

Article

Optimal Market Share in Automobile Insurance Auction Markets

Manuel Rodríguez , Rolando Rubilar-Torrealba * , Cristóbal Fernandez-Robin , Diego Yáñez 
and Bernardo Pincheira

Departamento de Industrias, Universidad Técnica Federico Santa María, Valparaíso 2520000, Chile; cristobal.fernandez@usm.cl (C.F.-R.); diego.yanez@usm.cl (D.Y.); bernardo.pincheira@usm.cl (B.P.)

* Correspondence: rolando.rubilar@usm.cl

Abstract

The automobile insurance industry plays a pivotal role in the financial system, fostering economic stability through effective risk management and consumer confidence. Continuous enhancement in price optimisation not only ensures the sustainability of insurers but also fosters a more competitive, fair, and balanced market, which is vital for a country's economic development. The objective of this research is to develop a methodology for determining the optimal price offered by insurance firms for automobile policies in an industry where a First Price Sealed Bid auction system operates. A statistical methodology is employed to ascertain the expected value and standard deviation of the policies on offer in the public domain, whereby these values are calculated using a heteroskedastic linear regression estimation methodology. Furthermore, the aforementioned expected values and standard deviation enable the calculation of the value of the cumulative distribution for an optimal price set within the public offer. This study demonstrates that identifying the optimal price that maximizes profits is analogous to establishing an expected market share for each niche automobile policy market. Moreover, the market share can be calculated through a straightforward heteroskedastic linear regression estimation for instances where market shares are below 50%.

Keywords: insurance; auction; market share; policy market

MSC: 91B26; 91B24



Academic Editor: Kamer-Ainur Aivaz

Received: 5 January 2026

Revised: 6 February 2026

Accepted: 9 February 2026

Published: 11 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

The problem of insurance markets has been a topic of extensive investigation for several decades. The pioneering contributions of [1,2] illustrate the influence of moral hazard effects on the probabilities of the occurrence of events, highlighting the challenge of information asymmetries that can reduce market efficiency [3].

The insurance market problem has both microeconomic and macroeconomic implications due to the importance of the insurance industry in economic development. Research by [4–7] shows the effects of the insurance industry on economic growth and the sophistication of the economy. In other contexts, Ippolito [8] show the effects of regulating the insurance industry and [9,10] analyze how firm-level and macroeconomic variables affect profitability in the insurance industry. Furthermore, the evidence presented in [11] indicates that unemployment insurance can influence the entrepreneurial landscape by stimulating

the emergence of new ventures. These studies provide only a partial view of the influence of the insurance market on a range of economic activities.

The industrial organization of insurance markets has also been widely studied. For instance, Joskow [12] argues that deregulation can increase efficiency, particularly when there are many market participants. However, alternative research, as exemplified by the work of [13,14], indicates that the implementation of deposit regulations engenders enhanced financial system performance, including increased liquidity and risk taking, among other effects.

The global insurance market is a vast and dynamic sector that plays a crucial role in the world economy. It provides financial security and protection against various risks for individuals, companies, and organizations. The case study presented in [15] illustrates the manner in which financial risk is transferred from insurers to policyholders. The work of [16] examines the nexus between insurance and financial markets, identifying gaps in the existing insurance literature. In particular, it highlights the role of externalities arising from endogenous effects that result in underinsurance even for the riskiest individuals.

The insurance industry is a highly complex entity that comprises a multitude of subtypes. As illustrated in [17], there are considerable discrepancies between the various types of insurance, underscoring the need to understand the nuances of each sub-industry.

In the field of vehicle insurance, both theoretical and empirical research has been conducted to develop this body of knowledge. Theoretical research is exemplified by the works of [18,19], while empirical research is exemplified by [20]. The articles in this group present models focused on the insurance industry, with a particular emphasis on agents' decision-making processes and estimation methods that facilitate empirical analysis. The assumption made in [21] that the market for insurance companies approximates a competitive market is based on the premise that there are a large number of companies in the USA. However, the existence of smaller markets may result in a competitive market that differs from the norm and, therefore, requires an alternative paradigm of industrial organization to be applied.

As stated by [22], the capacity to calculate suitable premiums is fundamental to the adaptation of insurance policies to the inherent risks. This requires the use of a sophisticated statistical modeling approach, enabling the identification of patterns and trends within historical claim data. The analysis of such data is pivotal to the calibration of premiums, ensuring the fair and accurate representation of the underlying risks.

Optimal pricing of insurance products represents a significant strategic challenge, given the inherent complexities of the market and the constant interaction between insurers competing for the same customer base under conditions of incomplete information. In this context, the study of pricing in competitive markets, such as the one developed by [23], employs a Bayesian game-theoretic approach to model the decision-making processes of insurers operating within an uncertain and dynamic environment. This approach considers uncertainty about market characteristics, such as demand elasticity and competitors' strategies, and proposes methods to maximize expected utility in stochastic environments. These contributions are particularly relevant for price bidding systems such as the First Price Sealed Bid, where optimal pricing strategies must balance competitiveness with financial sustainability. This paper aligns with the present study by highlighting the importance of advanced analytical methodologies for optimizing market share and maximizing profitability in the insurance industry.

The calculation of the premium has traditionally been performed using classical methods such as Generalized Linear Models (GLM). McCullagh [24] propose that these models adopt a conventional statistical approach, which assumes that the data emerge from a specific stochastic model, such as a Gaussian, Gamma or Poisson distribution.

This method provides a robust foundation for accurate risk estimation and premium determination. It is therefore unsurprising that there is a substantial body of literature on pricing in the insurance sector, which emphasizes the use of statistical models. Of these, GLMs are particularly noteworthy for their robustness and flexibility. As stated by [25], these models are appealing due to their capacity to estimate parameters in a readily comprehensible manner and to combine in a multiplicative fashion in order to accurately predict future costs associated with a customer seeking automobile insurance.

The premium calculated using GLM models represents the portion intended to cover expected losses and expenses related to potential claims, excluding the insurance company's operating expenses, such as taxes, contingencies, and profit margins. However, as highlighted by [26], there are certain limitations associated with linear models. For example, they are unable to capture non-linear relationships, are highly sensitive to outliers, and require a substantial amount of data to train accurate models with reliable predictions. Conversely, Nicholson [27] illustrates the intricacies inherent in the estimation of premiums in the context of catastrophic events, thereby underscoring the inherent complexities of the challenges confronting the insurance industry and the imperative to address the insurance problem through the implementation of novel pricing strategies.

The recent developments in the field of telematics have had a profound impact on the motor insurance industry, enabling insurers to make real-time adjustments to policy premiums based on the analysis of driving behavior data. The study by [28] illustrates the additional value of integrating telematic data into pricing strategies. It demonstrates how the continuous monitoring of driving behavior can result in more precise risk assessments, which in turn leads to more accurate premium calculations. By incorporating these data, insurers can better align premiums with individual risk profiles, enhancing pricing precision and fairness. This approach is particularly beneficial in a competitive market, where pricing flexibility is essential for insurers to maintain profitability while offering competitive rates.

This research illustrates the close interconnection between the automobile insurance industry and auction theory, specifically in relation to the pricing of policies displayed in insurance firms. Auction theory represents a significant and extensively researched field within the discipline of economics (see details in [29]). In the context of a market operating via auction, Attar [30] demonstrates the equilibrium conditions that prevail when the vendor is aware of the product's characteristics and prospective purchasers present non-exclusive contract offers. One of the key outcomes is that the vendor may either sell their entire endowment or forfeit the opportunity entirely. This aspect requires further investigation through contemporary research on markets utilizing auctions.

We present a novel approach to the estimation of optimal insurance policy pricing for insurance firms. Our methodology demonstrates that this pricing process is analogous to that of securing a market share in an industry comprising a limited number of market participants. This is based on an expected value model of the risks assumed by the insured, using a straightforward strategy of standardizing the prices offered at public auctions for the sale of an insurance policy in the automobile industry.

The present research is developed in Section 2 with the Materials and Methods, wherein the economic model that underlies the research, the research methodology, and the descriptive analysis of the database are delineated. Section 3 presents the findings of the research from an econometric and simulation perspective. Section 4 is the discussion section, and Section 5 contains the conclusions of the research.

2. Materials and Methods

The automobile insurance market employs an auction mechanism to determine policy prices, thus giving rise to a competitive problem of pricing. In this context, several insurers offer policies of comparable coverages, and the customer selects the lowest acceptable offer. In order to motivate this environment, a parsimonious expected-utility framework is employed to formalize two objects that are central to the empirical strategy: The customer's reservation premium, and the insurer's expected profit conditional on winning the contract.

Utilizing standard concavity assumptions on utility, the customer's optimization problem is shown to imply a well-defined willingness to pay that is contingent on wealth, coverage, and expected losses. This provides a microfoundation for treating the customer's acceptance threshold as a random variable in the market. From the insurer's perspective, pricing is a trade-off between margin and the probability of winning the contract.

In a first-price sealed-bid auction, an insurer sells the policy only if its bid is lower than those of competitors and the customer's reservation premium. Otherwise, the contract is not accepted. Incorporating unobservable preventive effort in a reduced-form way reflects the standard moral-hazard feature of automobile insurance, while keeping the model intentionally simple because the objective is not to estimate preferences or effort choices. Instead, the model's role is to justify the expected-profit representation and the auction-like win indicator that later allows optimal pricing to be linked to an expected market share.

