j associated with these characteristics. The term γ_j captures the mean utility of insurer j that is common across policyholders, while ε_{ij} is the idiosyncratic consumer–insurer match utility.

We allow both α_i and β_i to vary across policyholders. Under the sign convention in equation (3), $-\alpha_i$ corresponds to the marginal utility of income. The parameter β_i captures the utility associated with choosing the incumbent insurer, conditional on premiums and observed characteristics. We refer to β_i as the incumbency effect. It may reflect switching frictions, persistent consumer

preferences, prior service experiences, or other forms of state dependence.

Let $\theta_i = (\alpha_i, \beta_i)$ collect the policyholder-specific preference parameters. Assuming that ε_{ij} is independently and identically distributed Type-I extreme value, the probability that policyholder i chooses insurer j, conditional on the observed offer set S_i and on θ_i, is

$$P_{ij}(\theta_i) = \frac{\exp\left(V_{ij}(\theta_i)\right)}{\sum_{k \in S_i} \exp\left(V_{ik}(\theta_i)\right)}. \tag{4}$$

where

$$V_{ij}(\theta_i) = \alpha_i p_{ij} + \beta_i I_{ij,t-1} + W_i' \delta_j + \gamma_j. \tag{5}$$

is the systematic component of utility conditional on θ_i.

To express the incumbency effect in monetary terms, we define its premium equivalent as the change in premium that produces the same change in indirect utility as incumbency. For policyholder i,

$$PE_i = \frac{\beta_i}{-\alpha_i}. \tag{6}$$

The empirical analysis compares the estimated premium effect, incumbency effect, and premium-equivalent incumbency effect across alternative representations of premiums and offer availability.

6.2 Empirical Specification

We estimate the model on the Rio de Janeiro and São Paulo sample described in Section 4, comprising 42,069 renewal queries and 252,648 insurer offers.

In the main specification, we allow the premium coefficient and the incumbency effect to vary systematically with observed policyholder and vehicle characteristics. We parameterize

$$\alpha_i = \alpha_0 + \alpha_A Age_i + \alpha_V Value_i, \tag{7}$$

and

$$\beta_i = \beta_0 + \beta_C Claim_i, \tag{8}$$

where Age_i is policyholder age, $Value_i$ is the insured vehicle's value, and $Claim_i$ indicates whether the policyholder filed a claim in the previous year. Therefore, α_i is allowed to vary with policyholder age and vehicle value, and β_i, with recent claims experience. In addition, we allow bonus class, month, age, age squared, gender, and claimant status to have insurer-specific effects on utility and include insurer-specific constants.

Because both α_i and β_i may vary with observed characteristics, we evaluate both effects at the query-level sample means. The premium-equivalent incumbency effect reported in the main tables is the ratio of the evaluated incumbency effect to the absolute value of the evaluated premium coefficient. We use the same evaluation rule across the different price and offer-set constructions.

As a robustness check, we also allow for unobserved heterogeneity in the incumbency effect. In this specification,

$$\beta_i = \beta_0 + \beta_C Claim_i + \sigma_\beta \nu_i, \qquad \nu_i \sim N(0,1), \tag{9}$$

while the premium coefficient continues to vary with observed policyholder age and vehicle value as in equation (7). This specification allows policyholders with the same observed characteristics to differ in the utility they attach to remaining with the incumbent insurer.

Because β_i now depends on the unobserved draw ν_i, the unconditional choice probability is obtained by integrating the conditional logit probability over the distribution of ν_i:

$$P_{ij} = \int \frac{\exp\left(V_{ij}(\nu_i)\right)}{\sum_{k \in S_i} \exp\left(V_{ik}(\nu_i)\right)} \phi(\nu_i) d\nu_i, \tag{10}$$

where $\phi(\cdot)$ denotes the standard normal density. We use this random-coefficients specification as a robustness check; the main comparisons across price and offer-set constructions use the main observed-heterogeneity specification described above.

