Asia Pacific Management Review 21 (2016) 231–238

HOSTED BY

ELSEVIER

Contents lists available at ScienceDirect

Asia Pacific Management Review

journal homepage: www.elsevier.com/locate/apmrv

An examination of the relationship between vehicle insurance purchase and the frequency of accidents

Yung-Ching Hsu [a], Pai-Lung Chou [b], Yung-Ming Shiu [c, *]

[a] Investigation Section, Civil Service Ethics Office, Kaohsiung City Government, Kaohsiung, Taiwan, ROC
[b] Department of Risk Management and Insurance, National Kaohsiung First University of Science and Technology, Kaohsiung, Taiwan, ROC
[c] Department of Risk Management and Insurance, College of Commerce, National Chengchi University, Taipei, Taiwan, ROC

ARTICLE INFO

Article history:
Received 27 June 2016
Accepted 19 August 2016
Available online 5 December 2016

Keywords:
Vehicle accident occurrences
Insurance coverage
Moderating effects
Vehicle age

ABSTRACT

The relationship between insurance, occurrences of road traffic accidents (RTAs) and general traffic safety has received growing attention over recent years among academics, industry practitioners and government policymakers. Using data on vehicle damage insurance in Taiwan, we examine whether drivers with higher insurance coverage are more likely to be involved in RTAs, and whether the relationship is moderated by the gender of the insured party as well as the age of both the vehicle and the insured party. Using a probit regression, we identify a positive relationship between coverage and claims and find that an insured party with a poor claims history has a higher probability of being involved in RTAs. Although our findings provide support for adverse selection theory, when considering the moderating effect of vehicle age, the positive relationship between coverage and claims becomes insignificant; indeed, vehicle age weakens the positive influence of coverage on claims. Our results suggest that drivers with a poor driving record purchasing higher insurance coverage for their new vehicle tend to be involved in more RTAs and submit more insurance claims.

Introduction

Annual statistics on road traffic accidents (RTAs) in the US show that the number of RTA fatalities in 2010 fell to an all-time low since records were first collated in 1950 (NHTSA, 2010); however, this is not necessarily the case in other countries, particularly those countries that are rapidly becoming motorized. A very recent study reported that in 1990, road traffic injuries (RTIs) were ranked the ninth leading cause of the 'global disease burden' (Chekijian et al., 2014), and indeed, the World Health Organization has forecasted that by 2030, RTIs will become the fifth leading cause of this burden (WHO, 2012).

Public attention to the remarkably unacceptable death toll from RTAs has grown over recent decades, with the '*Global Status Report on Road Safety, 2013*' reporting that the total number of RTA fatalities across the world currently stands at 1.24 million per year (WHO, 2014). As a result, governments across the world have been placing considerable effort into enhancing road safety by imposing relevant laws and investing in highway capital (Nguyen-Hoang & Yeung, 2014), whilst vehicle manufacturers have also placed emphasis on improvements in vehicle design, with the common goal of reducing the frequency and severity of traffic accidents. However, regardless of the amount of effort expended, RTAs cannot be completely avoided; thus, the goal must essentially be to reduce their frequency and magnitude.

In order to finance the monetary losses arising from RTAs, drivers may consider purchasing vehicle insurance, such as covering for physical damage or losses resulting from collisions, theft or other unfortunate events. It is, however, a legal requirement in many countries, including Taiwan, for drivers to have valid liability insurance coverage protecting the insured party against any legal liability arising from accidents causing bodily injury and/or damage to others persons or their vehicles.

This study has a twofold purpose. Firstly, we examine the ways in which the purchase of vehicle insurance affects occurrences of RTAs. Two prominent theories, 'adverse selection' and 'propitious selection', are proposed within the insurance literature as the means of predicting and explaining the relationship between the

* Corresponding author.
Peer review under responsibility of College of Management, National Cheng Kung University.

http://dx.doi.org/10.1016/j.apmrv.2016.08.001
1029-3132/

purchase of vehicle insurance and RTA occurrences. In the present study, we aim to investigate which theory is supported by data on collision insurance in Taiwan. Secondly, we investigate the effects on this relationship that may be attributable to the gender of the insured party and the age of both the vehicle and the driver. To the best of our knowledge, no prior empirical research has been undertaken with the aim of examining the possible moderating effects on the relationship between purchase of vehicle insurance and occurrences of RTAs.

The debate continues within the literature with regard to whether 'adverse selection' theory is actually supported by the available empirical data. As noted by Karagyozova and Siegelman (2012), the term 'adverse selection' was first coined in the nineteenth century, with the related theory having subsequently been proposed and formalized by Akerlof (1970) and Rothschild and Stiglitz (1976). