Proceedings of the 8th Symposium on Applied Science, Business & Industrial Research – 2016
ISSN 2279-1558, ISBN 978-955-4709-52-2

Factors Affecting the Claim Based on the Characteristics of User and Insured Vehicle

*Madushanka MHS, Dharmawardane PMN and Appuhamy PADAN

Department of Mathematical Sciences, Faculty of Applied Sciences, Wayamba University of Sri Lanka
*sam19901118@gmail.com

ABSTRACT

In the case of automobile insurance, it is common for insurers to use a number of prior classification variables such as the driver's age, gender, occupation, and vehicle's usage and type, to adequately and fairly differentiate risk levels among policyholders (Lamaire, 1995). This study is mainly focused on finding novice factors that are affecting a motor insurance claim. Ten variables were considered in this study and secondary data were obtained from one of the Sri Lankan insurance companies. Preliminary analysis was conducted to identify the relationship between claim occurrence and the characteristics of the vehicle and user. Since the dependent variable was the occurrence of claim during the study period, Binary logistic regression model was fitted. This model helped to determine the factors that affect the occurrence of claim. The study revealed that just above 25% of insured vehicles under this company had claims during the period of study. In addition to that the majority of claims were made by the users' age between 27 and 48. Moreover six novice factors were identified as the influential factors for claim occurrence and those were agency type, purpose of use of vehicle, policy type, cylinder capacity, age of user and age of vehicle. Since the variables identified in this study were significantly associated with claim occurrence, it is worth to rethink by the company to incorporate these findings to modify the current premium equation.

KEYWORDS: Claim, Motor Insurance, Premium Equation

1 INTRODUCTION

One of the main products in Sri Lankan insurance companies is motor insurance. It is an agreement between the policyholder and the insurance company. Motor insurance mainly gives three types of coverages to the policyholder depending on the contract. They are property, liability and medical. To obtain these benefits, the policyholder has to pay the premium. The main objective of this study is to identify the characteristics of user and insured vehicle that affect the occurrence of claims. Results can be used by the selected insurance company to build up more powerful premium equation in order to generate more profit.

2 LITERATURE REVIEW

As for the risk classification of vehicle insurance, it refers to questions oriented to practices. Each country's automobile insurance system often conducts risk classification through characteristics of vehicle usage, brand, style, and characteristics of policy holder insured's gender, age, and claim record. (Crocker & Snow, 2000).

Puelz & Kemmsies (1993) used data of three personal vehicle insurance policies in Georgia, USA, including vehicle collision coverage, full coverage, and liability coverage, to evaluate how gender and other demographic variables impact on premium pricing. Their empirical research results showed that gender significantly affects premium rates, yet its influential degree is relatively less than other variables such as driving record, age, location, and vehicle type.

Braver & Trempel (2008) identified higher accident tendencies for young and elder drivers. To put their findings into a figure with car accident loss versus age, then the figure shows a line close to a U shape. Such results respond to the rate regulations in practice in Taiwan, which apply higher rate coefficients upon young and elder insured.

3 RESEARCH PROBLEM

The premium is the important quantity of the Motor Department, which is calculated by using premium equation. Two terms can be observed in the right hand side of the premium equation. One is constant term, which cannot be changed by insurance companies but by the government. The other term is non-constant term, which can be changed by insurance companies. Therefore the insurance companies' always look forward to changing non-constant terms in order to achieve maximum profit. The main focus behind this research is to identify the new factors that affect profit of the company. After identifying these factors the company can modify a premium equation which is believed to be more effective and profitable than existing ones.

4 METHODOLOGY

In this study secondary data were obtained from one of the insurance companies in Sri Lanka and all insured vehicles during 2010 to 2011 were considered. Altogether there were 45,475 insurance policies, consisted of characteristics of user and insured vehicle.

