FAIR TRADE IN INSURANCE INDUSTRY: PREMIUM DETERMINATION OF TAIWAN AUTOMOBILE INSURANCE

Emilio Venezian
Venezian Associate, Taiwan

Chu-Shiu Li
Department of Economics, Feng Chia University, Taiwan
100 Wen Hua Rd., Taichung, 407 Taiwan
Tel: 886-424517250 ext. 4070

Chwen-Chi Liu
Department of Insurance, Feng Chia University, Taiwan
100 Wen Hua Rd., Taichung, 407 Taiwan
Tel: 886-424517250 ext. 4106

Chin-Ping Yu*
Graduate Institute of Business, Feng Chia University, Taiwan
100 Wen Hua Rd., Taichung, 407 Taiwan
Tel: 886-424517250 ext. 4071
Fax: 886-424512510

Abstract

This paper examines premium determination of voluntary automobile insurance policy and risk classification under a heavily regulated rating system in Taiwan. We investigate the distribution of actual premium and pure premium, based on unique data to test if premium reflect appropriate gender-age factor. The reasonableness of loading and the difference in driving exposure between policyholder and driver are investigated for three different types of policy. An adjustment of gender-age premium coefficients is called for.

Keywords: Automobile Insurance; Premium Determination

* Corresponding author: P9431775@fcu.edu.tw

1. Introduction

Premium determination is a fundamental task in the automobile insurance industry. To achieve the goal of solvency and fairness, insurers are required to perform ratemaking through categorization by using available information. However, premium discrimination based on immutable characteristics such as age, gender, nationality and race is a debatable equity issue in terms of the constitutional right and actuarial fairness. Butler et al. (1988), for example, provide evidence that auto insurers overcharge most women in U.S. insurance market. Alternatively, Harrington and Doerpinghaus (1993) indicate, if there is a restriction on rate classification in the auto insurance market, the cost of risk by distorting incentives for claim settlement might increase and the provision of coverage seriously weaken.

Contrary to the past debate of restriction on rate classification in other countries, Taiwan adopts stringent regulation on all the aspects of automobile insurance business, including the minimal capital requirement, designation and approval of policy, rating system and even the basic premium. Authority in Taiwan officially mandates rating formulas and standard of risk classification based on age and gender, in addition to the bonus-malus system. The insured generally face the choices of varied coverage- premium menu a la Rothschild and Stiglitz (1976), that is, higher coverage with higher premium or lower coverage with lower premium. Under such a special environment, what making an insurer different from others are competitions in services and underwriting in terms of loss ratio?

The objective of this paper is to explore the relationships among automobile insurance rating system and risk classification in Taiwan using voluntary automobile vehicle damage insurance data. We are attempting to examine if consumers pay rates that are causally related to losses. Specifically, we investigate premium determination in Taiwan auto insurance market by focusing on age factor, loading and pure premium.

2. Literature

The literature of risk classification and rating can be divided into two broad approaches: actuarial and economic analyses. The focus of actuarial literature is the choice of estimation method of pure premium. Chang and Fairley (1979), Sant (1980), Weisberg and Tomberlin (1982) are typical examples, among others. On the other hand, economic literature concentrates on the efficiency of classification and rating system. Early studies since Hoy (1982), risk classification issues have been focused on the classified cost and the existence of efficient equilibrium. Although Hoy (1982) provides ambiguous results concerning the efficiency consequences of permitting risk classification, Crocker and Snow (1986) show that any market equilibrium with costless categorization is Pareto dominating the equilibrium in which categorization is not allowed. However, when the acquisition of information is costly, the market can be an efficient mechanism without categorization.

Bond and Crocker (1991) claim that the attainment of first-best allocations may be permitted if insurers classify risks on the basis of endogenous categorization. Although they concern the insured's voluntary consumption of products correlated with underlying loss propensities, the same argument can be applied to the choice of insurance policy option.

In the early debate, Harrington and Doerpinghaus (1993) examine efficiency consequences of regulation for restricting market classification in the auto insurance market. They argue if different market premiums are equal to marginally expected costs of coverage for each risk, then these premiums would not only provide the insured with incentives for reducing the cost of risk but also insurers with incentives for accurate classification to obtain information. When risk classification is restricted, the average cost of coverage increases. Thus, the average premium without classification would exceed the premium with classification due to covering the expected costs for insurers.