2.1. Economic Model

We propose a basic economic model in which a rational economic agent is attempting to maximize a specified utility function. In accordance with [19], we consider that the economic agent is exposed to two risks with potential losses, L_1 and L_2 , associated with a probability of occurrence, p_1 and p_2 , which we associate with the risk of personal and material injury, respectively. In this sense, the probability $p_s = p_{s0} - p(\epsilon_s)$ is characterized by an intrinsic probability of occurrence of the loss p_{s0} that decreases as a function of the level of effort of the insurance customer. The insurance customer's maximization problem is as follows

$$\max_{\epsilon_1, \epsilon_2} \{ (1 - p_1 - p_2)U(W - \pi) + p_1U(W - \pi - w_1L_1) + p_2U(W - \pi - w_2L_2) - c_1(\epsilon_1) - c_2(\epsilon_2) \}, \quad (1)$$

where $U(\cdot)$ corresponds to the insurance customer's utility function, which is concave and twice differentiable; W corresponds to the insurance customer's original level of wealth; π corresponds to the value of the premium that can be contracted; w_1 and w_2 correspond to the percentage of the loss that the insurance customer must have available to obtain insurance; Finally, $c_1(\epsilon_1)$ and $c_2(\epsilon_2)$ correspond to the costs that the economic agent must incur as a function of the efforts to avoid the loss. The terms $c_1(\epsilon_1)$ and $c_2(\epsilon_2)$ represent reduced-form cost shifters on the insurer side associated with the two types of insurable events, such as personal injury and material damage. These components consist of expected indemnity payments and claim-handling expenses.

The insured's preventive efforts have been shown to affect claim probabilities through $p_s(\epsilon_s)$, though it is important to note that ϵ_s is not observable by insurers at the time of quoting. In the pricing stage, firms treat claim probabilities as random and use expected probabilities p_s^* to summarize behavioral risk. These are defined as the expectation of $p_s(\epsilon_s)$ conditional on observable risk characteristics, integrating over unobserved effort and severity components. In the context of the First-Price Sealed-Bid mechanism, each firm determines the value of π_i , representing a trade-off between margin and the probability of achieving the lowest bid. Consequently, riskier behavior shifts optimal bids upwards and

can increase dispersion in offers, resulting in the distribution of auction prices becoming an equilibrium outcome of both competitive bidding and behavior-driven expected costs.

Conversely, the firm is confronted with an asymmetric problem, whereby it is unaware of the level of effort exhibited by the economic agent. Furthermore, the firm is engaged in a competitive market environment with another firm, aiming to secure a sale. Initially, the information available to competing firms in the insurance market is identical, but the associated costs of the other market participants are not known. The firm's profit function for each policy is given by the following equation

$$\begin{aligned}\mathbb{E}[\Pi(\pi; p_1^*, p_2^*)] = & (1 - p_1^* - p_2^*)\pi + p_1^*(\pi - L_1 + w_1 L_1) \\ & + p_2^*(\pi - L_2 + w_2 L_2) - f,\end{aligned}\quad (2)$$

where $p_1^* = \mathbb{E}(p_1)$ and $p_2^* = \mathbb{E}(p_2)$ correspond to the probabilities assumed by the firm for the insured; f corresponds to the fixed costs assumed by the firm for generating the policy.

To make explicit how risk heterogeneity translates into bid heterogeneity, a reduced-form cost channel is adopted. Let x_i denote observable risk characteristics and let ϵ_i capture unobserved preventive effort and other latent behavioral factors affecting claim probabilities. Insurers cannot observe ϵ_i at the quoting stage, so they summarize behavioral risk through $p_{s,i}^*(x_i) = \mathbb{E}[p_{s,i}(\epsilon_{s,i}) \mid x_i]$. For firm j , define the expected cost component of policy i as

$$C_{j,i}(x_i) = p_{1,i}^*(x_i)(1 - w_1)L_1 + p_{2,i}^*(x_i)(1 - w_2)L_2 + f_j, \quad (3)$$

where f_j captures firm-specific cost. The effect of behaviour on bids is primarily through expected-loss costs, rather than through an explicitly parameterized bidding rule.

Nevertheless, in order to sell the policy, the firm needs to offer a price that is lower than that of the competition and lower than the maximum value that the economic agent is willing to pay for the insurance, which is equivalent to an auction of the type First Price Sealed Bid known in the economic literature.

Assuming the existence of a random variable that determines the competitive price within the industry and that the maximum price assumed by the insurance policy purchaser is also random, the expected value of the portfolio for the insurance firm is defined in Equation (4). These relationships serve to determine the value of the portfolio of policies held by the firm.

$$\begin{aligned}\mathbb{E}[\Pi_{Tot}(\pi_1, \pi_2, \dots, \pi_N; P_1^*, P_2^*)] = & \sum_{i=1}^N ((1 - p_{i,1}^* - p_{i,2}^*)\pi_i \\ & + p_{i,1}^*(\pi_i - L_1 + w_1 L_1) \\ & + p_{i,2}^*(\pi_i - L_2 + w_2 L_2) - f) \mathbb{1}_{\{\pi_i > 0\}},\end{aligned}\quad (4)$$

where $P_1^* = p_{1,1}^*, p_{2,1}^*, \dots, p_{N,1}^*$ and $P_2^* = p_{1,2}^*, p_{2,2}^*, \dots, p_{N,2}^*$ correspond to the vector of probability for each type of claim for each potential insured assumed by the firm; $\mathbb{1}_{\{\pi_i > 0\}}$ corresponds to an indicator function, the value of which is one when the sale is generated by the firm and zero otherwise ($\pi_i(\rho_i, \rho_i^*, \rho_i^{max}) \in \{\rho_i, 0\}$) where $\rho_i, \rho_i^* = \min(\rho_{1,i}, \rho_{2,i}, \dots, \rho_{N,i})$ and ρ_i^{max} correspond to the offers made by the firm, minimum value of competitors, and the maximum value that the insurance policy purchaser accepts, respectively. It should be noted that the value is zero in the event that the proposed price exceeds that of the competition or the offer exceeds the insurance policy purchaser's willingness to pay, which we represent by means of the following equation

$$\pi_i(\rho_i, \rho_i^*, \rho_i^{max}) = \begin{cases} \rho_i & \text{if } \min(\rho_i, \rho_i^*, \rho_i^{max}) = \rho_i \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

Equations (2) and (4) define the expected value of profit outcomes, not the choice variables. Specifically, Π is used to denote the profit associated with policy i , which is random due to claims, unobserved cost components, and acceptance outcomes. The firm's bid premium, denoted by π_i , is a deterministic decision variable that is determined at the time of bid submission. Consequently, Equation (2) should be interpreted as an expected-profit mapping, and the Equation (4) as outlined as the expected portfolio value.

The insurance firm's objective is to maximize the expected value of the benefits. When quoting a policy, it is reasonable to assume that each client is independent, which allows each policy to work independently of the other. This assumption leads to the following expression

$$\max_{\pi_1, \dots, \pi_N} \mathbb{E}[\Pi_{Tot}(\pi_1, \dots, \pi_N; P_1^*, P_2^*)] = \sum_{i=1}^N \max_{\pi_i} \mathbb{E}[\Pi(\pi_i; p_{i,1}^*, p_{i,2}^*)] \mathbb{E}[\mathbb{1}_{\{\pi_i > 0\}}]. \quad (6)$$

2.2. Bid Price Distribution

It should be noted that the optimization problem is concerned with determining the optimal price for each of the policies considered independently. If we consider ρ_i^* and ρ_i^{max} as independent random variables, the probability that the firm is able to sell the policy and incorporate it into its portfolio is given by the following expression

$$\begin{aligned} \mathbb{P}(\pi_i \geq 0) &= \mathbb{P}(\rho_i \leq \rho_i^*, \rho_i \leq \rho_i^{max}) \\ &= \mathbb{P}(\rho_i \leq \rho_i^*) \mathbb{P}(\rho_i \leq \rho_i^{max}). \end{aligned} \quad (7)$$

The assumption of randomness in the estimation of bid prices can be supported by the findings of [31], who shows that there is always a unique symmetric equilibrium and firms adopt a randomized approach to their behavior. In this context, it is necessary to define the probability of obtaining the sale of the insurance policy, which is associated with the probability of the minimum bid value. The distribution of the minimum bid value is given by the following expression

$$\mathbb{P}(\rho_i \leq \min(\rho_{1i}, \rho_{2i}, \dots, \rho_{Ni})) = \prod_{j=1}^M [1 - F_j(\rho_i)], \quad (8)$$

where $F_j(\rho_i)$ is used to denote the cumulative probability function of ρ_i for each firm j ; M is used to represent the number of participants in each of the public policy offerings. Using Equations (7) and (8), and $G(\rho_i) = \mathbb{P}(\rho_i \leq \rho_i^{max})$, we have the following definition of the probability of winning the policy auction

$$\mathbb{P}(\pi_i \geq 0) = \left(\prod_{j=1}^M [1 - F_j(\rho_i)] \right) [1 - G(\rho_i)]. \quad (9)$$

This methodology enables the formulation of an algorithmic approach to estimating the price and market share based on a process of maximizing firms' profits through the implementation of different types of policies. The procedure is defined in the following algorithmic representation of Algorithm 1. In this context, the vector *Sample* refers to a vector that accumulates the price of the other firms participating in the auction market. The value *MaxAccept* refers to the maximum value that a client would accept on the value

of its policy. The vector Value refers to the value of each policy for the firm, which is dependent on the simulation process. $\Pi_{mean}[\pi]$ refers to the average value of the policy at each possible value of π , which gives us its expected value. Share $[\pi]$ refers to the average participation of the firm for each value of π . Ultimately, the maximum profit value for the firm, which corresponds to a specific market share (Share*), is represented by π .

Algorithm 1 Estimation of optimal price and market share.

Require: $p_1, p_2, L_1, L_2, w_1, w_2, \pi_{grid}$, Number of firms (M), Sample of prices distribution (Φ_{prices}), Maximum accepted value distribution (Φ_{max}), Number of repetitions (R), fixed cost (c).

```

for each  $\pi \in \pi_{grid}$  do
  count  $\leftarrow$  0
  for  $i = 1$  to  $R$  do
    for  $j = 1$  to  $M$  do
      Sample[j]  $\leftarrow$   $\Phi_{prices}$ 
    end for
    MaxAccept  $\leftarrow$   $\Phi_{max}$ 
    if min(Sample, MaxAccept,  $\pi$ ) =  $\pi$  then
      Value[i] =  $(1 - p_1 - p_2)\pi + p_1(\pi - L_1 + w_1L_1) + p_2(\pi - L_2 + w_2L_2) - f$ 
      count  $\leftarrow$  count + 1
    else
      Value[i] = 0
    end if
  end for
   $\Pi_{mean}[\pi]$  = mean(Value)
  Share $[\pi]$  = count/M
end for
 $\pi^* \leftarrow$  max( $\Pi_{mean}$ )
Share*  $\leftarrow$  Share $[\pi]$ 

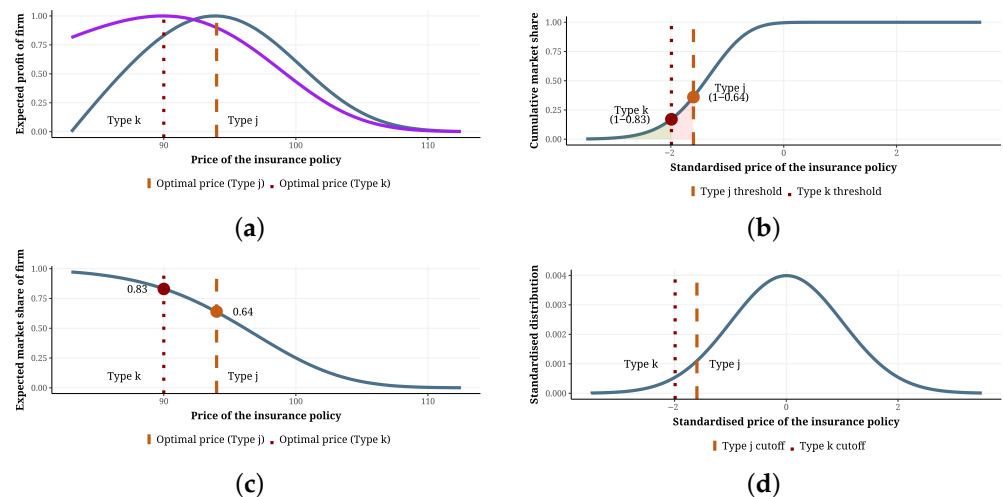
```

Table 1 presents the parameters of two distinct policy types for use in a simulation that involves a vector of 1000 for each auction policy and 10,000 repetitions, using a Monte Carlo simulation strategy. The simulation is based on the assumption that there are a total of eight firms participating in the market, with the prices bid by each firm following a normal distribution. Additionally, the customer's maximum acceptable value is generated from a non-negative distribution for the simulation, to ensure economic plausibility. These baseline distributions were used exclusively for the Monte Carlo simulation in the implementation of Algorithm 1. The Normal specification (truncated at zero) is adopted as a benchmark to generate competitor bids and customers' maximum acceptable values, and it is not essential for the empirical implementation, where we work with the observed bid distributions after standardization rather than imposing a parametric family. Furthermore, empirical evidence suggests that the standardized densities of the logarithmic transformation of bids are approximately unimodal and symmetrical in our data. Throughout this study, the normal distribution is parameterized as $N(\mu, \sigma^2)$, where the second parameter denotes the variance.

Figure 1 illustrates the profit-maximizing price of the firm, conditional on the required market share that maximizes the expected value of its portfolio for each policy type (a). The market share data (b) enable us to ascertain the maximum value of the distribution associated with a standard normal distribution (c). Ultimately, the price can be recovered for inhomogeneous price distributions through the inverse of the standardization process (d).

Table 1. Simulation parameters.

Parameter	Value	Parameter	Value
Number of firms	8	$L_{1,j}$	300
Sample distribution	$N(110, 100)$	$L_{2,j}$	400
Maximum accepted value distribution	$N(130, 100)$	$L_{1,k}$	400
$p_{1,j}$	0.2	$L_{2,k}$	400
$p_{2,j}$	0.1	$w_{1,j}$	0.2
$p_{1,k}$	0.1	$w_{2,j}$	0.2
$p_{2,k}$	0.1	$w_{1,k}$	0.2
c	3	$w_{2,k}$	0.2

**Figure 1.** Representation of the optimal response of prices offered based on industry standardised prices. The segmented line corresponds to type j and the dotted line to type k . (a) Profits of the firm as a function of price. (b) Cumulative market share distribution. (c) Market share in terms of policy price. (d) Optimal response.

2.3. Data and Summary Statistics

To verify the theoretical approach, we provide the details of the data used for the analysis. For this research, we used data from vehicle insurance quotes made in a massive distribution channel. The end customer requests a quote and can observe multiple price offers for comparable coverages across insurers. Our sample therefore represents the competitive pricing environment within a single broker distribution channel. The data used are anonymous and were obtained by obtaining car insurance quotes on an insurance broker's website between 1 January and 31 December 2024 (Public access to databases and technical information is available in the following repository <https://github.com/rolando-rubilar/Insurance-Automobile>).

A summary of the descriptive statistics of the variables used in this research is presented in Tables 2 and 3.

Table 2. Descriptive statistics for continuous variables of the sample of risks.

Continuous Variables			
Variable	Number of Risks	Mean	St. Dev.
Antiquity	40.168	5.89	3.24
Valuation	40.168	13,487,928.72	8,445,219.82

Based on the descriptive statistics of the sample, the average age of the vehicles is 5.89 years. The average valuation of the vehicles is 13,487,928.72 U.M (U.M. corresponds to the local currency units used, where 1000 U.M. $\approx$ 1 USD).

Regarding vehicle age, older vehicles are associated with lower quoted prices. In the case of the commercial value variable, vehicles with higher commercial values tend to receive higher quoted premiums, consistent with higher expected repair or replacement costs and, more generally, greater potential claim severity for newer and more valuable cars.

Table 3. Descriptive statistics for discrete variables of the sample of risks.

Categorical Variables			
Variable		Number of Risks	%
Age			
A1		2	0.005%
A2		2.124	5.288%
A3		8.663	21.567%
A4		16.512	41.107%
A5		12.007	29.892%
A6		339	0.844%
A7		521	1.297%
Gender			
Female		25.449	63.36%
Male		14.719	36.64%
Civil Status			
CS1		1.205	3.00%
CS2		27.193	67.70%
CS3		3.147	7.83%
CS4		4.530	11.28%
CS5		4.093	10.19%
Credit Score			
C1		28.867	71.87%
C2		11.301	28.13%
Properties			
P1		21.186	52.74%
P2		11.518	28.67%
P3		7.464	18.58%
Engine Capacity			
EC1		11.385	28.34%
EC2		20.871	51.96%
EC3		7.912	19.70%
Vehicles			
V1		15.927	39.65%
V2		14.239	35.45%
V3		10.002	24.90%
Claims History			
CH1		18.939	47.15%
CH2		10.709	26.66%
CH3		5.497	13.69%
CH4		5.023	12.50%

Categorical variables are coded as mutually exclusive groups. Age is coded as A1 (date of birth not recorded), A2 (21–30), A3 (31–40), A4 (41–50), A5 (51–60), A6 (61–70), and A7 (71–80). Civil Status is coded as CS2 (married), CS3 (divorced), CS4 (single), CS5 (widowed), and CS1 (not available in the broker database, hence “unknown”). Credit Score is coded as C1 (“good payers”) and C2 (“bad payers”). Claims History is coded as CH1 (no claims history available), CH2 (one claim in the last four years), CH3 (two or three claims), and CH4

(four or more claims). Properties is coded as P1 (not available/unknown), P2 (one property), and P3 (more than one property). Vehicles is coded as V1 (not available/unknown), V2 (one vehicle), and V3 (more than one vehicle). We treat “unknown” categories (e.g., CS1, P1, V1) as an explicit level rather than dropping observations, and their prevalence is reported in Table 3 (The term ‘unknown’ is recorded in the database when the policy applicant does not declare that variable in the application. In the numerical estimates, the designation “unknown” is considered an additional category).

With regard to the age variables used in the analysis, the most represented groups in the sample are Groups A5 and A6, corresponding to persons aged 51–60 and 61–70, respectively. It is also observed that 63.36% of the insured correspond to women and that 67.70% belong to the marital status Group CS2, which comprises those insured who are married.

We can also observe that the sample contains approximately 71.87% of policyholders belonging to Group C1 of the Credit Score, which corresponds to “good payers”, while only 28.13% correspond to “bad payers”, which means that they can obtain lower rates, meaning that they represent a lower risk of unpaid policies.

Regarding the variable of number of properties, we have that the majority of policyholders belong to Group P1, which indicates those policyholders who are not in the database to obtain the number of properties.

The Engine Capacity variable indicates that the sample is mainly divided between Groups EC1 and EC2, which correspond to vehicles with low and medium engine capacity, respectively. The remaining Group EC3 corresponds to those vehicles with a high engine capacity. For this variable, prices tend to be lower for low and medium engine capacity vehicles, while they tend to be higher for high engine capacity vehicles, since they represent a higher risk.

The number of vehicles variable indicates that most of the sample is concentrated in Group V2, which corresponds to those policyholders who have 1 vehicle, while the second largest group is Group V1 which shows those policyholders who are not in the database to obtain the number of vehicles. The remaining but not smaller group Group V3 indicates those policyholders who have more than one vehicle.

Looking at claims history, a key variable in assessing future risk, the majority of the sample belongs to the Group CH1. This category represents policyholders for whom no claims history is available in the database. Group CH2 includes those with one claim in the last four years, while Group CH3 includes those with two or three claims in the same period. Finally, Group CH4 consists of policyholders with four or more claims in the last four years. In particular, premiums tend to be lower for people with a confirmed history of one claim in four years. Conversely, premiums are higher for those with no claims history in the database, although not as high as for those with two or more claims in the specified period.

2.4. Estimation Methods

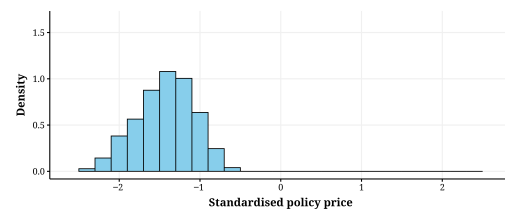
For the purposes of this research, we will not assume a specific price distribution of the policies on offer. If we assume that, for each policy quote, competitors offer a price independent of each other, then we can standardize the offers made by the competing firms, knowing the mean (μ) and variance (σ^2) of each policy. However, in the empirical case it is normal that we do not know the values of mean and variance, so it is necessary to use estimated values of mean and variance, which allows us to generate the following definition

$$X_{j,i} = \ln \rho_{j,i},$$

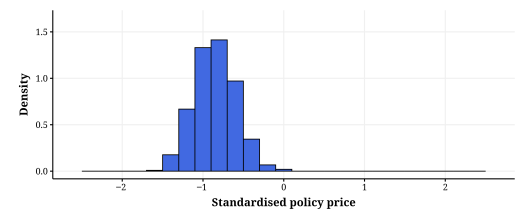
$$StdPrice_{j,i} = \frac{X_{j,i} - \bar{X}_i}{s_i}, \quad (10)$$

where $\bar{X}_i$ represents the average value of the logarithm of all auction participants in i insurance type, s_i represents the standard deviation of the logarithm of the participants in auction of i insurance type. Equation (10) enables us to undertake an empirical research to ascertain the optimal price above which the maximum profit can be obtained, on the assumption that the minimum price offered is the price that the economic agent is willing to accept.

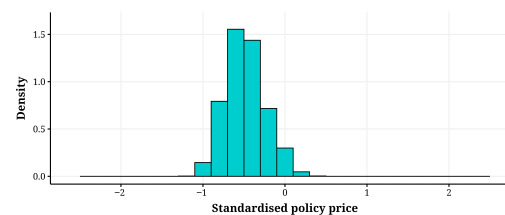
As illustrated in Figure 2, the figure displays the standardize price densities of the eight firms operating within the insurance market. It is observed that certain firms maintain a consistent pricing strategy of offering lower prices, a practice that has been demonstrated to affect their market share. In contrast, other firms adopt a higher pricing approach on average, a strategy that often leads to the emergence of niche markets characterized by a comparatively modest market share. This finding indicates that certain firms specialize in specific policies, while others employ a pricing strategy with the aim of capturing a greater market share.



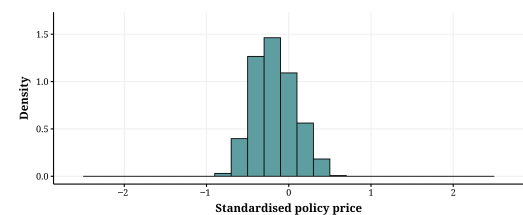
(a) Density of firm 1.



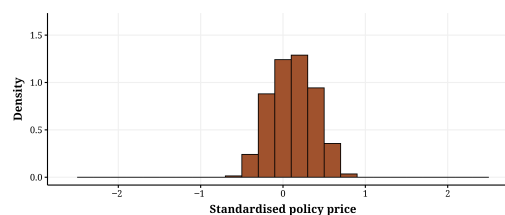
(b) Density of firm 2.



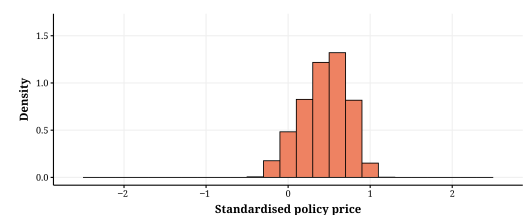
(c) Density of firm 3.



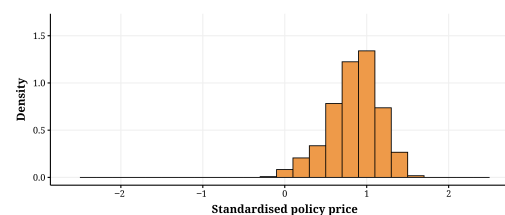
(d) Density of firm 4.



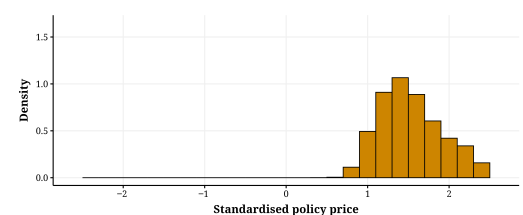
(e) Density of firm 5.



(f) Density of firm 6.



(g) Density of firm 7.



(h) Density of firm 8.

Figure 2. Standardised price density of market participating firms. Differing histogram colours represent different firms in the figure.

As can be seen in Figure 3, the distribution of the lowest bid value in each auction is shown (a), as well as the distribution of the second lowest bid value (b). Furthermore, the distribution of the mean value between the minimum value and the second minimum value is displayed. The dotted line corresponds to the 80th percentile of the distribution, which attempts to define a strategy in which 20% of the market share is obtained.

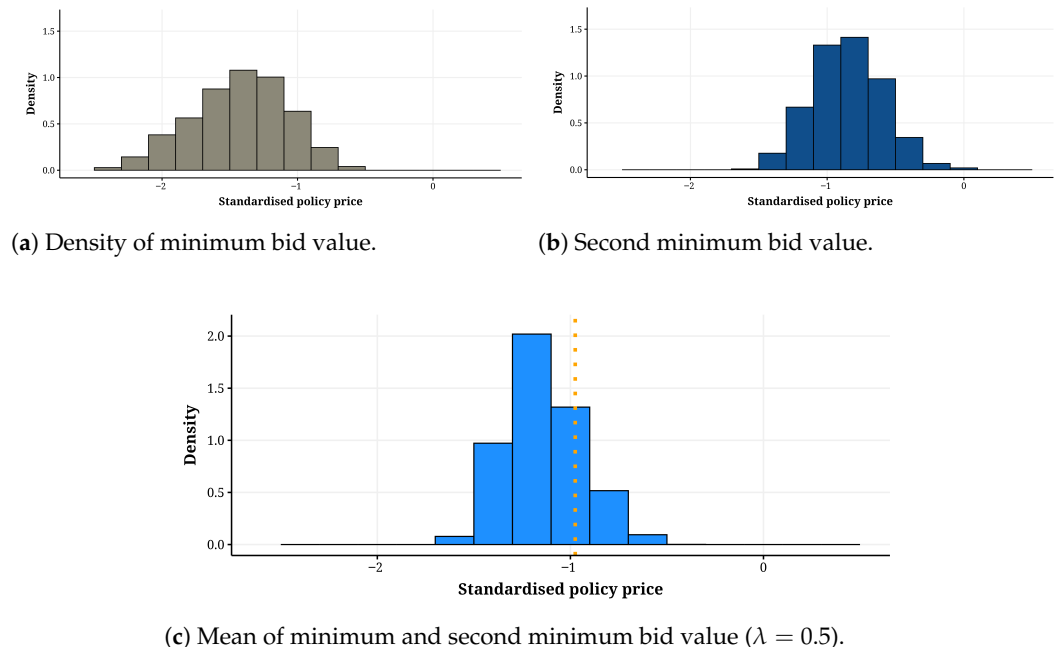


Figure 3. Standardised price density of value of winning firms' bids. Different colours in the histograms represent the minimum bid value, the second minimum value, and a linear combination of the two.

