Behavioral Professionals: Evidence From the Commercial Auto Insurance Industry

Alon Y. Rubinstein*

University of California, Berkeley

February 15, 2023

(Click for latest version)

Abstract

A cornerstone of the IO study of *selection markets* is that competition disciplines sellers to customize coverage and premiums optimally. But is this the case? Using data from one of the largest Israeli commercial auto insurance providers, an affiliate of a multinational insurance company, I find there is too little adjustment in the intensive margin. Premiums barely change with expected costs as projected by pre-determined factors (vehicle age) and signals (claim history). At the same time, I find there is too much adjustment in the extensive margin, with an excessive denial of insurance in response to recent claims. Using unique grading documents, I integrate the insurer's subjective risk assessment into the study of insurance markets. I find that the insurer's risk assessment outweighs recent claims and misevaluates vehicle age. Structural model estimates suggest that insurers enjoy incumbency advantages over their own customers, and clients are rationally inattentive to competitors' pricing unless they are faced with a price increase. Both channels allow sub-optimal behavior to persist. Finally, I find that supply-side behavioral frictions, which result in excessive denial, mainly harm *disadvantaged customers*—single-fleet clients of old vehicles—and diminish with the client's fleet size.

*I would like to thank Stefano DellaVigna and Benjamin Handel for their invaluable mentorship and advice. I would like Ned Augenblick, Nano Barahona, Kei Kawai, Jonathan Kolstad, Sebastian Otero, Ricardo Perez-Truglia, Frank Schilbach, Carolyn Stein, Steve Tadelis, Dmitry Taubinsky, Matthew Rabin, and participants in the U.C. Berkeley Industrial Organization and Psychology and Economics seminars, and Russell Sage Foundation Summer Institute in Behavioral Economics, for their thoughtful feedback, comments and suggestions. All remaining errors are my own. E-mail: arubinstein@berkeley.edu.

1 Introduction

Perhaps the key feature of *selection markets*, such as the private market for insurance, is that consumers vary not only in their willingness to pay but also in how costly they are to the seller. Therefore, insurance providers care about both the *quantity* of policies they sell and the *quality* of the clients they cover. Market forces can fail to achieve efficiency if buyers know better than sellers how risky they truly are (Akerlof (1970), Rothschild and Stiglitz (1976)).

Thus, it is not surprising that much attention in the IO literature is devoted to studying the demand side in insurance markets and quantifying the implications of selective sorting on policy and welfare (e.g., Cutler and Reber (1998), Chiappori and Salanie (2000), Cardon and Hendel (2001), Cohen and Einav (2007), Fang et al. (2008), Cutler et al. (2008), Carlin and Town (2009), Lustig (2010), Einav et al. (2010a), Einav et al. (2013), Starc (2014), Finkelstein and Poterba (2014), Hackmann et al. (2015), Handel et al. (2015), Cabral (2016), Cabral et al. (2018) and Einav et al. (2019)).[1] A cornerstone of these studies is that sellers customize coverage and premiums optimally. The view is that while market forces might fail to discipline buyers, competition disciplines similar, equally informed insurers; providers offer optimal coverage and premiums. In reality, however, insurers are not identical and might not be equally informed. For example, customers know their insurer better than other providers, and insurers know their clients better than their competitors. Due to these relational asymmetries, market forces might fail to eliminate mispriced risk, suggesting that supply-side frictions may affect premiums and coverage.

In this paper, I offer an alternative perspective on the insurance market. As in recent studies, I allow for behavioral frictions in demand for insurance. However, in contrast to the literature on selection markets, I recognize that insurers might fail to assess risk accurately and to customize coverage plans and prices optimally.

With this perspective in mind, this paper studies the pricing and coverage behavior in private insurance markets. Specifically, I address four main questions. First, do large sellers in private insurance markets customize offers and prices as the IO literature predicts?

[1] See handbook chapters by Einav et al. (2021) and Handel and Ho (2021) describing studies on IO of selection markets, in general, and insurance markets, in particular.

If not, does it reflect their biased beliefs? Third, why do market forces fail to compete away sellers with biased beliefs? Fourth, what are the welfare implications of supply-side behavioral frictions, and do these vary if clients are covered by individual or fleet base contracts?

To address these questions, I study the Israeli commercial auto insurance market. This market provides an excellent laboratory to study supply-side frictions for two reasons. First, we expect professional buyers to choose carefully between insurance plans and discipline sellers to customize offers accordingly. Second, the market is limited to unregulated property coverage; insurers can charge any price and deny coverage without constraints. I use comprehensive data from one of the largest commercial auto insurance providers in Israel, an affiliate of a large multinational insurance company. The data includes *all* the information available to the insurer: (i) premiums, coverage, and claim expenses by policy and client; and (ii) internal policy pre-renewal assessments, known as the *"Go—No Go" grades*. Furthermore, I obtain data on the market competitors' premiums by generating fictitious policy applications. These datasets allow me to portray the gap between premiums and expected costs by pre-determined factors and claim history, identify the gap between objective and subjective risk assessment, and quantify its impact on coverage, pricing, profits, and welfare.

I start by providing evidence of the gap between premiums charged by the insurer and the cost of providing coverage as a function of pre-determined and stochastic factors. I find there is too little adjustment in the intensive margin. The insurer barely adjusts premiums per value with determinants predicting higher expected cost per value, such as vehicle age and claim history. Consequently, the insurer profits by providing coverage to new vehicles and clients with favorable past performance.

Interestingly, I find that the insurer's adjustment of premiums per value regarding claim history is based solely on recent claims while putting no emphasis on augmented past performance, despite aggregate claim history serving as a predictive signal of future claims. In contrast, recent performance has no additional predictive power.

Next, I study whether these pricing patterns are specific to this particular insurer or apply to other market competitors. Specifically, I generate fictitious policy applications using an Israeli insurance agency and examine how the premiums vary by vehicle characteristics and claim history. I find that the market-wide price patterns are comparable to those of the insurer. Moreover, the analysis of market-wide premiums for new policies indicates a substantial adjustment on the extensive margin. It is impossible to generate a premium offer for a new policy if a customer has been involved in at least two claim events in the three preceding years.

After providing evidence indicating that both the insurer and its competitors do not adjust premiums based on a customer's observable characteristics, I turn to the internal grading documents. Despite their richness, the observed prices are insufficient to identify the insurer's beliefs, as both supply and demand factors determine equilibrium premiums. I exploit the variation in the "Go—No Go" grades and policies' observable characteristics to identify the impact of pre-determined and stochastic factors in determining the insurer's subjective risk assessment. Internal grading recommends no change in premiums—"Go"—for most policies, ignoring the predictive power of vehicle age and claim history on costs. The lack of recommended price adjustment spills over to coverage. Internal grading data recommends denying comprehensive coverage to almost half of the "No Go" graded policies rather than increasing their premiums. This is especially relevant for old vehicles and costly clients. This strategy also reflects a biased risk assessment as signaled by recent and augmented claim history. The "Go—No Go" grades are too sensitive to recent claims with almost no predictive power of future costs, conditional on the augmented history of claims.

Internal grading and premiums reflect demand and supply factors. To distinguish between these forces, I develop and estimate a structural model that allows customized prices and coverage to reflect the insurer's subjective risk assessment and commonly used demand and supply factors. In its simplified version, the model consists of two periods. In the first period, customers self-sort to sellers. In the second period, a wedge emerges between their insurer and other providers. Customers decide whether to renew their policies or search for an outside offer. The decision to renew depends not only on their private information and search costs but also on the supply side, that is, their insurer's private information, subjective risk assessment, and customized offers.

I take advantage of the panel structure and internal grading to identify and quantify a client's willingness to pay and the insurer's subjective risk assessment. Regarding the willingness to pay, a key concern is that premiums are subject to strategic considerations. I use the panel structure of my data, which follows many clients with large fleets over multiple coverage periods, to identify an external source of variation in premiums and estimate clients' willingness to pay. The across-client variation permits conditioning out the client-specific effect on premiums. The within-client variation allows for identifying exogenous shocks in price adjustments by using predicted —rather than actual—adjustments in premiums for those who renew their coverage (Bundorf et al. (2012)) and those who do not (Crawford et al. (2018)).

As for the insurer's subjective risk assessment, this is identified by decomposing expected profits into premiums and expected costs using a two-step procedure. First, I take advantage of the informational symmetry between the insurer and the econometrician to nonparametrically identify the expected profits for each policy by inversion of the share of recommendations (Berry (1994)). Then, using policies for which no change in premiums is recommended ("Go"), I identify the insurer's beliefs of expected costs by subtracting the previous year's premiums from expected profits. To further account for possible latent strategic considerations, I focus on nonfleet clients.

Three main demand side findings emerge. First, customers adversely select to renew coverage. Second, new customers are adversely selected; they cost more, conditional on observables. Last, customers are rationally inattentive to premiums unless they incur a price increase. Both the adverse selection of new customers and the rational inattention of renewing consumers point to asymmetries between incumbent insurer-insuree pairs and others that allow the insurers room for error in customized prices and coverage.

In terms of the supply side, the insurer's subjective risk assessment, two main findings emerge. First, the insurer gives more weight to recent claims without predictive power of future costs. The law of large numbers implies that demand exhibits increasing returns to fleet size, as large fleets are less likely to be affected by the insurer's biased risk assessment. Second, the insurer erroneously evaluates common predetermined factors such as vehicle age.

Using the estimated demand and supply parameters, I analyze the impact of the insurer's biased risk assessment on coverage, premiums, profits, and welfare using a set of counterfactuals. I find that supply-side frictions mainly harm *disadvantaged customers*—single-fleet clients of old vehicles. A profit-maximizing firm does not deny coverage as informational asymmetries between the customer and the insurer are modest. In contrast, the insurer denies coverage to old vehicles and clients with poor recent performance, which results in lower profits. Furthermore, the clients face a substantial reclassification risk, diminishing with customers' fleet size; volatility in recent performance drops with the number of insured vehicles. As a result, customers benefit from purchasing coverage as a group.

Finally, do insurers adjust premiums once they learn those might be mispriced? During my study, I shared my preliminary stylized findings that point to possible mispricing by vehicle age with the managerial team. To assess the impact of information on pricing, I compare premiums by vehicle age over the covered period. I find almost no change in premiums and profits in consecutive years before the managerial team learned about my findings. In contrast, I find a moderate increase in premiums between periods once they were informed of my findings. The adjustment of prices upon learning is consistent with my findings that much of the mispricing reflects the insurer's biased assessment of risk rather than strategic considerations.

This paper contributes to the literature on firm behavior. A growing body of research documents that large suppliers in nonselection markets customize prices too little based on observable demand factors (Orbach and Einav (2007), McMillan (2007), Cho and Rust (2010), Shiller and Waldfogel (2011), Cavallo et al. (2014), DellaVigna and Gentzkow (2019)). My paper shows that sellers in selection markets fail to customize prices also on expected cost, a cornerstone of the study of selection markets.

This paper also adds to the literature on reclassification risk and market unraveling (Cutler and Reber (1998), Hendel and Lizzeri (2003), Koch (2014), Finkelstein et al. (2005), Handel et al. (2015), Hendren (2017), Fleitas et al. (2020), Ghili et al. (2021), Cuesta and Sepúlveda (2021)). IO theory attributes both aspects to asymmetric information and regulation. I find that insurers do not adjust premiums and excessively deny customers, despite modest asymmetric information and no regulations. For individual customers, the insurer amplifies welfare loss from the reclassification of risk. This relates to group insurance (Bundorf et al. (2012), Tilipman (2022)), as client's size dilutes insurers' mise-valuation of risk and the consequences of overdenial.

The paper also relates to the literature on imperfect competition in selection markets (Veiga and Weyl (2016), Mahoney and Weyl (2017), Lester et al. (2019), Cuesta and Sepúlveda (2021), Tebaldi (2022)). I show two channels generating market power: information asymmetries among insurers (Jin and Vasserman (2020)) and behavioral demand-side frictions (Sydnor (2010),Abaluck and Gruber (2011),Barseghyan et al. (2013), Handel (2013), Handel and Kolstad (2015), Spinnewijn (2017), Bhargava et al. (2017), Brot-Goldberg et al. (2017),Ho et al. (2017), Handel et al. (2019), Gottlieb and Smetters (2021)).[2] Imperfect competition allows imperfect behavior by insurers, which might drastically change the welfare consequences of imperfect competition.

Last, my paper also contributes to the literature on managerial practices, highlighting the impact of monitoring, feedback, and on-the-job training (Bloom and Van Reenen (2007) and Bloom et al. (2013)). My findings indicate that these elements improve profits even among professional sellers in a big-data industry.

The remainder of the paper is organized as follows. In section 2, I describe the setting. Section 3 provides descriptive statistics regarding the data I exploit in the empirical analysis. In Section 4, I provide evidence of the gap between insurer pricing and expected cost. Section 5 presents evidence of the gap between objective and subjective expected costs using the "Go—No Go" grades. Section 6 then develops and estimates the demand for policy renewal, and section 7 develops and estimates the insurer's subjective costs and supply of insurance. In section 8, I conduct a counterfactual analysis to study the implications of supply side frictions. Finally, Section 9 concludes.

2 Setting

In this paper, I take advantage of proprietary for the years 2013 to 2020 obtained from a large Israeli company (with an annual average revenue during the sample period of $37.5 million in 2020 terms) operating in the commercial auto-insurance market to examine the

[2]Due to informational asymmetries across insurance providers, a perfectly competitive outcome is implausible even when considering a frictionless economy with homogeneous products.

relationship between insurer pricing, perceived costs, and customer's realized costs. [3] The provided dataset includes all data the insurer has from 2013 to 2020. In the empirical application, I take advantage of the information symmetry between the insurer and the econometrician in terms of the determinants of costs and pricing. In this section, I characterize in detail the insurer's affiliation with an international insurance company, its portfolio (in terms of both customers and products), data, and business operations.

2.1 Vertical Relationship

The insurance company operates under a unique vertical relationship, compared with the standard market structure in the insurance literature in general and in the auto insurance literature in particular. The Israeli insurance company is an affiliate of a large international insurance company (henceforth, IIC) with asset value of over $50 billion as of December 2020. The IIC provides capacity, which allows the Israeli insurer to sell policies. This is a result of regulation in Israel, which sets reserve requirements per premiums charged to avoid the failure of insurers to repay claims.[4] In terms of the division of cost and revenue, the IIC pays all claim damages (net of deductibles), while the Israeli insurer pays all additional operational costs. Revenue is split between the IIC and the Israeli insurer based on yearly agreed-upon shares.

A possible concern is that the distorted incentives might lead the Israeli insurer to oversupply insurance, as claims are paid by the IIC. Thus, the IIC provides guidance on pricing by setting a lower bound on premiums charged, conditional on vehicle and client observable characteristics. As a result of repeated interactions with the IIC, the Israeli insurer puts emphasis on the portfolio's return, as negative outcomes often lead to a lower share of revenue in succeeding years. Throughout the paper, I consider the joint profits of both the Israeli insurer and the IIC from operations in this market as a whole.

[3] Throughout the paper, I use and report monetary values in nominal New Israeli Shekels (ILS) to avoid creating artificial variation in the data. Annual inflation between 2013 and 2020 ranged from 0.84% to -0.63%, and the value of 1 ILS ranged from $0.26 to $0.29.

[4] Throughout the sample years, the capacity constraint was not binding. Therefore, I consider the opportunity cost of providing insurance to a different customer to be zero.

2.2 Insurer Portfolio

The insurer provides three types of coverage: (i) third-party Coverage, which only covers the cost of damage to third-party property; (ii) comprehensive coverage, which includes all damages to a vehicle in addition to damages covered by third-party coverage; and (ii) partial coverage, which covers the same types of damage as comprehensive coverage, excluding theft. None of these types of policies cover bodily injuries to the policyholder (or to third parties). Israeli regulations mandate that all vehicle owners purchase special coverage for bodily injury through a separate, heavily regulated policy. The vast majority of the insurer's portfolio—over 87%—consists of comprehensive coverage policies (see Appendix Figure A.1, Panel A).

The insurers provide coverage to commercial vehicles, including trucks, buses, minibuses, trailers, and heavy equipment (e.g., tractors, bulldozers, cranes). Approximately two-thirds of the insurer's portfolio consists of trucks (see Appendix Figure A.1, Panel B). Therefore, I focus mainly on comprehensive coverage policies for trucks throughout the empirical application.

The relationship between the insurer and the clients differs in four ways from the common relationship in previous studies on insurance markets. First, clients typically own a fleet of vehicles. The insurer's clients are quite diverse in terms of their fleet size. More than 10% of policies are of a single-vehicle client, a quarter of policies are of clients who insure a fleet of fewer than five vehicles, and more than a quarter of the policies are of clients who insure at least 100 vehicles (see Appendix Figure A.2).

Furthermore, unlike some markets in which firms offer "take-it-or-leave-it" prices, equilibrium premiums are the result of bargaining between the insurer and the client (especially when they are fleet owners). Equilibrium price-setting has both favorable and unfavorable consequences. On the one hand, premiums are endogenous and might be correlated with unobservable (to the insurer and the econometrician) components of the demand for insurance. On the other hand, exogenous variation in pricing as a result of a firm experiment or pilot does not allow examination of its behavior relative to profit-maximizing behavior, as these prices are, by definition, nonoptimal off-equilibrium premiums.

As in other markets, premiums are prorated. Yet, unlike standard insurance markets, 8 comprehensive and partial coverage policies are priced in terms of *premium per value*. For instance, a premium per value of 4% implies that the customer pays 4,000 ILS for a 6-month policy for a vehicle valued at 200,000 ILS.[5]

Finally, unlike the common setting in the selection market literature, there is no regulation of pricing and coverage provided by the insurer. The lack of regulation is consequential; the insurer can provide a customer any coverage at any premium, and deny coverage if it wishes to do so. Since premiums are not regulated, IO theory suggests that customer denial can only be explained by excessive adverse selection (Akerlof (1970)).[6]

2.3 Business Operations: "Go-No Go"

The Israeli insurer, operating since the 1950s, employs hundreds of workers. These include employees in the analytical team, overseen by the Chief Operating Officer (henceforth, COO), and underwriters, who are in contact with the customers, either directly or through their agents. Over the sample period, the insurer sold approximately 175,000 policies to over 13,000 different customers. Due to the high volume of customers, and the differentiated occupational requirements and skills of employees, the firm operates in an orderly, systematic structure.

On a monthly basis, the COO provides the underwriters a document, which is named "Go—No Go". The document includes a grade for each policy that is about to end (usually a month or two before the end of the policy coverage contract). The grading system is defined as follows: a "Go" grade implies that the analytical team recommends renewing the customer's policy at the same premium per value and terms (deductibles). A "No-Go" grade implies that the analytical team recommends nonrenewal of the customer's policy at the same premium per value. Typically, a "No-Go" grade will include a recommendation on how to continue the relationship with the customer, if at all. There are four common recommendations: (i) renew the policy without increasing premiums (i.e., increase deductibles), (ii) increase the premium per value of the policy, (iii) do not provide

[5]Vehicle values are usually determined by the Levi Itzhak vehicle price list, which is the standard practice by both commercial and noncommercial auto insurance markets.

[6]An alternative explanation could be high operational costs.

comprehensive coverage (i.e., third-party only), and (iv) do not provide any coverage (deny). An example is provided in Appendix Figure A.3.

At first glance, the complexity of the firm's operations might seem disadvantageous. Yet, this complexity provides additional information which otherwise could have been obtained only by a survey of the analytical team employees. I take advantage of the insurer grading in the empirical application to extract insurer beliefs. Observed prices are insufficient to identify the insurer's beliefs, as they are determined in equilibrium by both supply and demand forces; an insurer might increase prices either because of high expected costs or due to high demand for insurance. Additional assumptions are required to differentiate between the two. In this paper, I identify the insurer's beliefs by exploiting the variation in insurer grading without any structural assumption regarding insurer behavior.

In this section, I describe the data and provide descriptive evidence of insurer pricing and realized costs (I refer to the "Go—No Go" grades in section 4). As mentioned earlier, the main dataset includes all the insurer's data from 2013 to 2020. The data include (i) contract characteristics (premium, coverage type, deductibles, duration, and an indicator on whether a driver under the age of 24 is allowed to operate vehicles), (ii) vehicle characteristics (vehicle value, vehicle age, vehicle type, vehicle weight, and vehicle model), (iii) customer characteristics (claim history, zip code, and fleet size), (iv) costs (commission and claim damages), and (v) identifying information (policy id number, vehicle license number, and client id number).[7] Using the identifying information, it is possible to track the clients and their vehicles over time.

In the empirical analysis, I focus on comprehensive insurance policies. Column 1 in Table 1 presents the summary statistics of comprehensive coverage policies for trucks (Column 1 in Appendix Table A.1 presents the summary statistics for all vehicles).[8] The sample consists of 51,684 policies. Claims are reported for roughly a quarter of the poli-

[7]In this market, insurance policies are not tied to a specific driver but rather to a specific vehicle. In general, each insured vehicle can be operated by any driver over the age of 24 with a valid license to operate a vehicle of that class. A client can extend the policy coverage for young drivers (between the age of 21 to 24), which in general increases the premium charged.

[8]In the empirical framework, I mainly focus on trucks, as the premiums charged by the market competitors are available for trucks only. I also repeat the entire empirical analysis for all vehicle as well. The results, which are consistent with those analyzing insurance for trucks, are reported in the appendix.

cies. The insurer enjoys a mean profit of 1,587 ILS and a profit margin of 16%, as the mean premium and costs are 9,938 (or 3.33% of the vehicle value) and 8,361 ILS, respectively. From the insurer's perspective, the portfolio's performance (a profit margin of 16%) is satisfactory. The insurer's positive performance is of key importance, as the systematic mispricing of risk cannot be concluded by observing a failing firm. The profitable performance supports the notion that this particular insurer is not *competed away* from the market. This is complemented by the fact that the insurer is one of the largest insurers in the Israeli commercial auto-insurance market and is affiliated with one of the largest multinational insurance companies.

In this paper, I examine whether the insurer assesses risk correctly in both the intensive and the extensive margins. I examine the intensive margin by estimating the gap between the actual pricing and the expected cost as a function of a customers' observable factors. In the case of a recurring customer, the insurer should adjust premiums or possibly even deny insuring the customer based on changes in the observable characteristics. The evolution of observable factors can be divided into two groups: predetermined changes (i.e., vehicle age) and stochastic shocks (i.e., claim history).

In Figure 1, I provide a first glance at the relationship between the premium, costs, and vehicle age. Figure 1, presents the mean premium per value and the cost per value by vehicle age (0 to 10) for comprehensive insurance policies for trucks. As is apparent from the figure, the premiums do not adjust optimally over vehicle age, as the cost-per-value curve is, on average, a counter-clockwise rotation of the premium-per-value curve. Therefore, the insurer generates its profits by providing insurance for new vehicles and incurs losses on old vehicles.[9]

Column 2 in Table 1 presents the summary statistics for a subset of the comprehensive coverage policies for trucks: those with a vehicle age of six years or above.[10] This subset represents about one-fifth of this sample.[11] In general, the share of policies involved in a claim is not higher than that of the entire sample (Column 1). Yet, the mean damage (i.e.,

[9]The nominal relationship between mean premium, cost, and vehicle age is of a similar nature and is reported in Appendix Figure A.4.

