CHAPTER 4

RESEARCH METHODOLOGY

4.1 Introduction

This chapter consists of two sections. The first section describes the data and the type of date set used in this analysis. The second section explains how the employed models were used to work out the data through computer software, E-view version 3.1.

In this chapter, we attempt to test for the presence of adverse selection and market signaling in the Malaysian automobile insurance market by estimating the premium-deductible schedule and the demand function for deductible.

Our main focuses are testing for separating or pooling equilibrium and nonlinearity of pricing system in the insurance market. Therefore, we employ some methods for these purposes.

4.2 Data Description

For the estimation of model, we use a cross sectional data set from a representative insurer domiciled in Malaysia for 2001. The sample contains data on 345 insured parties under comprehensive insurance scheme. For each data, we have information not only regarding the gender and age for an individual insured, but also his claim history. Besides that, we also know the characteristics of the automobile such as age, cubic capacity and so on.

We assume that each insured in the data set is homogenous with respect to his or her loss producing characteristics relative to any other insured in the category, from the perspective of the underwriting insurer, who controls for an individual's age and gender, age and cubic capacity of automobile.

4.3 The Model

In this section, we employ the empirical model in order to test for separation and nonlinearity of pricing system in Malaysian automobile insurance market. The premium-deductible and the demand function for a deductible are interpreted as a standard hedonic system as discussed by Rosen (1974). Then, we estimate the empirical counterparts to these models by using a procedure described by Bartik (1987) and Epple (1987).

4.3.1 Testing for nonlinearity of the premium-deductible schedule

The first step, in attempting to test for nonlinearity of a pricing system, we estimate the premium-deductible schedule as a reduce form hedonic premium which is postulated in model 1. The description of variables is shown in the TABLE 4a.

We assume that it is determined by the market interaction of insurer and customers. After the estimation of model has been done, the determination of the reliability of result is performed. The evaluation would show whether the estimates of the parameters obtained are theoretically meaningful and statistically satisfactory (Kew, 1996).

$$P = \alpha_0 + \alpha_1 Perage + \alpha_2 Male + \alpha_3 Age + \alpha_4 Cubic + \alpha_5 D_1 + \alpha_6 D_2 + \alpha_7 CD_1 + \alpha_8 CD_2 + \mu \quad (1)$$

P	Gross premium of insurance coverage
Perage	Age of the insured
Male	=1 for a male =0 for otherwise
Age	Age of the automobile
Cubic	Cubic capacity of automobile
D_1	= 1 for RM 251- 500 deductible = 0 for otherwise

D_2	= 1 for more than RM 500 deductible = 0 for otherwise
CD_1	Interaction variables between cubic capacity and D_1 dummy variable
CD_2	Interaction variables between cubic capacity and D_2 dummy variable
μ	Error term

TABLE 4a Variables Description for Model 1

A priori, we expected that there is positive relation between gross premium and male insured (Male = 1). However, the control variables of perage and age are related negatively to the gross premium. These control variables reflect the prospective loss of insured. However, we also include a set of interaction terms of deductible choice with cubic capacity of automobile that reflects our suspicions that the marginal premium of insurance coverage varies with the automobile's characteristic.

In model 1, the dummy variables, D_1 and D_2 represent deductibles categories as described in TABLE 4a. D_1 and D_2 are equal to zero when compared with the benchmark, that is the RM100-250 deductible. Hence, we can test for the nonlinearity of the premium-deductible schedule by using this benchmark and the fact that the deductible variables interact with other explanatory variables.

The estimates from model 1 is very important for us to estimate the marginal effects on the premium of moving from the RM100-250 deductible group to the RM251-500 deductible group and more than RM500 deductible group as follow (Puelz and Snow, 1994) respectively,

$$B1 = \alpha_5 + \alpha_7 C \tag{2}$$

and

$$B2 = \alpha_6 + \alpha_8 C \tag{3}$$

Over the range of deductibles in our sample, nonlinearity is present when the marginal premium between the RM100-250 and RM251-500 deductible differs from the marginal premium between the RM251-500 and more than RM500 deductible.

Therefore, we estimate

$$B3 = \frac{1135.8643B2 - 169.9411B1}{965.9232} \tag{4}$$

Since three deductible categories in the data are: RM100-250, RM251-500 and more than RM500, thus we use the mean values for each level of deductibles in order to obtain a more representative result. Then, we compare the value of *B3* with the value of *B1*. There is a nonlinear premium-deductible schedule in the automobile insurance market if the value of the calculated *B3* is different from the calculated *B1*. However, a calculated value for *B3* which equals the calculated *B1* is evidence of a linear pricing system in the market.

4.3.2 Testing for separation

We test for separating or pooling equilibrium in the in insurance market by estimating the demand function for a deductible in the model 5 via ordered logit.

$$\overline{D} = \beta_0 RT + \beta_1 Male + \varepsilon \qquad (5)$$

$\overline{D}$	is 0 for D = RM 100- 250 deductible is 1 for D = RM 251- 500 deductible is 2 for D = more than RM 500 deductible
RT	= 1 for those who filed a claim or more than one claim during year 2001 = 0 for otherwise
Male	= 1 for a male = 0 for otherwise
ε	Error term

TABLE 4b Variables Description for Model 5

In ordered dependent variable models, the observed outcomes representing ordered or ranked categories. Thus, by implementing the ordered logit in E-view version 3.1 programme, our ordered dependant variable that is $\overline{D}$, deductible variable must be integer valued. Otherwise, we cannot estimate the model 5.

Then, we regress the choice of deductible, $\overline{D}$ against two explanatory variables namely, risk types, RT and gender dummy variable, Male. We use the claim history as the proxy for risk types because it is unobservable. This identification supported by Boyer and Dionne (1989). They showed that the past accident record is good predictor of an individual's loss probability. However, the variable of Male plays the role to control for variation in risk preferences.

We must interpret with care to the estimated coefficients of this ordered logit model.[21] Therefore, we are interested in whether the coefficient of risk types, RT variable is statistically significant different from zero or not in the model 5. If the sign for RT variable is negative and the riskiness has a statistically significant influence on deductible choice, thus we conclude that the insurance market is in separating equilibrium with presence of asymmetric information. The customer of different risk types are separated by deductible choices: the high-risk type chooses a lower deductible, while a higher deductible is more likely chosen by low-risk customer.

Inversely, we have a pooling equilibrium or equilibrium with perfect categorization in the insurance market if RT variable shows an insignificant influence

[21] Greene, William H. (1997), "Econometric Analysis", Third Edition, Prentice-Hall, Inc.

on deductible choice. In other words, there is no separation by risk type when the market is in equilibrium.

4.4 Conclusion

The chapter discussed estimation of the premium-deductible schedule (refer to section 4.3.1) and the demand function for a deductible (refer to section 4.3.2) in order to test for the presence of adverse selection and market signaling in the Malaysian automobile insurance market.