3. Automobile Insurance Rating System in Taiwan

Comparing with the long history of insurance industry in Western countries, automobile insurance in Taiwan, first introduced in 1968, has not been well established until the recent decades. Since the mid-1990, the government has heavily regulated the rating system of automobile insurance in Taiwan. For example, a new rating system for vehicle damage coverage was launched on July 1, 1996 and the law of compulsory automobile liability insurance adopted in 1998. Authority determines all the pricing formulas and the weight of rating factors. (The Non-life Insurance Association conducted the work under the authorization of Ministry of Finance.) Even the loading factor was not permitted to be freely adjusted based on insurer's condition until 2002. Since all the insurers comply with the same operating rules and assume the obligation to submit monthly report on every automobile insurance transaction, the available data are consistent and abundant for our study.

In this paper, we focus on the analysis of private automobile insurance for vehicle damage coverage. There are three major types of coverage for damages to vehicles: comprehensive (form A), comprehensive (form B), and moving collision coverage. As shown in Table 1, the comprehensive (form A) policy, sold with compulsory increasing per-claim deductibles, covers all of the specified risks; the moving collision policy, sold with no deductibles, covers only one specific risk. In the last row, the values show the basic premiums corresponding to the appropriate coverage. As expected, more coverage (or no deductible) requires higher premium.

In fact, the premiums of vehicle damage coverage for all the policy options are calculated by using an official formula:

$$P = L * BP * M * I \tag{1}$$

where
P: premium
BP: basic premium
M: manufactured coefficient
I: inured coefficient
L: the ceiling of loading factor permitted by authority (= 1.35)

In the above formula, manufactured coefficient is closely related to the vehicle age and type. The insured coefficient is composed of two parts: immutable characteristics (gender and age) and driving (claim) records. Table 2 shows the gender-age coefficients of various ages for male and female, respectively. For the same age, male has higher pricing coefficient than female, ranging from 9% (elder) to 19% (younger).

4. Data and Methodology

We obtained the entire data set of private automobile vehicle damage insurance contracts in Taiwan for years 2003 and 2004. These two full calendar years of transaction records allow us to build a data set showing claim distribution in 2003 policy year. There were 403,979 policies with 69,367 claims including three major types of coverage for damages to vehicles as described in Table 1.

According to gender-age coefficients in Table 2, for a given age, males have a higher pricing coefficient than females, with the difference ranging from 9% for older drivers to 19% for younger drivers. As a matter of fact, most automobile insurance policies are purchased under the name of a female family member to echo the difference in gender coefficients for males and females. This phenomenon brings us the motivation to test the following hypotheses: (a) does premium reflect appropriate gender-age factor? (b) Is loading reasonable for all types of coverage? (c) Is there any difference in driving exposure between policyholder and driver?

To investigate the appropriateness of premium by age, we first calculate pure premium by multiplying average loss ratio with claim amount for each age. The pure premium is then compared with actual premium. Similar procedures are applied to demonstrate loading distribution and driving exposure between policyholder and driver.

From the viewpoint of traditional insurance theories, the premium charged by insurers from policyholders could be broadly divided into pure premium and loading expense, as shown in (1). The "loading" in (1) represents a multiple of loading over pure premium. On the other hand, pure premium is the claim each policy incurred on insurers, which is also the product of claim frequency per policy and claim severity.

$$\begin{aligned} \text{Premium} &= \text{Pure Premium } (1+\text{loading}) \\ &= \text{Incurred Claims per policy } (1+\text{loading}) \\ &= (\text{Claim Frequency} * \text{Claim Severity})(1+\text{loading}) \qquad (2) \end{aligned}$$

To address the pricing strategies adopted by insurers for different policyholders in terms of coverage type, age, and gender, this article incorporates the data of insurance policies and posterior claims to generate average premium, pure premium, and loading for different policyholders.

In the part of analysis, extra attention should be paid to the fact that pure premium and loading in equation (2) are insurers' anterior estimation based on past experiences when they price insurance policies. The pure premium and loading analyzed in the article is the posterior actuarial figures calculated by the actual policy data, such as the amount of premium paid by and claim requested by different policyholders, in order to reflect the pure premium and loading actually incurred on different policyholders after they make their insurance claims (or no any reimbursement is made).

According to equation (2) and the above theoretical bases, the formulas of actuarial average premium, pure premium, and loading of any policyholders would be as follows:

Average Premium = Average Pure Premium (1 + loading) (3)

where

Average Pure Premium = [Sum of (Paid Claims and Estimated Unpaid Claims)] / Number of Policies (4)

Loading = [Average Premium/ Average Pure Premium] – 1 (5)

5. Empirical Results – Individual Policy

Table 3 shows average premium, average pure premium, and loading generated from different coverage types. To analyze the variance in statistic distribution of average premium, average pure premium, and loading from various coverage types, we thereby divide these policies into four major categories and incorporate the data of insurance policies and posterior claims from the four different coverage types into the above equations (2)~(5). As shown in Table 2, there is a significant difference among the average premium, average pure premium, and loading in the actuarial results generated, with the biggest variance in Coverage A and Coverage C. The average premium in Coverage A is 47441.64, roughly five times as much as that of Coverage C, 10146.93. Meanwhile, the average pure premium in Coverage A is 15569.32, roughly nine times as much as that of Coverage C, 1750.57. However, the loading in Coverage A is 2.04712, only half as much as that of Coverage C, 4.79634.