[10]Column 2 in Appendix Table A.1 presents the summary statistics for the equivalent subset of all vehicles.

[11]Appendix Figure A.5 depicts the distribution of policies by vehicle age.

cost of claims) per value is substantially higher, by more than 60%. The mean damage is 14% lower, yet the vehicle value depreciates by 47%. This suggests that, conditional on a claim, the expected damage is not proportional to vehicle value. Unlike the mean damage per value, the mean premium per value increases by only 20%. As a result, providing comprehensive coverage policies to old vehicles (six years and above) on average does not benefit the insurer but instead generates losses. Consequently, the insurer's satisfactory profit margin is derived by mostly providing coverage to relatively new trucks.

Regarding past performance, Column 3 in Table 1 presents the summary statistics for a subset of the comprehensive coverage policies for trucks: those that reported a claim in the previous period.[12] This subsample consists of 16% of the whole sample and does not differ substantially from that of the entire sample in terms of premium per value (3.56% relative to 3.33%), vehicle age (5.00 relative to 4.18), or vehicle value (275,954 relative to 298,659).[13] Yet, policies with past realized claims incur higher costs. Among policies with a reported claim in the previous period, 34.82% report a claim also in the current period (relative to 23.98%), and the mean damage is 9,345 ILS, which is substantially higher than the entire sample (6,794 ILS). Consequently, on average, the insurer exhibits losses for providing coverage in policies with a reported claim in the previous period.

Thus far, I have divided the sample based on whether a policy incurred a claim in the previous period. In the insurance market in general, and with this insurer (and its competitors) in particular, it is customary to measure the client's performance based on the aggregate loss ratio. The aggregate loss ratio is defined as the ratio of damages (i.e., net cost of claims) to revenue (i.e., premiums) with respect to all of the customer's past policies. In Panel A of Figure 2, I divide the sample into four groups based on the level of loss ratio (at the start of the policy) and depict the relationship between the premium per value, cost per value, and the client's aggregate loss ratio. The mean cost per value in-

[12]It should be noted that "reported claim" does not necessarily imply that the customer reported the claim, as third parties usually report claims on customers that generate third-party damages. Furthermore, throughout the analysis, I do not consider the "at-fault" side. A reported claim is defined as an event in which the insurer exhibits costs as a consequence of providing coverage to the client. In addition, Column 3 in Appendix Table A.1 represents the same segmentation of all vehicles.

[13]Since the sample begins in 2013, I am unable to observe reported claims in the previous period. I take this into account in the empirical framework and omit the 2013 policies, or policies of new clients, when conducting comparative statics regarding past performance. About one-fifth of the policies with at least one year of documented history reported a claim in the previous period.

creases with the loss ratio, which suggests a positive and persistent relationship between past and future performance, consistent with the relationship between past and future claims. Yet, the mean premium per value does not adjust accordingly. It is flat in both relative and absolute terms. As with claims at the policy level, on average, the insurer exhibits losses when providing coverage to customers with poor past performance, while it enjoys profits by providing insurance to customers that performed well in the past, as they also tend to perform well in the future. As with vehicle age and claims, the insurer's portfolio is profitable as the overwhelming majority of the insurer customers are beneficial —with a loss ratio of under 1, greater than 82% (Panel B of Figure 2). This composition is not exogenously determined; the insurer denies customers with poor performance at a higher rate (see section 4).

To summarize, the statistics presented in this section raise two opposing findings. On the one hand, the insurer is profitable. The profit margin of comprehensive coverage policies for all vehicles is 19.65%, and 15.97% in the case of trucks. As mentioned above, these margins are satisfactory from the insurer's perspective. On the other hand, the data presented suggest that the insurer can do better. Both predetermined (age) and stochastic factors (past performance) are correlated with a higher cost per value in the future, yet, the relationship between those factors and premium per value is quite flat. On average, the insurer's profits are generated by a specific segment of customers and vehicles. New vehicles and customers with adequate past performance are beneficial, while old vehicles and customers with poor past performance are costly to insure.

3 Insurer Pricing and Costs

In this section, I estimate the gap between actual pricing and the expected cost of policies as a function of predetermined and stochastic factors. This section is ordered as follows. First, I analyze the relationship between premium per value, cost per value, and vehicle age. Then, I examine the predictive power of past performance and its relationship with current costs, premiums, and profits. I differentiate between recent and overall past performance by considering the aggregate loss ratio and the previous year's loss ratio. I study whether the recent claim history is more predictive of future costs and the relationship of both with pricing. In addition, I examine the competitors' pricing schemes to assess whether the documented insurer behavior is unique or similar to market-wide patterns. To further establish that the insurer misprices risk, I conduct a few robustness tests to rule out alternative channels. Finally, I provide evidence of firm learning by examining insurer pricing after providing information on the flat pricing scheme over the vehicle life cycle.

3.1 Predetermined Changes

In this part, I examine whether the insurer adjusts premiums according to the evolution of costs by vehicle age. Specifically, I quantify the relationship between premium per value, cost per value, and vehicle age. I do so by estimating the following fixed-effect models.

$$\frac{\text{Premium}_{\ell t}}{\text{Value}_{\ell t}} = \sum_{a=1}^{A} \beta_a^p \mathbb{1}\{\text{Vehicle Age}_{\ell t} = a\} + \eta_\ell^p + \varepsilon_{\ell t}^p$$
$$\frac{\text{Cost}_{\ell t}}{\text{Value}_{\ell t}} = \sum_{a=1}^{A} \beta_a^c \mathbb{1}\{\text{Vehicle Age}_{\ell t} = a\} + \eta_\ell^c + \varepsilon_{\ell t}^c,$$

where ℓ and t index the license (vehicle) and period, respectively. $\mathbb{1}\{\text{Vehicle Age}_{\ell t} = a\}$ is an indicator of a specific vehicle age, a, and (η_ℓ^p,η_ℓ^c) are vehicle fixed effects (with respect to pricing and realized costs). I estimate the premium and cost-per-value trends over the vehicle life cycle using a saturated model with fixed effects at the license level. That is, I identify and quantify the trends using within-license variation in the premium and cost per value.

The results are depicted in Figure 3. The patterns are consistent with the summary statistics provided in the previous section. The premium per value does not change substantially over the vehicle's life cycle; it increases by less than 0.5 p.p over the first seven years of the vehicle's age, and by less than 1 p.p over the first ten years. In contrast, costs increase considerably over the vehicle's life cycle. The cost per value increases by more than 3 p.p over the first seven years of the vehicle's age, and by more than 5 p.p over the first ten years. These patterns are inconsistent with *perfect* insurer behavior; the optimal pricing strategy suggests that premiums should adjust according to changes in the expected cost of providing insurance. Since the expected cost per value increases with vehicle age, so should the premium per value; however, the observed premiums per value are quite flat.[14]

The result described in Figure 3 suggests that there is limited variation in pricing within vehicles, but does not imply limited variation in premium per value *across* vehicles. Appendix Figure A.7 shows that this is not the case, however. There is substantial variation in premium per value, as expected when (i) equilibrium premiums are determined in a bargaining process between the insurer and the client and (ii) there is substantial heterogeneity in bargaining power, possibly due to the considerable variation in clients' fleet size.

A possible explanation for the lack of price variation over the vehicle's life cycle is related to client characteristics. As noted earlier, a substantial portion of clients purchase insurance coverage for multiple vehicles. It could be the case that the observed flatness in premiums is artificial. When providing insurance coverage to a large fleet, premiums per value are not expected to change if the client's vehicle age distribution does not vary over time, as fleet owners purchase new vehicles to replace the old ones. Consequently, the lack of variation over vehicle age does not necessarily reflect a lack of adjustment in pricing, as fleet price adjustments might not be necessary. Cross-subsidization within a fleet is an alternative mechanism that can explain the documented trends. An optimally behaving insurer might not change premiums. The cross-subsidization results in clients artificially overpaying to insure new vehicles and underpaying to insure old ones.

I test whether the observed lack of price adjustment is solely driven by fleet cross-subsidization. I do so by examining nonfleet customers. I re-estimate the models considering only nonfleet customers. The results are reported in Figure 4. The reported premium per value and cost per value are similar in spirit to those in Figure 3. Although the standard errors are larger relative to the entire sample (as expected when considering a smaller sample), the patterns are quite similar. The premium per value barely changes

[14] A variant of the model in terms of nominal ILS (instead of per value) is conducted as well. Results are presented in Appendix Figure A.6, Panel A. Furthermore, a replication of the model with regard to the entire sample (all vehicle types) is reported in Appendix Figure A.6, Panel B. The estimated patterns in both variants are consistent with this figure.

over the vehicle's life cycle, increasing by less than 1 p.p over the first ten years. In contrast, the cost per value increases substantially over the vehicle's life cycle by more than 6 p.p over the first ten years. Therefore, although fleet cross-subsidization might be a complementing factor, it is certainly not the sole determinant generating the patterns in the data.

3.2 Cost Shocks

After documenting price misadjustments with respect to a deterministic factor, I examine whether the insurer adjusts premiums optimally when faced with a stochastic cost. Specifically, I quantify the relationship between different current period outcomes and previous period performance.[15] I do so by estimating the following model.

$$Y_{jt} = \beta^c \mathbb{1}\{\text{Claim}_{jt-1} \geq 1\} + X_{jt}\delta + \varepsilon_{jt},$$

where j and t index policy and period, respectively.[16] $\mathbb{1}\{\text{Claim}_{jt-1} \geq 1\}$ is an indicator for whether policy j was involved in at least one claim event in the previous period. Y_{jt} indicates the current period's four outcomes in question. Specifically, (i) $\mathbb{1}\{\text{Claim}_{jt} \geq 1\}$, an indicator for whether at least one claim was reported in the current period, (ii) current period damage (net claim cost) per value, (iii) change in premium per value (that is, the ratio of current to previous premium per value, minus one), and (iv) current period policy loss ratio (damage over premium). I control for vehicle characteristics (value, age, weight, type, and a young driver indicator).[17]

The results are reported in Table 2, Panel A.[18] Column 1 reports the relationship be-

[15] Therefore, I exclude new policies in the following empirical analysis.

[16] I conduct the analysis at the policy level and not the license level (as before). Not doing so would result in a selected sample. Intuitively, an insurance policy that covered a vehicle that incurred a total loss claim in the previous period would be renewed (if at all) in the current period with respect to coverage of a different vehicle.

[17] I do not control for vehicle value when considering the outcome variables damage per value or change in premium per value.

[18] A replication of the model with regard to the entire sample (all vehicle types) in reported in Appendix Table A.2, Panel A. The results are similar.

tween the previous and current claim outcomes. The probability of at least one claim in the current period is 12 p.p higher if a claim was reported in the previous period.[19] The results demonstrate that claim history serves as a persistent signal of current performance (t-stat = 15.22), even when considering a relatively naive measure of past performance.[20] Column 2 reports the estimated model with regards to damage per value. Consistent with the findings in Column 1, the damage per value ratio is on average 1.7 p.p higher, compared with a policy that was not involved in a claim event in the previous period.

Column 3 reports the estimated model of the change in premium per value. If the insurer adjusts premiums correctly, the standard model predicts that under optimal pricing, the premiums should increase with claim history, as past performance serves as a predictive signal of current claims. However, the results suggest that this is not the case. Premiums do not significantly differ, and the coefficient is of the wrong sign; the premium per value drops by -0.1 percent when a claim is reported in the previous period. As a result, the policy is less profitable. As expected, given the results on damages and premiums, the loss ratio associated with a reported claim in the previous period is 41.8 p.p point higher (column 4).

Similar to the examination of predetermined changes, fleet cross-subsidization may be an alternative mechanism giving rise to artificial noncorrelation between the policy's current premium and past performance; clients overpay for policies with good performance and underpay for policies with poor performance. Furthermore, an additional channel that might explain the lack of correlation is that the insurer adjusts prices based on the customer's overall performance—with regard to all policies—and does not assess each policy separately. Regarding vehicle age, I test whether large fleets might give rise to the observed noncorrelation by re-estimating the model when considering solely nonfleet customers. The results are reported in Table 2, Panel B. The relationship between current costs (claim indicator and damage per value) and past claims is similar to those reported for the entire sample. In contrast, the results reported in Column 3 indicate that premiums are adjusted based on past performance. When considering nonfleet policies, the

[19]This is an interpretation based on the linear probability modeling assumption. The relationship is robust with regard to other specifications, such as a logistic and probit model (not reported).

[20]In the next part, I consider the customer's performance (aggregate loss ratio), which takes into account both the cost of damages of all of the customer's policies.

correlation between past claims and current premiums is significant and positive. Policies that incurred a claim in the previous period face a 1.7 percent increase in premiums, compared with policies that were not involved in a claim event. Despite the price increase following poor performance, the relationship between current period loss ratio and past performance suggests that the price adjustment is inadequate. The loss ratio associated with a reported claim in the previous period is 54.7 p.p higher. These results illustrate that the insurer is aware of the persistence of claim history, yet does not adjust premiums sufficiently. Thus, fleet cross-subsidization cannot explain the observed patterns.

3.3 Recent vs. Older Claim History

In the previous subsection, I study how the insurer adjusts prices with regards to previous period claims. In this part, I examine how the insurer adjusts prices when considering both new information and past signals. In particular, I consider the following two signals: (i) the client's aggregate loss ratio over time and (ii) the client's previous year loss ratio.[21] Optimal insurer behavior implies that premiums should adjust with respect to each of these signals based on their relative predictive power: the signal-to-noise ratio. I quantify the relationship between the two signals and different current period outcomes using the following regression model:

$$Y_{jt} = \alpha \text{Aggregate LR}_{jt} + \beta \text{Prev. Yr. LR}_{jt} + X_{jt}\delta + \varepsilon_{jt},$$

where j and t index policy and period, respectively. Aggregate LR_{jt} and Prev. Yr. LR_{jt} are the two signals: the clients' aggregate loss ratio over time and the clients' previous year loss ratio, respectively. It is important to note the previous year's performance is reflected in both variables. Yet, Aggregate LR_{jt} weighs previous data equally, without considering the recency of previous period outcomes. Y_{jt} indicates the four outcomes of the current period in question.

The results are reported in Table 3. Columns 1 and 2 describe the relationship between the client's aggregate loss ratio over time and the client's previous year loss ratio with current period damages: the probability of at least one claim in the current period and

[21] Similar to previously, I exclude policies of customers with less than one year of observed history.

current period damage per value, both at the policy level. The customer's aggregate loss ratio serves as a predictive signal of future claims. The aggregate loss at the start of the policy is positively correlated (statistically significant) with (i) the indicator of whether the policy incurs a claim in the current period (column 1: t-st=5.95) and (ii) the policy's damage per value (column 2: t-st=4.23). In contrast, the customer's previous year loss ratio is not correlated with either cost variable. The coefficient is either in the *wrong* sign, in the case of the indicator of at least one claim in the current period (column 1:-0.0004), or substantially smaller in order of magnitude relative to the aggregate loss ratio, as is the case when considering the damage per value (column 2: 0.0003 relative to 0.0055); in both cases, the coefficients are not statistically significant. The lack of additional information does not imply that recent performance is not informative, but rather that it is not more informative than older claim history. Consequently, a *perfect* insurer should only consider the aggregate loss ratio when adjusting premiums.

In Column 3, I report the estimated model, which quantifies the relationship between the two variables measuring the previous loss ratio and the change in premium per value. The result indicates a deviation from optimal pricing. Unlike an optimally price-setting perfect insurer, the insurer does not increase premiums when facing a high aggregate loss-ratio client. Furthermore, the insurer reacts to negative results, but considers the wrong signal. Premiums increase when considering a customer with poor performance in the previous year, controlling for aggregate performance over time, despite (i) recent performance not incorporating any additional information relative to aggregate claim history, and (ii) the insurer not adjusting premiums based on the more informative signal—the aggregate loss ratio. Column 4 reports the estimated model with regard to the policy's loss ratio. Consistent with misadjustments, policies of clients with higher aggregate loss ratios are associated with adverse results. Yet, conditional on aggregate loss ratios, policies of clients with higher aggregate loss ratios in previous year are not associated with these results. The insurer overreacts to recent noisy shocks (previous year loss ratio) and underweighs the predictive power of the augmented claim history (aggregate loss ratio).

3.4 Market Behavior

Thus far, I have provided evidence that the insurer misprices risk. Premiums insufficiently adjust to predetermined changes and stochastic shocks. A possible concern is the external validity of these results. The findings are based on the pricing data of one insurer (although it is one of the largest insurers in the market and affiliated with a large multinational insurance company). Observing that one insurance company systematically misprices risk does not imply that the *market* is *imperfect*. If the competitors price risk correctly, we expect that in the long run, an imperfect insurer would be competed out of the market.[22]

Unlike many cases where it is difficult to observe the pricing of all firms in the industry, I am able to extract prices for a large number of trucks at different ages and in different conditions (e.g., claims) for all major competitors. I address this issue by examining the market competitors' pricing schemes. I do so with data from the Israeli insurance agency *Orlan Insurance Agency, Ltd. (henceforth, Orlan) (1994)*. As part of its business operations, Orlan has ties to the largest insurance companies in Israel (including the insurer from which I obtained the data). To provide competitive insurance premiums to its clients, Orlan's agents can compare premiums (and coverage terms) for new policies across all insurers (in contact with Orlan), as a function of their characteristics. Orlan's agents access the data using the Orlanet Calculator (henceforth, calculator), which provides information regarding offered pricing and terms from each insurance provider.[23]

Using the calculator, I generate fictitious offers for comprehensive insurance policies for 2,041 distinct trucks model-value-year triads insured between January and March 2020.[24] I use standard insurance coverage and vehicle characteristics as additional inputs necessary to generate an offer.[25] I generate two observations for each distinct vehicle

[22]This statement is true if (i) customers treat insurance coverage as an homogeneous good, (ii) customer search does not incur any costs and (iii) incumbent insurers do not possess an informational advantage over their competitors.

[23]Orlan state that agents should not use the provided dataset in order to price renewing policies, but rather use the calculator solely for new policies.

[24]The data generating process was conducted in the beginning of March 2020, before Israel began enforcing social distancing and other rules to limit the spread of COVID-19.

[25]I.e., vehicles without heavy equipment, default driver characteristics (any driver over the age of 24, excluding individuals with a criminal record or a revoked license), and no additional coverage (e.g., extensive

model-year-value triad: (i) no claims in the last 3 years and (ii) one claim in the last 3 years, which occurred last year. I focus on the four largest insurers in this market: the insurer that provided me the data (denoted as "the insurer"), and its three main competitors (denoted as "rival 1", "rival 2", and "rival 3").

I examine the external validity of my findings in two steps. First, I assess the validity of the calculator. I do so by conducting a within-insurer comparison between the offered premium for coverage through the insurer (using the calculator) and the actual premium charged, to verify that the calculator offers' premiums match the data provided by the insurer.[26] After verification of the calculator's validity, I conduct an across-insurer comparison of the premium offers and examine pricing trends of vehicle age and claim history across the market.[27]

I use the calculator's generated offers to conduct a comparison between the market's insurance providers. I examine the market premium trends of both predetermined changes and stochastic shocks. Figure 5 graphs the premium per value trend throughout the age distribution for the four insurers.[28] The results demonstrate that not only is the insurer imperfect, but rather the *market* is *imperfect*. The price trend of the insurer and rivals 1 and 2 is remarkably similar; no trend in premium per value almost throughout the entire age distribution.[29] In contrast, rival 3 raises the premium per value across the age distribution, which suggests that rival 3 adjusts premiums similarly to an optimally

legal defense, riots, earthquakes).

[26]The offered premiums do not need to be identical to those provided for a few reasons. Mainly, the calculator suits nonfleet truck owners as it does not take into account customer's fleet size.

[27]Panel A in Appendix Figure A.9 depicts the within comparison. The correlation is very high (R^2=0.90). Policies with a higher offered premium (using the calculator) are on average charged a higher price in practice (using the insurer's dataset). Yet, the coefficient is not 1 (0.8). Panel B in Appendix Figure A.9 graphs the difference between the premiums generated by the calculator relative to the data. The graph indicates that the calculator premiums trend over the vehicle life cycle is steeper than the one shown in my data. This implies that the results in this part might overemphasize the steepness trend in market premiums over vehicle life cycles. As I show, this is not a concern as the across-insurer comparison suggests that the market premiums per value are quite flat with regard to vehicle age.

[28]The sample consists of 876 observations, as some of the rival insurers do not offer coverage to vehicles of specific types and weights.

[29]A possible concern is that the similarities in pricing suggest that the insurer and the other two rivals coordinate premiums. I examine this issue by considering the variation in pricing of the three insurers across different observations. The results are provided in Appendix Figure A.10. The figure demonstrates that, although the trends over vehicle age are similar, the premiums differ substantially within the truck model-value-year triads.

pricing firm. Yet, rival 3 is charging higher premiums relative to other competitors. The minimal premium per value charged by any insurance provider (not restricted to the four insurers) is quite flat throughout the age distribution.

In terms of claim history, Table 4 presents the relationship between the premium per value offered by each of the four insurers and the minimum premium per value in the market as a function of claim history. The results suggest that, conditional on offering coverage, the insurer is more sensitive to the previous claim history than rivals 1 and 2, as they offer the *same* premium per value, regardless of whether the customer reported a claim in the previous year. Such behavior diverges from optimal pricing insurers; the claim history serves as a precise signal for future performance (see Tables 2 and 3). Thus, the competition also appears to diverge from optimal pricing behavior.

Furthermore, it is impossible to generate a policy offer from rival 3—which is the only insurer to substantially increase premiums per value over the vehicle life cycle—for a customer with a single claim last year (and none in the two years before that). The lack of provided coverage is market wide when considering a customer with at least two claims in the last three years.[30] This pattern suggests that the insurer's adjustment to customer's risk is mostly on the *extensive margin* (i.e., whether to provide coverage at all) and less so on the intensive margin (i.e., increasing premiums); a pattern which the insurance literature has attributed to adverse selection. I document a similar pattern by the insurer in section 4.

3.5 Robustness

The findings in this section indicate that the insurer (and its market competitors) misprice risk for a significant segment of its customers. Specifically, the insurer does not adequately adjust premiums when considering changes in predetermined characteristics (vehicle age) or stochastic factors (claim history). A potential alternative explanation to the observed patterns in the data is fleet cross-subsidization. As I show, the documented patterns hold when considering nonfleet customers, suggesting that cross-subsidization is not the only mechanism generating the documented patterns. Furthermore, the analy-

[30]Regulation in Israel limit insurers acquirement of information by allowing them to require new customers to provide information with regard to claim history from the last three years.

sis of market premiums complements this notion, as the offered premiums are generated for nonfleet customers. In this part, I examine two additional alternative explanations: (i) weak predictive power and (ii) dynamic pricing strategies.