It is important to note that the difference between the lowest and second lowest bid value is beneficial to the economic agent. This is because, assuming that the second lowest price is lower than the maximum acceptance price, the economic agent would be willing to accept a higher price to buy the policy. Therefore, the optimal pricing strategy should correspond to a convex combination between the minimum value and the second minimum value, as represented by the following relationship

$$\rho^{\text{opt}} = \lambda^* \rho^{\text{min}} + (1 - \lambda^*) \rho^{\text{2nd}} \quad (11)$$

where ρ^{min} corresponds to the empirically observed minimum standardized price; ρ^{2nd} corresponds to the second minimum standardized value; λ^* corresponds to the optimal value of parameter to be empirically obtained. The parameter λ is instrumental in regulating the premium update in Algorithm 1. A higher λ implies larger adjustments in the quoted premium in response to the estimated win probability trade-off. This can be interpreted as placing relatively more weight on market gains versus short margin. In practice, the value of the parameter λ does not have to be chosen ad hoc. Insurance companies are able to estimate or calibrate it using historical bid data by matching observed premium adjustment dynamics and target win rates, or selecting λ to maximize performance according to a business objective.

We will use a heteroskedastic linear regression estimation method (see [32,33]) to estimate the mean and standard deviation of logarithm of bid price as a preliminary strategy for validating the elements necessary for standardizing the policies offered in the market. The conditional mean of log prices is modeled linearly in covariates, while the conditional variance is specified as an exponential function of selected regressors. In this

context, we will proceed to project the mean value of the prices of the policies offered in the market, which will then be used as a variable for the estimation of the standard deviation. The Normality assumption applies to the conditional regression disturbances used for likelihood-based estimation, not to the unconditional market distribution of bids. We define the specific model of the mean of industry prices for policy as follows

$$Y_i = \beta_0 + \beta_1 Antiquity_i + \beta_2 \log(Valuation_i) + \beta_3 Gender_i + \beta_4 Score_i + \beta_5 Claims_i + \sum_{s=6}^M \beta_s x_i + \epsilon_i, \quad (12)$$

$$\sigma_{i,Y}^2 = \exp(\alpha_0 + \alpha_1 Antiquity_i + \alpha_2 \log(Valuation_i)),$$

where Y_i corresponds to the natural logarithm of mean price of policy i across the industry, with the $\bar{X}_i = \hat{Y}_i$ value representing the expected value of the natural logarithm of the specific policy; $\sigma_{i,Y}$ corresponds to the modeling of variance as a function of explanatory variables α 's; $\beta = \beta_0, \beta_1, \dots, \beta_M$ corresponds to the vector of parameters of the model, considering the time fixed effects and the respective control variables; $\alpha = \alpha_0, \alpha_1, \alpha_2$ corresponds to the vector of parameters of the variance; ϵ_i corresponds to the error term of the estimated model in Equation (12). The following specification is presented for the purpose of calculating the standard deviation

$$Z_i = \gamma_0 + \gamma_1 Antiquity_i + \gamma_2 \log(Valuation_i) + \gamma_3 Gender_i + \gamma_4 Score_i + \gamma_5 Claims_i + \sum_{s=6}^M \gamma_s x_i + \epsilon_i, \quad (13)$$

$$\sigma_{i,Z}^2 = \exp(\delta_0 + \delta_1 Antiquity_i + \delta_2 \log(Valuation_i)),$$

where Z_i correspond to standard deviation of the natural logarithm of price of policy i across the industry; $\gamma = \gamma_0, \gamma_1, \dots, \gamma_M$ corresponds to the vector of parameters of the model, considering the time fixed effects and the respective control variables; $\delta = \delta_0, \delta_1, \delta_2$ corresponds to the vector of parameters of the variance; ϵ_i corresponds to the error term of the estimated model in Equation (13). The variance function includes only Antiquity and $\log(Valuation)$ in order to keep the scale equation parsimonious and easily interpretable, and because these are the main continuous covariates available in the present setting. From an economic perspective, older vehicles and higher valuations are naturally associated with greater uncertainty in terms of claim severity and repair costs. The remaining controls are kept in the mean equation and standard deviation equation to keep the regression compact and readable. This is done in order to avoid over-parameterization of the heteroskedastic component, and to ensure the preservation of estimation stability and transparency. The following variables have been integrated into both models:

- *Antiquity_i* corresponds to the difference between the year of the quotation date and the year of manufacture of the insured vehicle in policy i .
- $\log(Valuation_i)$ corresponds to the natural logarithm of the commercial value of the insured vehicle in policy i .
- *Gender_i* corresponds to the gender of the insured of policy i . For gender modeling it takes value 1 when the insured is male and 0 otherwise.
- *Score_i* corresponds to the credit rating of the economic agent in the financial system of policy i . In particular, economic agents are divided into two groups, where Group 1 (C1) corresponds to “good payers” and Group 2 (C2) corresponds to “bad payers”.

- $Claims_i$ corresponds to a variable that divides economic agents into 4 groups according to their accident and insurance history. Group 1 (CH1) corresponds to persons with no recorded insurance or traffic accident history; Group 2 (CH2) corresponds to policyholders who had 1 claim in the last 4 years; Group 3 (CH3) indicates those policyholders who had between 2 and 3 claims in the last 4 years; Group 4 (CH4) indicates those policyholders who had 4 or more claims in the last 4 years
- The variables designated as x are those that will be employed in the various specifications for each policy i . These include variables such as:
 - Civil status of persons separated into 4 groups.
 - The amount of properties owned by the insurance customer separated into 3 groups.
 - Engine capacity of the insured vehicle.
 - Total number of vehicles of the insurance customer.

The results yielded by the models specified in Equations (12) and (13) allows us to establish the following relationships, which are used to standardize the estimated prices described in Equation (10).

3. Results

Two models will be used for the estimation of the mean value of the logarithm policy price, as well as two models for the estimation of the standard deviation value of the logarithm policy price. In both the initial model of the mean value of the logarithm policy price and the first model of the standard deviation value of the logarithm policy price, the variables of Antiquity, Valuation, Gender, Credit Score, and Claims are considered as independent variables, without control variables. The second model of the mean price and standard deviation prices considers the same variables, but with the control variables, in order to ascertain the model's robustness in the context of incorporating control variables. In order to estimate the parameters of the model, a sample will be drawn from the original data set. This sample will be equivalent to 80% of the original data, with the remaining 20% being reserved for the evaluation of the empirical strategy, considering the bids in which the eight insurance companies participated, which ensures that winning probabilities and market-share mappings are computed on a comparable competitive set. To mitigate sensitivity to a single split, we also report results under repeated splits and summarize out-of-sample metrics across repetitions. This approach will allow for the observation of the robustness of the empirical strategy to different values of λ .

Table 4 reports the regression results for the mean value of the logarithm policy price (column 1 and column 2), exhibiting high levels of significance for almost all variables in the model, except for Gender in the regression in column 1, whether or not controls are considered. Conversely, the null hypothesis of constant variance is rejected, indicating that modeling with variance dependent on the defined variables is preferable to a model without such consideration.

The LR tests reported at the bottom of Table 4 reject the homoskedastic linear model in all cases, supporting the use of a heteroskedastic (log-location-scale) specification. In the scale equation (Columns 3–4), Antiquity and log(Valuation) exhibit statistically significant effects on dispersion, with magnitudes and significance levels that may change once controls are included, consistent with heterogeneity in pricing uncertainty across observable segments.

Similarly, columns 3 and 4 of Table 4 show the results of the standard deviation of the logarithm of prices. It can be observed that in model 3, neither the antiquity nor the Claims (NH) variable is significant, which changes when considering the control variables in model 4, showing a high level of significance. Additionally, the null hypothesis of

constant variance is rejected, indicating that a model that incorporates variance modeling is preferable.

Table 4. Regression of the mean and standard deviation of the offered price.