Fig.1 shows the difference of average premium and average pure premium among different coverage types in terms of policyholders' age. After confirming the existence of significant difference in average premium, average pure premium, and loading among four different types of coverage, we further divide the policies under the same formula by age and gender. We then incorporate the classification of age and gender into the data of policies and posterior claims to explore if there is any pricing difference. As shown in Fig. 1, the average premium of four types of coverage are all in an L shape, suggesting that the average premium paid by the policyholders under the age of 35 is relatively higher and those over 35 is relatively lower and stable. In addition, only those elder over 60 have a downward trend in Coverage C and Coverage B with or without deductible. Besides, after comparing average premium with average pure premium, we find there is an overcharge of average premium paid by the policyholders under 35, with Coverage A and Coverage C having the most overcharge, which is roughly seven times as much as that of average pure premium.

Fig. 2 shows the difference in average premium and average pure premium among different coverage types at different gender and age. The results indicate that the average premium paid by policyholders under the age of 35 is relatively higher and apparently male policyholders incur more unreasonable overcharge of premium. Fig. 2 presents the examples of two different coverage types, Coverage B no deductible and Coverage C.

In addition to the above analysis of average premium and average pure premium, Fig. 1 demonstrates the loading paid by policyholders at different age for different coverage types based on the formula of equation (5). Fig. 3 shows that the highest loading is around 500% for Coverage C and around 200% for other coverage. It is easy to see that all of the actual loadings are significantly higher than the maximum value (135%) which is permitted by the government authority.

6. Empirical Results -- Grouped Policies

The analytic results of premium and loading in the first part of empirical study in this paper prove that the premium and insurers' loading expense indicate significant difference by coverage types, policyholders' age and gender. In other words, coverage types, age, and gender are the major factors that affect premium determination in Taiwan's vehicle insurance market. Such actuarial outcome could raise the question whether the pricing of insurance policy is reasonable. We will examine this question in the second part of empirical study.

From equation (1), we take logarithms to obtain regression model in equation (6). We use the group data to do the test in order to solve the problem of policies with no claims, where each group has the same number of claims. Thus, the explanatory variables contain the average manufactured coefficient of the group and average insured coefficient of group, while the explained variables are the average pure premium of the group. The descriptive statistics of grouping data is shown in Table 4.

Regression Model：

$$\ln(PP) = a + b\ln(AI) + c\ln(AM) \quad (6)$$

PP = pure premium

AI = average insured coefficient of group

AM = average manufactured coefficient of group

From the equation (6), the intercept indicates the size of basic premium. The intercepts shown in Table 5 is 7.8704 for Coverage C and significantly exceed 9 for other coverage types. Thus, these results imply that the basic premium for these four coverage types is away above a normal range, the same result as the above analysis concerning loading factor.

For the second step, we then examine whether the coefficients of lnAI and lnAM are equal to 1 simultaneously. From Table 5, we find the coefficients of lnAI and lnAM for Coverage A, Coverage B with no Deductible, and Coverage C are significantly different from 1, suggestting that the factors of manufactured coefficient and insured coefficient fail to reflect the real situation of claims of policyholders. In other words, the premium based on mandatory formula (1) for varies coverage do not consistent with the actual observations. However, the coefficients of lnAI an lnAM for Coverage B with Deductible are insignificantly different from 1, implying that the mandatory formula set by the government authority is applicable for Coverage B with Deductible.

7. Conclusion

To investigate the appropriateness of premium determined by official formula, this paper compare the official parameters with observed ones. We first calculate pure premium by multiplying average claim frequency times the average severity of the claims for each policy type and age. The pure premium is then compared with actual premium. Similar procedures are applied to demonstrate loading distribution and driving exposure between policyholder and driver.

Our empirical evidence indicates that the youngest policyholders do not have high loss exposure that would be anticipated by considering the table of gender-age coefficients. In other words, the youngest are overcharged for more than four times of the pure premium. In addition, the mandatory formula for premium is not designed for the coverage types except for Coverage B with Deductible. Thus, to fully reflect the real situation of claims of policyholders, an adjustment of gender-age premium coefficients is needed.

References

Bond, E. and K. Crocker, 1991, "Smoking, Skydiving, and Knitting: The Endogenous Categorization of Risks in Insurance Markets with Asymmetric Information," *Journal of Political Economy*, 99(1): 177-200.