3.5.1 Out-of-Sample Prediction

The results so far suggest that the insurer (and market) misprices risk and underestimates the value of both vehicle age and claim history. In this part, I quantify the predictive power of vehicle characteristics and past performance, as weak predictive power suggests that disregarding the information might be optimal.

The analysis is conducted in two steps. First, I use data on comprehensive insurance policies from 2014 to 2018 to estimate a cost function. The cost function is constructed using a regression analysis without any customer or license fixed effects. Although it is possible to generate a more precise cost function, I use a simple regression analysis method to demonstrate that even simple methods can generate beneficial cost estimates for out-of-sample observations. The regression model includes observable vehicle characteristics (age, value, weight, class, and type) and claim history (aggregate loss ratio).

Using the estimated cost function, I divide the 2019–2020 sample into 25 groups based on projected damage per value. Appendix Figure A.11 reports the relationship between the mean predicted damage per value for each of the 25 groups, the premium per value charged, and the actual damage per value. As the figure illustrates, the relationship between the predicted and actual damage per value is almost one-to-one, which illustrates that vehicle and customer observable covariates serve as predictive signals of future claims. The figure further demonstrates the lack of adjustment in premium per value. Consistent with earlier findings, the correlation between premium per value and damage per value is less than one.

3.5.2 Dynamic Complementary

Thus far, I have considered a static framework in analyzing the insurer's mispricing of risk. Yet, the relationship between the insurer and its customers is not static. It could be the case the insurer is aware of the incurred losses yet continues to provide coverage since it believes the customer is profitable in the long run. Excluding coverage or increasing premiums from a costly customer today might lead the insurer to not enjoy future profits when the customer becomes profitable.

In this part, I examine whether dynamic considerations can explain the flat pricing patterns (relative to expected cost) using a relatively simple method. First, using data from 2014 to 2015, I divide the insurer's clients into two groups: the first consists of customers with both an average loss ratio of at least 2 and an average vehicle age of at least 5, with 166 customers in all; and the second consists of all other customers who purchased at least one policy from 2014 to 2015 and are not part of the first group, with 3,170 customers. After classifying the costly clients, I examine how the profits are affected when these customers are dropped. Specifically, I analyze the insurer's ex-post performance if it did not provide any coverage to those customers from 2016 to 2020.

In Appendix Table A.3 I provide summary statistics on the policies of both types of customers from 2016 to 2020. The classified "drop" group consists of 1.5% of the relevant policies. These policies exhibit an average loss ratio of 1. The insurer incurs losses from providing them coverage. Their average profit margin is approximately –14%. Hence, dynamic complementary cannot be the sole channel generating these patterns.

3.6 Learning

Would the insurer change its policy if it was informed of these patterns? Is the insurer aware of these patterns? During my interaction with the insurer, I provided findings similar to those presented in Figures 1 and 3. In this part, I examine how the insurer reacts when informed that premiums do not sufficiently adjust over the vehicle life cycle.

An analysis of the firm pricing scheme before and after I presented the findings to the insurer serves two purposes. First, this analysis provides an additional robustness test, as it further examines whether the documented pricing scheme reflects optimal behavior *from the insurer's perspective*. If not fully adjusting premiums over the vehicle life cycle is the optimal pricing scheme, then the information is already incorporated into the firm's decision-making. Therefore, information on flat premiums per value should not cause the insurer to react. However, if premiums adjust, then it must be the case that the insurer was not fully aware of the trends, and, thus, did not price optimally. Second, this part is of interest as pre- and post-price trend analysis allows examination of managerial practice in general, and firm learning, on-the-job training, monitoring, and feedback, in particular. I examine how the insurer reacted to the provided information using the following event study model:

$$\frac{\text{Premium}_{\ell t}}{\text{Value}_{\ell t}} = \sum_{a=1}^{A} \Big(\sum_{t=-2}^{-1} \beta_a^t \mathbb{1}\{\text{Vehicle Age}_{\ell t} \in G_a\} \Big) + \alpha_a \mathbb{1}\{t > 0\} \mathbb{1}\{\text{Vehicle Age}_{\ell t} \in G_a\} + \eta_\ell + \varepsilon_{\ell t},$$

where ℓ and t index the license and period, respectively. $\mathbb{1}\{\text{Vehicle Age}_{\ell t} \in G_a\}$ is an indicator of a specific vehicle age group, G_a (sample is divided to five groups: 0-1,2-4,5-7,8-10, and >10). β_a^t represents the offered premium per value, by age group, before information was given to the insurer, while α_a indicates the premium per value after information was given to the insurer. As before, the analysis includes a license fixed effect.

The results are depicted in Figure 6. The patterns for the two years before the event indicate that there is no pretrend. The premium per value does not change substantially over the vehicle's life cycle; it increases by less than 0.2 p.p over the first ten years of the vehicle's life, and by less than 1 p.p over the entire vehicle life cycle. Interestingly, the premium per value after I provided the information differs from beforehand. The insurer increases premium per value by approximately 0.5 p.p over the first ten years of the vehicle's age, and by more than 1 p.p over the entire vehicle life cycle. Yet, the increase in price trend over vehicle age is inadequate. The premium per value trend over the vehicle life cycle is still flatter than that of the cost per value; The insurer should have increased the premium per value for age groups 8–10, and for those $>$10 by an additional 1 p.p.

The results are mixed. On the one hand, the significant change in premium per value indicates that the insurer mispriced risk beforehand and that the flat pricing scheme was inadequate. On the other hand, the relatively small adjustment indicates that even when the insurer is aware of the patterns, it does not fully adjust premiums. As I further demonstrate in section 4, this does not indicate that the insurer misprices risk, but rather that it prefers not to implement its beliefs on the intensive margin. Instead, the insurer prefers to exclude costly policies.

4 Insurer Grading

In section 3, I show that the insurer (and market competitors) does not adjust premiums optimally when faced with predetermined changes and stochastic shocks. In spite of their richness, the observed prices are insufficient to identify the insurer's beliefs for two reasons. First, the observed premiums are determined in equilibrium by both supply and demand factors. The insurer might not raise the premium charged either because it believes that the expected cost of providing insurance did not change, or because an increase in premiums substantially increases the probability the customer will not renew the policy. In addition, the analysis of premiums and costs is based on a selected sample; the subsample of customers who choose to renew the policy and the policies the insurer selects to renew. Separation might not be exogenous. The sample potentially consists only of those customers who are perceived as profitable by the insurer.

I use the insurer's internal grading documents, the "Go—No Go" grading (see section 2) to extrapolate the insurer's beliefs, as "Go" grades are assigned to profitable policies, regardless of demand factors. Furthermore, the "Go—No Go" documents grade all customers, regardless of their decision to renew.

In this section, I describe the "Go—No Go" grading data, provide descriptive evidence on the relationship between insurer grading and observable predetermined and stochastic factors, and estimate how these factors affect insurer's grading.

4.1 Summary Statistics

The "Go—No Go" grading documents are provided on a monthly basis by the COO to the insurer's underwriters, who are in contact with the customers. The decoded documents consist of 14,288 grades of policies, providing comprehensive or partial insurance to all vehicle types. The dataset consists of a subsample of the insurer's customers. Summary statistics are provided in Appendix Table A.4. In general, the customers for which I observe grades are more than half of the insurer's portfolio. These customers are different from those with no grade. They are substantially smaller in terms of fleet size (average size of the fleet is 6.09), they are charged a lower premium per value (average of 3.03% relative to 2.32%), they are more profitable (average profit margin of 24.42% relative to 13.01%), and they have a lower loss ratio (average loss ratio of 59.98% relative to 77.10%)

Table 5 presents summary statistics on both graded policies and the distribution of insurer grading. In general, the sample consists of profitable policies (the average loss ratio is below 50%). Among the sample, 87% of the policies are given a "Go" grade (column 2). These policies are quite similar in observable characteristics to policies that face a term adjustment or rejection (premium per value, vehicle age, and vehicle value). The substantial difference is the previous year performance; 12% of the policies reported a claim during the policy duration, and the average loss ratio is 29%, while policies that face an increase in deductibles, premiums or were denied incurred losses in the previous year (i.e., loss ratio above 100% in the previous year). These statistics indicate that the overwhelming majority of policies are profitable and do not require adjustments in pricing. The rest of the portfolio requires further consideration.

Among the policies that received a "No-Go" grade, 35% required only an increase in deductibles (column 3), and 21% of the policies required an increase in premiums (column 4). For the remaining 44%, the analytical team recommends not providing comprehensive coverage; either to provide only third-party coverage (14% - Column 5) or to provide no coverage at all (30% - column 6). These statistics clearly demonstrate that a substantial portion of the insurer's adjustment is on the extensive margin, by not providing comprehensive coverage instead of increasing prices.

The statistics also demonstrate that the insurer's grades are based on shocks and predetermined changes. In general, the insurer appreciates the importance of performance and vehicle age. The average loss ratio of policies that the analytical team recommends increasing deductibles or premiums is above 100% and the average loss ratio of policies that the analytical team recommends denying is over 300%. As for vehicle age, the analytical team recommends offering only third-party coverage to old vehicles, regardless of their performance. The average loss ratio among the policies that the analytical team recommends only third-party coverage is 40%—profitable policies—yet the average vehicle age is close to 10 years. Overall, the insurer's beliefs are not entirely misguided: it is aware of the importance of taking into account vehicle age and loss ratio when renewing a policy.

4.2 Regression Analysis

In this part, I examine how the insurer grades customers based on their observable characteristics. Specifically, I examine how a "Go" grade is assigned using the following probit model.

$$GO_{ij} = \sum_{a=1}^{A} \beta_a \mathbb{1}\{\text{Vehicle Age}_{ij} \in a\} + \alpha_{lr}\text{Aggregate LR}_{ij} + \alpha_{recent}\text{Prev. Yr. LR}_{ij} + \delta X_{ij} + \eta_i + \varepsilon_{ij},$$

where i and j index the customer and the policy, respectively. GO_{ij} is an indicator of whether the policy is assigned a "Go" grade. The variables Vehicle Age$_{ij}$, Aggregate LR$_{ij}$, and Prev. Yr. LR$_{ij}$ denote vehicle age, the client's aggregate loss ratio over time, and the client's previous year loss ratio, respectively. Furthermore, I consider η_i to denote an unobserved random error at the client level. I divide the sample into four groups based on vehicle age: (i) up to 1 year old, (ii) age 2 to 4 years old, (iii) age 5 to 7 years old, and (iv) age 8 years and above. Controls include observable characteristics, client fleet size, vehicle weight, vehicle type, and underage driver indicator.

The results are presented in Table 6. In Column 1, I examine how vehicle age determines a "Go" grade assignment. The insurer discontinuously reacts to vehicle age. Vehicle age affects grading mainly when the vehicle age at least 8 years. This is in contrast to the findings in section 4, which suggest that damage per value increases continuously throughout the vehicle life cycle. Therefore, the results indicate that the insurer erroneously reacts to predetermined factors using simplistic rule-of-thumb rules. Regarding claim history, the insurer considers past performance in grading. The results in Column 2 demonstrate that a client's aggregate loss ratio over-time is significantly correlated with the "Go" assignment. Interestingly, the client's previous year loss ratio has an additional effect on the "Go" assignment, despite it not having any predictive power of future claims, conditional on augmented claim history. This is consistent with the relationship with premiums as well (as documented in section 4). These patterns do not change when estimating the effect of all variables jointly (column 3). To further examine whether the recent year loss ratio has a substantial and persistent role in grading, I re-estimate the model while considering a subsample of policies: those with at least five years of data observed by the insurer on the client. The results are presented in Column 4.

Even though considerable information on the customer is available, the insurer does not underweigh the importance of the previous year loss ratio. This indicates that the recency bias persists and does not diminish with the insurer's data history on the client.

To conclude, the results in this section suggest that the insurer does not change terms for the overwhelming majority of policies, while extensively denying comprehensive coverage from a substantial portion of policies requiring an adjustment in terms. In general, the insurer is aware of the impact of vehicle age and claim history in predicting future performance. However, it mistakenly regards vehicle age in a discontinuous fashion and outweighs the importance of recent performance.

5 Demand for Renewal

In this section, I develop and estimate a model of customer policy renewal and cost realization. The goal of this section is to identify and quantify (i) the cost of comprehensive provided coverage, (ii) customers' willingness to pay for policy renewal, and (iii) customers' private information on risk and its relation to demand.

5.1 Model Setup

At the end of the policy contract, a customer can choose whether to renew a comprehensive coverage policy. The customer's net utility from policy renewal equals the difference between the utility from renewing the policy with the insurer, u^1, and the outside option, u^0:

$$U = \underbrace{v(d) - \alpha p}_{u^1} - \underbrace{\max\{v(d) - \tilde{\alpha}\tilde{p} - k + \lambda \mathcal{I}, 0\}}_{u^0}, \tag{1}$$

where $v(d)$ is the utility from comprehensive coverage, which depends on the policy's expected damage, d. $v'(d) > 0$ implies that consumers are adversely selected. p is the premium per value charged by the insurer, while $\tilde{p}$ is the expected premium per value charged by the market competitors. k reflects the costs incurred by the customer when searching for offers among market competitors. $\mathcal{I}$ is a dummy variable that equals 1 if the customer faces a price increase. $\lambda\mathcal{I}$ represents a customer's rational inattention. Intuitively, the consumer becomes aware of the outside option and might decide to search and reoptimize when faced with a price increase.[31] Rational inattention is modeled as an additional component of search cost. The standard search cost generates power for the insurer market. This is true for rational inattention as well. Unlike the standard search cost, rationally inattentive consumers who become attentive following a price increase represent a demand-side friction that might drive a profit-maximizing insurer to provide a uniform pricing scheme.

5.2 Econometric Model

I specify the net utility from policy renewal as a linear function of customer and vehicle observable characteristics, premium per value, the previous period premium per value, and expected damage per value.

$$U_{ijt} = -\alpha p_{ijt} - \lambda \mathcal{I}_{ijt} + \beta x_{ijt} + \gamma d_{ijt} + \psi_{ij} + \omega_i + \varepsilon_{ijt},$$

where i, j, and t index customer, vehicle, and period, respectively. p_{ijt} is the premium per value. $\lambda\mathcal{I}_{ijt} = \lambda \times \mathbb{1}\{p_{ijt} > p_{ijt-1}\}$ represents customer sensitivity to price increase. x_{ijt} are vehicle and client observable characteristics. d_{ijt} is the policy's expected damage per value. γ represents customer selection based on expected damage. Specifically, $\gamma > 0$ implies adverse selection, as customers with a higher utility from renewing their comprehensive insurance policy tend to cost more. ω_i and ψ_{ij} represent the customer-level and license-level unobserved demand components which are fixed over time, respectively. ψ_{ij} is the unobserved license-level demand component, which might be correlated with premiums, while ω_i is an exogenous unobserved client-level demand component; ω_i is normally distributed, $\omega_i \sim N(0, \sigma^2_\omega)$. ε_{ijt} represents variation in unobserved demand fac-

[31] In this model, I consider rational inattention, but unlike Ho et al. (2017), I do not estimate the increase in premiums necessary to cause the customers to be attentive. Instead, I set that increase to zero, which match both the uniform pricing in the data and the recommendation regarding a "Go" grade assignment.

tors across the client's vehicles and over time, which follows a logistic distribution.

Equation 1 suggests that the net utility from policy renewal is non-linear. Some customers search among market competitors, while others decide not to purchase comprehensive insurance at all. Data limitations do not permit determining whether the customer decides to purchase a policy from a competitor. Furthermore, I do not obtain competitor's pricing scheme for all types of vehicles (specifically, all non-truck vehicles). Nevertheless, competitor-offered premiums are a function of observable characteristics, all observed by the econometrician. Therefore, βx_{ijt} takes into account search costs k, utility from renewal $v(d)$, and competitor's pricing scheme $\tilde{p}$.

Regarding the expected damage per value, I adopt a specification similar in spirit to Einav et al. (2013). Damage per value follows a pseudo-Poisson distribution according to the following exponential expected value:

$$d_{ijt} = \exp(\delta x_{ijt} + \nu_i), \tag{2}$$

where ν_i represents customers' private information regarding risk, which is constant over time.[32] ν_i is normally distributed, $\nu_i \sim N(0, \sigma_\nu^2)$, independent of observable characteristics x_{ijt}. This is an appropriate modeling fit as the damage per value is heavily skewed, conditional on claim, and characterized by a significant number of small-scale claims in terms of damage per value (see Appendix Figure A.12). The unobserved constant structure implies that the damage per value varies over time, yet the unobserved component is fixed, as in Einav et al. (2013). This modeling assumption allows taking into account policies that are not up for renewal (for instance, ones that exhibited a total loss event) when estimating Equation 2.

The logistic distribution assumption implies that the probability of renewal, denoted

[32] Cohen and Einav (2007) use a Poisson distribution to fit the policy's number of claims. This implies that the claim process is both state-independent and independent of conditional damage. Since the main goal is to estimate the relationship between consumer demand and the insurer's cost of providing coverage, I use a pseudo-Poisson distribution which accommodates both (i) the possibility of no damages at all, which occurs quite frequently, and (ii) possible dependence between the number of claims and the conditional damage of claims. In practice, I take into account the duration of each policy by estimating a pro-rated variant of Equation 2: $d_{ijt} = \exp(\delta x_{ijt} + \nu_i) \times \tau_{ijt}$, where τ_{ijt} is the duration of the policy. See Appendix B for extensive discussion of the modeling assumptions and estimation process.

by $R_{ijt} = 1$, is described using the standard logit model.

$$\Pr(R_{ijt} = 1 | X_{ijt}, d_{ijt}, \omega_i, \psi_{ij}) = \frac{\exp(-\alpha p_{ijt} - \lambda \mathcal{I}_{ijt} + \beta x_{ijt} + \gamma d_{ijt} + \psi_{ij} + \omega_i)}{1 + \exp(-\alpha p_{ijt} - \lambda \mathcal{I}_{ijt} + \beta x_{ijt} + \gamma d_{ijt} + \psi_{ij} + \omega_i)}$$

Identification

I discuss how variation in the data identifies the model. Identification of damage per value is established in the literature.[33] The effect of observable characteristics on expected damage per value—δ—is identified by variation in the realized damage per value across clients' and vehicles' characteristics, as in Cohen and Einav (2007), Bundorf et al. (2012). Identification of heterogeneity in clients' private information of cost—σ_ν—is established by the within-client correlation across the client's different policies, as in Einav et al. (2013).

The main identification challenge is to estimate consumers' willingness to pay with observational data. I observe adjustments in premiums between consecutive coverage periods for a non-random sample of vehicles. Premiums are determined in equilibrium by both supply and demand forces. I address this issue in two steps. First, premiums might be correlated with the vehicle-specific unobserved component, ψ_{ij}. I take advantage of the population of clients, that contains a large number of fleets with multiple vehicles over multiple coverage periods and treat license-specific unobserved determinants of premiums as fixed rather than random. In the spirit of Crawford et al. (2018), I estimate the following (log) premium equation.

$$\log(P_{ijt}) = \Omega X_{ijt} + F_{ij} + \zeta_{ijt}, \tag{3}$$

where F_{ij} reflects the license-specific determinant of the log premium charged by the insurer. I use the license fixed-effect estimated in the premium model above as a proxy for the constant demand unobservable, $\psi_{ij} = \rho F_{ij}$, to control for the license-specific constant-term which endogenously sets the premium.

Another concern is that the change in premium charged over time is correlated with unobserved demand factors that vary within client or vehicle, over time. I deal with this

[33] See Einav et al. (2010b).

challenge using the fact that premiums are explained remarkably well by the premium equation above—Equation 3—R^2=0.98. The panel structure allows me to use predicted, rather than actual, premiums and impute premiums for those who renew their policy, as in Bundorf et al. (2012), as well as for those who do not, as in Crawford et al. (2018). Specifically, I consider the following renewal probability:

$$\Pr(R_{ijt}=1|X_{ijt},d_{ijt},\hat{p}_{ijt},\hat{f}_{ijt},\omega_i)=\frac{\exp(-\alpha\hat{p}_{ijt}-\lambda\hat{\mathcal{I}}_{ijt}+\beta x_{ijt}+\gamma d_{ijt}+\rho\hat{f}_{ijt}+\omega_i)}{1+\exp(-\alpha\hat{p}_{ijt}-\lambda\hat{\mathcal{I}}_{ijt}+\beta x_{ijt}+\gamma d_{ijt}+\rho\hat{f}_{ijt}+\omega_i)}, \quad (4)$$

where $\hat{f}_{ijt}=\frac{\hat{F}_{ij}}{V_{ijt}}$ and $\hat{p}_{ijt}=\frac{\hat{\Omega}X_{ijt}}{V_{ijt}}$ are the predicted premium license-specific constant term and time-varying term, respectively. Identification of both price sensitivity, α, and sensitivity to price increase, λ, are established using variation in the predicted time-varying premium component, $\hat{p}_{ijt}$, and its relationship to the previous period predicted value term, $\hat{\mathcal{I}}_{ijt}=\mathbb{1}\{\hat{p}_{ijt}-\hat{p}_{ijt-1}>0\}$. As with cost, the effect of observable characteristics on renewal—β—is identified by variation across clients' and vehicles' characteristics, and heterogeneity in clients' demand component (which is independent of premiums)—σ_ω—is established by within-client correlation across the client's different policies, as in Einav et al. (2013). Last, the selection parameter, γ, is identified using within-client variation in observable characteristics that determine the damage per value and explain renewal. To illustrate this, consider a vehicle characterized by a low mean expected damage per value at the current period and a high mean expected damage per value at the next period. Adverse selection implies that the willingness to pay for insurance will increase with private information regarding cost; and the increase in damage per value between periods for costly clients with a high probability of a claim event will be higher than customers with low probability of a claim.

Estimation proceeds in three steps. Initially, since some policies are not renewed, I do not observe their vehicle value. I predict the (log) vehicle value as a function of the previous period value and vehicle type, and treat it as data; both for those that renewed and those that did not. The new vehicle value is well explained by the preceding one (see Appendix Table A.6). Then, I generate the predicted premium per value and predicted increase in premium per value using the estimates of Equation 3 - (see Appendix Table A.7). Finally, I jointly estimate Equations 2 and 4 via Maximum Simulated Likelihood, similar to the approach in Train (2009). I estimate the parameters in the policy renewal and damage equations maximum likelihood using the observable client and vehicle characteristics x, predicted premium per value $\hat{p}$, dummy indicator for an increase in predicted premium per value $\mathcal{I}$, and realizations regarding damage per value and policy renewal. To estimate σ, I use 200 Halton draws for each client: 100 with respect to the unobserved demand component and 100 for private information regarding cost (as in Train (2000)). I then exploit the normal distribution density to derive the likelihood function. See Appendix B for more details.

5.3 Results

The estimation results are provided in Table 7.[34] The cost estimates regarding vehicle age and claim history are similar to those documented in the reduced-form analysis. Expected damage per value increases sharply with vehicle age, and the aggregate loss ratio also plays a significant role in predicting future claims.

In contrast, the previous year loss ratio does not provide any additional information, conditional on augmented data. In addition, new customers are adversely selected; customers who joined the insurer in the last year tend to cost more, conditional on observables. These results demonstrate that the incumbent insurer enjoys a comparative information advantage, as past performance is not shared across the market and competitors observe only recent claims.