	Mean of $\log(\text{price})$		SD of $\log(\text{price})$	
	Model 1	Model 2	Model 3	Model 4
Antiquity	−0.0292 *** (−19.39)	−0.0235 *** (−27.04)	−0.000317 (−0.65)	0.00101 * (1.83)
$\log(\text{Valuation})$	0.280 *** (30.19)	0.304 *** (42.42)	−0.0134 *** (−4.31)	−0.00792 * (−1.69)
Gender (male)	0.0111 (1.24)	0.0293 *** (6.44)	0.00424 (1.46)	0.00953 *** (3.33)
Score	−0.241 *** (−25.15)	−0.219 *** (−41.78)	−0.0651 *** (−21.32)	−0.0646 *** (−20.13)
Claims (G2)	0.0544 *** (4.29)	0.0536 *** (8.66)	−0.00494 (−1.19)	−0.00578 (−1.48)
Claims (G3)	0.108 *** (7.85)	0.0974 *** (14.55)	0.0110 ** (2.45)	0.00944 ** (2.24)
Claims (NH)	−0.0706 *** (−7.09)	−0.0808 *** (−14.68)	−0.00403 (−1.24)	−0.00738 ** (−2.12)
Constant	−3.934 *** (−25.11)	−3.538 *** (−24.94)	0.508 *** (9.60)	0.510 *** (5.61)
	$\ln(\sigma^2)$			
Antiquity	−0.0134 * (−1.81)	0.0172 ** (2.28)	−0.0494 *** (−6.80)	−0.0469 *** (−6.34)
$\log(\text{Valuation})$	0.126 *** (2.93)	0.381 *** (8.95)	−0.271 *** (−6.31)	−0.199 *** (−4.50)
Constant	−4.461 *** (−6.18)	−10.22 *** (−14.35)	−0.0342 (−0.05)	−1.372 * (−1.86)
	LR test vs. linear model			
Controls	No	Yes	No	Yes
Observations	5419	5419	5419	5419
$\chi^2(2)$	22.06	84.39	58.42	42.98
$\text{Prob} > \chi^2$	<0.001	<0.001	<0.001	<0.001

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

As the dependent variable is $\log(\text{price})$, the coefficients in the location equation can be directly interpreted in economic terms. In the context of continuous regressors in levels, β is defined as a semi-elasticity, implying that a one-unit increase in the regressor results in an approximate change of $100 \times \beta$ percent in the expected price. For $\log(\text{Valuation})$, the coefficient is defined as an elasticity, indicating that a 1% change in valuation is associated with a $\beta\%$ change in the quoted premium. In the case of categorical variables, percentage differences are reported relative to the omitted baseline category, which represent percentage variations in price. This approach integrates statistical significance with economically meaningful magnitudes, thereby offering a comprehensive framework for analyzing economic phenomena.

Some covariates include an “unknown” level because the broker platform does not record the attribute for a subset of quotes. We retain “unknown” as an explicit category to avoid sample selection. Consequently, these indicators capture the component related to limited information, while any residual unobserved heterogeneity associated with missingness is absorbed by the intercept and the disturbance. This may result in a slight increase in uncertainty in predictions.

In order to evaluate multicollinearity in the covariates of the mean of $\log(\text{price})$ equation, variance inflation factors (VIF) are computed, with time fixed effects excluded. The findings suggest that there is minimal collinearity for the primary continuous predictors, with VIF values of 3.79 for $\log(\text{Valuation})$ and 1.95 for Antiquity. The VIF values for most of the categorical controls are less than 7. Higher VIF values are predominantly observed between the age-group and civil-status indicators, which is expected given the overlap in information between demographic groupings and the baseline category. It is important to note that, while multicollinearity may inflate standard errors for some categorical coefficients, it does not appear to compromise the identification of the key continuous variables. Furthermore, coefficient magnitudes are interpreted with this in mind. In order to verify that the parsimonious variance specification is not driven by collinearity, the same VIF diagnostic is performed for the standard deviation of $\log(\text{price})$ equation. The VIF values remain low for the variables that enter the variance function, with Antiquity presenting a VIF of 1.95 and $\log(\text{Valuation})$ a VIF of 3.79, suggesting that the component is not materially affected by multicollinearity.

Table 5 summarizes the out-of-sample predictive performance across the Monte Carlo replications for both components of the log-location–scale specification. For the location (mean) equation, the MAE averages 0.112 (S.D. 0.002), with RMSE averaging 0.143 (S.D. 0.003), and relatively tight ranges across replications, indicating stable predictive accuracy under resampling. For the scale (standard deviation) equation, performance is similarly stable, with mean MAE 0.067 (S.D. 0.001) and mean RMSE 0.088 (S.D. 0.002). Negative log-likelihood values for both components also vary modestly across replications, suggesting that the fitted conditional distributions remain well-behaved out of sample. Overall, these results provide evidence that the estimated location and scale components generalize consistently beyond a single split, supporting the robustness of the empirical implementation.

Table 5. Summary statistics of performance metrics (Monte Carlo replications).

Metric	Mean	S.D.	Min	Max
MAE_media	0.1124480	0.0021003	0.1063342	0.1197531
RMSE_media	0.1430223	0.0028706	0.1353635	0.1530711
NLL_media	−0.5324092	0.0213740	−0.5887069	−0.4554694
MAE_sd	0.0672068	0.0013578	0.0631012	0.0715035
RMSE_sd	0.0883735	0.0021128	0.0817736	0.0952797
NLL_sd	−1.0102120	0.0251761	−1.0885630	−0.9280388

Table 6 reports out-of-sample experiments using test data that were not used to estimate the regression parameters. The columns correspond to the target market share levels $s \in \{0.20, 0.30, 0.40, 0.50\}$, while the rows correspond to the candidate values of the convex-combination parameter λ . Each cell reports the market share achieved $\hat{s}(\lambda)$ implied by the auction rule in the test sample. We highlight the values of λ that minimize $|\hat{s}(\lambda) - s|$ for each target share, consistent with the selection rule in Equation (14). As expected, the optimal λ increases with the target market share. Performance deteriorates for higher target shares (above 50%), where the mapping from premiums to win probabilities becomes more sensitive to tail estimation error.

$$\lambda^*(s) = \arg \min_{\lambda \in \Lambda} |\hat{s}(\lambda) - s|, \quad (14)$$

Table 6. Results comparing market shares and a lambda value. The value in bold represents the λ value closest to the theoretical result.

	Market Share			
	0.20	0.30	0.40	0.5
$\lambda = 0.00$	0.11	0.14	0.17	0.20
$\lambda = 0.25$	0.19	0.21	0.24	0.26
$\lambda = 0.50$	0.25	0.30	0.33	0.36
$\lambda = 0.75$	0.30	0.34	0.39	0.45
$\lambda = 1.00$	0.32	0.39	0.45	0.50

4. Discussion

The framework developed in this paper is intended for use in quotation channels where price competition is transparent and data-rich, such as online broker markets. In such markets, customers can observe multiple competing offers for comparable insurance coverages, and insurers can routinely approximate competitors' quote distributions through periodic quote sampling. In such an environment, price standardization is naturally implemented within relatively homogeneous product segments. This is precisely the setting in which bid distributions are meaningful and comparable. The conditioning information used in our empirical implementation corresponds to variables that are typically available in this channel, but not necessarily for other industries.

However, in market structures with lower price transparency or a smaller number of active competitors, observed price distributions may be less comparable across firms and influenced by pricing frictions. In regulated environments or those with restrictions on price reassessment, strategic adjustment is limited. Our framework should be used as a benchmarking tool for competitive positioning. The approach in this article is most directly applicable to transparent multiple-quotation channels. In that sense, other structures require consideration of differences in competition intensity, information frictions and regulatory constraints on pricing.

In contrast to [30], who demonstrates that the seller either sells their entire endowment or not at all, this research demonstrates that the optimal response of auto insurance market players is not total market capture. Rather, it is partial market share that is optimal from a short-run resource maximization perspective.

It is crucial to consider the approximation of an industrial organization problem with a limited number of bidders in the market, which contrasts with other research such as that conducted by [21], which is based on markets with a substantial number of market participants. Furthermore, our methodological approach allows for the generalization to a competitive market by considering markets with a larger number of participants, thereby enabling us to ascertain the number of participants that generates zero profits in the insurance market.

The methodology presented in this research provides an implementation of a strategy for the determination of a price that maximizes the firm's profit, conditioned on the firm's internal costs that are unknown to all other firms. Nevertheless, the results presented in Table 6 indicate that the heteroskedastic regression procedure encounters difficulties when estimating large market shares, thereby limiting its applicability to relatively competitive scenarios with relatively low market shares of market participants.

Furthermore, the estimation considered time effects as a fixed effect in models 2 and 3. Although this captured the time effects, it also masked some autoregressive effects that could be addressed through ARIMAX (see [34,35]), GARCH (see [36–38]), and other methodologies that allow the time effect to be a variable in the model, thereby capturing the effects of changes in monetary policy or time effects of the capital markets. In the

context of our estimation, observations are conducted over a short period, which presents challenges in the estimation of factors associated with economic cycles or structural changes. Consequently, these behaviors fall outside the scope of the research objectives.

The findings of this study demonstrate the efficacy of employing statistical techniques to determine the most advantageous pricing in an automobile insurance market that operates under a first-price sealed-bid setting. In this context, the use of models based on heteroskedastic regression estimation to calculate the expected value and standard deviation of the policies offered has been demonstrated to be an effective strategy for the maximization of profits. The results indicate that optimal pricing is analogous to estimating the expected market share for each policy segment, particularly in niches where market shares are less than 50%. In this sense, the results shown in Table 6 should be interpreted as an empirical guideline rather than a theoretical threshold. In order to predict a large market share, it is necessary to operate on the lower tails of the competitors' bid distribution. In this area, small location-scale estimation errors can have a significant impact on win probabilities. This approach emphasizes the necessity for the application of other robust models, such as generalized linear models (GLMs), which can more accurately describe the relationships between explanatory variables and pricing outcomes. Furthermore, they provide a robust foundation for optimizing pricing strategies in the competitive context of the insurance market.