6.3 Interpretation and Identification

Conditional on each policyholder's observed offer set, the premium coefficient is identified from how insurer choice varies with relative quoted premiums within renewal queries. The incumbency effect is identified from variation across policyholders in whether a given insurer is the incumbent, conditional on premiums, insurer-specific mean utility, and the other observed determinants of utility. Variation in these relationships with policyholder age and vehicle value, and with recent claims experience, identifies the corresponding interaction terms.

Our demand specification treats quoted premiums as conditionally exogenous after controlling for the observed utility shifters, insurer effects, and the observed offer set. This assumption would be violated if insurers' pricing decisions depend on information that is unobserved to us and correlated with consumers' preferences for particular insurers. Industry interviews indicate, however, that premiums are determined primarily by policyholder and vehicle characteristics used to assess expected risk, including demographics, claims history, bonus class, and vehicle characteristics observed in our data. This institutional evidence supports the conditional-exogeneity assumption, although it does not establish it.

We interpret β_i as an incumbency effect rather than a pure switching cost. It may reflect switching frictions as well as persistent consumer–insurer preferences, prior service experiences, or other forms of state dependence. The insurer-specific mean utilities and observed preference shifters capture systematic differences

in insurer attractiveness, while the random-coefficients specification allows for additional heterogeneity in the incumbency effect. Accordingly, we interpret the estimated premium-equivalent incumbency effect as a composite measure of persistence and focus on how it changes across the different price and offer-set constructions, rather than attributing it to a specific mechanism.

Insurers also decide whether to return a quote based on risk assessment, proprietary information, and business strategy. Our estimation does not require quote availability to be random. Instead, conditional on the observed determinants of utility, selection into the observed offer set should not be systematically related to the idiosyncratic consumer–insurer utility component ε_{ij}. The institutional setting limits one potential source of such correlation: before quotes are returned, the platform does not allow brokers to selectively route a renewal request to preferred insurers or privately communicate policyholder-specific switching intentions or insurer preferences.

Finally, the observed offer set may differ from the set of alternatives ultimately evaluated by the policyholder. Brokers may present or emphasize only some returned quotes, and consumers may attend to only a subset of those presented. Section 6.6 examines the sensitivity of the demand estimates to smaller effective choice sets and to price-dependent broker presentation.

6.4 Price and Offer-Set Reconstruction

We re-estimate the same demand model under progressively coarser representations of premiums and offer availability. In the first exercise, we replace the quoted premiums of non-chosen offers with predicted premiums while preserving each policyholder's observed offer set and chosen premium. In the second, we additionally replace consumer-specific offer availability with a common offer set for policyholders renewing in the same state and month. The sequential construction shows how the demand estimates change first when exact alternative premiums are replaced and then when consumer-specific offer availability is also

replaced.

We begin with the benchmark specification, which uses each policyholder's observed offer set and the insurer-specific premiums quoted in that renewal. Price imputation for alternative insurers has precedent in insurance demand applications. Honka (2014), for example, observes which insurers policyholders considered but imputes premiums for considered insurers that were not selected. Our first exercise uses a related idea: we preserve the observed offer set and the chosen premium but replace the quoted premiums of non-chosen insurers with predicted prices.[7] Because the quoted premiums are observed in our data, we can directly compare the distribution of the predicted premiums with that of the observed quotes.

The price prediction model is estimated out of sample using other renewal queries. It uses policyholder and vehicle characteristics, location and month, the identities of the chosen and alternative insurers, and the chosen premium. Because all offers in a renewal query price the same policyholder and vehicle, the chosen premium provides additional information about the overall price level of the query. We then add a draw from the empirical distribution of prediction errors so that the predicted premiums retain residual dispersion. To avoid in-sample prediction, we divide renewal queries into five groups and predict premiums for each group using a model estimated on the other four groups. The resulting predicted premiums closely reproduce the marginal distribution of the observed alternative premiums: the predicted mean and standard deviation are € 711.1 and € 509.4, compared with € 702.1 and € 498.5 in the observed data. Thus, the exercise preserves substantial price variation while replacing the consumer–insurer-specific premiums of non-chosen offers with predicted values.

In the second exercise, we additionally replace each policyholder's observed

[7] Our prediction procedure is tailored to our data and differs from the imputation procedure in Honka (2014).

offer set with a common offer set for policyholders renewing in the same state and month. For each state–month, this set contains every insurer chosen by at least one policyholder in that state–month. For insurers that belong to both the policyholder's observed offer set and the common set, we retain the price used in the first exercise: the observed premium for the chosen insurer and the predicted premium for non-chosen insurers. Insurers added through the common offer set receive premiums from the same prediction procedure. Average offer-set size rises from 6.01 to 11.54, and 99.97% of the offers actually returned are contained in the corresponding common sets. This construction is not intended to recover a policyholder's latent consideration set. Rather, it removes consumer-specific variation in insurer availability while preserving geographic and temporal variation.

Appendix B provides implementation details for the main price reconstruction and reports an alternative construction using type-averaged premiums with broad market sets; Appendix C considers an additional reconstruction preserving within-market price dispersion.

6.5 Demand estimation results

Table 6 reports the fixed-coefficient estimates across the three specifications. Using the observed offer sets and quoted premiums, the premium coefficient evaluated at sample-mean characteristics is -0.0237 and the corresponding incumbency effect is 2.525. The premium-equivalent incumbency effect is € 106.5 (95% CI: € 103.8–€ 109.1).

Holding the observed offer sets fixed but replacing the premiums on non-chosen offers with predicted values reduces the magnitude of the premium coefficient evaluated at sample-mean characteristics from 0.0237 to 0.0126, a decline of roughly 47%. The incumbency effect changes much less, from 2.525 to 2.307. As a result, the premium-equivalent incumbency effect rises to € 183.3 (95% CI: € 178.4–€ 188.3), about 1.72 times the benchmark.

The similarity between the marginal distributions of predicted and observed alternative premiums is important for interpreting this change. The predicted premiums retain approximately the same overall dispersion as the observed quotes, so the lower estimated price sensitivity is not driven by mechanically compressing the distribution of alternative prices. Because insurer choice in the model responds to relative premiums within each policyholder's offer set, replacing observed alternative quotes with predicted premiums changes the consumer–insurer price rankings and gaps that link relative prices to choices. This provides a natural explanation for the smaller estimated price response.

In the final specification, we retain the predicted-price construction and replace the policyholder-specific observed offer set with the common state–month offer set. The premium coefficient evaluated at sample-mean characteristics changes only slightly further, from -0.0126 to -0.0124, while the incumbency effect rises from 2.307 to 2.868. The premium-equivalent incumbency effect therefore rises to € 231.6 (95% CI: € 226.5–€ 236.7), about 2.17 times the observed-offer benchmark.

This second step operates through a different margin. The common offer set adds alternatives that were chosen by at least one policyholder in the same state and month but did not necessarily return a quote to the focal policyholder. The model must therefore rationalize why those additional alternatives were not chosen. Because the premium coefficient evaluated at sample-mean characteristics changes little between the second and third columns, most of the adjustment appears in the incumbency effect. Taken together, the two steps produce different patterns in the estimates: replacing observed alternative premiums with predicted prices primarily reduces the magnitude of the premium coefficient, whereas replacing consumer-specific offer availability produces the larger change in the incumbency effect.