As for demand, the estimation of the structural model indicates that customers adversely select to renew policies. Private information regarding customer cost of coverage is associated with a higher renewal probability. However, it should be pointed out that asymmetric information is quite modest.

Price elasticity estimates suggest that customers are sensitive to premiums charged. Moreover, they are very sensitive to an increase in premiums per value (relative to previous period). The coefficient suggests that an increase in premium per value is equivalent to an additional increase of 0.4 p.p in premium per value, which is about 10 percent of the

[34]The estimated coefficient fit both the damage per value distribution and renewal probability quite well. See Appendix Figure A.13.

average premium per value charged. The demand-side friction both generates the insurer market and incentivizes a uniform pricing scheme.

6 Insurer Pricing and Cost

In this part, I develop and estimate a model of insurer policy assessment and supply of coverage. The goal is to identify and quantify the insurer's subjective expected cost of providing coverage and the process of choosing between increasing premiums and rejection, when an adjustment in terms is perceived as necessary.

6.1 Model Setup

When the policy contract is about to end, the insurer has three options with regard to policy i: (i) "Go" - the analytical team recommends renewing the policy with the same premium as previous period; (ii) "Adjust" - the analytical team recommends renewing the policy but not in the current terms, by either increasing premiums and/or deductibles; and (iii) "Reject" - the analytical team recommends not to provide comprehensive coverage. The profit margin, denoted by Π_i under each alternative, denoted by A_i, is defined as follows:

$$\Pi_i = \begin{cases} \Pi_i^g = 1 - \frac{d_i}{p_{-1i}} & \text{if } A_i = \text{Go} \\ \Pi_i^a = q_i(\Delta p_i) \times \left(1 - \frac{d_i}{\tau(\Delta p_i) \times (p_{-1i} + \Delta p_i)}\right) & \text{if } A_i = \text{Adjust} \\ 0 & \text{if } A_i = \text{Reject}, \end{cases}$$

where d_i is the insurer's perceived damage per value of policy i. p_{-1i} is policy i's previous period premium per value. Δp_i is the optimal increase in premium per value by the insurer when adjusting prices. $q_i(\Delta p_i) \in [0, 1]$ is the renewal probability in case of an increase in premium of $\Delta p_i > 0$.[35] $\tau(\Delta p_i) \in [0, 1]$ represents exerted costs incurred by the insurer in case of an adjustment in terms.

The recommendation is intended to maximize the expected profit margin. Therefore,

[35] This implies that I am normalizing the probability of renewal to be one in case a "Go" grade is given. Allowing the renewal probability to be lower than one does not change the analysis.

the optimal strategy is equivalent to selecting the recommendation that maximizes π_i, defined as:

$$\pi_i = \begin{cases} \pi_i^g = \log(p_{-1i}) - \log(d_i) & \text{if } A_i = \text{Go} \\ \pi_i^a = -\log(z_i) & \text{if } A_i = \text{Adjust} \\ 0 & \text{if } A_i = \text{Reject}, \end{cases} \tag{5}$$

where $z_i = 1 - q_i(\Delta p_i) \times (1 - \frac{d_i}{\tau(\Delta p_i)\times(p_{-1i}+\Delta p_i)})$.

Before describing the econometric model, I illustrate the implications of a high adjustment cost on the supply of insurance. Adjustment costs affect customers as an increase in $\tau(\Delta p_i)$ reduces π_i^a. Therefore, adjustment cost both reduce the probability of a price increase, and increase the probability of denial of coverage. These outcomes jointly determine the effect of supply-side frictions on consumer surplus. When π_i^g is sufficiently large, consumers would benefit from these frictions. In contrast, customers who the insurer perceives as costly, i.e., low π_i^g, are less likely to be offered coverage. Furthermore, supply-side frictions might negatively affect low-cost customers as well; that is, a good customer might be involved in a claim event. As a result, customers with high volatility in outcomes are prone to denial, while customers with modest volatility are not.

6.2 Econometric Model

The insurer's decision can be re-expressed as determined by the customer's observable factors $X_i = (p_i, x_i)$.

$$\pi_i = \begin{cases} \pi_i^g = \underbrace{\log(p_{-1i}) - \log(d(X_i))}_{\pi^g(X_i)} + \varepsilon_i^g & \text{if } A_i = \text{Go} \\ \pi_i^a = \underbrace{\log(z(X_i))}_{\pi^a(X_i)} + \varepsilon_i^a & \text{if } A_i = \text{Adjust} \\ 0 & \text{if } A_i = \text{Reject}. \end{cases} \tag{6}$$

Identification of the model requires independence between the observed and unobserved factors. Formally:

$$\varepsilon \perp X.$$

This assumption holds if the insurer does not have any informational advantage relative to the econometrician, i.e., no omitted variables. The rationale for the identification strategy of the structural model is based on the fact that I observe *all* of the information documented by the insurer, which implies informational symmetry. The main challenge in identifying π is that unobserved demand factors might be correlated with the client's and vehicle's observable characteristics. These include undocumented "soft information" (Crawford et al. (2018)) and strategic factors. Specifically, (i) the insurer might know more about the customer's willingness to pay than the econometrician—as reflected by client's and vehicle's observable characteristics (including premiums)—and (ii) strategic considerations, which include fleet cross-subsidization and possible marketing incentives, as providing coverage to a large fleet might serve as an advertisement that attracts new customers. As in the reduced-form analysis, I address these challenges by focusing on non-fleet customers.

As a result, $\pi^g(X)$ and $\pi^a(X)$ are nonparameterically identified (Berry (1994)). Furthermore, since a "Go" grade assignment implies a recommendation to renew a policy with the same premium per value, identification of $\pi^g(X)$ permits identification of the insurer's risk assessment, both in terms of level (group) and slope (selection).

$$\begin{aligned}\pi^g(X) &= \log(p_{-1}) - \log(d(x,p)) \\ &= \log(p_{-1}) - \underbrace{\log(d(x,\bar{p}(x)))}_{\text{Group cost}} - \Big(\underbrace{\log(d(x,p)) - \log(d(x,\bar{p}(x)))}_{\text{Selection}}\Big),\end{aligned}$$

where $\bar{p}(x)$ is the average premium per value charged by customers of observable characteristics x. Given the model specification, the perceived cost function is identified using variation in characteristics across groups and within variation in premiums. Variation in group observable characteristics, x and $\bar{p}(x)$, identifies the insurer's perceived cost of providing comprehensive coverage to the mean customer of group x. Variation within groups in premium charged identifies the perceived selection by the insurer. The one-to-one relationship between $\pi^g(X)$ and $\log(d(x,p))$ demonstrates that identification of the perceived cost function depends on the significant proportion of policies assigned a "Go" grade, for which demand forces do not play a role. If all policies are assigned either an "Adjust" or "Reject" grades, it is not possible to identify the perceived cost. To illustrate this, consider two policies: one that is assigned an "Adjust" grade, and the other a "Reject" grade. I cannot identify whether the rejected policy is denied coverage due to supply forces, i.e., a higher cost of providing coverage, or demand, i.e., a lower willingness to pay. The same holds for two policies with different recommendations regarding the magnitude of the increase in premiums. The offset of demand forces when a "Go" grade is assigned allows for extraction of the insurer's subjective risk assessment.

While the data is rich enough to extract the insurer's perceived costs, it is insufficient to separately identify demand forces and supply frictions determining denial. To illustrate this, consider that for any price increase from p to $\tilde{p}$, the insurer's profit margin, including adjustment cost, is $q(\tilde{p} - p) \times (1 - \frac{d}{\tau(\tilde{p}-p)\times\tilde{p}})$. Identification of adjustment costs, $\tau(\cdot)$, is only possible when the renewal probability is (or approaches) one. Yet, both data limitations, specifically as the number of recommended adjustments are quite scarce and consumers are sensitive to a price increase, do not allow quantifying both channels using, for instance, an identification at infinity approach.[36]

Despite the non-parametric identification, the estimation is made using some parametric assumption. Similar to demand estimation, I assume that the expected damage per value of policy i is defined using an exponential term.

$$d_i = \exp(\tilde{\delta} x_i + \tilde{\rho}(p_i - \bar{p}(x_i))$$

An examination of both $\tilde{\delta}$ and δ permits a comparison between the objective and perceived determinants of damage per value. As a result, $\pi^g(X)$ is expressed using a linear term, while $\pi^a(X_i)$ is estimated using a linear function. Furthermore, I assume $\varepsilon_i^g = \sigma(\epsilon_i^g - \tilde{\epsilon}_i)$ and $\varepsilon_i^a = \sigma(\epsilon_i^a - \tilde{\epsilon}_i)$. $(\epsilon^g, \epsilon^a, \tilde{\epsilon})$ are i.i.d and follow a Type-I extreme value distribution. The

[36]The recommended increase in premiums is quite discrete as well. Distribution of recommended price increases can be found in Appendix Figure A.14.

model is estimated using a multinominal logit model.[37]

$$\pi_i^g = \frac{1}{\sigma}(\log(p_{-1}) - \tilde{\delta}x_i + \tilde{\rho}(p_i - \bar{p}(x_i)) + \epsilon_i^g$$
$$\pi_i^a = \frac{1}{\sigma}(\beta_0 \log(p_{-1}) + \beta_x x_i + \beta_p \bar{p}(x_i)) + \epsilon_i^a$$
$$\pi_i^r = \epsilon_i^r$$

6.3 Results

The results are presented in Table 8. In the first column, I examine how vehicle age determines a "Go" grade assignment. The insurer discontinuously evaluates vehicle age. The insurer does not consider vehicle age, as long as it is less than 8 years, while not assigning a "Go" grade for a substantial portion of old vehicles, defined as age 8 and above. This result is consistent with reduced-form analysis findings of the "Go" grade assignment (see Table 6), while it is in contrast to the estimation results of the cost function in Table 7, which suggests that the cost of providing coverage increases throughout the vehicle life cycle.

With regard to past performance, the insurer perceives claim history as a predictive signal regarding future claims. The probability of a "Go" grade assignment is less for policies with a reported claim in the previous period and a higher aggregate loss ratio. As documented in Table 6, the insurer places additional emphasis on the previous year's performance, despite the fact that estimation results of the cost function (see Table 7) indicate that this is not a predictive signal. The overweighing of the previous year loss ratio affects customers heterogeneously, depending on their fleet size. Single-fleet customers are exposed to substantial volatility in their performance. A good driver might incur high damages. Large fleets are less exposed to this risk; as fleet size increases, the probability of extreme events diminishes, suggesting that the insurer assessment and pricing scheme might be advantageous for large fleets, yet disadvantageous for a single-fleet customer.

Lastly, the coefficient regarding selection $\tilde{\delta}$ indicates that the insurer perceives substantial adverse selection. Customers paying more than the average premium paid, based on

[37] This implies that identification of the subjective cost function components is independent of $\pi^a(X_i)$'s function form.

its observable characteristics, have a lower probability of their policy being assigned a "Go" grade, although cost estimates suggest modest private information.

In Figure 7, I present a comparison between the insurer's subjective determinants of cost, $\tilde{\delta}$, and the objective cost function δ. The figure indicates that the insurer underweighs factors such as new customer indicator and vehicle age groups below 8 (age 2–4 and 5–7). In contrast, the insurer overweighs other factors, including the previous year's loss ratio (although it has no predictive power), the aggregate loss ratio, especially that of a loss-generating clients (i.e., loss ratio above one), and vehicle age, if it is at least 8 years old.

To summarize, the insurer is aware of the importance of vehicle age and claim history as determinants of future performance. Yet, it misevaluates vehicle age in a discontinuous fashion and overweighs the importance of recent performance. Two key conclusions emerge. First, the results of this analysis further demonstrate the importance of considering not only the intensive margin but also the extensive margin. Second, the insurer's biased assessment harms disadvantaged customers; that is, those that purchase comprehensive coverage for a single, old vehicle.

7 Counterfactual Analysis

Using the structural estimates of the demand for policy renewal and the insurer's behavior implied by the "Go—No Go" grading, I conduct a few counterfactuals to assess the implications of supply-side behavioral frictions. The analysis is conducted under the assumption that market competitors do not respond to changes in the insurer's behavior, as I do not estimate the cross-substitution patterns between the market competitors' pricing schemes and net utility from policy renewal. The counterfactual analysis is conducted by drawing 200 Halton values: 100 values for the private information regarding cost and 100 values for unobserved demand factors.

I start by examining the premium per value charged for providing coverage for the average truck owner, who was charged a premium per value of 3.5 p.p in the previous period. The results are presented in Panel A of Figure 8. I present three different pricing schemes. The blue curve is the optimal pricing scheme, that is, the profit-maximizing premiums. The red curve is the optimal pricing scheme, while restricting the premiums per value does not change throughout the possible states. The green curve is implied by a structural estimation of the insurer's behavior (see Table 8). The results in Figure 8, Panel A indicate that asymmetric information is quite modest, as the average truck owner is not denied coverage by either an unrestricted or uniform-price restricted profit-maximizing insurer throughout the vehicle life cycle. As with expected damage per value, a profit-maximizing insurer increases premiums per value with vehicle age. In addition, a profit-maximizing insurer does not increase premium per value for a new vehicle. This demonstrates how rationally inattentive consumers might lead profit-maximizing insurers to provide a somewhat uniform pricing scheme.

The insurer behavior implied by the "Go—No Go" grades differs substantially from that of a profit-maximizing insurer. The insurer does not change premiums for vehicles up to the age of 8, and then rejects the policy. The rejection occurs despite the limited selection, suggesting that the insurer excessively denies customers, relative to a rational insurer facing customers with private information of cost. It is not adverse selection that generates rejection, but rather firm practices; underadjustment of the intensive margin spills over to the extensive margin. The insurer forgoes profits. In particular, the insurer could have increased profits by 7 percent if it had acted as a profit-maximizing firm.

In Panel B of Figure 8, I examine the pricing scheme for a single-fleet truck owner following different claim realizations. I consider the observable characteristics of the average customer, characterized by a loss ratio of 70 percent. Note that a single-vehicle customer with an expected loss ratio of 70 percent, who is a profitable customer on average, exhibits a loss ratio of 200 percent or above at a probability of 11 percent. That probability diminishes substantially when considering a fleet of vehicles with the same loss ratio.

As in the case of a new vehicle, optimal pricing does not change following a significant positive realization (loss ratio below 50 percent). The premium per value does change substantially, however, after a negative realization. The premium almost doubles after an outcome of loss ratio of 200 percent, or above. As mentioned above, asymmetric information is quite modest, as the average single-fleet truck owner is not denied coverage by an unrestricted or uniform-price restricted profit-maximizing insurer following any claim realization.

As with vehicle age, the insurer behavior implied by the "Go—No Go" grades extensively deviates from that of a profit-maximizing insurer. The insurer does not change premiums for a loss ratio below 200 percent, yet rejects the policy when it exceeds that loss ratio, despite the limited selection, suggesting that the insurer excessively denies customer, as with old vehicles. The insurer forgoes profits of 16 percent by deviating from profit-maximizing behavior. Furthermore, the net consumer surplus from facing a behavioral insurer, relative to a profit-maximizing firm, is negative. This is not surprising as the probability of denial by the behavioral insurer is higher than the probability that only the profit-maximizing firm would increase premiums. The insurer's biased assessment regarding the recent claim history and the lack of adjustment on the intensive margin harms disadvantage customers—those that purchase comprehensive coverage for a single vehicle. The probability of facing a denial drops substantially with fleet size, suggesting the demand is increasing return to scale. Identical single-fleet customers benefit from purchasing coverage as a whole, independent of price bargaining or risk pooling incentives.

8 Conclusion

A cornerstone in the research on risk and insurance is that providers price correctly. In this paper, I inquire whether this is the case. Using data from the one of the largest Israeli commercial auto insurance providers, I find there is too little adjustment in the intensive margin. Premiums barely change with expected costs as projected by predetermined factors (vehicle age) and signals (claim history). Furthermore, I find there is too much adjustment in the extensive margin; that is, an excessive denial of insurance following a negative realization. Using unique grading documents, I integrate the insurer's subjective risk assessment into the study of selection markets, in general, and insurance markets, in particular. I find that the insurer's risk assessment overweighs recent claims and misevaluates vehicle age. Structural model estimates suggest that insurers enjoy incumbency advantages over their own customers, and clients are rationally inattentive to competitors' pricing unless they are faced with a price increase. Both channels allow suboptimal behavior to persist. Finally, I find that supply-side behavioral frictions, which result in excessive denial, diminish with a client's fleet size. This implies that disadvantaged single-vehicle owners are harmed by supply-side frictions, while purchasing insurance coverage as a whole dilutes those losses and might even generate benefits. The spillover of the lack of intensive-margin adjustment on the extensive margin raises important concerns regarding policy intervention.[38]

The results in this paper document the importance of considering both behavioral frictions in selection market analysis and implementing IO structural tools in behavioral economics. The insurer's subjective beliefs regarding the cost of providing coverage differ from an objective assessment. Moreover, the implied behavior suggests overadjustment on the extensive margin and underadjustment on the intensive margin, relative to that implied by state-of-the-art IO analysis. With regard to behavioral economics, examining solely premiums might be quite misleading. When only considering the intensive margin, one might erroneously conclude that the insurer does not take into account observable characteristics when assessing risk. The IO setting, which considers both the intensive and the extensive margins, is essential for identifying and quantifying the effect of biased beliefs.

References

Abaluck, J. and Gruber, J. (2011). Choice inconsistencies among the elderly: evidence from plan choice in the medicare part d program. *American Economic Review*, 101(4):1180–1210.

Akerlof, G. A. (1970). The market for "lemons": Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3):488–500.

Barseghyan, L., Molinari, F., O'Donoghue, T., and Teitelbaum, J. C. (2013). The nature of risk preferences: Evidence from insurance choices. *American economic review*, 103(6):2499–2529.

[38]See Einav et al. (2021) regarding the equity and fairness of price variation.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, pages 242–262.

Bhargava, S., Loewenstein, G., and Sydnor, J. (2017). Choose to lose: Health plan choices from a menu with dominated option. *The Quarterly Journal of Economics*, 132(3):1319–1372.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does management matter? evidence from india. *The Quarterly journal of economics*, 128(1):1–51.

Bloom, N. and Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *The quarterly journal of Economics*, 122(4):1351–1408.

Brot-Goldberg, Z. C., Chandra, A., Handel, B. R., and Kolstad, J. T. (2017). What does a deductible do? the impact of cost-sharing on health care prices, quantities, and spending dynamics. *The Quarterly Journal of Economics*, 132(3):1261–1318.

Bundorf, M. K., Levin, J., and Mahoney, N. (2012). Pricing and welfare in health plan choice. *American Economic Review*, 102(7):3214–48.

Cabral, M. (2016). Claim timing and ex post adverse selection. *The Review of Economic Studies*, 84(1):1–44.

Cabral, M., Geruso, M., and Mahoney, N. (2018). Do larger health insurance subsidies benefit patients or producers? evidence from medicare advantage. *American Economic Review*, 108(8):2048–87.

Cardon, J. H. and Hendel, I. (2001). Asymmetric information in health insurance: evidence from the national medical expenditure survey. *RAND Journal of Economics*, pages 408–427.

Carlin, C. and Town, R. (2009). Adverse selection, welfare, and optimal pricing of employer sponsored health plans. *U. Minnesota Working Paper*, 1(2).

Cavallo, A., Neiman, B., and Rigobon, R. (2014). Currency unions, product introductions, and the real exchange rate. *The Quarterly Journal of Economics*, 129(2):529–595.

Chiappori, P.-A. and Salanie, B. (2000). Testing for asymmetric information in insurance markets. *Journal of political Economy*, 108(1):56–78.

Cho, S. and Rust, J. (2010). The flat rental puzzle. *The Review of Economic Studies*, 77(2):560–594.

Cohen, A. and Einav, L. (2007). Estimating risk preferences from deductible choice. *American economic review*, 97(3):745–788.

Crawford, G. S., Pavanini, N., and Schivardi, F. (2018). Asymmetric information and imperfect competition in lending markets. *American Economic Review*, 108(7):1659–1701.

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

Cutler, D. M., Finkelstein, A., and McGarry, K. (2008). Preference heterogeneity and insurance markets: Explaining a puzzle of insurance. *American Economic Review*, 98(2):157–62.

Cutler, D. M. and Reber, S. J. (1998). Paying for health insurance: the trade-off between competition and adverse selection. *The Quarterly Journal of Economics*, 113(2):433–466.

DellaVigna, S. and Gentzkow, M. (2019). Uniform pricing in us retail chains. *The Quarterly Journal of Economics*, 134(4):2011–2084.

Einav, L., Finkelstein, A., and Cullen, M. R. (2010a). Estimating welfare in insurance markets using variation in prices. *The quarterly journal of economics*, 125(3):877–921.

Einav, L., Finkelstein, A., and Levin, J. (2010b). Beyond testing: Empirical models of insurance markets. *Annu. Rev. Econ.*, 2(1):311–336.

Einav, L., Finkelstein, A., and Mahoney, N. (2021). The io of selection markets. In *Handbook of Industrial Organization*, volume 5, pages 389–426. Elsevier.

Einav, L., Finkelstein, A., Ryan, S. P., Schrimpf, P., and Cullen, M. R. (2013). Selection on moral hazard in health insurance. *American Economic Review*, 103(1):178–219.

Einav, L., Finkelstein, A., and Tebaldi, P. (2019). Market design in regulated health insurance markets: Risk adjustment vs. subsidies. *Unpublished mimeo, Stanford University, MIT, and University of Chicago*, 7:32.

Fang, H., Keane, M. P., and Silverman, D. (2008). Sources of advantageous selection:

Evidence from the medigap insurance market. *Journal of political Economy*, 116(2):303–350.

Finkelstein, A., McGarry, K., and Sufi, A. (2005). Dynamic inefficiencies in insurance markets: Evidence from long-term care insurance. *American Economic Review*, 95(2):224–228.

Finkelstein, A. and Poterba, J. (2014). Testing for asymmetric information using "unused observables" in insurance markets: Evidence from the uk annuity market. *Journal of Risk and Insurance*, 81(4):709–734.

Fleitas, S., Gowrisankaran, G., and Sasso, A. L. (2020). Reclassification risk in the small group health insurance market.

Ghili, S., Handel, B., Hendel, I., and Whinston, M. D. (2021). Optimal long-term health insurance contracts: characterization, computation, and welfare effects.

Gottlieb, D. and Smetters, K. (2021). Lapse-based insurance. *American Economic Review*, 111(8):2377–2416.

Hackmann, M. B., Kolstad, J. T., and Kowalski, A. E. (2015). Adverse selection and an individual mandate: When theory meets practice. *American Economic Review*, 105(3):1030–66.

Handel, B., Hendel, I., and Whinston, M. D. (2015). Equilibria in health exchanges: Adverse selection versus reclassification risk. *Econometrica*, 83(4):1261–1313.

Handel, B. and Ho, K. (2021). The industrial organization of health care markets. In *Handbook of Industrial Organization*, volume 5, pages 521–614. Elsevier.