Preliminary analysis was conducted to explore the characteristics of vehicle and user. Pearson's chi-squared test was carried out to find the associations between variables and advanced analysis was carried out to find the associations between more than two variables. Binary logistic regression models were used as an advanced analysis tool and the dependent variable was the occurrence or non-occurrence of claim during the period considered. It was used to identify the relationship between the dependent variable and independent variables.

The following factors of both user and vehicle were considered as independent variables that affect the occurrence of claim.

- Agency Type
- Policy Type
- Purpose of Use
- Age of User
- Parts Availability
- Vehicle Age
- Cylinder Capacity

5 RESULTS AND DISCUSSION

The following Fig. 1 shows the percentage of claims made during the study period.

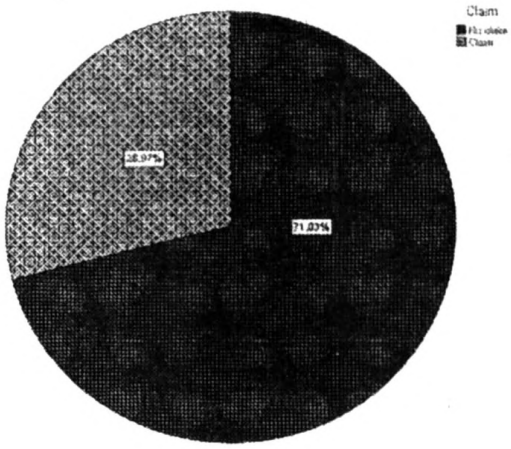

Figure 1: Percentage of Claims Made

According to Fig. 1, it can be observed that 71.03% percentage of vehicles have no claims and 28.97% percentage of policy holders made claims during the period considered. So the majority of vehicles had no claims.

The relationship between Claim Count and the Agency Type is illustrated in the following cluster bar chart given in Fig. 2.

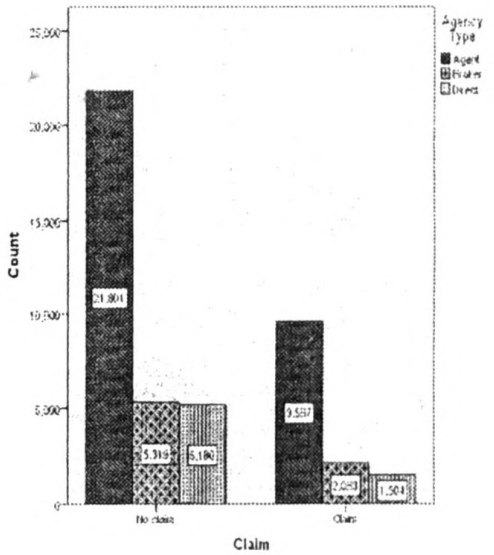

Figure 2: Claim Count vs. Agency Type

As shown in Fig. 2 when the insured did the businesses through an agent, 21.08% of policies had claimed. On the other hand 4.58% and 3.31% of policies made claims through broker and direct contacts respectively.

The Fig. 3 shows the relationship between part availability and the occurrence of claims.

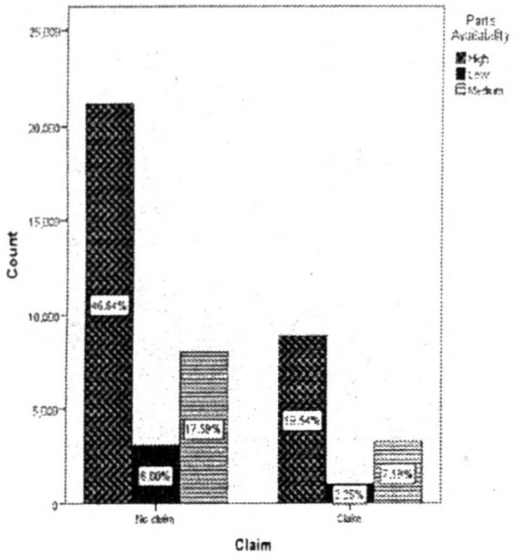

Figure 3: Claim Count vs. Part Availability

The cluster bar chart given in Fig. 3 indicates that 19.54% of policies claimed parts which were in higher grades.

