Evidence for Adverse Selection in the Automobile Insurance Market

Rachel J. Huang*

Assistant Professor, Finance Department

Ming Chuan University, Taipei, Taiwan

Larry Y. Tzeng

Professor, Department of Finance

National Taiwan University, Taipei, Taiwan

Jennifer L. Wang

Associate Professor, Department of Risk Management and Insurance

National Chengchi University, Taipei, Taiwan

Kili C. Wang

Instructor, Department of Risk Management and Insurance

Shih Chien University, Taipei, Taiwan

*Corresponding Author. Finance Department, Ming Chuan University, 250 Chung Shan N. Rd. Sec. 5, Taipei, 111, Taiwan. E-mail address: rachel@mcu.edu.tw; tel: 886-2-28824564 ext 2871.

Evidence for Adverse Selection in the Automobile Insurance Market

Abstract

This paper proposes a new method to directly test the existence of adverse selection in the automobile insurance market. By tracing the renewal decisions of the insured, we find that, in Taiwan automobile insurance market, the choice of insurance coverage and the previous-year claim records are positively correlated, which can be a result of adverse selection but not of moral hazard. We further find that the loss ratio is positively correlated to the choice of insurance coverage. This indicates that cross-subsidization may exist in the market. The finding of cross-subsidization further strengthens the evidence to support the existence of adverse selection.

JEL classification: D80, G22, C30

Keywords: insurance market, asymmetric information, adverse selection.

1. Introduction

Asymmetric information has been widely discussed in the insurance literature since Rothschild and Stiglitz (1976) and Shavell (1979) first pioneered this line of research. Following their footsteps, many researchers (e.g., Stiglitz, 1977; Wilson, 1977; Miyazaki, 1977; Spence, 1978; Dionne and Lesserre, 1985, 1987; and Arnott and Stiglitz, 1988) have provided insightful theoretical analysis on the market with asymmetric information problems. More recently, several papers have used empirical data to further examine the theoretical predictions on asymmetric information in real world.

However, the empirical evidence in the literature has not yet reached a determinant conclusion on whether asymmetric information problems do exist in the insurance market. Some researches (e.g., Puelz and Snow, 1994; Dionne and Gagne, 2002; and Finkelstein and Poterba, 2004) have found evidence to support the existence of asymmetric information in insurance markets, whereas other researches (e.g., Chiappori and Salanie, 1997, 2000; Cawley and Philipson, 1999; Cardon and Hendel, 2001; Dionne, Gourieroux, and Vanasse, 2001) have found no such evidence.

It is important to find out the existence of asymmetric information in reality. It is also essential to further identify its source, if asymmetric information exists. Two types of asymmetric information problems, adverse selection and moral hazard, specifically received a lot of attention in the insurance literature. Each of these problems has a different impact on insurance operation and should be remedied by a different solution. Therefore, it is important for insurers to further identify the source of the asymmetric information.

However, it is not an easy task to separate adverse selection from moral hazard. Chiappori and Salanie (2000) alleged that both adverse selection and moral hazard problems would cause the choice of coverage and the occurrence of the claim to be positively correlated. Thus, they examined the existence of asymmetric information by testing the conditional dependence between the choice of coverage and the occurrence of the claim. Unfortunately, under their approach, we cannot further identify whether asymmetric information is caused by the adverse selection or by moral hazard problem, because both adverse selection and moral hazard problems would cause the same result.

Among those researchers who found evidence to support the existence of asymmetric information, only a few have identified its source. Dionne and Gagne (2002) separated moral hazard from adverse selection by analyzing the effect of the replacement cost endorsement. Finkelstein and Poterba (2004) used data in the annuity market to provide a direct test for adverse selection under the assumption that moral hazard problems are limited in that market. Chiappori and Heckman (2000) and Abbring, Heckman, Chiappori and Piquet (2003) proposed employing dynamic data to verify the moral hazard problem. By adopting dynamic data, Abbring, Heckman, Chiappori and Piquet (2003) did not find strong evidence of moral hazard in France's automobile insurance market.

This paper intends to propose a new approach to directly trace adverse selection by using the insured's renewal decisions. We first use previous-year claim records as a proxy to classify the insured's risk type. We then examine whether the previous-year claim records are related to the choice of insurance coverage in the following year. Since the previous-year claims happened prior to the insured choosing their insurance coverage, only adverse selection, as argued by Abbring, Heckman, Chiappori and Piquet (2003), can explain the relationship between the previous-year claim records and the insurance coverage choices. Therefore, we can examine the existence of adverse selection by testing the correlation between the choice of insurance coverage and the previous-year claim records for the continuing renewal contracts.[1]

Furthermore, we intend to test whether cross-subsidization exists in the market. Unlike most papers in the literature generating their empirical hypotheses by Rothschild and Stiglitz (1976), we test whether the market equilibrium is separating equilibrium with cross-subsidization as proposed by Miyazaki (1977) and Spence (1978).[2] It should

[1] This argument rests on the assumption that an individual's risk preference does not change because of an accident. If the individual becomes more risk averse after experiencing an accident, he/she will choose a policy with higher coverage in the renewal decision, no matter which risk type the insured belongs to.
[2] Only a few of the researches support cross-subsidization. By analyzing the U.S. crop insurance market, Makki and Somwaru (2001) find that high risks are under charged whereas low risks are over charged.

be noted that both Rothschild and Stiglitz (1976) and cross-subsidization equilibrium proposed by Miyazaki (1977) and Spence (1978) predict separating equilibrium, that high risks choose high coverage and low risks choose low coverage. On the other hand, Rothschild and Stiglitz (1976) predict that both high risks and low risks are charged on basis of their own loss probabilities, whereas cross-subsidization equilibrium proposed by Miyazaki (1977) and Spence (1978) predict that high risks receive subsidization from low risks.