Handel, B. R. (2013). Adverse selection and inertia in health insurance markets: When nudging hurts. *American Economic Review*, 103(7):2643–82.

Handel, B. R. and Kolstad, J. T. (2015). Health insurance for" humans": Information frictions, plan choice, and consumer welfare. *American Economic Review*, 105(8):2449–2500.

Handel, B. R., Kolstad, J. T., and Spinnewijn, J. (2019). Information frictions and adverse selection: Policy interventions in health insurance markets. *Review of Economics and Statistics*, 101(2):326–340.

Hendel, I. and Lizzeri, A. (2003). The role of commitment in dynamic contracts: Evidence from life insurance. *The Quarterly journal of economics*, 118(1):299–328.

Hendren, N. (2017). Knowledge of future job loss and implications for unemployment insurance. *American Economic Review*, 107(7):1778–1823.

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

Jin, Y. and Vasserman, S. (2020). Buying data from consumers: The impact of monitoring in us auto insurance.

Koch, T. G. (2014). One pool to insure them all? age, risk and the price (s) of medical insurance. *International Journal of Industrial Organization*, 35:1–11.

Lester, B., Shourideh, A., Venkateswaran, V., and Zetlin-Jones, A. (2019). Screening and adverse selection in frictional markets. *Journal of Political Economy*, 127(1):338–377.

Lustig, J. (2010). Measuring welfare losses from adverse selection and imperfect competition in privatized medicare. *Manuscript. Boston University Department of Economics*.

Mahoney, N. and Weyl, E. G. (2017). Imperfect competition in selection markets. *Review of Economics and Statistics*, 99(4):637–651.

McMillan, R. S. (2007). Different flavor, same price: The puzzle of uniform pricing for differentiated products. *Same Price: The Puzzle of Uniform Pricing for Differentiated Products (January 17, 2007)*.

Orbach, B. Y. and Einav, L. (2007). Uniform prices for differentiated goods: The case of the movie-theater industry. *International Review of Law and Economics*, 27(2):129–153.

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

Shiller, B. and Waldfogel, J. (2011). Music for a song: an empirical look at uniform pricing and its alternatives. *The Journal of Industrial Economics*, 59(4):630–660.

Spinnewijn, J. (2017). Heterogeneity, demand for insurance, and adverse selection. *American Economic Journal: Economic Policy*, 9(1):308–43.

Starc, A. (2014). Insurer pricing and consumer welfare: Evidence from medigap. *The RAND Journal of Economics*, 45(1):198–220.

Sydnor, J. (2010). (over) insuring modest risks. *American Economic Journal: Applied Economics*, 2(4):177–99.

Tebaldi, P. (2022). Estimating equilibrium in health insurance exchanges: Price competition and subsidy design under the aca. Technical report, National Bureau of Economic Research.

Tilipman, N. (2022). Employer incentives and distortions in health insurance design: Implications for welfare and costs. *American Economic Review*, 112(3):998–1037.

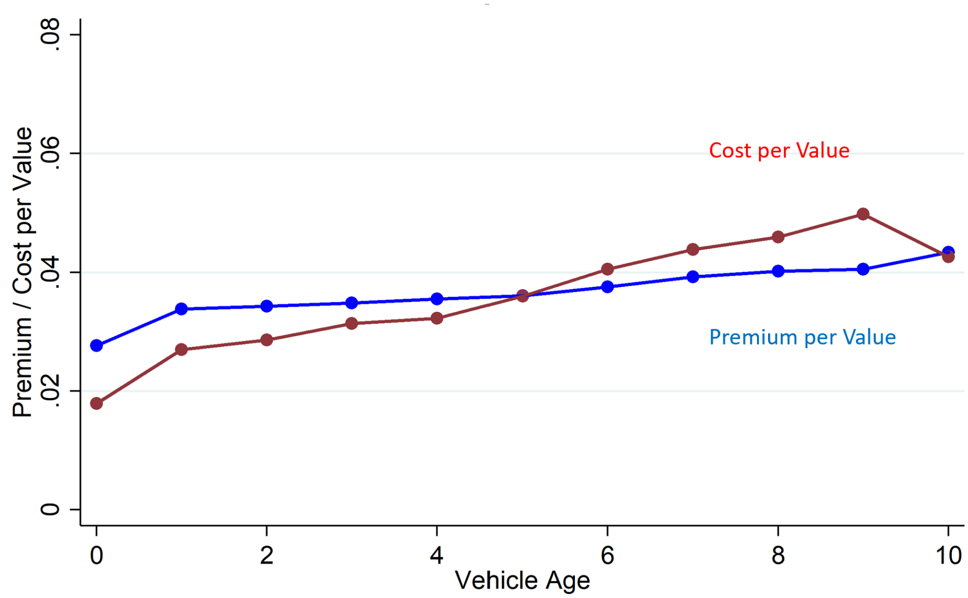
Train, K. (2000). Halton sequences for mixed logit.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge university press.

Veiga, A. and Weyl, E. G. (2016). Product design in selection markets. *The Quarterly Journal of Economics*, 131(2):1007–1056.

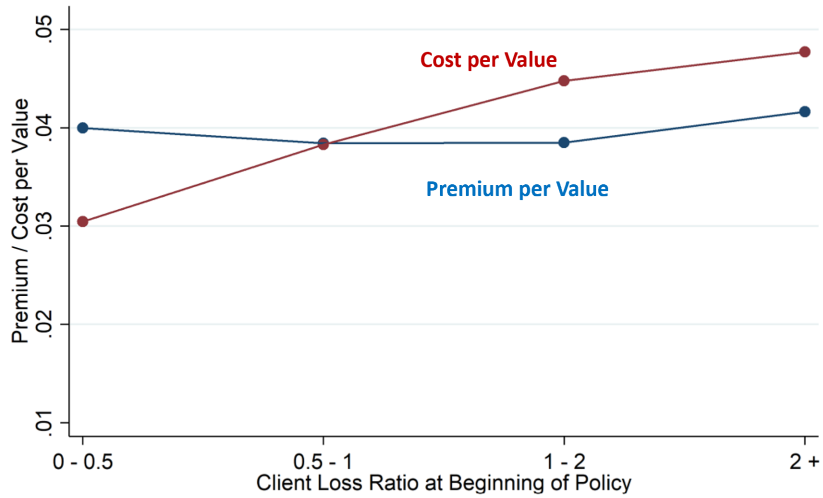
Figures

Figure 1: Summary statistics of premium and cost per value by vehicle age

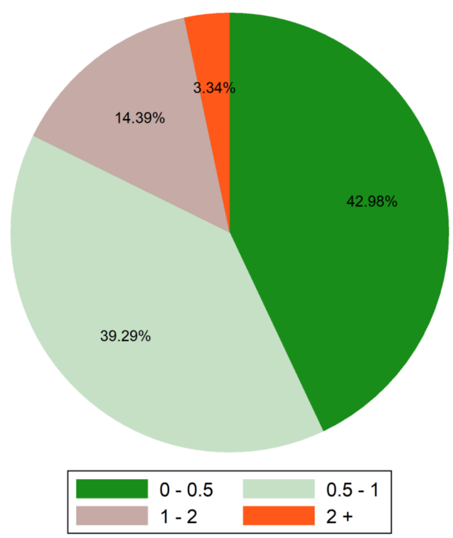


Notes: The figure describes the relationship between premium per value, cost per value, and vehicle age. The vertical axis depicts the mean premium per value (in blue) and costs per value (in red) of comprehensive coverage policies for trucks from 2013 to 2020. The horizontal axis depicts the vehicle's age. No controls are added. Both variables are standardized to an annual term policy. Premiums, costs, and vehicle values are measured in New Israeli Shekel (ILS).

Figure 2: Summary statistics of client loss ratio



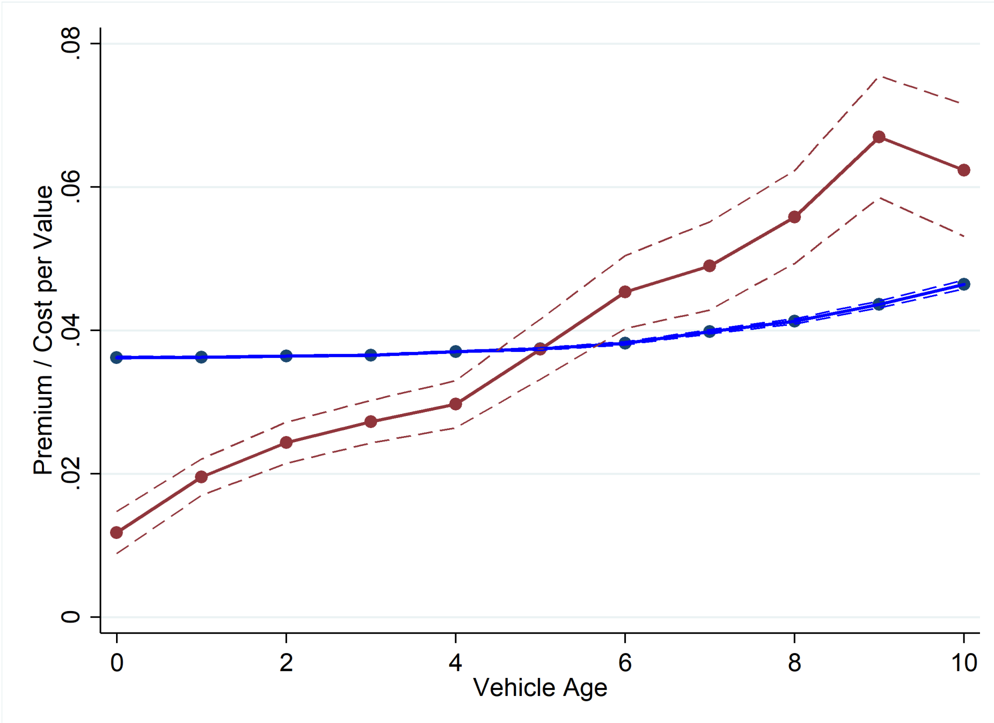
Panel A: Premium and cost per value by client loss ratio



Panel B: Distribution of policies by client loss ratio

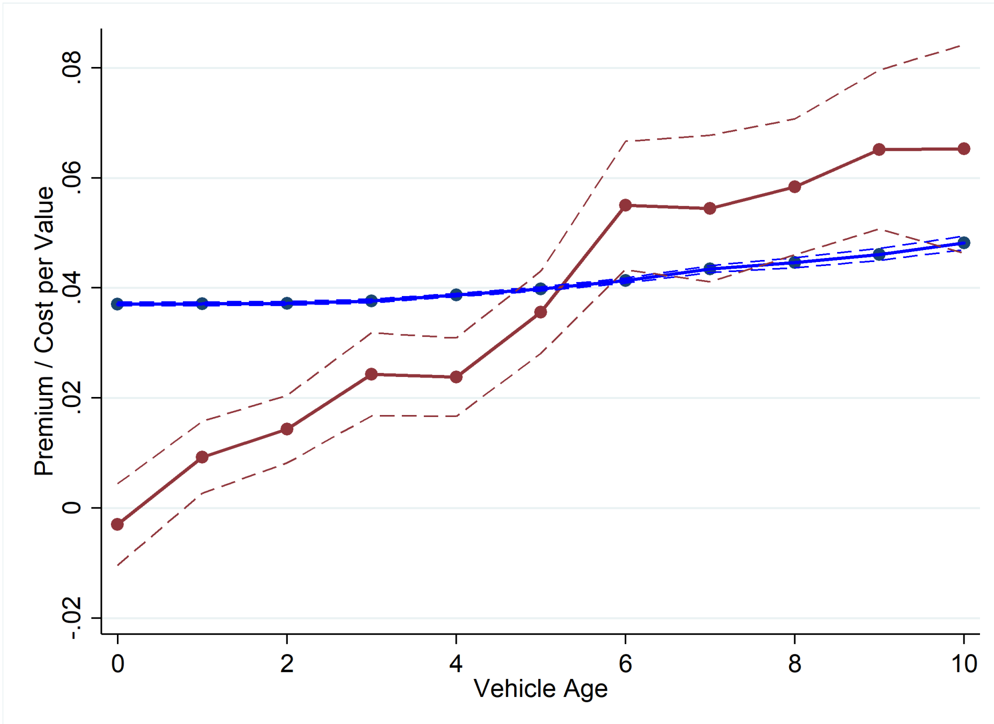
Notes: The figure describes premium per value and cost per value by client loss ratio and the sample distribution of client loss ratio. In Panel A, the vertical axis depicts the mean premium per value (in blue) and cost per value (in red) of comprehensive coverage policies for trucks from 2014 to 2020. The horizontal axis depicts the vehicle's age. No controls are added. Both variables are standardized to an annual term policy. Premiums, costs, and vehicle values are measured in New Israeli Shekel (ILS). The sample is divided into four groups: "0–0.5" client loss ratio (aggregate damage over aggregate premium) is up to 0.5, "0.5–1" client loss ratio is above 0.5 and below 1, "1–2" client loss ratio is above 1 and below 2. and "2+" client loss ratio is above 2. Panel B reports the distribution of the sample among the four groups of clients past performance.

Figure 3: Premium and cost per value by vehicle age



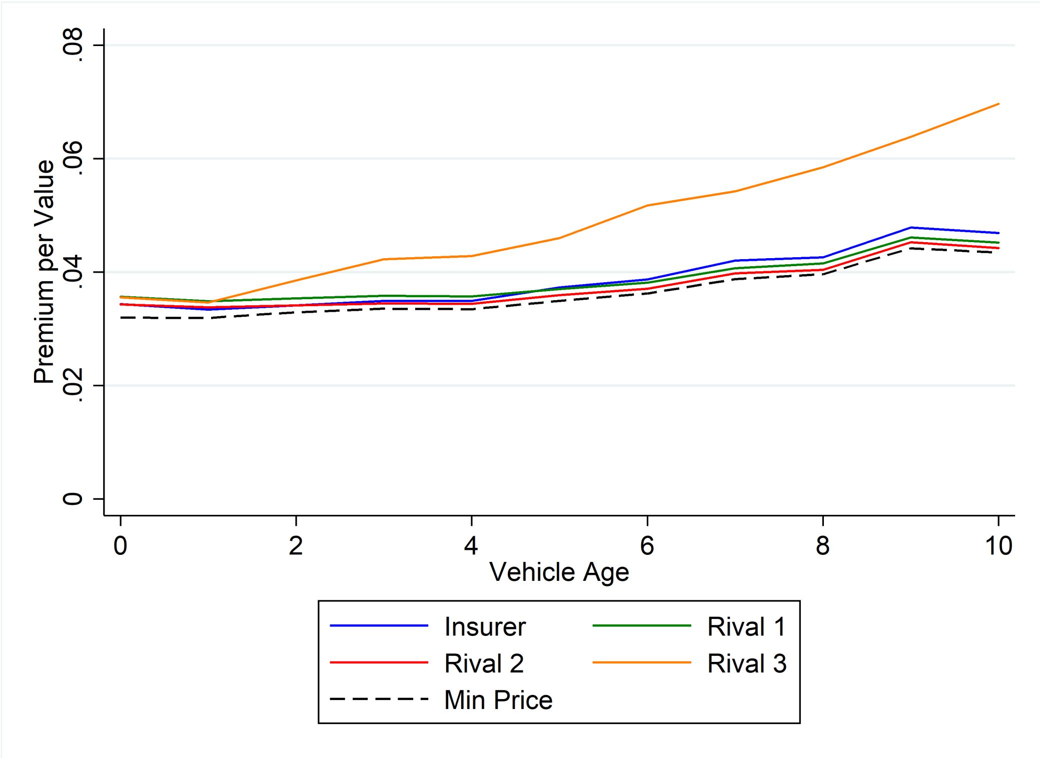
Notes: The figure reports the estimation results of a fixed-effect (license level) saturated regression of premium per vehicle value (in blue) and cost per vehicle value (in red) on vehicle age. Each vehicle age, from 0 to 10, has a unique coefficient. The vertical axis depicts the two dependent variables. The horizontal axis depicts vehicle age. The solid lines represent the regression coefficients. The dashed lines depict the 95% confidence interval. The confidence interval is constructed using robust standard errors clustered at the client level. The sample includes comprehensive insurance coverage policies for trucks from 2013 through 2020. Premiums and costs are normalized to an annual policy length. Premiums, costs, and vehicle values are measured in New Israeli Shekel (ILS).

Figure 4: Premium and cost per value by vehicle age for nonfleet clients



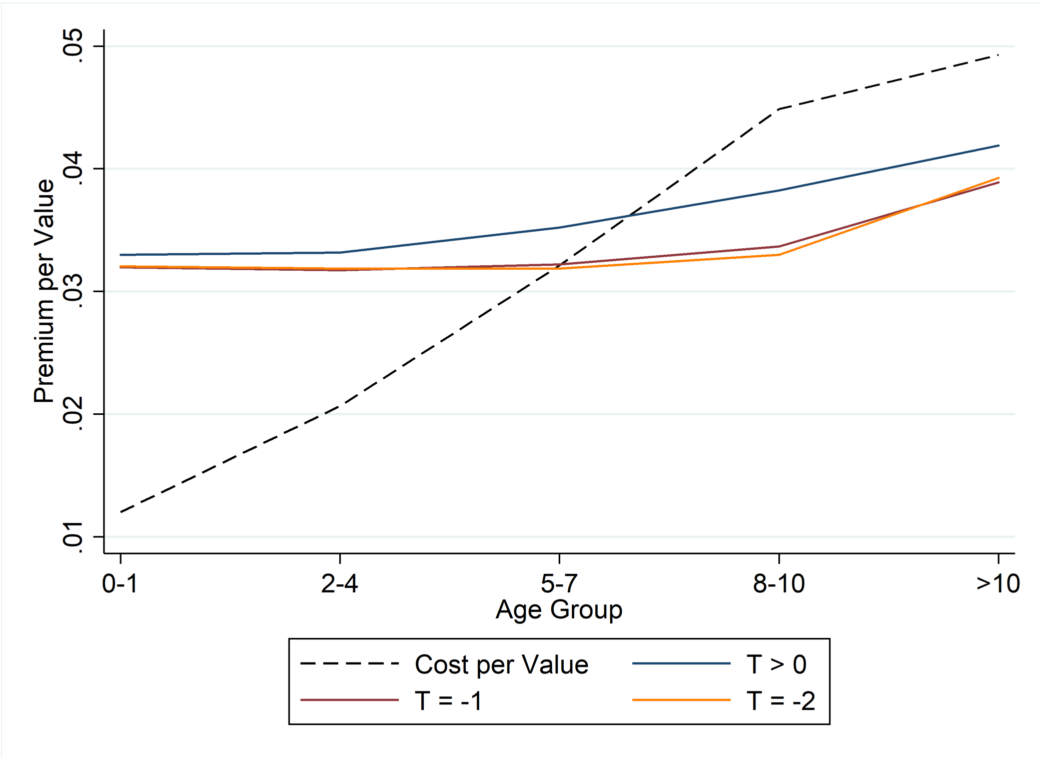
Notes: The figure reports a robustness regression estimation results of the model presented in Figure 3, for non-fleet customers, as defined by the insurer. The number of vehicles insured via any type of coverage by client in a given year is less than five.

Figure 5: Market-wide premiums per value by vehicle age



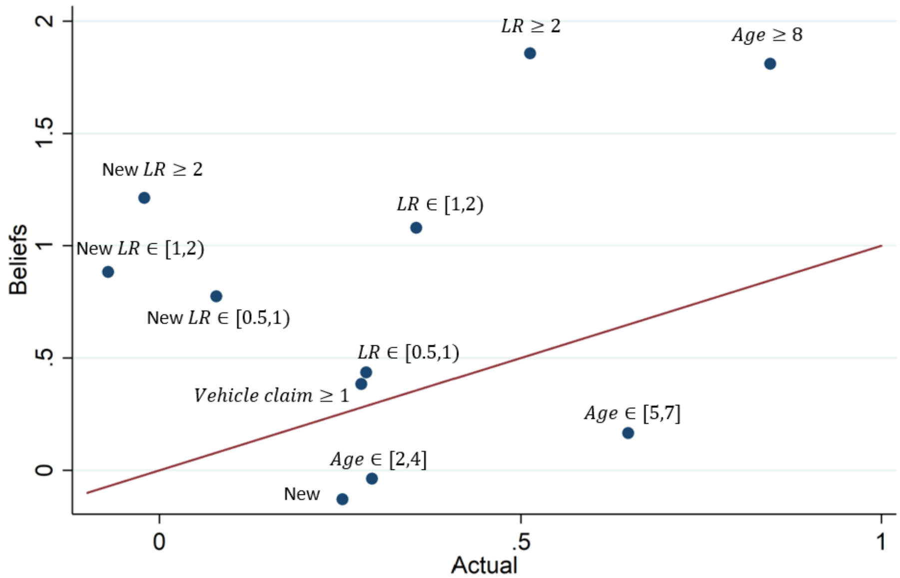
Notes: The figure reports the relationship between market premiums per value and vehicle age. Market-wide premiums are collected via fictitious policy offers generating via Orlan insurance agency's platform (Orlanet Calculator). The sample consists of 876 distinct vehicle model-age values for the top four insurers in the market (the insurer that provided the data and its three main competitors), without any reported claim in the last three years. The horizontal axis depicts premiums per value. The vertical axis depicts vehicle age. The curves are the coefficients of a saturated regression of premiums per value on vehicle age. The dashed line depicts the minimum premium per value in the market (not restricted to the four insurers). Premiums and values are measured in New Israeli Shekel (ILS).

Figure 6: Premium adjustment based on new information



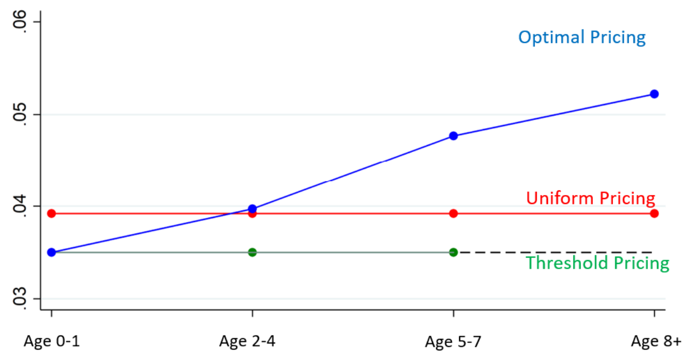
Notes: The figure reports the estimation results of a fixed-effect regression of premium per vehicle value and cost per vehicle value on vehicle age. Observations are divided into five groups based on vehicle age. $T = 0$ indicates the timing at which the insurer was provided information regarding the misadjustment in pricing over the vehicle life cycle. The vertical axis depicts both the premium and cost per value variables. The horizontal axis depicts vehicle age. The red and orange lines represent the estimated premium per value before information was given. The blue line represents the estimated premium per value after information was given. The black dashed line represents the cost per value. The sample includes comprehensive insurance coverage policies for trucks from 2013 through 2020. Premiums and costs are normalized to an annual policy length. Premiums, costs, and vehicle values are measured in New Israeli Shekel (ILS).

Figure 7: Comparison of objective and subjective cost estimates

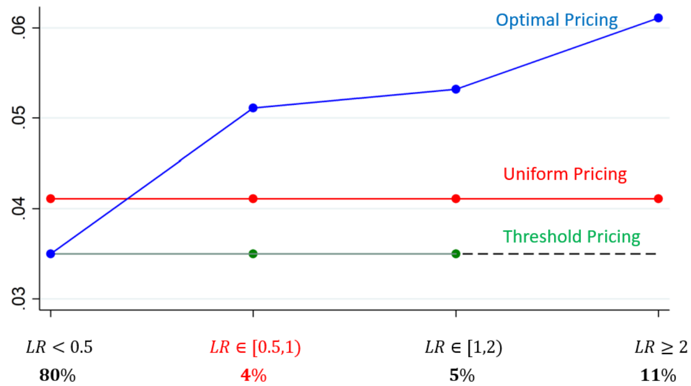


Notes: The figure reports the estimation results of both the objective cost function, as reported in Table 7 and the insurer's perceived cost, as reported in Table 8. The vertical axis measures the coefficient of the subjective cost components, while the horizontal axis measures the coefficient of the objective cost components.

Figure 8: Counterfactual analysis



Panel A: Vehicle age



Panel B: Claim

Notes: The figures above provide two counterfactuals. In Panel A, I examine the trend in premium per value (measured on the vertical axis) over the vehicle life cycle (measured on the horizontal axis). The counterfactual analysis is conducted for the average truck owner, who paid 3.5 p.p premium per value. In Panel B, I examine the trend in premium per value (measured on the vertical axis) over different realizations of current year loss ratio (measured on the horizontal axis). The counterfactual analysis is conducted for a single vehicle truck owner with an average loss ratio (0.7) who paid 3.5 p.p premium per value. The blue curve indicates optimal pricing, the red curve indicates optimal pricing conditional on uniform pricing, and the green curve indicates the pricing strategy based on the behavior implied by the "Go—No Go" grades, as presented in Table 8.

Tables

Table 1: Summary statistics of comprehensive coverage policies for trucks

	(1) All	(2) Vehicle Age$\geq$ 6	(3) Claim$_{t-1}$ $\geq$1
Policies	51,684	15,506	8,358
Share	100%	30.00%	16.17%
Weighted Share (by Premium)	100%	18.86%	15.96%
Mean Premium	9,938	6,246	9,811
At least 1 claim	23.98%	23.78%	34.82%
Mean Damage	6,794	5,861	9,345
Mean Commission	1,557	1,012	1,577
Mean Profit	1,587	-627	-1,111
Profit Margin	15.97%	-10.04%	-11.32%
Mean Vehicle Age	4.18	8.93	5.00
Mean Vehicle Value	298,659	160,383	275,954
Mean Premium per Value	3.33%	4.01%	3.56%

Notes: The table reports summary statistics of comprehensive coverage policies for trucks between 2013 and 2020. The first column reports statistics for all policies, the second column describes the statistics for policies with a vehicle age of six or above, and the third column describes the statistics for policies with at least one claim in the previous period (Claim$_{t-1}$ $\geq$1). Profit margin is defined as mean profit (=premium−damage−commission) over mean premium. Mean premium per value is defined as premium over vehicle value. Mean damage is the mean damage of customers' claims (net of deductibles). Vehicle value, premium, commission, paid claims, and profit are measured in New Israeli Shekel (ILS). Vehicle age is measured in years. I exclude from the sample observation with an error , a change in vehicle within the policy, a change in coverage terms over the policy, and policies that did not end or that lasted for less than 30 days (without a claim).

Table 2: Policy outcomes and past performance

	Panel A: Entire Sample			
	(1)	(2)	(3)	(4)
	$\text{Claim}_t \geq 1$	$\frac{\text{Damage}_t}{\text{Value}_t}$	$\%\Delta\frac{\text{Premium}_t}{\text{Value}_t}$	Loss Ratio_t
$\text{Claim}_{t-1} \geq 1$	0.122***	0.017***	-0.001	0.418***
	(0.008)	(0.002)	(0.002)	(0.049)
log(Value)	Y	N	N	Y
Vehicle Age - 2^{nd} order	Y	Y	Y	Y
Vehicle Weight Class	Y	Y	Y	Y
Driver Underage Indicator	Y	Y	Y	Y
Observations	32,870	32,870	32,870	32,870
R-squared	0.022	0.009	0.034	0.006
	Panel B: Non-Fleet Policies			
	(1)	(2)	(3)	(4)
	$\text{Claim}_t \geq 1$	$\frac{\text{Damage}_t}{\text{Value}_t}$	$\%\Delta\frac{\text{Premium}_t}{\text{Value}_t}$	Loss Ratio_t
$\text{Claim}_{t-1} \geq 1$	0.118***	0.024***	0.017***	0.547***
	(0.013)	(0.004)	(0.005)	(0.101)
log(Value)	Y	N	N	Y
Vehicle Age - 2^{nd} order	Y	Y	Y	Y
Vehicle Weight Class	Y	Y	Y	Y
Driver Underage Indicator	Y	Y	Y	Y
Observations	8,372	8,367	8,367	8,367
R-squared	0.023	0.009	0.014	0.006

Notes: The table reports the relationship between previous claim history and current outcomes. Panel A's sample includes all non-new comprehensive insurance policies for trucks. Panel B's sample includes only comprehensive insurance policies for trucks of nonfleet clients (number of vehicles insured via any type of coverage by client in a given year is less than five year). $\text{Claim}_{t-1} \geq 1$ is an indicator that equals one if at least one claim has been reported with regard to the policy in the previous period. $\text{Claim}_t \geq 1$ is an indicator that equals one if at least one claim has been reported with regard to the policy in the current period. $\frac{\text{Damage}_t}{\text{Value}_t}$ denotes damage (net claim expenses) per value in the current period, $\%\Delta\frac{\text{Premium}_t}{\text{Value}_t}$ denotes the percent change in premium per value in the current period relative to the previous period ($\frac{p_t - p_{t-1}}{p_{t-1}}$, where p is the premium per value) and Loss Ratio_t denotes the current period policy's loss ratio measured as damage over premiums. Controls include (log) vehicle value, vehicle age (2^{nd} order), vehicle type, vehicle weight class, and driver underage indicator. Vehicle value, premium, damage, and loss ratio are measured in New Israeli Shekel (ILS). Vehicle age is measured in years. Robust standard errors, clustered at the client level, are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3: Policy outcomes and past performance: recent vs. older

	Panel A: Entire Sample			
	(1)	(2)	(3)	(4)
	$\text{Claim}_t \geq 1$	$\frac{\text{Damage}_t}{\text{Value}_t}$	$\%\Delta \frac{\text{Premium}_t}{\text{Value}_t}$	Loss Ratio_t
Client's Aggregate Loss Ratio	0.0357***	0.0055***	-0.0023	0.1483***
	(0.0060)	(0.0013)	(0.0022)	(0.0314)
Client's Previous Yr. Loss Ratio	-0.0004	0.0003	0.0033**	0.0067
	(0.0017)	(0.0006)	(0.0017)	(0.0152)
log(Value)	Y	N	N	Y
Vehicle Age - 2^{nd} order	Y	Y	Y	Y
Vehicle Type	Y	Y	Y	Y
Vehicle Weight Class	Y	Y	Y	Y
Driver Underage Indicator	Y	Y	Y	Y
Observations	35,765	35,765	35,765	35,765
R-squared	0.009	0.007	0.009	0.004

Notes: The table reports the relationship between previous claim history and current outcomes. The sample includes all comprehensive insurance policies for trucks from 2014 to 2020, with at least one year of performance history. Client's aggregate loss ratio is the client's ratio of total damages (starting 2013) per total revenue (starting 2013). The client's previous year loss ratio is the client's ratio of previous year's total damages (net claim expenses) over the previous year's total revenue (paid premiums). $\text{Claim}_t \geq 1$ is an indicator that equals one if at least one claim has been reported with regard to the policy at the current period. $\frac{\text{Damage}_t}{\text{Value}_t}$ denotes damage (net claim expenses) per value at the current period, $\%\Delta \frac{\text{Premium}_t}{\text{Value}_t}$ denotes the percent change in premium per value at the current period, relative to the previous period ($\frac{p_t - p_{t-1}}{p_{t-1}}$, where p is the premium per value) and Loss Ratio_t denotes the current period policy's loss ratio, measured as damage over premiums. Controls include (log) vehicle value, vehicle age (2^{nd} order), vehicle type, vehicle weight class, and driver underage indicator. Vehicle value, premium, damage, and loss ratio are measured in New Israeli Shekel (ILS). Vehicle age is measured in years. Robust standard errors, clustered at client level, are reported in parentheses. *** p<0.01, ** p<0.05, *p<0.1.

Table 4: Market premiums by claim history

	Dependent Variable: Premium per Value				
	(1)	(2)	(3)	(4)	(5)
	Insurer	Rival 1	Rival 2	Rival 3	Min. Price
Constant	0.0379***	0.0380***	0.0369***	0.0476***	0.0358***
	(0.0003)	(0.0002)	(0.0002)	(0.0005)	(0.0002)
1 Claim Last Yr (3 yrs)	0.0028***	0.0000	0.0000	·	0.0004
	(0.0004)	(0.0003)	(0.0003)		(0.0003)
$\geq$ 2 Claims Last 3 Yrs	·	·	·	·	·
Observations	1,752	1,752	1,752	876	1,752
R-squared	0.0270	0.0000	0.0000	0.0000	0.0009

Notes: The table reports the relationship between market premiums per value and claim history. Market-wide premiums are collected via fictitious policy offers generating via Orlan insurance agency's platform (Orlanet Calculator). The sample consists of 876 distinct vehicle model-age -values for the top four insurers in the market (the insurer that provided the data and its three main competitors). For each one, I generate two observations: one without any claim in the last three years and one with one claim in the last three years, which occurred last year. Note that the Orlanet Calculator does not generate policy offers for the case of at least two claims in the last three years. Columns 1 through 4 present the relationship between claim history and premium per value offered by the top four insurers in the market, while column 5 describes the relationship with regard to the minimum premium per value in the market (not restricted to the four insures). Premiums and values are measured in New Israeli Shekel (ILS). Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Summary statistics of the Go—No Go grading

	(1) All	(2) Go	(3) Inc. Ded.	(4) Inc. Prem.	(5) TP Only	(6) Deny
Policies	14,288	12,478	625	386	258	541
Share	100%	87.33%	4.37%	2.70%	1.81%	3.79%
Premium	9,028	8,965	8,762	10,809	7,251	10,368
Policies With Claim	15.39%	12.29%	36.32%	37.82%	20.54%	44.18%
Damage	4,320	2,558	9,635	12,208	2,924	33,845
% Loss Ratio	47.85%	28.53%	109.96%	112.94%	40.32%	326.45%
Vehicle Age	4.16	4.02	4.39	4.12	9.80	4.35
Vehicle Value	260,601	261,628	242,907	305,377	167,203	269,947

Notes: The table reports summary statistics of insurer grading for all comprehensive and partial coverage policies between 2018 and 2020. The first column reports statistics for all policies, the second column describe the statistics for policies that received a Go grade. Columns 3 through 6 describe the statistics for policies that received a No-Go grade. Column 3 describe the statistics for policies that the operational team recommends a change in terms without increasing premiums (increase deductibles), column 4 describes the statistics for policies that the operational team recommends a price increase, column 5 describes the statistics for policies that the operational team recommends to offer only third-party coverage (i.e., not to provide comprehensive coverage), and column 6 describes the statistics for policies that the operational team recommends denying. Variables are defined as in Table 1.

Table 6: Probability of a Go grade

	Probit Model. Dependent Variable: $Go = 1$			
	(1)	(2)	(3)	(4)
Age 2-4	-0.023		-0.020	-0.024
	(0.014)		(0.014)	(0.017)
Age 5-7	-0.026*		-0.023*	-0.026
	(0.014)		(0.013)	(0.017)
Age 8+	-0.093***		-0.086***	-0.094***
	(0.017)		(0.016)	(0.020)
Client's Aggregate Loss Ratio		-0.032***	-0.032***	-0.072***
		(0.010)	(0.010)	(0.014)
Client's Prev. Yr. Loss Ratio		-0.025***	-0.024***	-0.026***
		(0.005)	(0.005)	(0.007)
Fleet Size	Y	Y	Y	Y
Vehicle Type	Y	Y	Y	Y
Vehicle Weight Class	Y	Y	Y	Y
Driver Underage Indicator	Y	Y	Y	Y
Sample	All	All	All	History ≥ 5
Observations	14,288	14,288	14,288	9,586
Pseudo R-squared	0.03	0.12	0.13	0.12

Notes: The table reports the relationship between the previous claim history, vehicle, and assignment of a Go grade. The sample includes all comprehensive and partial coverage policies with an assigned insurer grading between 2018 and 2020. The dependent variable is equal if the operational team assigned the policy with a Go grade. The explanatory variables consist of three age variables: (i) Age 2-4, a dummy variable that equals one if the vehicle age is between 2 and 4, (i) Age 5-7, a dummy variable that equals one if the vehicle age is between 5 and 7 and (i) Age $8+$, a dummy variable that equals one if the vehicle age is 8, or above. There are two explanatory variables with regard to claim history: (i) client's aggregate loss ratio, which is the client's ratio of total damages (starting in 2013) per total revenue (starting in 2013) and (ii) client's previous year loss ratio is the client's ratio of previous year total damages (net claim expenses) over the previous year total revenue (paid premiums). Both are measured in New Israeli Shekel (ILS). Controls include fleet size of client, which is defined as the number of vehicles insured by the client at a given year, vehicle type, vehicle weight class and an indicator for permitted underage driver. The estimation is conducted using a probit model. Coefficients reported are marginal effect at mean. Columns 1 through 3 include the entire sample, while column 4 includes the sum sample of policies with at least 5 years of recorded history (starting in 2013). History is measured as the sum of the years each of the client's policies are observed. Robust standard errors, clustered at client level, are reported in parentheses. *** p<0.01, ** p<0.05, *p<0.1.

Table 7: Structural estimation - demand side

	Damage per Value		Renew Comp. Coverage	
Premium per Value			-162.991***	(8.834)
License Avg. Premium per Value			123.330***	(7.288)
Vehicle Value (in 100,000 ILS)			0.124***	(0.013)
Price Increase			-0.718***	(0.032)
Age groups:				
0-1	(omitted)		(omitted)	
2-4	0.219***	(0.070)	-0.594***	(0.037)
5-7	0.522***	(0.073)	-1.188***	(0.045)
≥ 8	0.725***	(0.074)	-1.423***	(0.049)
Client's Aggregate Loss Ratio:				
$[0.5, 1)$	0.348***	(0.070)	-0.237***	(0.049)
$[1, 2)$	0.436***	(0.089)	-0.681***	(0.066)
≥ 2	0.600***	(0.147)	-0.723***	(0.109)
Client's Prev. Yr. Loss Ratio:				
$[0.5, 1)$	0.108	(0.069)	0.230***	(0.034)
$[1, 2)$	-0.046	(0.085)	0.334***	(0.042)
≥ 2	-0.005	(0.132)	0.418***	(0.072)
$\mathbb{1}(\text{Claim}_{i\ell t-1}) \geq 1$	0.311***	(0.058)	0.652***	(0.031)
Comp. Coverage	0.202*	(0.112)	0.332***	(0.056)
Fleet Size	-0.0004**	(0.0001)	-0.0019**	(0.0002)
Underage Driver	0.056	(0.076)	0.081*	(0.044)
Joined last yr.	0.196***	(0.071)	0.433**	(0.204)
History (in 1,000 yrs)	-0.017	(0.011)	0.001	(0.013)
Selection	1.079***	(0.284)		
Client unobs. s.e.	0.061	(0.056)	1.741	(0.041)
Observations	91,603		73,171	
Log Likelihood	-8696		-33756	

Notes: The table reports the results of the structural estimation of the demand for insurance. The left column presents the main estimates of damage per value. The right panel presents the main estimates of renewal. Joined last year is an indicator that equals to 1 if the client purchased its first policy from the insurer in the last year. Client's aggregate loss ratio is the client's ratio of total damages (starting in 2013) per total revenue (starting in 2013). Client's previous year loss ratio is the client's ratio of previous year total damages (net claim expenses) over the previous year total revenue (paid premiums). $\mathbb{1}(\text{Claim}_{i\ell t-1}) \geq 1$ is an indicator that equals one if at least one claim has been reported with regard to the license in the previous period. Price increase is an indicator that equals one if the policy faces an increase in premium per value. Selection measures the relationship between the unobserved client-level damage component and the unobserved client-level demand component. Premium, damage, and vehicle value are measured in New Israeli Shekel (ILS). Client unobserved s.e. measures the magnitude of heterogeneity (standard errors) in unobserved client-level damage and demand components. *** p<0.01, ** p<0.05, *p<0.1.

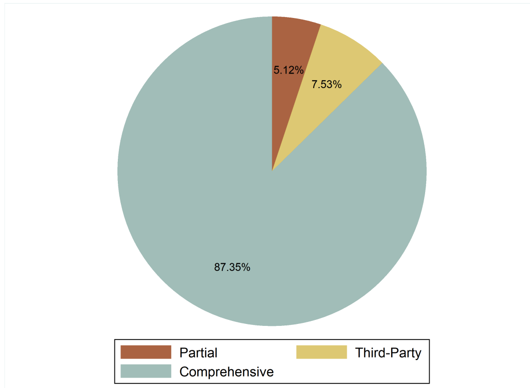
Table 8: Structural estimation - supply side

	Go		Adjust Terms		Deny
$\log(p_i)$	1.223***	(0.461)	1.549***	(0.568)	
$\log(p_i) - \log(\bar{p}(x))$	-1.555***	(0.531)	-2.566***	(0.652)	
Age groups:					
0-1	(omitted)		(omitted)		
2-4	-0.015	(0.209)	0.074	(0.245)	
5-7	-0.282	(0.244)	0.427	(0.280)	
≥ 8	-2.297***	(0.267)	-2.457***	(0.351)	
Client's Aggregate Loss Ratio:					
$[0.5, 1)$	-0.552**	(0.220)	0.258	(0.267)	
$[1, 2)$	-1.393***	(0.204)	-0.205	(0.258)	
≥ 2	-2.476***	(0.220)	-0.984***	(0.283)	
Client's Prev. Yr. Loss Ratio:					
$[0.5, 1)$	-0.964***	(0.286)	0.341	(0.335)	
$[1, 2)$	-1.211***	(0.261)	0.232	(0.309)	
≥ 2	-1.544***	(0.230)	-0.098	(0.285)	
$\mathbb{1}(\text{Claim}_{it-1}) \geq 1$	-0.466***	(0.175)	0.326	(0.209)	
Comp. Coverage	0.051	(0.206)	0.115	(0.270)	
Fleet Size	0.035	(0.070)	0.184**	(0.086)	
Underage Driver	-0.501*	(0.272)	-0.694**	(0.338)	
New Client	0.196	(0.171)	0.433**	(0.204)	
History	-0.017	(0.011)	0.001	(0.013)	
Constant	8.943***	(1.659)	4.818**	(2.054)	0
Observations	6,347				
Log Likelihood	-1954.2				
Pseudo R-squared	0.266				

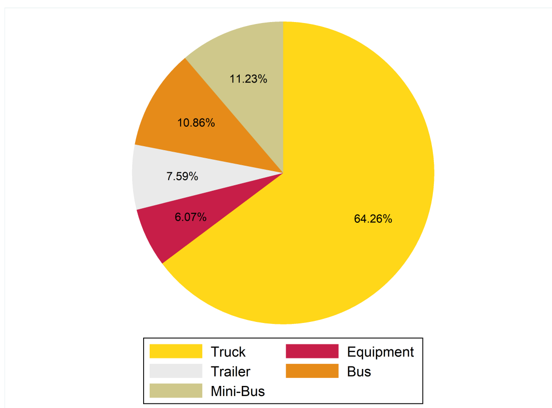
Notes: The table reports the relationship between the policy's observable characteristics and insurer grading. The sample includes all comprehensive and partial coverage policies with an assigned insurer grading between 2018 and 2020 for nonfleet policies (i.e., policies for clients with a fleet size below 5 during the relevant year). The insurer's alternatives are (i) "Go", which means the operational team recommends renewing policy with the same terms, (ii) "Adjust", which means the operational team recommends offering a policy with increased premiums or deductibles, and (iii) "Deny", which means the operational team recommends denying comprehensive coverage. The explanatory variables consist of vehicle age variables, client's aggregate loss ratio, and client's previous year loss ratio, as defined in Table 6. $\log(p_i)$ is the policy's log premium per value and $\log(p_i) - \log(\bar{p}(x))$ is the difference between the policy's log premium per value and the log average premium per value paid for a policy with the same observable characteristics. Additional explanatory variables include, $\mathbb{1}(\text{Claim}_{it-1}) \geq 1$, an indicator for a claim event at the policy level in the previous period, comprehensive coverage dummy variable, underage driver indicator, fleet size, history of client with insurer, and new client indicator (joined last year). Estimation includes controls for vehicle type and weight class and year. History and vehicle age are measured in years. Premiums, damages, and vehicle values are measured in New Israeli Shekel (ILS). Estimation is conducted using a multinominal logistic regression model. Analytical asymptotic standard errors are reported in parentheses. *** p<0.01, ** p<0.05, *p<0.1.

A Additional Figures and Tables

Figure A.1: Policies by coverage type and vehicle type



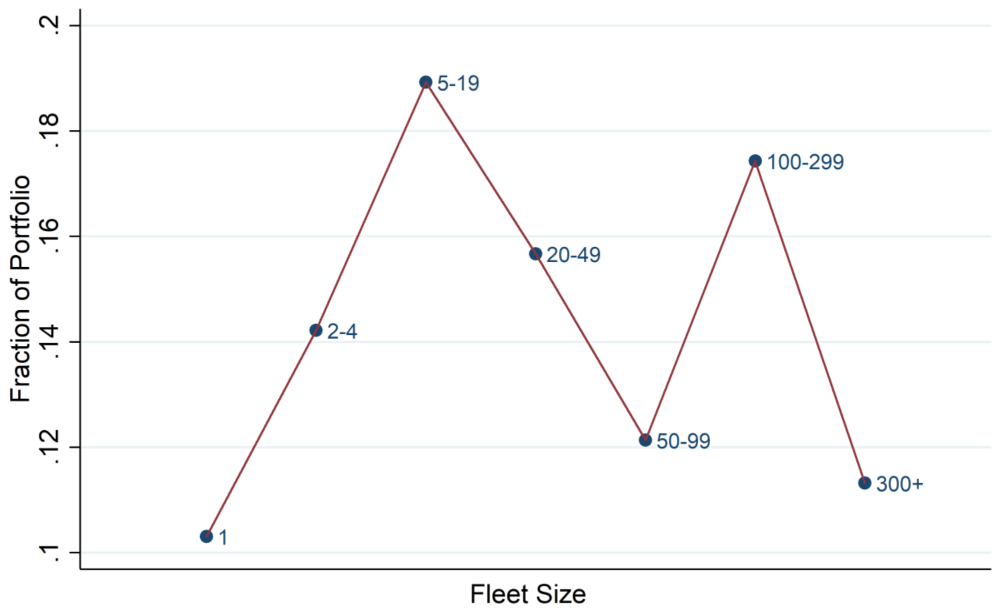
Panel A: Coverage type



Panel B: Vehicle type

Notes: The figure depicts the distribution of policies by coverage type (Panel A) and vehicle type (Panel B). The sample include all insurance policies for all vehicles from 2013 through 2020. Policies are weighted by premiums, measured in New Israeli Shekel (ILS).