Butler, P., T. Butler, and L. Williams, 1988, "Sex-Divided Mileage, Accident, and Insurance Data Show that Auto Insurers Overcharge Most Women," *Journal of Insurance Regulation*, 6: 243-84, 372-420.

Chang, L. and W. Fairley, 1979, "Pricing Automobile Insurance under Multivariate Classification of Risks: Additive versus Multiplicative," *Journal of Risk and Insurance*, 46: 73-96.

Crocker, K. and A. Snow, 1986, "The Efficiency Effects of Categorical Discrimination in the Insurance Industry," *Journal of Political Economy*, 94(2): 321-344.

Harrington, S. and H. Doerpinghaus, 1993, "The Economics and Politics of Automobile Insurance Rate Classification," *Journal of Risk and Insurance*, 60(1): 59-84.

Hoy, M., 1982, "Categorizing Risks in the Insurance Industry," *Quarterly Journal of Economics*, 97(2): 321-336.

Rothschild, M. and Stiglitz, J. E. 1976, "Equilibrium in Competitive Insurance Market: The Economics of Markets with Imperfect Information," *Quarterly Journal of Economics*, 90: 629-650.

Sant, D., 1980, "Estimating Expected Losses in Auto Insurance," *Journal of Risk and Insurance*, 49: 133-151.

Weisberg, H. and T. Tomberlin, "A Statistical Perspective on Actuarial Methods for Estimating Pure Premiums from Cross-Classified Data," *Journal of Risk and Insurance*, 49: 539-563.

Table 1: Four Types of Coverage for Property Damages to Vehicles

Insured Perils	Comprehensive (Form A)	Comprehensive (Form B)	Comprehensive (Form B) – No deductible	Two-Car Collision – No deductible
Rollover	Yes	Yes	Yes	
Lightening	Yes	Yes	Yes	
Fire (Explosion)	Yes	Yes	Yes	
Flying Objects	Yes	Yes	Yes	
Moving Collision	Yes	Yes	Yes	Yes
Other Collision	Yes	Yes	Yes	
Vandalism	Yes			
Non-Excluded Unknown Perils	Yes			
Deductibles (1000 NT$)	3/5/7	3/5/7	0	0
Basic Premium (NT$) 2001	47,096	23,119	25,433	11,918

Table 2 Gender-Age Coefficients Used in Rating System

Gender Pricing Coefficient Age	Male	Female
Age under 20	1.89	1.70
Age 20 above but under 25	1.74	1.57
Age 25 above but under 30	1.15	1.04
Age 30 above but under 60	1.00	0.90
Age 60 above but under 70	1.07	0.96
Age 70 above	1.07	0.96

Table 3 : Average Premium and Pure premium by Coverage Type

Coverage Type		Coverage A		Coverage B With deductible		Coverage B No deductible		Coverage C	
Number of Policies		14835		38101		146592		109368	
Male	Female	5244	9591	12621	25480	36447	110145	38397	70971
Number of Claims		5013		11781		50205		7394	
Male	Female	1750	3263	3880	7901	11953	38252	2645	4749
Average Premium		47441.64		24006.25		24503.15		10146.93	
Average Pure Premium		15569.32		8344.79		7725.20		1750.57	
Loading		2.04712		1.87680		2.17185		4.79634	

Table 4: The Descriptive Statistics of Grouping Data

Coverage Type		Coverage A	Coverage B No Deductible	Coverage B With Deductible	Coverage C
Number of Policies		14835	146592	38101	109368
Number of Groups		109	1004	256	148
Number of Policies per Group	Max	314	2228	579	1573
	Min	76	24	31	39
	Ave	136.10	146.01	148.83	738.97
Number of Claims per Group	Max	49	54	51	51
	Min	43	47	44	44
	Ave	46	50	46	50

Table 5: The Results of Regression Analysis by Grouping Policies

Variable	Coverage A	Coverage B No Deductible	Coverage B With Deductible	Coverage C
Intercept (It)	9.8693	9.5833	9.3433	7.8704
Theoretical Value (TV)	11.0600	10.4439	10.3490	9.6859
P{It=TV}	0.00***	0.00 ***	0.00 ***	0.00 ***
Coeff (ln_AI)	0.7672	1.1757	1.2063	3.1441
Coeff (ln_AM)	0.9672	0.3801	0.8082	-0.0647
P{Coeff (ln_AI) = Coeff (ln_AM) =1 }	0.0161**	0.00 ***	0.3896	0.00 ***
R-squared	0.8155	0.0974	0.4591	0.3989

*** denote a significant at the 1 %

** denote a significant at the 5 %

Fig. 1: Average Premium and Average Pure Premium by Age