In alignment with this perspective, generalized linear models (GLMs) have been extensively acknowledged for their ability to encapsulate intrinsic risk heterogeneity within the insurance market, as evidenced by prior studies [24]. By incorporating key variables such as driver age, claims history and geographic factors, GLMs allow for accurate estimates of claims frequency and severity, which is crucial for adjusting rates in a competitive and sustainable manner. The results of this research reinforce the importance of addressing technical challenges, such as extra dispersion, by using appropriate distributions (e.g., Poisson or Gamma), demonstrating that classical methodologies remain applicable and effective even in markets with complex dynamics. This contributes significantly to improving price accuracy and risk management.

Nevertheless, the findings also underscore a significant drawback of the heteroskedastic regression, particularly in contexts where price estimation is conducted in markets with a high degree of market concentration. The modeling of large market share values, where a single firm dominates the market, presents a significant challenge in terms of accuracy. This issue can be attributed to the challenges associated with accurately estimating low-value policies within the sample and to the assumptions of independence and absence of intertemporal considerations of insurance companies, which do not align with the characteristics of a competitive market. These limitations are particularly evident in markets where market shares exceed 50%, resulting in a significant discrepancy between the heteroskedastic regression model estimates and the empirical observations. Conversely, when market shares are below 50%, the experimental model aligns more closely with empirical observations, indicating that the heteroskedastic regression approach may be applicable in less concentrated markets.

Conversely, the findings underscore the need to transcend conventional methodologies and embrace hybrid approaches that integrate machine learning models. This approach is consistent with the one proposed by [26], in which the authors emphasize how contemporary machine learning techniques, such as boosting or automatic feature selection, can reinforce traditional methodologies, markedly enhancing predictive capability in highly competitive contexts. The integration of these approaches could assist in overcoming the constraints of the model, thereby enhancing insurers' capacity to predict optimal prices and respond more efficiently to emerging risk dynamics.

The findings of this study are in accordance with the principles of game theory as outlined by [23], particularly with regard to the complexities of pricing strategies in markets characterized by incomplete information. The application of this framework demonstrates the crucial role of uncertainty in insurers' strategic decision-making processes, including the estimation of competitors' pricing strategies and the assessment of customers' demand elasticity. The results provide further evidence of the value of probabilistic and game-theoretic approaches, indicating that optimal pricing in first-price sealed-bid systems necessitates a sophisticated balance between competitiveness and profitability. By employing analogous analytical techniques, this study demonstrates that aligning pricing strategies with the objective of optimizing market share and financial sustainability is not only theoretically sound, but also practically applicable in real-world contexts. These findings emphasize the need for insurers to adopt advanced strategic frameworks that take into account both stochastic market conditions and the dynamic interaction between market participants.

Furthermore, in the context of first-price auction systems, the practice of bid shading is emerging as a pivotal competitive strategy, offering a sophisticated approach to optimizing pricing in competitive markets. By strategically reducing premiums and maintaining a surplus above break-even level, insurers can secure market share without compromising financial sustainability. These findings underscore the value of integrating theoretical auction models with empirical tools to enhance insurers' responsiveness and optimize their market share in dynamic and competitive environments.

As [28] notes, the incorporation of telematic data into pricing strategies represents a significant advancement in the motor insurance industry. By leveraging real-time driving behavior data, insurers can improve their ability to assess risk, leading to more precise and equitable premium calculations. This approach is aligned with the increasing need for adaptive pricing mechanisms that respond dynamically to market conditions and customers' specific risk profiles. In highly competitive markets, such as those utilizing auction-based systems, the incorporation of telematic data could further refine risk estimates, conferring an advantage upon insurers. These developments underscore the need to integrate emerging technologies with established methodologies to effectively navigate the complexities of modern insurance markets.

In this context, Nicholson [27] identifies several critical issues facing the insurance industry, including increasing competitive pressures, the effects of digitalization and the need to adapt to evolving risk patterns. These challenges emphasize the importance of developing advanced and flexible analytical tools that allow insurers to dynamically adjust their pricing. In the context of the first price sealed auction, the utilization of such tools could facilitate the expeditious response of insurers to market movements, thereby enabling the optimization of their participation in specific niches. The combination of traditional statistical methodologies with modern technological tools enables insurers to better meet the major challenges they face. This improves both the accuracy of their pricing and their ability to compete effectively in an increasingly digitalized and dynamic environment. It also enables them to achieve their sales and profitability targets. In this way, insurers can ensure the financial sustainability of the auto line of business.

The pricing environment considered in this study should be interpreted within the regulatory and competition-policy constraints that typically govern automobile insurance markets. In many legal systems, premiums are subject to regulations designed to protect consumers, requirements regarding the filing of rates, and restrictions on pricing. These factors may limit the degree of flexibility that a fully dynamic "auction-like" interpretation would imply. Consequently, the present framework is best regarded as a tool for the structuring of competitive quoting and benchmarking within regulated posted-price channels, rather than as a prescription for unconstrained bidding. Furthermore, the mechanism is

predicated on independent pricing decisions based on observed quotes, thus obviating the need for coordination or information exchange across firms. However, it is important to note that practical deployment in concentrated markets must remain consistent with applicable antitrust and regulatory standards. In more restrictive regimes, the approach may be used primarily for scenario analysis and risk-adjusted pricing guidance.

It should be noted that this study is subject to a further limitation, namely that it focuses on a short-term model, with the evaluation of each policy conducted within the same period. This restricted approach precludes the possibility of considering the impact of longer-term pricing strategies, such as those employed in policy renewal games, which could have a considerable impact on the competitiveness of insurers. Moreover, practical implementation must take into account the regulatory and competition policy constraints specific to each jurisdiction, such as filed rate requirements, consumer protection rules, oligopolistic conditions, and restrictions on price adjustments, which may limit the degree of flexibility implied by a purely auction-based approach. As a potential solution, we propose extending the model to incorporate intertemporal discounting, which could facilitate a more precise evaluation of pricing decisions over time. This extension would be beneficial not only for the automobile insurance sector, but also for other sectors, such as interest rate pricing in annuities, where long-term dynamics play a pivotal role.

A second limitation is related to the tractability assumption adopted in Section 2.2, where the minimum competitor bid and the consumer's maximum acceptable price are treated as independent variables when computing acceptance and winning probabilities. In practice, these objects may be correlated because both can be jointly driven by observed and unobserved factors. In the event of such dependence, the independence-based decomposition may have a significant impact on the implied win probabilities and, consequently, the quantitative mapping from premiums to expected market shares. It is suggested that a natural extension for future research would be to relax this assumption by specifying a joint distribution. This could be achieved via a copula or by conditioning on a richer set of observables.

Closely related, our framework is inherently static. The model's emphasis on within-period pricing has the effect of abstracting from key long-run drivers of profitability in insurance markets. This includes factors such as customer retention, switching costs and customer lifetime value. Furthermore, the model does not take into account evolving market conditions over time, such as changes in risk composition, competition intensity and macro shocks. Additionally, it does not encapsulate the dynamics inherent in the renewal process of insurance policies, wherein insurers engage in recurrent interactions with policyholders, and the determination of optimal pricing is contingent upon the value of the ongoing relationship. The extension of this approach to a multi-period setting, incorporating state variables to track customer history and market conditions, as well as retention probabilities and an intertemporal objective with discounting, would facilitate the selection of premiums through a trade-off between current margins and expected future value and market share.

5. Conclusions

This research presents a methodology for estimating the profit-maximizing price for firms participating in the automobile insurance market in the context of First Price Sealed Bid. The objective of the firm is to maximize profits, which is achieved through the implementation of a short-term model that considers the probability of occurrence of personal and material accidents.

In addition to its methodological contributions, this study underscores the pivotal role of the automobile insurance industry as a foundational element of the financial sector in a

multitude of economies. Automobile insurance serves not only to protect individuals and businesses against potential financial losses, but it also makes a significant contribution to overall economic stability by enabling risk sharing and building consumer confidence. In many countries, the automobile insurance market constitutes a substantial component of the general insurance sector, serving as a primary driver of premium income and employment. As vehicles remain a fundamental aspect of personal mobility and commercial activity, optimizing pricing strategies in this sector have significant implications. Improved pricing methodologies ensure the sustainability of insurers and promote fair competition, which is essential to maintain policyholder confidence and the overall health of the financial ecosystem. This emphasizes the crucial necessity of cultivating a culture of continuous innovation and utilizing sophisticated analytical tools to address the evolving challenges confronting this vital sector.

Moreover, it was demonstrated that the issue of identifying the maximum bid price that yields the greatest profits for the firm is equivalent to that of determining the proportion of the market that maximizes the firm's profits. This was achieved by standardizing the prices of public offers made by insurance companies. The objective was to identify the value that could be offered in the market that would maximize the firm's total profits. It was assumed that public offers respond to a random variable that is independent of the historical behavior of the firms. Should the assumption of independence between companies' policy offerings be invalidated, a potential line of research emerges, linking elements of industrial organization with elements of stochastic processes. This could have significant implications for practitioners in the automobile insurance industry, as well as in other industries exhibiting analogous behavior, such as annuities.

The strategy of profit maximization through market share pricing was successful in that the simulation allows for standardized pricing of the automobile insurance policy, based on public independent variables of the ordering customers. This pricing strategy is valid for market shares of less than 50%. It should be noted that all policy segments were assumed to have an equivalent cost structure, which may not be entirely accurate. This limitation suggests a potential avenue for further research, whereby specific market niches could be segmented using private information from insurance companies.

Author Contributions: Conceptualization, M.R. and R.R.-T.; methodology, M.R., R.R.-T. and C.F.-R.; validation, D.Y., B.P. and R.R.-T.; formal analysis, M.R. and R.R.-T.; investigation, M.R., R.R.-T. and C.F.-R.; resources, B.P.; data curation, R.R.-T.; writing—original draft preparation, M.R. and R.R.-T.; writing—review and editing, M.R., R.R.-T., C.F.-R. and D.Y.; visualization, B.P.; supervision, R.R.-T. All authors have read and agreed to the published version of the manuscript.

Funding: Rolando Rubilar-Torreálba was funded by Agencia Nacional de Innovación y Desarrollo Fondecyt Iniciación grant number 11250084.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data presented in the study are openly available on GitHub at <https://github.com/rolando-rubilar/Insurance-Automobile> (accessed on 1 September 2025).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Arrow, K.J. Uncertainty and the welfare economics of medical care. In *Uncertainty in Economics*; Elsevier: Amsterdam, The Netherlands, 1978; pp. 345–375.
2. Zeckhauser, R. Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives. *J. Econ. Theory* **1970**, *2*, 10–26. [[CrossRef](#)]

3. Akerlof, G.A. The market for 'lemons': Quality uncertainty and the market mechanism. In *Uncertainty in Economics*; Academic Press: Cambridge, MA, USA, 1978; Volume 66, pp. 235–251.
4. Kessler, D.; de Montchalin, A.; Thimann, C.; Hufeld, F.; Kojien, S. The macroeconomic role of insurance. *Econ. Regul. Syst. Risk Insur. Mark.* **2017**, *2*, 20–54.
5. Christophersen, C.; Jakubik, P. Insurance and the macroeconomic environment. *EIOPA-Financ. Stab. Rep.* **2014**, 44–55.
6. Horvey, S.S.; Odei-Mensah, J.; Liebenberg, A.P. Insurance market development and economic growth in Africa: Contingencies and thresholds of structural transformation. *Heliyon* **2024**, *10*, e40046. [[CrossRef](#)]
7. An, C.; He, X.; Zhang, L. The coordinated impacts of agricultural insurance and digital financial inclusion on agricultural output: Evidence from China. *Heliyon* **2023**, *9*, e13546. [[CrossRef](#)] [[PubMed](#)]
8. Ippolito, R.A. The effects of price regulation in the automobile insurance industry. *J. Law Econ.* **1979**, *22*, 55–89. [[CrossRef](#)]
9. Lee, C.Y. The effects of firm specific factors and macroeconomics on profitability of property-liability insurance industry in Taiwan. *Asian Econ. Financ. Rev.* **2014**, *4*, 681–691.
10. Eling, M.; Schaper, P. Under pressure: How the business environment affects productivity and efficiency of European life insurance companies. *Eur. J. Oper. Res.* **2017**, *258*, 1082–1094. [[CrossRef](#)]
11. Hombert, J.; Schoar, A.; Sraer, D.; Thesmar, D. Can unemployment insurance spur entrepreneurial activity? Evidence from France. *J. Financ.* **2020**, *75*, 1247–1285. [[CrossRef](#)]
12. Joskow, P.L. Cartels, competition and regulation in the property-liability insurance industry. *Bell J. Econ. Manag. Sci.* **1973**, *4*, 375–427. [[CrossRef](#)]
13. Ni, T.; Ren, Y.; Choi, J. Deposit insurance system and commercial bank profitability: A quasi-natural experiment based on the deposit insurance regulations. *Financ. Res. Lett.* **2024**, *67*, 105763. [[CrossRef](#)]
14. Nguyen, L.; Worthington, A.C. Industry regulation, fund characteristics, and the efficiency of Australian private health insurers. *Account. Financ.* **2021**, *61*, 781–801. [[CrossRef](#)]
15. Cooper, Z.; Craig, S.V.; Gaynor, M.; Van Reenen, J. The price ain't right? Hospital prices and health spending on the privately insured. *Q. J. Econ.* **2019**, *134*, 51–107. [[CrossRef](#)]
16. Carvajal, A.; Thereze, J. Insurance contracts and financial markets. *Math. Soc. Sci.* **2023**, *121*, 8–19. [[CrossRef](#)]
17. Lim, Q.M.; Lee, H.S.; Har, W.M. Efficiency, productivity and competitiveness of the Malaysian insurance sector: An analysis of risk-based capital regulation. *Geneva Pap. Risk Insur.-Issues Pract.* **2021**, *46*, 146–172. [[CrossRef](#)]
18. Posey, L.L.; Thistle, P.D. Automobile insurance and driver ability: Contract choice as a screening mechanism. *Geneva Risk Insur. Rev.* **2017**, *42*, 141–170. [[CrossRef](#)]
19. Zheng, H.; Yao, Y.; Deng, Y.; Gao, F. Information asymmetry, ex ante moral hazard, and uninsurable risk in liability coverage: Evidence from China's automobile insurance market. *J. Risk Insur.* **2022**, *89*, 131–160. [[CrossRef](#)]
20. Cummins, J.D.; Tennyson, S. Moral hazard in insurance claiming: Evidence from automobile insurance. *J. Risk Uncertain.* **1996**, *12*, 29–50. [[CrossRef](#)]
21. Cummins, J.D.; Tennyson, S. Controlling automobile insurance costs. *J. Econ. Perspect.* **1992**, *6*, 95–115. [[CrossRef](#)]
22. Brockman, M.J.; Wright, T. Statistical motor rating: Making effective use of your data. *J. Inst. Actuar.* **1992**, *119*, 457–543. [[CrossRef](#)]
23. Mourdoukoutas, F.; Boonen, T.J.; Koo, B.; Pantelous, A.A. Optimal premium pricing in a competitive stochastic insurance market with incomplete information: A Bayesian game-theoretic approach. *Insur. Math. Econ.* **2024**, *119*, 32–47. [[CrossRef](#)]
24. McCullagh, P. *Generalized Linear Models*; Routledge: Abingdon, UK, 2019.
25. Haberman, S.; Renshaw, A.E. Generalized linear models and actuarial science. *J. R. Stat. Soc. Ser. D Stat.* **1996**, *45*, 407–436. [[CrossRef](#)]
26. Maulud, D.; Abdulazeez, A.M. A review on linear regression comprehensive in machine learning. *J. Appl. Sci. Technol. Trends* **2020**, *1*, 140–147. [[CrossRef](#)]
27. Nicholson, J.E. Challenges for the Insurance Industry in the Future. *J. Insur. Regul.* **2019**, *38*. [[CrossRef](#)]
28. Henckaerts, R.; Antonio, K. The added value of dynamically updating motor insurance prices with telematics collected driving behavior data. *Insur. Math. Econ.* **2022**, *105*, 79–95. [[CrossRef](#)]
29. Chatterjee, R. A brief survey of the theory of auction. *South Asian J. Macroecon. Public Financ.* **2013**, *2*, 169–191. [[CrossRef](#)]
30. Attar, A.; Mariotti, T.; Salanié, F. Nonexclusive competition in the market for lemons. *Econometrica* **2011**, *79*, 1869–1918. [[CrossRef](#)]
31. Farinha Luz, V. Characterization and uniqueness of equilibrium in competitive insurance. *Theor. Econ.* **2017**, *12*, 1349–1391. [[CrossRef](#)]
32. Harvey, A.C. Estimating regression models with multiplicative heteroscedasticity. *Econom. J. Econom. Soc.* **1976**, *44*, 461–465. [[CrossRef](#)]
33. Greene, W.H. *Econometric Analysis*, 8th ed.; Pearson Education Limited: London, UK, 2018.
34. Feizabadi, J. Machine learning demand forecasting and supply chain performance. *Int. J. Logist. Res. Appl.* **2022**, *25*, 119–142. [[CrossRef](#)]

35. Bonga-Bonga, L.; Mwamba, M.J. Multivariate models for the prediction of stock returns in an emerging market economy: Comparison of parametric and non-parametric models. *Macroecon. Financ. Emerg. Mark. Econ.* **2024**, *17*, 25–41. [[CrossRef](#)]
36. Chaudhary, R.; Bakhshi, P.; Gupta, H. Volatility in international stock markets: An empirical study during COVID-19. *J. Risk Financ. Manag.* **2020**, *13*, 208. [[CrossRef](#)]
37. Apostolakis, G.N.; Floros, C.; Gkillas, K.; Wohar, M. Financial stress, economic policy uncertainty, and oil price uncertainty. *Energy Econ.* **2021**, *104*, 105686. [[CrossRef](#)]
38. Umar, M.; Mirza, N.; Rizvi, S.K.A.; Furqan, M. Asymmetric volatility structure of equity returns: Evidence from an emerging market. *Q. Rev. Econ. Financ.* **2023**, *87*, 330–336. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.