Figure A.2: Distribution of policies by client's fleet size



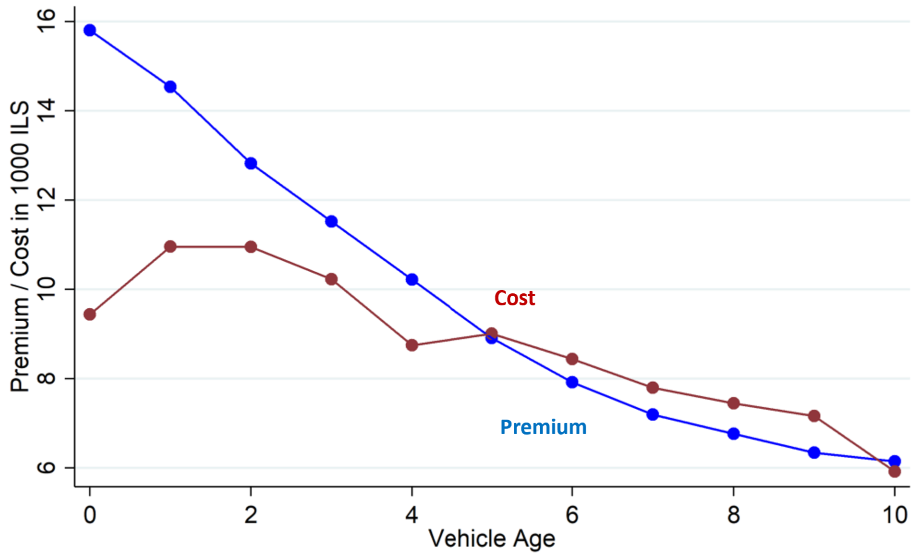
Notes: The figure reports the distribution of all insurance policies from 2013 to 2020 by client's fleet size. Client's fleet size is defined by the number of total insurance policies purchased by the client in a given year. The distribution of policies is weighted by policy premiums, which are measured in New Israeli Shekel (ILS).

Figure A.3: Example of a "Go—No Go" grade document

	הערות	תאריך סוף פוליסה	תאריך תחילת פוליסה	פוליסה
Check	✓	2/29/2020	3/1/2019	390342117419
Same as last year	חידוש ללא שינוי	2/29/2020	3/1/2019	390344197219
Do not decrease	אין לרדת מתעריף\תנאים קיימים	2/29/2020	3/1/2019	390344983119
Increase third party deductible to 7000	יש להעלות אקסס צד ג' ל-7,000 ₪	2/29/2020	3/1/2019	390342188019
Increase premiums by 7.5%	יש להעלות תעריף ב-7.5% מאשתקד	2/29/2020	3/1/2019	390880039719
Offer third party coverage only	צד ג' בלבד	2/29/2020	3/1/2019	390343082519
Due to claims do not renew	לאור מצב תביעות לא ניתן לחדש באמצעותנו	2/29/2020	3/1/2019	390342820819

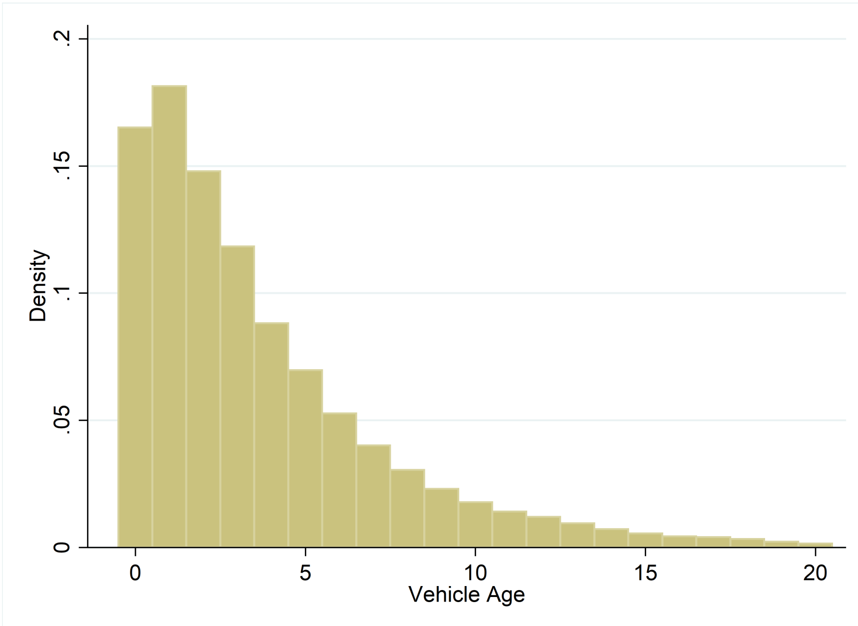
Notes: The figure reports an example of a "Go—No Go" grade document. The first column (from right) indicates the policy id number. The second column is the date at which the policy began. The third column is the end date of the policy. The fourth column is the "Go—No Go" grade. In the fifth row I provide a translation of a "Go—No Go" grade from Hebrew. The top three rows are policies that received a "Go" grade, while the bottom four rows are policies that received a "No-Go" grade.

Figure A.4: Summary statistics of premium and cost in nominal value by vehicle age



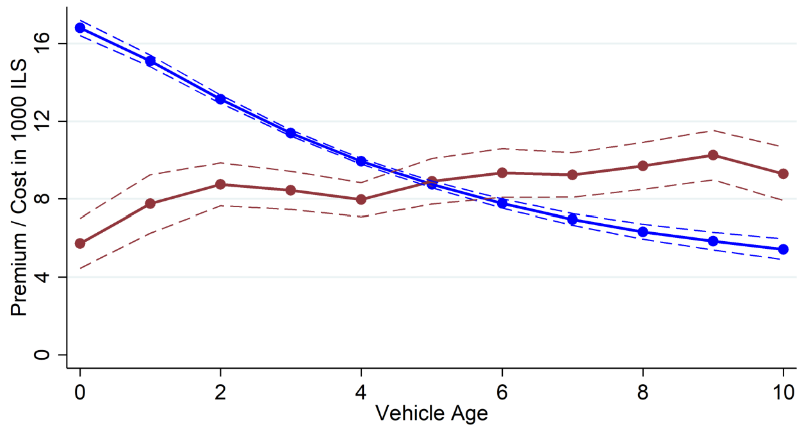
Notes: The figure depicts the premium (in blue) and cost (in red) in nominal values for comprehensive coverage policies for trucks from 2013 to 2020. The vertical axis depicts premiums and costs in 1000 New Israeli Shekel (ILS). No controls are added. Both variables are standardized to an annual term policy. Vehicle values are measured in New Israeli Shekel (ILS).

Figure A.5: Distribution of trucks by vehicle age

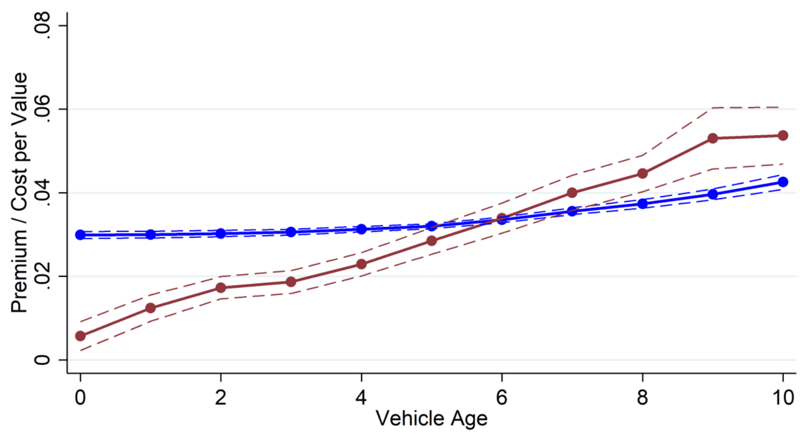


Notes: The figure depicts the distribution of comprehensive coverage policies for trucks by vehicle age. The sample includes insurance policies for trucks from 2013 through 2020. Vehicle age is measured in years.

Figure A.6: Alternative specifications of premium and costs by vehicle age



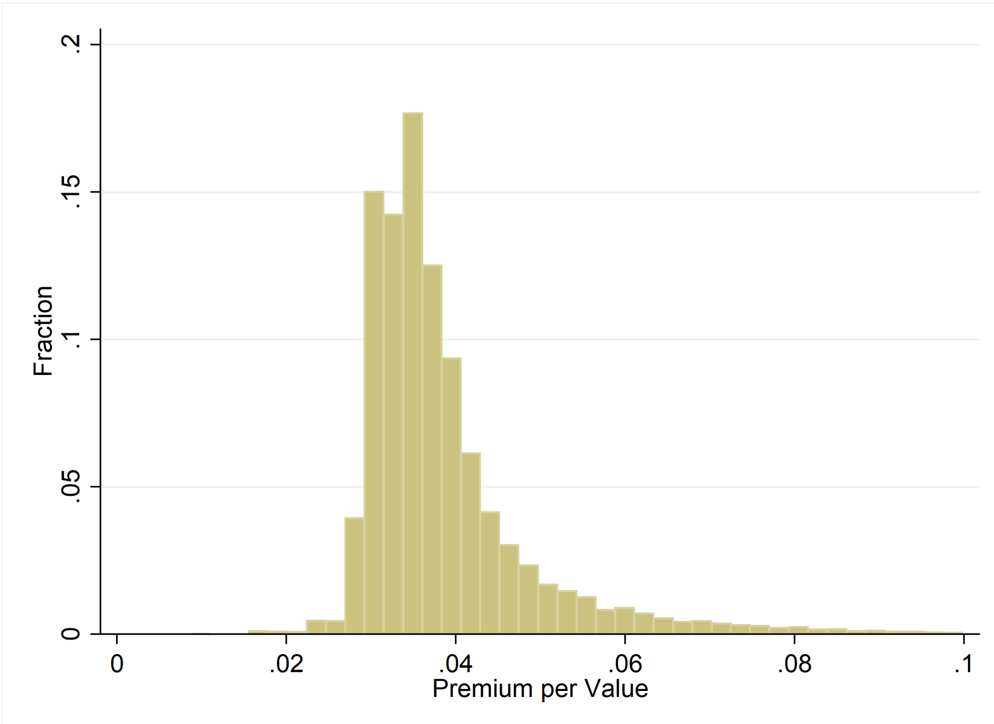
Panel A: Nominal Values



Panel B: All Vehicle Type

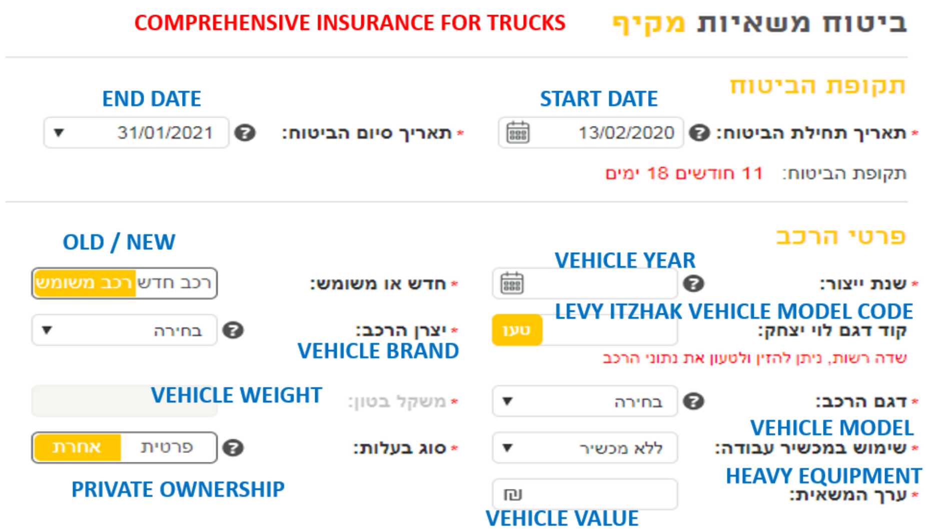
Notes: The figure depicts a variation of the analysis conducted presented in Figure 3. Panel A depicts premiums (in blue) and costs (in red) in nominal values, instead of normalized by vehicle value. The vertical axis is measured in 1,000 ILS. Panel B depicts a model identical to that of Figure 3, but includes all vehicles in the sample, rather than only trucks.

Figure A.7: Distribution of premium per value

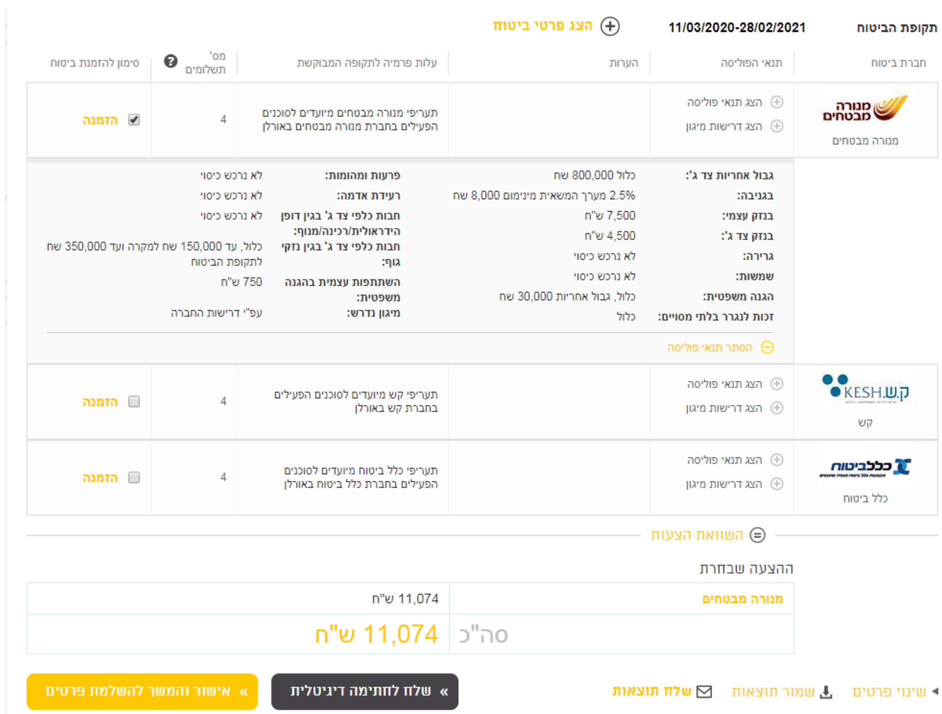


Notes: The figure depicts the distribution of premium per value paid for trucks with comprehensive coverage policies. The sample includes all trucks with comprehensive coverage from 2013 through 2020. Premiums, which are normalized to an annual policy length, and vehicle values are measured in New Israeli Shekel (ILS).

Figure A.8: Orlanet Calculator



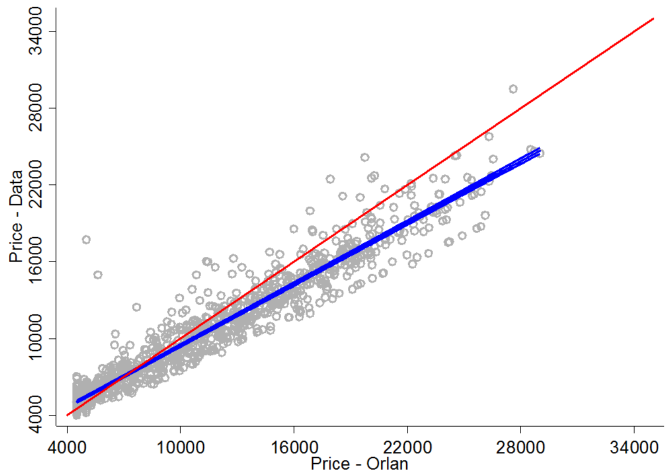
Panel A: Inputs



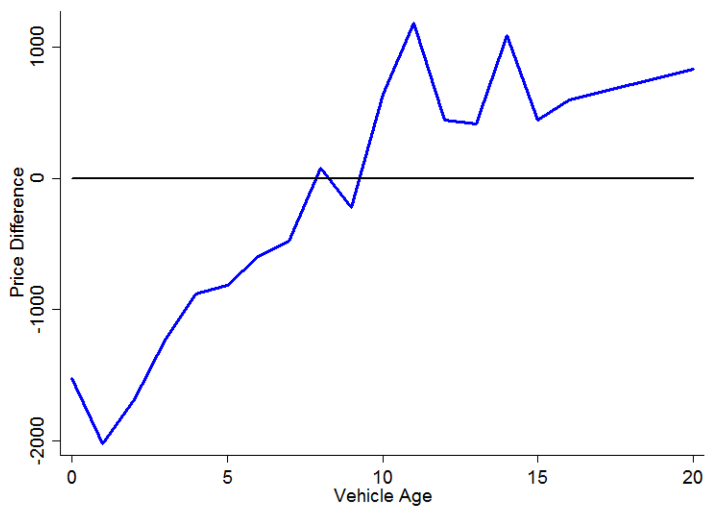
Panel B: Outputs

Notes: The figure depicts the process of generating fictitious comprehensive policy coverage for trucks using the *Orlanet Calculator*. Panel A describe the input process, and panel B illustrates the outputs.

Figure A.9: Premium comparison of Orlan pricing of insurer vs. actual premiums



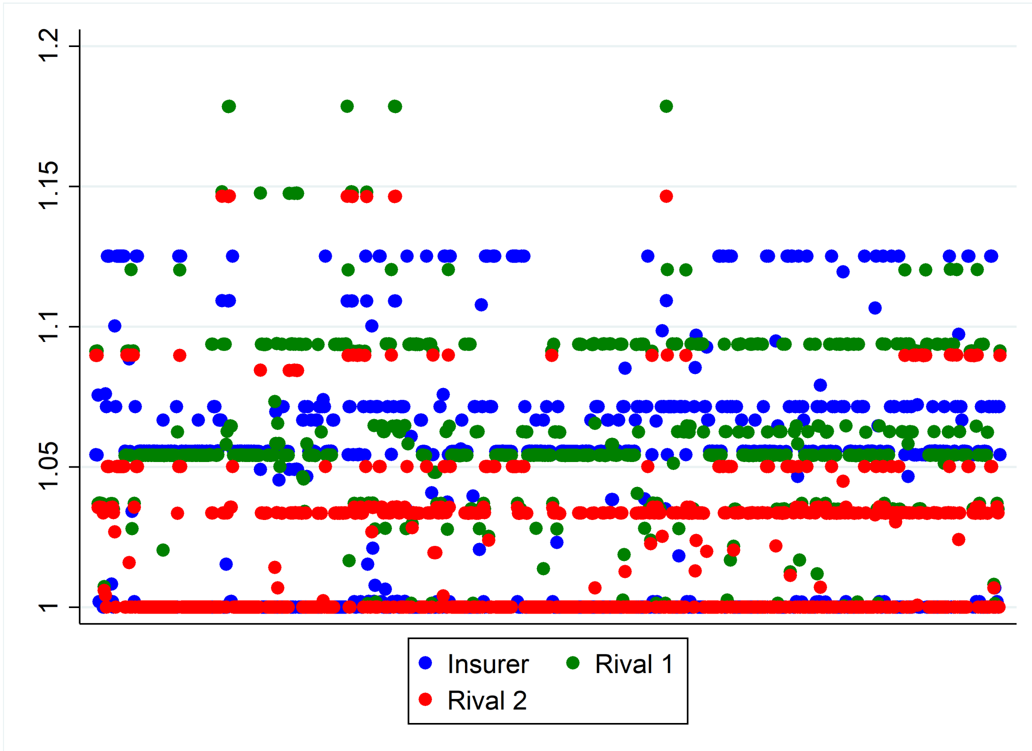
Panel A: Orlan and Actual Pricing



Panel B: Difference by Vehicle Age

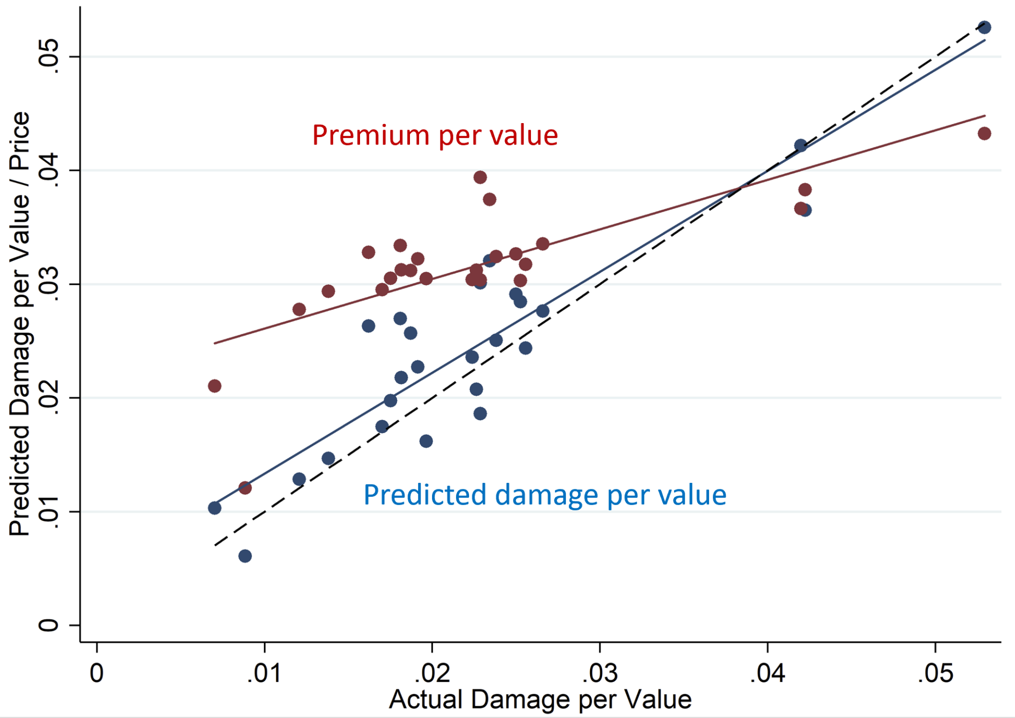
Notes: The figure in Panel A presents scatter and regression coefficients and 95% confidence interval of a within-insurer comparison in order to validate that the Orlanet Calculator pricing offers match the data provided by the insurer. A total of 2,041 observations are included. For each observation, I calculate the Orlan pricing using the average of both no claims in the last 3 years and one claim in the last 3 years, which occurred last year. The estimated slope equals 0.80 (0.01). R-square = 0.90. The red curve is the 45-degree line. Prices (premiums) are measured in New Israeli Shekel (ILS). The figure in Panel B depicts the mean difference between Orlan premiums and actual premiums by vehicle age.