Table 6: Demand Estimates with Observed and Reconstructed Offer Sets

	Observed offers	**Reconstructed alternative premiums**	**Reconstructed premiums + common offer set**
Premium coefficient at sample means	-0.024***	-0.013***	-0.012***
	(0.0002)	(0.0001)	(0.0001)
Incumbency effect at sample means	2.525***	2.307***	2.868***
	(0.028)	(0.026)	(0.025)
Premium-equivalent incumbency (€)	106.5***	183.3***	231.6***
	(1.35)	(2.52)	(2.60)
As % of mean premium (€ 564)	18.9%	32.5%	41.0%
Average offer-set size	6.01	6.01	11.54
Offer observations	252,648	252,648	485,342
Renewal queries	42,069	42,069	42,069

Notes: The same fixed-coefficient utility specification is estimated in all columns. The sample consists of renewal queries from the Rio de Janeiro and São Paulo metropolitan markets. Column 1 uses each policyholder's observed offer set and quoted premiums. Column 2 preserves the observed offer set and chosen premium but replaces the premiums of non-chosen offers with the out-of-sample predictions described in the text. Column 3 additionally replaces consumer-specific offer availability with the common state–month offer set described in the text. Because the premium coefficient varies with age and vehicle value and the incumbency effect varies with claimant status, the reported premium and incumbency effects are evaluated at query-level sample means. Premium-equivalent incumbency is the incumbency effect evaluated at sample means divided by the absolute value of the premium coefficient evaluated at sample means; standard errors are computed by the delta method. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

6.6 Robustness to Broker Filtering and Limited Attention

The observed offer set may be larger than the set a policyholder ultimately evaluates. Even if the broker presents all returned quotes, limited attention may lead consumers to focus on only a subset of them. To assess whether the benchmark is sensitive to a moderately smaller effective choice set, we randomly remove 10%, 15%, and 20% of alternative returned offers and re-estimate demand. The premium coefficient is essentially unchanged across these samples; the base incumbency coefficient is also stable at the 10% and 15% reductions and declines at 20%. This exercise examines the robustness of the benchmark to smaller effective choice sets, as might arise under limited consumer attention, without modeling the attention process itself.

Random removal does not address the possibility that brokers selectively emphasize particular quotes. We therefore consider a second exercise in which presentation depends on price. We always retain the chosen contract, so the observed choice remains feasible, and the incumbent quote, so the incumbency comparison is preserved. Among the remaining offers, we retain either the cheapest or the most expensive half.

Under the competitively priced rule, the average effective set contains about 3.8 alternatives and the premium-equivalent incumbency effect evaluated at sample-mean characteristics is € 108.1, compared with € 106.5 in the benchmark. Under the alternative rule that retains the most expensive half, the estimate falls to € 28.9. Thus, the price-based filtering rules we examine do not produce a materially larger premium-equivalent incumbency effect than the benchmark.

Appendices B and C report two additional offer-set and price reconstructions. They produce substantially smaller base price coefficients than the observed offer set specification, reinforcing the qualitative result that demand estimates are sensitive to how missing offer information is reconstructed.

7 Conclusion

This paper uses auto insurance renewal data that record the chosen contract together with the full set of insurer quotes returned to the broker. We exploit this information by estimating the same demand model under progressively coarser representations of the offer environment: first replacing the premiums on non-chosen offers with predicted prices while preserving each policyholder's observed offer set, and then replacing consumer-specific offer availability with a common state–month offer set.

The differences are economically meaningful. At sample-mean policyholder characteristics, the premium-equivalent incumbency effect is about € 106 with the observed offer sets and quoted premiums. Replacing alternative premiums with predicted prices raises it to about € 183: the estimated premium coefficient becomes substantially smaller in magnitude, while the incumbency effect changes comparatively little. Replacing the consumer-specific offer set with the common state–month offer set raises the measure further to about € 232, with little additional change in the premium coefficient but a substantial increase in the incumbency effect.

The two steps therefore reveal distinct measurement consequences of missing offer-level information. When exact alternative premiums are not observed, the relationship between within-set relative prices and choices is weakened even when the marginal price distribution is reproduced closely. When consumer-specific availability is not observed, the model must explain choices against alternatives that may not have been available to the policyholder. In our setting, these two information losses increase measured inertia through different components of the demand model.

More broadly, the results highlight the importance of accounting for heterogeneity in the alternatives and prices consumers face. In markets where these objects are not observed, recovering or otherwise accounting for the consumer-

specific choice environment can matter not only for modeling search or consideration, but also for the measurement of price sensitivity and consumer inertia. Our setting provides a direct benchmark for the magnitude of this measurement problem in a market where the underlying offer environment is observed.

References

Abaluck, J. and A. Adams-Prassl (2021). What do consumers consider before they choose? identification from asymmetric demand responses. *Quarterly Journal of Economics 136*(3), 1611–1663.

Barseghyan, L., M. Coughlin, F. Molinari, and J. C. Teitelbaum (2021). Heterogeneous choice sets and preferences. *Econometrica 89*(5), 2015–2048.

Conlon, C. T. and J. H. Mortimer (2013). Demand estimation under incomplete product availability. *American Economic Journal: Microeconomics 5*(4), 1–30.

Crawford, G. S., R. Griffith, and A. Iaria (2021). A survey of preference estimation with unobserved choice set heterogeneity. *Journal of Econometrics 222*(1), 4–43.

Dubé, J.-P., G. J. Hitsch, and P. E. Rossi (2010). State dependence and alternative explanations for consumer inertia. *RAND Journal of Economics 41*(3), 417–445.

Ho, K., J. Hogan, and F. Scott Morton (2017). The impact of consumer inattention on insurer pricing in the medicare part d program. *RAND Journal of Economics 48*(4), 877–905.

Honka, E. (2014). Quantifying search and switching costs in the us auto insurance industry. *RAND Journal of Economics 45*(4), 847–884.

Honka, E. and P. Chintagunta (2017). Simultaneous or sequential? search strategies in the us auto insurance industry. *Marketing Science 36*(1), 21–42.

Honka, E., A. Hortaçsu, and M. A. Vitorino (2017). Advertising, consumer awareness, and choice: Evidence from the us banking industry. *RAND Journal of Economics 48*(3), 611–646.

Honka, E., A. Hortaçsu, and M. Wildenbeest (2019). Empirical search and consideration sets. In *Handbook of the Economics of Marketing*, Volume 1, pp. 193–257. Elsevier.

Israel, M. (2005). Services as experience goods: An empirical examination of consumer learning in automobile insurance. *American Economic Review 95*(5), 1444–1463.

Jõeveer, K. and K. Kepp (2023). What drives drivers? switching, learning, and the impact of claims in car insurance. *Journal of Behavioral and Experimental Economics 103*, 101993.

Kepp, K. and K. Männasoo (2025). Explaining switching behavior: Consumer attention and choice in car insurance market. *Journal of Economic Behavior & Organization 238*, 107233.

Kim, M., D. Kliger, and B. Vale (2003). Estimating switching costs: the case of banking. *Journal of Financial Intermediation 12*(1), 25–56.

Klemperer, P. (1995). Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade. *Review of economic studies 62*(4), 515–539.

Laas, D., H. Schmeiser, and J. Wagner (2016, January). Empirical findings on motor insurance pricing in germany, austria and switzerland. *Geneva Papers on Risk and Insurance : Issues and Practice 41*(3), 398–431.

Moraga-González, J. L., Z. Sándor, and M. R. Wildenbeest (2023). Consumer search and prices in the automobile market. *Review of Economic Studies 90*(3), 1394–1440.

Shum, M. (2004). Does advertising overcome brand loyalty? evidence from the breakfast-cereals market. *Journal of Economics & Management Strategy 13*(2), 241–272.

Sovinsky Goeree, M. (2008). Limited information and advertising in the us personal computer industry. *Econometrica 76*(5), 1017–1074.

SUSEP (2006). Guia de orientação e defesa do segurado. Rio de Janeiro.

SUSEP (2022). Sistema de estatísticas da SUSEP (SES). Market share and premium data for the Brazilian insurance sector.