To provide evidence for the cross-subsidization hypothesis,[3] we further test whether the loss ratio is related to the choice of coverage. When facing the adverse selection problem described by Rothschild and Stiglitz (1976), an insurer could design break-even policies with different ranges of coverage in a competitive insurance market. The insured would then self-select insurance policies according to their own risk types. The high-risk insured will choose high coverage and be charged a high premium rate, while the low-risk insured will choose low coverage and be charged a low premium rate. Thus, in Rothschild and Stiglitz's separating equilibrium, the expected loss ratio is not related to the choice of insurance coverage, as shown in appendix A. On the other hand, the loss ratio could be positively correlated to the choice of insurance coverage if there is a cross-subsidization in the market, as shown in Appendix B. It is important to note

[3] According to Crocker and Snow (1985), the consumers' welfare can be improved if cross-subsidization is allowed in the insurance market with adverse selection.

that the loss ratio could be unrelated to the choice of insurance coverage if only a moral hazard problem exists in the market, as shown in Appendix C. Thus, we may use cross-subsidization evidence to further separate adverse selection from moral hazard.

Our empirical results find strong evidence to support that asymmetric information problems do exist in Taiwan's market. Second, we find a positive relationship between the choice of insurance coverage and the previous-year claim records, which indicates the existence of adverse selection. Furthermore, we find that the loss ratio is significantly positively correlated to the choice of insurance coverage, which indicates the existence of cross-subsidization. Thus, our empirical results support that adverse selection problems exist in Taiwan's automobile insurance market and the market equilibrium is a separating equilibrium with cross-subsidization.

The rest of this paper is organized as follows: Sections 2 introduce the automobile insurance in Taiwan and explain why contracts with different features could screen individuals. Section 3 describes the data and the summary statistics. Section 4 provides evidence for asymmetric information. Section 5 confirms that adverse selection exists in Taiwan automobile insurance market. Section 6 further demonstrates the evidence of cross-subsidization. Section 7 concludes and suggests some further extension.

2. Automobile insurance in Taiwan and the contract screening

In 2003, approximately 5.6 million car owners purchased automobile insurance from 25 insurance companies in Taiwan. Automobile insurance accounts for about 50 percent of the insurance premium volume in most property-liability insurance companies and has occupied the largest market share of the property-liability insurance market.

Three types of automobile insurance have been observed in the market: compulsory liability, supplementary liability, and comprehensive coverage for damage. Compulsory liability covers only the injuries to a third party, no matter if the insured is responsible. Supplementary liability, which is purchased voluntarily, covers both property damage and bodily injury. Comprehensive coverage, which is also purchased voluntarily, provides coverage for property damage to the driver's automobile. In 2003, the total premiums paid for comprehensive coverage insurance were about US$0.37 billion and the incurred losses were about US$0.23 billion.

There are three types of contracts in comprehensive coverage insurance: Type A, B and C. Type A coverage covers all risks, including all kinds of collision and non-collision losses, which may be caused by missiles or falling objects, fire, explosion, windstorm, intentional body damage, malicious mischief, and any unidentified reasons other than the exclusions in the policy. Type B coverage selected risks. It also covers collision and non-collision losses as Type A does. However, the non-collision losses caused by intentional body damage, malicious mischief, and the unidentified reasons covered under type A are specifically excluded from type B. Type C covers only damage in a collision involving two or more vehicles. Collision losses caused by hitting other objects—such as a telephone pole, a tree, or a building—and non-collision losses that used to be covered under types A and B are specifically excluded from type C.

To examine the asymmetric information problems in Taiwan's automobile insurance market, we focus on comprehensive coverage insurance. Most of the empirical papers examine asymmetric information problem in insurance market by limiting their focus to whether individuals purchase greater payment exhibit higher risk. From a different angle to provide evidence of asymmetric information, Finkelstain and Poterba's (2004) paper indicates that insurance contracts with different features could screen individuals directly. For example, an individual with a higher life expectation will purchase back-loaded annuity, which has a higher payment in later periods than an annuity with flate payment. An individual with a lower life expectation will demand an annuity with guarantee period. Following their paper, we argue that different risk types of individuals will self-select among the three types of comprehensive coverage insurance.

In the present of adverse selection, there will be two types of individuals: high risk type and low risk type. In automobile insurance, a high risk driver has a higher probability causing a car accident, no matter the accident involves other vehicles. Thus, type A and B contracts could be referred as "high coverage" contracts, and type C could be referred as a "low coverage" contract. Therefore, in the present of adverse selection, high risk individuals will purchase type A or B contracts and have a higher probability to file a claim, where as low risk individuals will purchase type C contract and have a lower probability to file a claim.

However, types A and B cover more risks than type C and obviously could have more claims. To control the effect caused by the broader coverage of types A and B, we use only claim data of collision losses with more than two cars involved,[4] since non-collision losses covered under types A and B are specifically excluded from type C.

It is worth noting that the above prediction is also observable in the present of moral hazard. When a market suffers moral hazard problem, the individuals covered by higher coverage will be less careful. Thus, if an individual purchase type A or B, he or she will drive less carefully because that the contract covers any damage that caused by any identified reason. Meanwhile, a driver becomes more careless will simultaneously increase both the probability of hitting another vehicle and other objects. Therefore, we predict that the individuals with type A or B contracts will have a higher probability to file collision losses with more than two cars involved.