The relationship between user's age and the occurrence of claims was depicted in Fig.4.

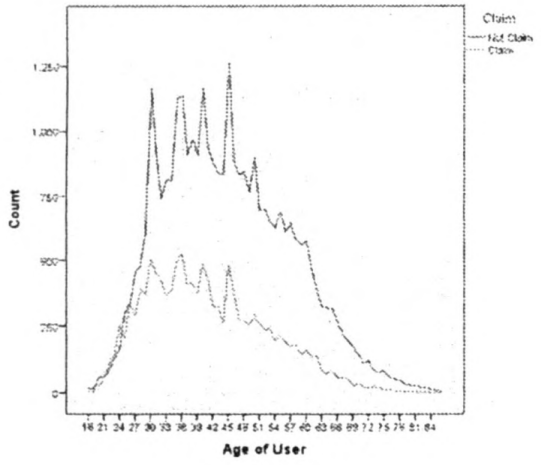

Figure 4: Claim Count vs. User's Age

Fig. 4 indicates that majority of claims were made by the age group, 27 to 48. Moreover the relationship between vehicle's age and the Claim Count shows in Fig.5.

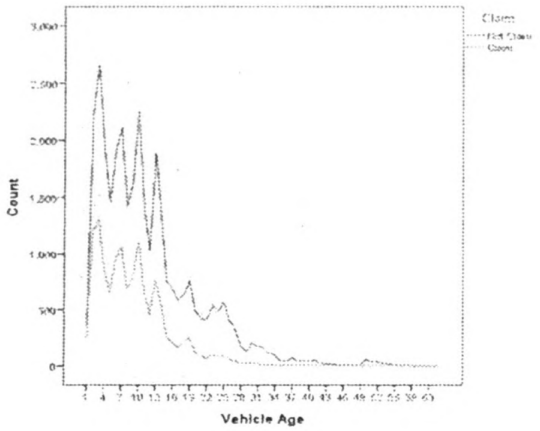

Figure 5: Claim Count vs. Vehicle's Age

Fig. 5, line chart, shows that most of the claims were made within the first 15 years.

The Chi-Square tests of associations with the corresponding hypotheses are given below.

Hypotheses:

H_0: There is no relationship between Claim occurrence vs. each of the categorical variable mention under Table 1.

H_1: There is a relationship between Claim occur vs. each of the categorical variable mention under Table 1.

Table 1: Chi-Square Test of Associations

Categorical Variable	P Value	Result
Agent Type	0.000	Reject H_0
Policy Type	0.000	Reject H_0
Purpose of Use	0.000	Reject H_0
Cylinder Capacity	0.000	Reject H_0
Age of User	0.000	Reject H_0

As shown in Table 1 there is no any variables having P-values greater than 0.05. Therefore the variables considered for Chi-Square test have a significant relationship with occurrence of claim.

The significance of the Binary Logistic regression model was tested and the hypotheses related to this test are mentioned below.

Hypotheses:

H_0: The model is not significant.

H_1: The model is significant.

Table 2: Omnibus Test of Model Coefficients

		Chi-square	df	Sig.
	Step	-3.469		.046
Step 2[a]	Block	2133.056	16	.000
	Model	2133.056	16	.000

Since 0.05>0.000, the null hypothesis was rejected. Therefore this model is significant at 5% level of significance. Table 3 shows the summary of the logistic model.

Table 3: Model Summary

Step	-2 Log likelihood	Cox & Snell R^2	Nagelkerke R^2
1	52603.843[a]	.460	.661
2	52607.312[a]	.466	.652

According to Cox and Snell method and Nagelkerke method the R^2 value of this model is 0.466 and 0.652, respectively, which are less than 1. Therefore, it can be considered as a suitable model for prediction. Further the adequacy of the model was tested using Hosmer & Lemeshow test and the results are given in Table 4. The corresponding hypotheses are as follows:

H_0: The model adequately fits the data.

H_1: The model does not adequately fit the data.