A Core Demand Coefficients

Table 7 reports the core utility coefficients from the fixed-coefficient specification estimated on the observed offer sets. Alternative-specific coefficients on bonus class, month, age, age squared, gender, and claimant status are included in the estimation but omitted from the table for brevity.

Table 7: Core Demand Coefficients: Observed Offer Sets

	Coefficient	Standard error
Premium	-0.03732***	0.00078
Premium × Age	0.000187***	0.000013
Premium × Vehicle Value	5.48×10^{-8}***	1.35×10^{-9}
Previous Insurer	2.5383***	0.0284
Previous Insurer × Claimant	-0.2841*	0.1475
Premium coefficient at sample means	-0.02371***	0.00021
Incumbency effect at sample means	2.5246***	0.0281
Premium-equivalent incumbency (€)	106.5***	1.35

Notes: Main fixed-coefficient specification estimated on the observed offer sets. Premium is measured in euros. Because premium is interacted with age and vehicle value and incumbency is interacted with claimant status, the marginal premium and incumbency effects are evaluated at query-level sample means. Premium-equivalent incumbency is the ratio of these evaluated effects; its standard error is computed by the delta method. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

B Alternative Offer-Set and Price Reconstructions

This appendix provides additional implementation details for the price and offer-set reconstructions used in Section 6.4. Section B.1 describes the main non-chosen-price reconstruction underlying Table 6. Section B.2 reports the alternative type-averaged reconstruction with broad market-level offer sets.

B.1 Price Reconstruction of non-chosen contracts

This section provides additional details on the price reconstruction used in the main information-deletion exercise in Section 6.4. The objective is to replace the

exact premiums on non-chosen offers while preserving as much of the observed price heterogeneity as possible. The procedure combines out-of-sample prediction with information from the policyholder's chosen premium and an empirical residual draw. All prediction steps are cross-fitted at the renewal-query level, so that no renewal query is used to predict its own alternative premiums.

We first divide renewal queries into five folds, assigning all offers belonging to the same query to the same fold. For each held-out fold, we estimate on the other four folds the log-premium equation

$$\log p_{ij} = x'_{ij}\theta + r_{ij}, \tag{11}$$

where p_{ij} is the premium quoted by insurer j to policyholder i. The vector x_{ij} includes insurer, state, and month fixed effects; policyholder age and age squared, gender, bonus class, and claimant status; vehicle age and vehicle value; and an indicator for whether insurer j is the incumbent. The model is estimated using all returned quotes in the training folds. Let $\widehat{\theta}_{-k(i)}$ denote the coefficients estimated excluding the fold containing query i. For every observed quote, we construct the out-of-fold residual

$$\widehat{r}_{ij} = \log p_{ij} - x'_{ij}\widehat{\theta}_{-k(i)}. \tag{12}$$

One insurer has only a single quote in the estimation sample and therefore does not provide sufficient support for a separate insurer fixed effect. For that observation, we use the same first-stage specification without insurer fixed effects; this exception has no material effect on the aggregate results.

The first-stage prediction captures observable variation in premiums but does not exploit the fact that all insurers in a renewal query are pricing the same policyholder and vehicle. We therefore use the observed chosen premium to anchor the overall price level of the query. Let $c(i)$ denote policyholder i's chosen

insurer and $\widehat{r}_{ic(i)}$ the corresponding out-of-fold residual from equation (12). For each non-chosen insurer j, define the residual difference

$$\Delta r_{ij} = \widehat{r}_{ij} - \widehat{r}_{ic(i)}. \tag{13}$$

We model this difference as

$$\Delta r_{ij} = z'_{ij}\gamma + u_{ij}, \tag{14}$$

where z_{ij} includes the chosen-offer residual $\widehat{r}_{ic(i)}$, fixed effects for the chosen and alternative insurers and their interaction, state and month fixed effects, and the same policyholder and vehicle characteristics used in equation (11). This equation is also estimated using five-fold cross-fitting, so that $\widehat{\gamma}_{-k(i)}$ is estimated without observations from query i's fold.