---

[4] Although types A, B, and C all cover collision losses with more than two cars involved, the policyholders of types A and B are more likely to be high risk drivers and could more likely cause collision losses with more than two cars involved.

3. Data and summary statistics

Our data come from a large insurance company that controls over 30 percent of the market share in Taiwan's automobile insurance market. Thus, we believe that the data should be representative for the entire automobile insurance market in Taiwan. In total, we assemble 185,704 observations, which can be distributed as 59,186, 61,627, and 64,891 observations in 1999, 2000, and 2001, respectively.[5] The dataset also includes information about age, gender, and marriage status for the insured's characters and age of car, brand, and registered location for the car's characters. The definitions of all the variables are displayed in Tables D1 in Appendix D. The summary statistics of all the variables for the whole sample are displayed in Table D2 in Appendix D.

To track the existence of adverse selection, we further use a sub-sample in which the insured purchased insurance in year t and renewed the policy in year $t+1$. In the sub-sample, there are 26,420, 28,986, and 31,305 observations in 1999, 2000, and 2001, respectively. The summary statistics of the renewal decision and all the other variables for the sub-sample are displayed in Table D3 in Appendix D.

To track cross-subsidization, we investigate whether the loss ratio of the individual who purchases types A and B is larger than that of the individual who purchases type C. The loss ratio of the individual is defined as the amount of the claim the individual files

[5] We start from year 1999, because that type C is launched from that year.

divided by the amount of the insurance premium the individual pays. The summary statistics of the loss ratio (LR) for the whole sample and for the renewal sample are respectively displayed in Tables D2 and D3 in Appendix D.

4. Evidence for asymmetric information

Our empirical analysis begins by testing the existence of asymmetric information in Taiwan's automobile insurance market. We follow the empirical model suggested by Chiappori and Salanie (2000), which used a pair of probit models to test the conditional dependence between the choice of coverage and the occurrence of the claim. The probit models are as the follows:

$$\text{Prob}(y_i = 1) = \Phi(X_i b_a)\text{, and} \tag{1}$$

$$\text{Prob}(z_i = 1) = \Phi(X_i b_b) \quad , \tag{2}$$

where X_i is the variable for individual i's information, b_a and b_b are the coefficient vectors of the regressor, and Φ is the cumulative distribution function of $N(0,1)$. y_i is a dummy variable that indicates individual i chooses high-coverage policies. Since both types A and B cover non-collision claims and type C covers only collision claims, we classify types A and B as high coverage and type C as low coverage. Thus, when an individual i chooses comprehensive automobile insurance coverage of type A or B, then $y_i = 1$; otherwise $y_i = 0$. z_i is another dummy variable that marks a claim. In this paper we do not use all the claims when defining the variable z_i.

Instead, we employ only those claims that involve a collision with at least two cars. To avoid a potential bias caused by unobservable accidents in type C, we employ the same criteria to identify a claim for all policies, i.e., $z_i = 1$ when an individual i files a claim caused by a collision with at least two cars; otherwise $z_i = 0$.

The estimated residuals of the above two probits can be calculated as follows:

$$\hat{e}_i = E(e_i \mid y_i) = \frac{f(X_i b_a)}{\Phi(X_i b_a)} y_i - (1 - y_i) \frac{f(X_i b_a)}{\Phi(-X_i b_a)}, \tag{3}$$

$$\hat{h}_i = E(h_i \mid z_i) = \frac{f(X_i b_b)}{\Phi(X_i b_b)} z_i - (1 - z_i) \frac{f(X_i b_b)}{\Phi(-X_i b_b)}, \tag{4}$$

where f is the density distribution function of $N(0,1)$, respectively.

To test the conditional dependent of $\hat{e}_i$ and $\hat{h}_i$, we follow Chiappori and Salanie (2000) and use a test statistic W[6]:

$$W = \frac{\left(\sum_{i=1}^{n} \hat{e}_i \hat{h}_i\right)^2}{\sum_{i=1}^{n} \hat{e}_i^2 \hat{h}_i^2}. \tag{5}$$

W is distributed asymptotically as a $c^2(1)$. We test its significance under the null hypothesis of $\operatorname{cov}(e_i, h_i) = 0$. If there is asymmetric information, the conditional dependence between the choice of coverage and the occurrence of the claim should be significantly positive.

---

6 In Chiappori and Salanie (2000) empirical dataset, the difference of the length of the policies comes from the mismatch of year and calendar year, so the W-statistic in their research needs a weight. Because our data are calculated on a calendar year basis, our W-statistic does not need the weight factor.

Using the whole sample, we report the conditional correlations between coverage and claim for each year in Table 1. Consistent with our hypothesis, we find that the correlation coefficients (r) between e_i and h_i in this table are all positive. Furthermore, the statistics (W) show that the correlations between the choice of coverage and the occurrence of a claim are significantly different from zero. Thus, the empirical evidence implies that asymmetric information problems do exist in Taiwan's automobile insurance market.

[Insert Table 1 about here]

It should be noted that many of the claims are generated by new cars and many new cars' owners who purchased the comprehensive insurance do not renew their policies for the following year. The means of z (the average claim frequency) for the sub-sample are 9.94%, 11.55%, and 4.65% in 1999, 2000, and 2001, respectively. However, the means of z for the entire sample are 27.44%, 31.01%, and 32.41% in 1999, 2000, and 2001, respectively. On the other hand, the means of $carage0$ (the proportion of new cars) for the sub-sample are 1.25%, 1.04%, and 0.83% in 1999, 2000, and 2001, respectively, whereas, the means of $carage0$ for the entire sample are 43.03%, 39.06%, and 31.24% in 1999, 2000, and 2001, respectively. It is generally believed in the insurance industry that new cars could contribute most of the asymmetric information problems. By comparing the empirical results from the entire sample and the sub-sample, we are able to further analyze whether the asymmetric information problems prevail in the market, no matter which policies result from new cars.

In Table 2, we perform Chiappori and Salanie's test by using the sub-sample data for the insured who continue to renew their policies in the following year. Consistent with the results in Table 1, all of the correlation coefficients (r) in Table 2 are still positive. The result once again confirms the existence of asymmetric information in the sub-sample group of the continuing renewal insured. However, we find all of the correlation coefficients (r) in Table 2 are respectively smaller than those in Table 1. It should be noted that the sub-sample consists of fewer new cars, whereas the entire sample includes many new cars. Thus, combing the results of Tales 1 and 2, our empirical evidence indicates that the insurance markets for both new cars and old cars could suffer from asymmetric information problems, although the insurance market for new cars could suffer more severe problems.

[Insert Table 2 about here]

5. Evidence for adverse selection

In this section, we directly track the existence of adverse selection. We use a sub-sample where the insured purchased insurance in year $t-1$ and renewed their policy in the following year t. In year t, the insurer could observe whether the individuals file a claim in the previous year $t-1$. Thus, this record could be indicated as a proxy of individual's risk type. The individual's contract choice at year t will be affected by all of the observable variables at year t and his or her risk type. To test whether a high risk type will choose high coverage, we use a probit model as the follows:

$$\text{Prob}(y_{it}=1)=\Phi(X_{it}b_c+b_d z_{it-1}), \tag{6}$$

where y_{it} is as defined in the previous section, X_{it} is the variable for individual i's information at year t, and b_c and b_d are the coefficient vectors of the regressor. When individual i files a claim caused by a collision with at least two cars in year $t-1$, then $z_{it-1}=1$; otherwise $z_{it-1}=0$.

In Equation (6), z_{it-1} is now as a proxy of individual's risk type. It does not mean that the individual's risk type is observable at year t. The insurer could only conclude that the insured has a high probability of being a high risk type when the insured files a claim at year $t-1$. If the sample of $z_{it-1}=1$ contains more high risk drivers, then the theory of adverse selection would predict a positive b_d. However, it is very important to recognize that there exists another force which could drive b_d being negative: experience rating. If the insured files a claim at year $t-1$, then he or she will face an increase of premium for all choice of contracts at year t. A high coverage contract would become less affordable to the insured who ever filed a claim in the previous year.

Thus, the effect of experience rating would lead the insured with previous claim records to choose a contract with less coverage at year t. Therefore, b_d would be negative.

In Table 3, we report the estimator between the choice of the renewal decision and the previous-year claim records. We find all the coefficients are significantly positive. This result confirms that a high risk type insured will choose high coverage contract. Our result support that the effect of experience rating is dominated by adverse selection. Combining the findings in Tables 2 and 3, our empirical results provide evidence to support that adverse selection contributes to asymmetric information problems in Taiwan's automobile insurance market.

[Insert Table 3 about here]

It should be noted that the evidence from Tables 1 and 2 supports the existence of asymmetric information problem in the insurance market, where evidence from Table 3 indicates that asymmetric information problems could be caused by adverse selection. However, we cannot exclude the possibility of the existence of moral hazard, since it is possible that both adverse selection and moral hazard might co-exist in the market.

6. Evidence for cross-subsidization

Finally, to track the existence of cross-subsidization, we construct a Tobit regression as follows:

$$LR_i = X_i b_e + b_f y_i + u_i , \quad (7)$$

where LR_i is the total claim amount divided by the premium for each individual. X_i and y_i are as defined above. It is important to recognize that LR_i has a mass on zero since most individuals do not file a claim. Thus, we use a Tobit regression to estimate the parameters in Equation (7). According to our model in Appendix B, we predict that b_f, the coefficient of y_i in Equation (7), should be significantly positive if cross-subsidization exists.

In Table 4, be using the renewal sample, we report b_f in the Tobit regression. If there is adverse selection but no cross-subsidization in the market, the choice of policy should not correlate to the loss ratio as shown in Appendix A. From Table 4, we find that all the coefficients are significantly positive, which implies that the insured who buy high coverage policies tend to have high loss ratios. The evidence from Tables 3 and 4 together imply that the insurance market may exist both adverse selection and cross-subsidization. Moreover, since moral hazard can not make individual's choice of the policy and the individual's loss ratio correlated as shown by Appendix C, evidence from Table 4 could further support the existence of adverse selection.

[Insert Table 4 about here]

7. Conclusion

Identifying the source of asymmetric information is an important issue in the insurance literature. Previous studies have separated moral hazard from adverse selection by using either dynamic data (Chiappori and Heckman, 2000) or the particularity of a sample set (Dionne and Gagne, 2002; and Finkelstein and Poterba, 2004). In this paper, we propose a new method to directly examine the existence of adverse selection. Using three-year individual data from a large insurance company in Taiwan, this paper analyzes whether adverse selection exists in the insurance market. We use renewal decisions to track adverse selection problems. By adopting claim records as a proxy of the insured's risk type, we allege that only adverse selection, not moral hazard, can explain the relationship between the previous-year claims and the choice of coverage in the following year. We further argue that the individual's loss ratio will be positively correlated to individual's choice of coverage if there exist both adverse selection and cross-subsidization in the market.