Table 4: Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	13	.8498	.096
2	15	.7818	.085

Since the significance value is 0.085>0.05, there is no enough evidence to reject null hypothesis at 5% level of significance. Therefore the model adequately fits the data. The model coefficients and its odds ratios are given in Table 5.

Table 5: Variables in the Equation

Variables	B	S.E.	Wald	df	Sig.	Exp(B)
Agent Type			64.223	2	.000	
Agent Type(1)	.186	.030	39.238	1	.000	1.204
Agent Type(2)	-.011	.041	.079	1	.009	.989
Policy Type(1)	-.154	.024	42.441	1	.000	.857
Purpose of Use			34.353	3	.000	
Purpose of Use(1)	-.423	.078	29.393	1	.000	.655
Purpose of Use(2)	-.489	.085	33.438	1	.000	.613
Purpose of Use(3)	-.356	.111	10.261	1	.001	.700
Cylinder Capacity			78.300	4	.000	
Cylinder Capacity(1)	-.308	.064	23.016	1	.000	.735
Cylinder Capacity(2)	-.119	.061	3.846	1	.050	.887
Cylinder Capacity(3)	-.350	.072	23.421	1	.000	.705
Cylinder Capacity(4)	-.358	.100	12.906	1	.000	.699
Age of User			655.264	3	.000	
Age of User(1)	.973	.046	443.614	1	.000	2.646
Age of User(2)	.483	.041	137.163	1	.000	1.621
Age of User(3)	.237	.043	30.511	1	.000	1.268
Age of Vehicle			765.835	3	.000	7.641
Age of Vehicle (1)	2.034	.127	254.588	1	.000	
Age of Vehicle (2)	1.751	.128	187.429	1	.000	5.762
Age of Vehicle (3)	.949	.134	50.078	1	.000	2.583
Constant	-2.608	.167	243.130	1	.000	.074

According to Table 5, it can be observed that six variables significantly contribute to the model. Those variables are agent type, policy type, purpose of use, cylinder capacity, user age and vehicle age.

Based on the identified variables, the regression equation can be written as,

Logit (P_{ijklmn}) = Constant + β_i (Agent Type)$_i$, (i=1,2) + β_j (Policy Type)$_j$,(j=1) + β_k(Purpose of Use)$_k$,(k=1,2,3) + β_l (Cylinder Capacity)$_l$, (l=1,2,3,4) + β_m (Age of User)$_m$, (m=1,2,3) + β_n (Age of Vehicle)$_n$,(n=1,2,3)

Where, P_{ijklmn} is the probability of claim occured and β_i, β_j,...., and β_n are corresponding constant coefficients.

6 CONCLUSION

This research was conducted using the secondary data set which consisted of 45,475 motor insurance policies to identify the novice factors of user and insured vehicle which affect the occurrence of the claim at one of the pioneer insurance companies in Sri Lanka.

According to the results of the study it can be concluded that, there are six significant factors, namely, agent type, purpose of use, policy type, cylinder capacity, age of user and age of vehicle which affect the occurrence of motor insurance claims. Therefore these factors may help to modify the current premium equation in order to maximize the profit.

REFERENCES

- Braver, E. R. and Trempel R. E., (2004), Are Older Drivers Actually at Higher Risk of Involvement in Collisions Resulting in Deaths or Non-Fatal Injuries among Their Passengers and Other Road Users?, *Injury Prevention*, 10(1): 27–32
- Crocker, K. and Snow A., (2000). The Theory of Risk Classification, in G. Dionne (ed.), *Handbook of Insurance*, 245–276, London: Kluwer.
- Lemaire, J., (1995). *Bonus-Malus Systems in Automobile Insurance*. Boston, MA: Kluwer Academic Publishers.
- Puelz, R., and Kemmsies W., (1993). Implications for Unisex Statutes and Risk-pooling: The Costs of Gender and Underwriting Attributes in the Automobile Insurance Market, *Journal of Regulatory Economics*, 5(3): 289-301.