For a non-chosen offer, the reconstructed premium is

$$\widetilde{p}_{ij} = \exp\left\{x'_{ij}\widehat{\theta}_{-k(i)} + \widehat{r}_{ic(i)} + z'_{ij}\widehat{\gamma}_{-k(i)} + \widetilde{u}_{ij}\right\}. \tag{15}$$

The first three terms recover the conditional mean of the alternative premium while anchoring its overall level to information contained in the actual chosen premium. The final term, $\widetilde{u}_{ij}$, restores residual price dispersion. It is drawn from the empirical distribution of out-of-fold residuals u_{ij} for the same alternative-insurer $\times$ chosen-insurer pair. To maintain adequate support, when fewer than 50 donor observations are available for a given pair, we draw from residuals for the same alternative insurer across chosen insurers; if fewer than 50 such observations are available, we draw from the pooled out-of-fold residual distribution. In all cases, observations from the held-out fold are excluded from the donor pool.

The premium of the chosen contract is never reconstructed and remains equal to its observed value. Thus, the exercise removes the exact consumer–insurer

price mapping for the non-chosen alternatives while retaining the realized chosen premium and substantial cross-sectional and within-query price dispersion. Among non-chosen offers, the reconstructed premiums have a mean of € 711.1 and a standard deviation of € 509.4, compared with € 702.1 and € 498.5 for the observed premiums. The reconstruction therefore does not mechanically compress the marginal distribution of alternative prices. In the broader-menu exercise, the same prediction procedure is used to assign premiums to insurer–policyholder pairs added to the reconstructed state–month offer set.

B.2 Type-Averaged Premiums with Broad Market Sets

This alternative reconstruction assumes consumers face all firms active in their market (i.e., firms chosen by at least one policyholder in that state-month) and imputes alternative prices from observationally similar consumers' accepted contracts. The construction procedure is as follows:

Step 1: Define consumer types

We stratify consumers into types based on observable characteristics used in risk assessment and pricing. Specifically, we create categorical variables for:

- **Bonus class:** Four categories (1–3, 4–6, 7–9, 10)
- **Age:** Five categories ($\leq$25, 26–35, 36–50, 51–65, >65)
- **Gender:** Male or Female
- **State:** Brazilian state (UF)
- **Vehicle age:** Four categories ($\leq$3, 4–6, 7–10, >10 years)

Step 2: Calculate type-specific average premiums

For each combination of consumer type (t), month (m), and firm (f), we calculate the average premium charged by that firm to consumers of that type in that

month:

$$\bar{p}_{fmt} = \frac{1}{N_{fmt}} \sum_{i \in \{f,m,t\}} p_i$$

where f indexes firms, m indexes months, t indexes consumer types, and N_{fmt} is the number of contracts sold by firm f to type t consumers in month m. This average is calculated using only observed chosen contracts.

Step 3: Construct reconstructed offer sets

For each consumer i who chose a contract from firm j in the observed data:

- Retain their actual chosen contract with the premium they paid: (j, p_{ij})
- Add all other firms $k \neq j$ that were active in consumer i's market (state $\times$ month)
- For each alternative firm k, assign the type-averaged premium $\bar{p}_{kmt_i}$, where t_i is consumer i's type and m is the month
- Drop firm-type-month combinations where no contracts were observed (i.e., where $\bar{p}_{kmt_i}$ is undefined).

The resulting reconstructed dataset gives each consumer all firms active in the market at type-averaged premiums. The number of alternatives is therefore substantially larger than in the observed data: approximately 11–12 rather than 6–7 on average. The reconstruction therefore replaces both consumer-specific offer availability and consumer-specific alternative premiums with broad market sets and type-level premiums.