Figure A.10: Market-wide premiums—Insurer and Rivals 1 and 2



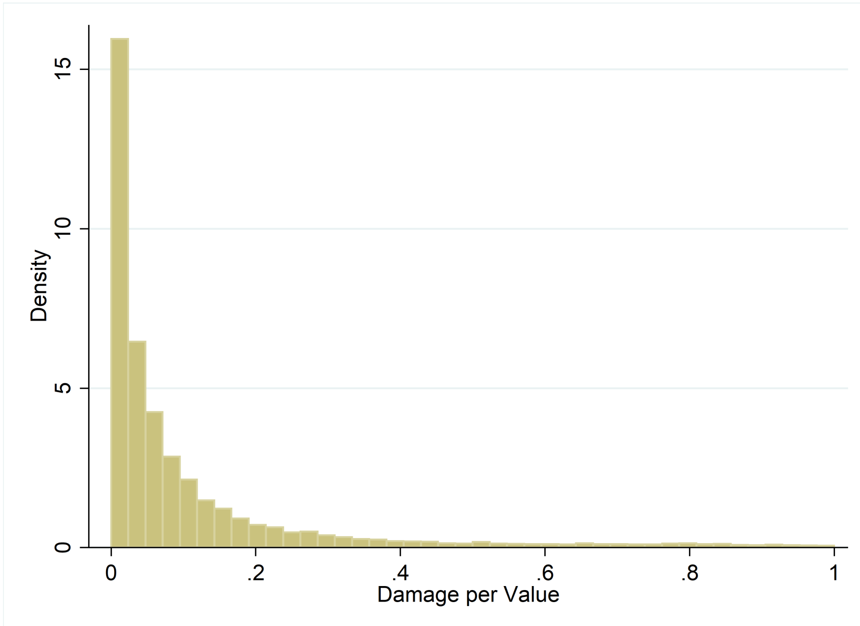
Notes: The figure reports the distribution of premiums offered by the insurer and its two main competitors. Premiums are calculated using the Orlanet Calculator. The horizontal axis depicts all 876 observations with distinct vehicle model-age value characteristics. I use the premium offered for the case of no claim in the last three years. The vertical axis depicts the premiums charged by each insurer, scaled by the lowest premium offered by the three competitors. The lowest premium offered is normalized to one. Prices (premiums) and vehicle values are measured in New Israeli Shekel (ILS).

Figure A.11: Out of sample prediction of costs



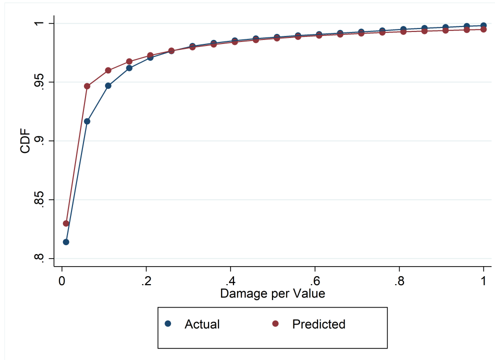
Notes: The figure reports the relationship between premium per value, predicted damage per value, and actual damage per value. Using data on policies from 2014 to 2018, I estimate a cost function (using regression analysis) by the following observable vehicle characteristics (age, value weight class, and type) and claim history (aggregate loss ratio). Based on the cost estimates, I divide my 2019–2020 sample into 25 groups based on projected damage per value. The vertical axis depicts predicted damage per value (in blue) and premium per value (in red). The horizontal axis depicts actual damage per value. The horizontal axis depicts the actual damage per value. The solid blue and red lines represent the regression coefficients of actual damage per value on predicted damage per value and premium per value, respectively. The dashed line is the 45-degree line. Premiums, damages, and vehicle values are measured in New Israeli Shekel (ILS).

Figure A.12: Distribution of conditional damage per value

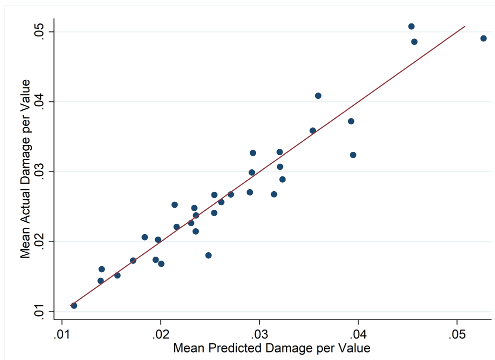


Notes: The figure reports the distribution of conditional damage per value; that is, damage per value if at least one claim occurred for comprehensive and partial coverage policies for all vehicles from 2013 to 2020. Damages and vehicle values are measured in New Israeli Shekel (ILS).

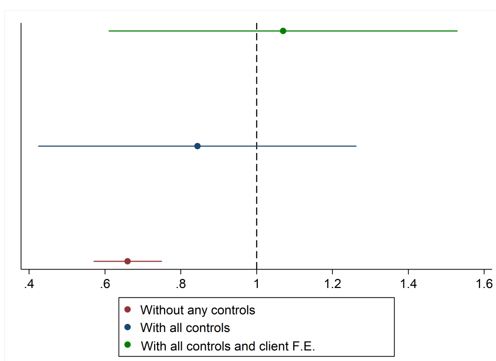
Figure A.13: Model fit



Panel A: Distribution of damage per value



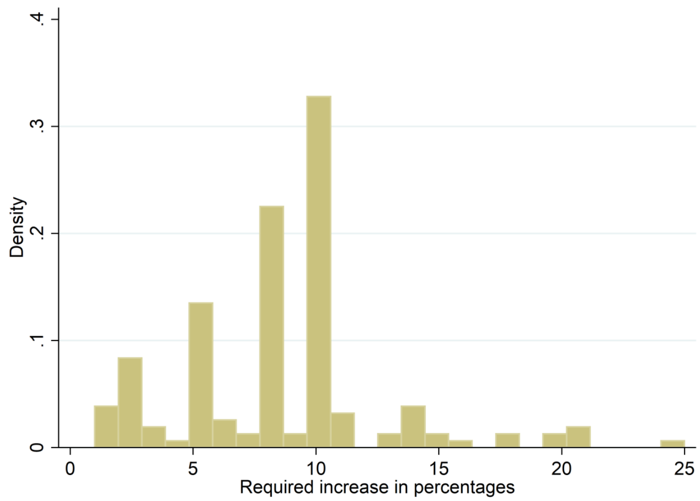
Panel B: Avg. damage per value by groups



Panel C: Relationship between predicted and actual renewal

Notes: The figure reports the model fit. Panel A depicts the predicted and actual distribution of damage per value. Panel B depicts the relationship between predicted and actual mean damage per value. The sample is divided into groups based on the category variables vehicle age, aggregate client loss ratio, client's previous year loss ratio, and new client indicator. Figure presents group with at least 300 observations. Panel C depicts the relationship between the predicted probability of renewal and the share of realized renewal. Controls are variables included in demand estimation.

Figure A.14: Distribution of recommended increase in premium per value



Notes: The figure reports the recommended increase in premium per value, for all policies for which the analytical team recommends a price increase. Premiums and values are measured in New Israeli Shekel (ILS).

Table A.1: Summary statistics of comprehensive coverage policies for all vehicles

	All	Vehicle Age≥ 6	$\text{Claim}_{t-1} \geq 1$
Policies	102,372	33,378	12,347
Share (by Premium)	100%	32.60%	12.06%
Weighted Share (by Premium)	100%	19.13%	13.95%
Mean Premium	7,463	4,710	8,635
At least 1 claim	18.26%	16.55%	30.63%
Mean Damage	4,998	4,021	8,062
Mean Commission	998	668	1,027
Mean Profit	1,467	21	-1,084
Profit Margin	19.65%	0.45%	-5.26%
Mean Vehicle Age	4.78	10.22	5.08
Mean Vehicle Value	270,456	151,123	279,721
Mean Premium per Value	2.76%	3.12%	3.09%

Notes: The table is a replication of Table 1, consisting of comprehensive coverage policies for all vehicle types.

Table A.2: Policy outcomes and past performance - all vehicles

	Panel A: Entire Sample			
	(1)	(2)	(3)	(4)
	$\text{Claim}_t \geq 1$	$\frac{\text{Damage}_t}{\text{Value}_t}$	$\%\Delta\frac{\text{Premium}_t}{\text{Value}_t}$	Loss Ratio_t
$\text{Claim}_{t-1} \geq 1$	0.103***	0.013***	-0.003	0.350***
	(0.007)	(0.002)	(0.002)	(0.038)
log(Value)	Y	N	N	Y
Vehicle Age - 2^{nd} order	Y	Y	Y	Y
Vehicle Type	Y	Y	Y	Y
Vehicle Weight Class	Y	Y	Y	Y
Driver Underage Indicator	Y	Y	Y	Y
Observations	65,031	65,031	65,031	65,031
R-squared	0.051	0.009	0.023	0.003
	Panel B: Non-Fleet Policies			
	(1)	(2)	(3)	(4)
	$\text{Claim}_t \geq 1$	$\frac{\text{Damage}_t}{\text{Value}_t}$	$\%\Delta\frac{\text{Premium}_t}{\text{Value}_t}$	Loss Ratio_t
$\text{Claim}_{t-1} \geq 1$	0.104***	0.019***	0.009***	0.463***
	(0.011)	(0.004)	(0.004)	(0.094)
log(Value)	Y	N	N	Y
Vehicle Age - 2^{nd} order	Y	Y	Y	Y
Vehicle Type	Y	Y	Y	Y
Vehicle Weight Class	Y	Y	Y	Y
Driver Underage Indicator	Y	Y	Y	Y
Observations	12,715	12,715	12,715	12,715
R-squared	0.029	0.009	0.029	0.005

Notes: The table reports the results of the estimation model presented in Table 2, with regard to all vehicle types.

Table A.3: Summary statistics by customer classification

	Drop	Keep
Policies	74,456	1,102
Customers	3,170	166
Mean Premium	5,843	5,274
Mean Damage	3,658	5,252
Mean Commission	768	752
Mean Profit	1,417	-730
Profit Margin	24.26%	-13.83%
Loss Ratio	62.60%	99.58%

Notes: The table reports summary statistics for all policies from 2016 to 2020. The policies are classified into two groups. "Drop" includes policies of clients that incurred a loss ratio of at least 2 between 2013 and 2015 and their average vehicle age is at least 5. "Keep" includes policies of clients that incurred a loss ratio of at least 2 between 2013 and 2015 or their average vehicle age is below 5. Profit margin is defined as mean profit (premium-damage-commission) over mean premium. Loss ratio is defined as the mean damage of customers' claims (net of deductibles) over mean premiums. Premiums, commissions, damages, and profits are measured in New Israeli Shekel (ILS).

Table A.4: Summary statistics of policies of graded customers vs. rest

	All	Graded Customer	Not
Policies	109,630	55,868	53,762
Share (by Premium)	100%	50.96%	49.04%
Weighted Share (by Premium)	100%	56.27%	43.73%
Mean Premium	7,371	8,138	6,573
Mean Damage	4,973	4,882	5,067
Mean Commission	965	1,269	650
Mean Profit	1,432	1,988	855
Profit Margin	19.43%	24.42%	13.01%
Loss Ratio	67.47%	59.98%	77.10%
% Comprehensive	93.38%	95.20%	91.49%
Mean Vehicle Age	4.83	4.95	4.72
Mean Vehicle Value	275,594	268,580	282,883
Mean Premium per Value	2.67%	3.03%	2.32%

Notes: The table reports summary statistic to all comprehensive and partial coverage policies for all vehicle types between 2013 and 2020. The first column reports statistics for all policies, the second column describes the statistics for a sub-sample of the data consisting of all clients for whom at least one of their policies has a documented grade. Column 3 describes all other policies with regard to the other customers. Profit margin is defined as mean profit (premium-damage-commission) over mean premium. Loss ratio is defined as mean damage of customers' claims (net of deductibles) over mean premiums. Vehicle value, premium, commission, damages, and profits are measured in New Israeli Shekel (ILS). Vehicle age is measured in years. I exclude from the sample observation with an error, change in vehicle within the policy, change in coverage terms over the policy and policies that did not end, or that lasted for less than 30 days (without a claim).

Table A.5: Insurer grading and policy renewal

	(1) $\Delta\frac{\text{Premium}}{\text{Value}}$	(2) ΔDed. TP	(3) $\Delta\frac{\text{Ded. Own}}{\text{Value}}$	(4) Third Party	(5) Renew
Recommendation:					
$\Delta\frac{\text{Premium}}{\text{Value}}$	0.8550**	-2.4199	0.4113*		
	(0.3496)	(18.210)	(0.2203)		
ΔDed. TP	-0.0003	0.7162***	0.0009		
	(0.0003)	(0.1135)	(0.0007)		
$\Delta\frac{\text{Ded. Own}}{\text{Value}}$	-0.0015	0.2434	1.0853***		
	(0.0022)	(0.2588)	(0.0080)		
Go	0.0015	0.1290	0.0016		
	(0.0009)	(0.1513)	(0.0015)		
Increase Ded. Only	0.0028**			-0.0273*	0.0310
	(0.0011)			(0.0159)	(0.0458)
Increase Premium		0.0822	-0.0005	-0.0155	-0.0223
		(0.1418)	(0.0016)	(0.0159)	(0.0506)
TP Only				0.1487***	0.0061
				(0.0493)	(0.0416)
Deny				0.0663	-0.6703***
				(0.0572)	(0.0205)
Vehicle Age Included	Y	Y	Y	Y	Y
Observations	8,652	8,652	8,652	10,080	14,282
R-squared	0.0630	0.1753	0.4894	0.1179	0.0822

Notes: The table reports the relationship between the analytical team's recommendation and the terms of renewal. The dependent variables are the change in premium per value (column 1), the change in deductible with regard to third-party property damage (column 2), the change in deductible with regard to own-property damage, normalized by vehicle value (column 3), an indicator as to whether the policy has been renewed with only third-party coverage (column 4), and whether the policy has been renewed at all (column 5). The explanatory variables, in order, are the recommended change in premium per value, the recommended change in deductible with regard to third-party property damage, the recommended change in deductible with regard to own property damage (normalized by vehicle value), an indicator as to whether the policy received a "Go" grade, an indicator as to whether the analytical team recommends to not renew policy at current terms but does not require a premium increase ("Increase Ded. Only"), an indicator as to whether the analytical team recommends not renewing the policy at the current terms and requires a premium increase ("Increase Premiums"), an indicator as to whether the analytical team recommends not renewing the policy with comprehensive coverage ("TP Only"), and an indicator as to whether the analytical team recommends denying coverage from customer ("Deny"). Premiums and deductibles are measured in New Israeli Shekel (ILS). Changes in third-party deductibles are measured in 1,000 ILS. Controls include vehicle age (saturated control). Robust standard errors are reported in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A.6: First stage prediction of log value based on previous year log value

Vehicle type	Coeff.	s.e.	R^2	Obs.
Truck	0.895	(0.001)	0.98	39,599
Heavy eq.	0.936	(0.002)	0.97	5,267
Trailer	0.923	(0.002)	0.90	17,058
Bus	0.898	(0.001)	0.99	7,378
Mini-bus	0.836	(0.002)	0.95	9,665
Heavy eq. add-on	0.929	(0.004)	0.92	4,270

Notes: The table reports the results of the first stage estimation: prediction of log value based on previous year log value. The vehicles' values are estimated separately for the five vehicle type groups. Furthermore, the values of heavy equipment add-ons are estimated separately. Vehicle values are measured in 100,000 New Israeli Shekels (ILS). History is measured in 1,000 years.

Table A.7: First stage prediction of log premium

IIC min price	0.496***	(0.056)	Claim Last Yr. (Vehicle)	-0.006***	(0.001)
Aggregate Loss Ratio:			Previous Yr. Loss Ratio:		
< 0.5	(omitted)		< 0.5	(omitted)	
≥ 0.5 and < 1	-0.003**	(0.002)	≥ 0.5 and < 1	0.002	(0.002)
≥ 1 and < 2	-0.004*	(0.002)	≥ 1 and < 2	0.021***	(0.003)
≥ 2	0.003	(0.004)	≥ 2	0.021***	(0.005)
Vehicle Age:					
0-1	(omitted)		log(Value)	0.222***	(0.013)
2-4	-0.017***	(0.001)	log(Total Value)	0.446***	(0.013)
5-7	-0.068***	(0.002)	Partial Coverage	-0.240***	(0.008)
8+	-0.065***	(0.003)	Underage driver	0.017***	(0.004)
			Joined last year	0.004**	(0.002)
History	-0.023	(0.002)	Joined 2 years ago	-0.004**	(0.001)
Observations	101,019				
R-squared	0.978				
Within R-squared	0.662				

Notes: The table reports the results of the first-stage estimation: prediction of log premium per value. The regression includes client and license fixed effects, in addition for year, vehicle type, and vehicle weight dummies. IIC min price is the international insurance company minimum price, based on observable characteristics. The sample excludes third-party coverage policies and 2013 policies. Premiums are measured in New Israeli Shekel (ILS) and vehicle value is measured in 100,000 New Israeli Shekels (ILS). History is measured in 1,000 years. *** p<0.01, ** p<0.05, * p<0.1.

B Estimation Details

B.1 Estimation Details

I discuss the joint estimation of the parameters of the demand model: consumers' willingness to pay for policy renewal and expected damage per value of policy. Estimation is conducted in three steps. First, since I do not observe vehicle values for policies that are not renewed, I predict the (log) vehicle value as a function of previous period value and vehicle type, and treat it as data; both for those that renewed and those that did not. Then, I generate the predicted premium per value and predicted increase in premium per value using vehicle and client covariates and license fixed effect. Lastly, I jointly estimate Equations B.1 and B.2 via Maximum Simulated Likelihood. In this section, I describe in detail the Maximum Simulated Likelihood estimation process.

The parameter of interest are the expected damage per value and policy renewal equations.

$$d_{ijt} = \exp(\delta x_{ijt} + \log(\tau_{ijt}) + \nu_i) \tag{B.1}$$

$$\Pr(R_{ijt} = 1 | x_{ijt}, d_{ijt}, \hat{p}_{ijt}, \hat{\mathcal{I}}_{ijt}, \hat{f}_{ijt}, \omega_i) = \frac{\exp(-\alpha \hat{p}_{ijt} - \lambda \hat{\mathcal{I}}_{ijt} + \beta x_{ijt} + \gamma d_{ijt} + \rho \hat{f}_{ijt} + \omega_i)}{1 + \exp(-\alpha \hat{p}_{ijt} - \lambda \hat{\mathcal{I}}_{ijt} + \beta x_{ijt} + \gamma d_{ijt} + \rho \hat{f}_{ijt} + \omega_i)} \tag{B.2}$$

where ν_i represents customers' private information regarding risk, which is constant over time, is normally distributed, $\nu_i \sim N(0, \sigma_\nu^2)$. ω_i is an exogenous unobserved client-level demand component, which is constant over time; ω_i is normally distributed, $\omega_i \sim N(0, \sigma_\omega^2)$. ν_i and ω_i are uncorrelated and independent of all other covariates. τ_{ijt} is the duration of the policy, which take into account the duration of each policy by estimating a pro-rated variant of the damage per value equation.

I start by describing the damage per value likelihood. Damage per value follows a pseudo-Poisson distribution defined by parameter (which also equals the expected value) described in Equation B.1. This choice has two implications. On the one hand, it accommodates both (i) the possibility of no damages at all, which occurs quite frequently, and (ii) possible dependence between the number of claims and the conditional damage of claims, (iii) implementing pro-rated policies is quite simple.

The log likelihood of observing damage per value D_{ijt} is given by:

$$\log(d_{ijt} = D_{ijt}|x_{ijt}, \tau_{ijt}, \nu_i) = D_{ijt}(\delta x_{ijt} + \ln(\tau_{ijt}) + \nu_i) - \exp(\delta x_{ijt} + \ln(\tau_{ijt}) + \nu_i) - \log(D_{ijt}!). \tag{B.3}$$

While $d_{ijt}!$ is not well-defined when it is a continuous number, it does not possess a challenge in estimation as eliminating the element does not change the maximum function. In order to calculate the implied probabilities for each event, I approximate $d_{ijt}!$ for continuous values using the Gamma alternative of Stirling's formula, which fits small values especially well.

With regards to renewal, the log likelihood of observing a renewal or not, denoted by $\mathcal{Y}_{ijt} = \{0, 1\}$ is given by Equation B.2 and can be re-expressed as:

$$\log Pr(R_{ijt} = \mathcal{Y}_{ijt}|x_{ijt}, d_{ijt}, \hat{p}_{ijt}, \hat{\mathcal{I}}_{ijt}, \hat{f}_{ijt}, \omega_i, \nu_i) = \mathcal{Y}_{ijt}\Big(- \alpha\hat{p}_{ijt} - \lambda\hat{\mathcal{I}}_{ijt} + \beta x_{ijt} + \gamma \exp(\delta x_{ijt} + \log(\tau_{ijt})$$
$$+ \nu_i) + \rho\hat{f}_{ijt} + \omega_i\Big) + \log(1 + \exp(-\alpha\hat{p}_{ijt} - \lambda\hat{\mathcal{I}}_{ijt} + \beta x_{ijt} + \gamma \exp(\delta x_{ijt} + \log(\tau_{ijt}) + \nu_i) + \rho\hat{f}_{ijt} + \omega_i))$$

Therefore, the joint log likelihood of observing for client i's damage per value D_{ijt} and renewal or non-renewal, $\mathcal{Y}_{ijt} = \{0, 1\}$ for each policy can be described as follows:

$$L_i = \sum_{jt \in \mathcal{J}_i} \int \int \log(d_{ijt} = D_{ijt}|x_{ijt}, \tau_{ijt}, \sigma_\nu \varepsilon_i^d)$$
$$+ \log Pr(R_{ijt} = \mathcal{Y}_{ijt}|x_{ijt}, d_{ijt}, \hat{p}_{ijt}, \hat{\mathcal{I}}_{ijt}, \hat{f}_{ijt}, \sigma_\omega \varepsilon_i^r, \sigma_\nu \varepsilon_i^d)\phi(\varepsilon_i^d)\phi(\varepsilon_i^r)d\varepsilon_i^d d\varepsilon_i^r$$

where $\mathcal{J}_i$ includes all of the policies (vehicle-period) of client i, and ϕ denotes standard Normal distribution.

In practice, I approximate both integral using 100 Halton draws for each unobserved term. 100 Halton draws achieve greater accuracy in some setting (for instance, mixed logit estimations) than 1,000 pseudo-random draws (Train (2000)). I follow the procedure as described in Train (2000). In order to take into account policies that are not up for renewal, I set the log renewal likelihood to zero.

The estimated parameters are the set of $\theta = (\alpha, \lambda, \beta, \gamma, \rho, \delta, \sigma_{nu}, \sigma_{omega})$ that maximize the likelihood for the N clients:

$$\theta = \arg\max \sum_{i=1}^{N} L_i$$