Applying the same approach as Chiappori and Salanie (2000), our empirical evidence confirms that asymmetric information problems do exist in Taiwan's comprehensive automobile insurance market. Furthermore, our results make a new contribution to the insurance literature by further directly tracking the existence of adverse selection. Our empirical evidence shows that the correlation between the choice of insurance coverage and the previous-year claim records is significantly positive, and the loss ratio is also significantly positively correlated to the choice of coverage. Thus, we conclude that adverse selection exists in Taiwan's comprehensive automobile insurance market.

It should be also noted that our findings actually do not refute but rather supplement those papers that found no evidence to support the existence of asymmetric information. First, we employ data from comprehensive automobile insurance, which was not yet used to examine asymmetric information problems in the literature. Second, we use data in the early stage of the developing insurance market, while most of the data used in the literature has come from a well-developed market. It is possible that asymmetric information exists in the early stage of the insurance market, as described by Rothschild and Stiglitz (1976). However, over the years, insurance companies collect more and more information to screen the insured and eventually eliminate asymmetric information problems in the market. Thus, our findings suggest that further investigation of asymmetric information problems by comparing different insurance markets or data from different countries is definitely needed.

References

Abbring, J.H, P. Chiappori, J.J. Heckman, and J. Pinquet, 2003, Adverse Selection and Moral Hazard in Insurance: Can Dynamic Data Help to Distinguish? *Journal of the European Economic Association*, 1: 512-521.

Abbring, J.H. P. Chiappori, and J. Pinquet, 2003, Moral Hazard and Dynamic Insurance Data, *Journal of the European Economic Association*, 1: 767-820.

Arnott, R., and J.E. Stiglitz, 1988, Randomization with Asymmetric Information, *Rand Journal of Economics*, 19: 344-362.

Cardon, J.H., and I. Hendel, 2001, Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey, *Rand Journal of Economics*, 32: 408-427.

Cawley, J., and T.J. Philipson, 1999, An Empirical Examination of Information Barriers to Trade in Insurance, *American Economic Review*, 89: 827-846.

Chiappori, P., and J. Heckman, 2000, Testing for Moral Hazard on Dynamic Insurance Data: Theory and Econometric Tests," working paper.

Chiappori, P., and B. Salanie, 1997, Empirical Contract Theory: The Case of Insurance Data, *European Economic Review*, 41: 943-950.

Chiappori, P., and B. Salanie, 2000, Testing for Asymmetric Information in Insurance Markets, *Journal of Political Economy*, 108: 56-78.

Crocker, K.J., and A. Snow, 1985, The Efficiency of Competitive Equilibria in Insurance Markets with Asymmetric Information, Journal of Public Economics, 26: 207-219.

Dionne, G., and R. Gagne, 2002, Replacement Cost Endorsement and Opportunistic Fraud in Automobile Insurance, *Journal of Risk and Uncertainty*, 24: 213-230.

Dionne, G., C. Gourieroux, and C. Vanasse, 2001, Testing for Evidence of Adverse

Selection in the Automobile Insurance Market: A Comment, *Journal of Political Economy*, 109: 444-455.

Dionne, G., and P. Lasserre, 1985, Adverse Selection, Repeated Insurance Contracts and Announcement Strategy, *Review of Economic Studies*, 52: 719-723.

Dionne, G., and P. Lasserre, 1987, Adverse Selection and Finite-Horizon Insurance Contracts, *European Economic Review*, 31: 843-862.

Finkelstein, A., and J.M. Poterba, 2004, Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market, *Journal of Political Economy*, 112: 183-207.

Makki, S.S., and A. Somwaru, 2001, Evidence of Adverse Selection in Crop Insurance Markets, *Journal of Risk and Insurance*, 68: 685-708.

Miyazaki, H., 1977, The Rate Race and Internal Labour Market, *Bell Journal of Economics*, 8: 394-418.

Puelz, R., and A. Snow, 1994, Evidence on Adverse Selection: Equilibrium Signaling and Cross-Subsidization in the Insurance Market, *Journal of Political Economy*, 102: 236-257.

Rothschild, M., and J.E. Stiglitz, 1976, Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information, *Quarterly Journal of Economics*, 90: 629-649.

Shavell, S., 1979, On Moral Hazard and Insurance, *Quarterly Journal of Economics*, 93: 541-562.

Spence, M., 1978, product Differentiation and Performance in Insurance Markets, *Journal of Public Economics*, 10: 427-447.

Stiglitz, J. 1977, Monopoly, Nonlinear Pricing, and Imperfect Information: The Insurance Market, *Review of Economic Studies*, 44: 407-430.

Wilson, C.A., 1977, A Model of Insurance Markets with Incomplete Information, *Journal of Economic Theory*, 16: 167-207.

Winter, R.A., 2000, Optimal Insurance under Moral Hazard, Handbook of Insurance, Edited by G. Dionne, Boston, Dordrecht and London: Kluwer Academic.