In the fixed-coefficient model, the base premium coefficient falls from -0.037 with observed offers to -0.003 under this reconstruction, while the base previous-insurer coefficient changes from 2.538 to 1.496. In the random-incumbency specification, the corresponding base coefficients change from -0.043 and 3.121 to -0.003 and 1.208. Because premium and incumbency are interacted with policy-

holder characteristics, ratios of these reported base coefficients do not correspond to the representative premium-equivalent incumbency effect reported in Table 6.

C Alternative Reconstruction with Stochastic Offer Sets

This additional reconstruction maintains within-firm price dispersion and introduces stochastic variation in the composition of reconstructed offer sets. This relaxes the assumption that all firms are available to every consumer while preserving more of the observed price distribution.

The construction procedure is as follows:

Step 1: Retain chosen contract

For each consumer, we keep their observed contract (chosen firm and actual premium paid).

Step 2: Randomly select alternative firms

We identify all other firms active in the consumer's market (defined as region $\times$ month) and randomly select a subset of these firms. The number of alternatives is set to match the maximum observed offer-set size in that market. The chosen firm is explicitly excluded from the random draw to avoid duplication.

Step 3: Assign premiums from observed price ranges

For each selected alternative firm k, we assign a premium drawn uniformly from the range of premiums observed for that firm in the market-month. Specifically, we calculate the minimum and maximum premiums offered by firm k in market m during month t:

$$p_{kmt}^{min} = \min_{i \in \{k,m,t\}} p_i, \quad p_{kmt}^{max} = \max_{i \in \{k,m,t\}} p_i$$

and draw uniformly from the range $[p^{min}_{kmt}, p^{max}_{kmt}]$.

Step 4: Construct reconstructed offer set

The reconstructed offer set contains the consumer's actual choice plus the randomly selected alternatives with sampled premiums. The total number of alternatives matches the maximum observed offer-set size in that market.

This reconstruction differs from the type-average specification in two ways. First, it allows for stochastic rather than deterministic offer availability because not all firms appear in every reconstructed offer set. Second, it maintains within-firm price dispersion rather than using type-averaged premiums, capturing the fact that observationally similar consumers may receive different quotes from the same firm.

Table 8 reports the coefficient estimates from this reconstruction alongside the observed-offer estimates. The base premium coefficient is much smaller in absolute value in the reconstructed data (-0.004 versus -0.037 in the fixed-coefficient specification), and the previous-insurer coefficient also changes. These estimates provide a second illustration that reconstructed offer sets and premiums can materially alter demand coefficients. Because premium and incumbency are interacted with policyholder characteristics, ratios of the reported base coefficients do not correspond to the representative premium-equivalent incumbency effect reported in Table 6.

Table 8: Demand Coefficients under a Stochastic Offer-Set Reconstruction

	Observed Offer Data		**Reconstructed Data**	
	Fixed Coefficients	Random Coefficients	Fixed Coefficients	Random Coefficients
Price Sensitivity				
Premium, base coefficient	-0.037***	-0.043***	-0.004***	-0.004***
	(0.001)	(0.001)	(0.000)	(0.000)
Choice Inertia				
Previous Insurer, base coefficient	2.538***	3.121***	1.167***	1.259***
	(0.028)	(0.055)	(0.028)	(0.020)
Previous Insurer (SD)	—	1.830***	—	0.002
		(0.075)		(0.283)
Observations	252,648	252,648	496,815	496,815

Notes: Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. The fixed columns use the main fixed-coefficient specification. In the random-coefficients columns, only the coefficient on the previous-insurer indicator is random; premium enters with the same observed interactions as in the fixed-coefficient specification. The table reports base coefficients. Because premium and incumbency are interacted with policyholder characteristics, ratios of the reported base coefficients do not correspond to the representative premium-equivalent incumbency effect reported in Table 6. The signed scale estimate for the reconstructed-data random coefficient is -0.002 (s.e. 0.283); because the sign of a normal scale parameter is not identified, the table reports its absolute value.