Table 1 The Conditional Correlation Between Coverage and Claim in Comprehensive Automobile Insurance in Years 1999 to 2001

Year	r	W
1999	0.2845***	24.3063***
2000	0.3863***	294.970***
2001	0.4506***	1743.49***

Note: The significant level of 99% is denoted by ***

Table 2 The Conditional Correlation Between Coverage and Claim in the Same Year (Sub-sample of Continued Insured)

Year	r	W
1999	0.2217***	295.973***
2000	0.3330***	10.1900***
2001	0.2880***	75.5288***

Note: The significant level of 99% is denoted by ***

Table 3 The relationship Between Coverage and previous Claim records (Sub-sample of Continued Insured)

Year	b_d
1999	0.8507***
2000	1.0275***
2001	0.8892***

Note: The significant level of 99% is denoted by ***

Table 4 The Coefficient of Risk Type for Loss Ratio in Tobit Regression

Year	b_f
1999	1.7613***
2000	1.8566***
2001	2.0730***

Note: The significant level of 99% is denoted by ***

Appendix A

Assume that individuals face a binominal property risk with either a fixed loss with a probability p_i or no loss with $1-p_i$, where i denotes risk types in the economy, $i \in \{h,l\}$ and $1 > p_h > p_l > 0$. Let the fraction of h risk type agents is q. The types are the insured' s private information and are unobservable to the insurer. For the sake of simplicity, assume all individuals have the same initial wealth w, loss amount L, and utility function U with $U' > 0$ and $U'' \leq 0$.

On the basis of Rothschild and Stiglitz's (1976) separating equilibrium, a risk-neutral insurer with a cost loading l in a competitive insurance market will offer two contracts to the market, and each contract in the equilibrium set makes zero profit. Points B and G in Figure A1 denote the optimal contracts for high and low-risk type individuals, respectively. A high-risk type insured is indifferent about choosing points B or G, whereas a low-risk type insured prefers point G.

Zero profit contracts in the equilibrium suggest that point B contains high coverage Q_h for a high insurance premium $(1+l)p_h Q_h$, and point G contains low coverage Q_l for a low insurance premium $(1+l)p_l Q_l$. Let E denote the expectation operator. Thus, the expected value of loss ratio LR_i for the i type insured will become $E(LR_i) = \dfrac{p_i Q_i}{(1+l)p_i Q_i} = \dfrac{1}{1+l}$, which is independent of risk type i.

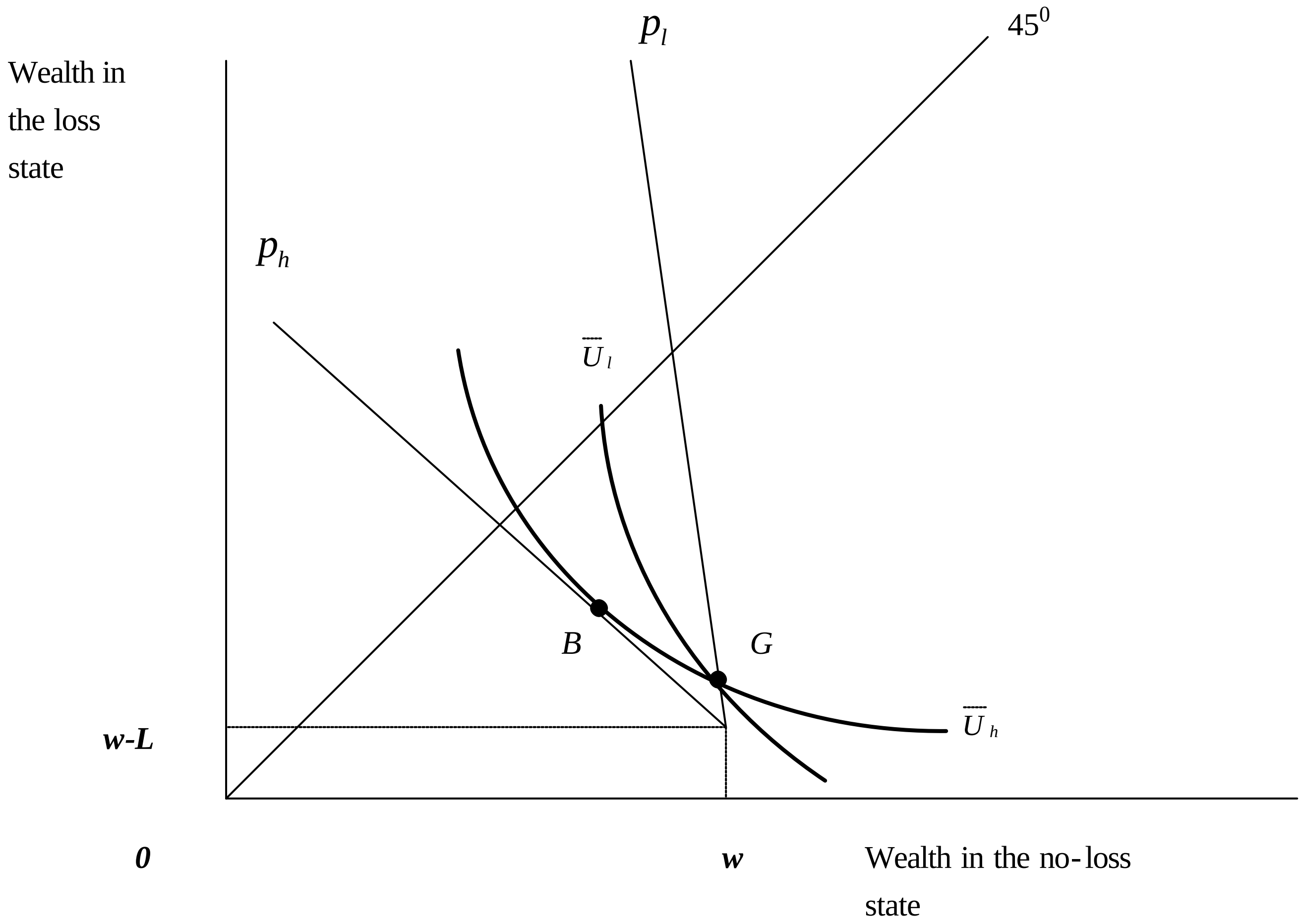

Figure A1 Separating Equilibrium

Appendix B

Based upon the model setting in Appendix A, we now consider Miyazaki's (1977) separating equilibrium with cross-subsidization. As proposed by Puelz and Snow (1994), under cross-subsidization, the insurance premium for the i type insured, $i \in \{h, l\}$, would be

$$P_i = (1 + l)\left[p_i Q_i + k_i\right], \tag{B1}$$

where k_i is a charge for cross-subsidization among contracts. Since the insurance market is under perfect competition, the insurer will earn zero expected profit, which implies that cross-subsidization across risk types will net out to zero, i.e., $\sum_i w_i k_i = 0$, where w_i denotes the proposition of the insured with risk type i. In Miyazaki's model, the insurer will earn profit on policies purchased by low risk type insured but incur losses on those purchased by high risk type insured. In other words, $k_h < 0 < k_l$. Therefore, the expected loss ratio of high risk types will be greater than that of low risk types, since

$$\begin{aligned} & E(LR_h) - E(LR_l) \\ &= \frac{p_h Q_h}{(1+l)\left[p_h Q_h + k_h\right]} - \frac{p_l Q_l}{(1+l)\left[p_l Q_l + k_l\right]} \\ &= \frac{p_h Q_h k_l - p_l Q_l k_h}{\left[p_h Q_h + k_h\right]\left[p_h Q_h + k_h\right]} \\ &> 0. \end{aligned} \tag{B2}$$

Appendix C

In this appendix, we adopt Winter's (2000) hidden action model of moral hazard to demonstrate that the expected loss ratio is uncorrelated to the insurance coverage. Assume that individuals face a binominal property risk with either a fixed loss L or no loss. Individuals could employ self-protection with cost x to reduce the loss probability. $p(x)$ with $p'(x) < 0$ and $p''(x) > 0$. Although the self-protection behavior could not be observed by the insurer, the insurer could still estimate the probability of risk occurrence. Thus, the risk neutral insurer would offer only a coverage Q for premium $(1 + l)p(x)Q$, where l denotes the insurance loading. The expected loss ratio LR --- $E(LR) = \dfrac{p(x)Q}{(1 + l)p(x)Q} = \dfrac{1}{1 + l}$ --- will be independent of the insurance coverage.

Appendix D

Table D1 Definitions of the Variables

Variables	Definition
y	a dummy variable equals 1 when an individual chooses a type A or B policy, otherwise equals 0
k	a dummy variable equals 1 when an individual chooses a type A or B policy in the next year, otherwise equals 0
z	a dummy variable equals 1 when an individual's claim is caused by a collision and the claim amount is above the threshold amount, otherwise equals 0
LR	loss ratio, equals loss amount divided by the premium
carage0	a dummy variable equals 1 when the car is new, otherwise equals 0
carage1	a dummy variable equals 1 when the age of the car is one year, otherwise equals 0
carage2	a dummy variable equals 1 when the age of the car is two years, otherwise equals 0
carage3	a dummy variable equals 1 when the age of the car is three years, otherwise equals 0
carage4	a dummy variable equals 1 when the age of the car is four years, otherwise equals 0
carage5	a dummy variable equals 1 when the age of the car is five years, otherwise equals 0
carage6	a dummy variable equals 1 when the age of the car is six years, otherwise equals 0
carage7	a dummy variable equals 1 when the age of the car is seven years, otherwise equals 0
carage8	a dummy variable equals 1 when the age of the car is eight years, otherwise equals 0
carage9	a dummy variable equals 1 when the age of the car is nine years, otherwise equals 0
carage10	a dummy variable equals 1 when the age of the car is ten years, otherwise equals 0
carage11	a dummy variable equals 1 when the age of the car is eleven years, otherwise equals 0
sexf	a dummy variable equals 1 when the owner of the car is female, otherwise equals 0
married	a dummy variable equals 1 when the owner of car is married
city	a dummy variable equals 1 when the owner of the car lives in the city, otherwise equals 0
arean	a dummy variable equals 1 when the car is registered in the north of Taiwan, otherwise equals 0
areas	a dummy variable equals 1 when the car is registered in the south of Taiwan, otherwise equals 0
areaeast	a dummy variable equals 1 when the car is registered in the east of Taiwan, otherwise equals 0
catpcd_1	a dummy variable equals 1 when the car is a sedan and is for non-commercial or long-term rental purposes, otherwise equals 0
catpcd_2	a dummy variable equals 1 when the car is a small freight-truck and is for non-commercial purposes or for business use, otherwise equals 0
tramak_i	*i=n,f,h,t,c*, dummy variable equals 1when the trademark of the car is the assigned brand, otherwise equals 0
age2	a dummy variable equals 1 when the age of the insured is between 30 and 25, otherwise equals 0
age3	a dummy variable equals 1when the age of the insured is between 60 and 30, otherwise equals 0
age4	a dummy variable equals 1 when the age of the insured is above 60, otherwise equals 0

Table D2 Summary Statistics for All Sample

year	1999		2000		2001	
Variable	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
z	0.2744	0.4462	0.3101	0.4625	0.3241	0.4681
y	0.5882	0.4922	0.6059	0.4887	0.5979	0.4903
LR	0.3769	1.2220	0.3964	1.1784	0.4113	1.2741
carage0	0.4303	0.4951	0.3906	0.4879	0.3124	0.4635
carage1	0.2441	0.4295	0.2049	0.4036	0.2122	0.4088
carage2	0.1431	0.3501	0.1595	0.3661	0.1446	0.3517
carage3	0.0746	0.2628	0.1020	0.3026	0.1245	0.3301
carage4	0.0486	0.2150	0.0605	0.2384	0.0804	0.2720
carage5	0.0302	0.1711	0.0395	0.1948	0.0483	0.2144
carage6	0.0157	0.1244	0.0231	0.1501	0.0321	0.1762
carage7	0.0076	0.0868	0.0111	0.1047	0.0178	0.1323
carage8	0.0029	0.0538	0.0054	0.0731	0.0081	0.0898
carage9	0.0014	0.0370	0.0021	0.0455	0.0037	0.0611
carage10	0.0008	0.0276	0.0008	0.0282	0.0016	0.0406
carage11	0.0002	0.0123	0.0005	0.0228	0.0006	0.0242
Sexf	0.6312	0.4825	0.6551	0.4753	0.6584	0.4742
Marria	0.5531	0.4972	0.6582	0.4743	0.7763	0.4167
City	0.5199	0.4996	0.5361	0.4987	0.5547	0.4970
arean	0.4635	0.4987	0.4752	0.4994	0.4998	0.5000
Areas	0.2737	0.4458	0.2695	0.4437	0.2508	0.4335
areaeast	0.0338	0.1806	0.0342	0.1816	0.0331	0.1789
catpcd_1	0.9717	0.1659	0.9744	0.1578	0.9777	0.1475
catpcd_2	0.0249	0.1559	0.0225	0.1482	0.0198	0.1393
Tramak_n	0.1478	0.3549	0.1277	0.3338	0.1346	0.3413
Tramak_f	0.1412	0.3482	0.1331	0.3397	0.1157	0.3198
Tramak_h	0.0817	0.2739	0.0902	0.2865	0.0873	0.2823
Tramak_t	0.3379	0.4730	0.3741	0.4839	0.3637	0.4811
Tramak_c	0.1159	0.3201	0.0890	0.2847	0.0827	0.2755
age2	0.1156	0.3197	0.1007	0.3009	0.0863	0.2809
age3	0.8254	0.3796	0.8684	0.3380	0.8602	0.3468
age4	0.0202	0.1405	0.0175	0.1312	0.0258	0.1584
Numbers of observations	59186		61627		64891	

Table D3 Summary Statistics for Renewal Sample

year	1999		2000		2001	
Variable	Mean	Std Dev	Mean	Std Dev	Mean	Std Dev
z	0.0994	0.2992	0.1155	0.3196	0.0465	0.2107
y	0.5523	0.4973	0.5421	0.4982	0.5305	0.4991
LR	0.4925	0.5000	0.4903	0.4999	0.4681	0.4990
carage0	0.0125	0.1111	0.0104	0.1013	0.0083	0.0908
carage1	0.2789	0.4485	0.2393	0.4266	0.2073	0.4054
carage2	0.2738	0.4459	0.2193	0.4138	0.2103	0.4075
carage3	0.1835	0.3871	0.2068	0.4050	0.1776	0.3822
carage4	0.1073	0.3095	0.1370	0.3439	0.1623	0.3688
carage5	0.0688	0.2531	0.0816	0.2738	0.1042	0.3055
carage6	0.0404	0.1969	0.0533	0.2246	0.0601	0.2377
carage7	0.0194	0.1379	0.0289	0.1675	0.0367	0.1881
carage8	0.0093	0.0960	0.0133	0.1146	0.0199	0.1397
carage9	0.0036	0.0599	0.0059	0.0768	0.0080	0.0890
carage10	0.0013	0.0364	0.0023	0.0484	0.0035	0.0589
carage11	0.0010	0.0320	0.0010	0.0311	0.0011	0.0334
Sexf	0.6183	0.4858	0.6362	0.4811	0.6566	0.4748
Marria	0.7472	0.4346	0.8318	0.3740	0.9281	0.2584
City	0.5535	0.4971	0.5640	0.4959	0.5720	0.4948
arean	0.4937	0.5000	0.5098	0.4999	0.5230	0.4995
Areas	0.2481	0.4319	0.2402	0.4272	0.2383	0.4261
areaeast	0.0338	0.1808	0.0310	0.1735	0.0292	0.1684
catpcd_1	0.9755	0.1547	0.9782	0.1459	0.9808	0.1372
catpcd_2	0.0210	0.1433	0.0185	0.1346	0.0165	0.1276
Tramak_n	0.1536	0.3606	0.1424	0.3495	0.1479	0.3550
Tramak_f	0.1214	0.3266	0.1118	0.3152	0.1045	0.3059
Tramak_h	0.0834	0.2765	0.0883	0.2837	0.0856	0.2797
Tramak_t	0.3250	0.4684	0.3446	0.4752	0.3462	0.4758
Tramak_c	0.1055	0.3071	0.0925	0.2897	0.0872	0.2822
age2	0.0800	0.2711	0.0627	0.2424	0.0569	0.2317
age3	0.8917	0.3107	0.8811	0.3237	0.9106	0.2853
age4	0.0215	0.1449	0.0331	0.1789	0.0285	0.1664
N*umbers of observations*	26420		28986		31305	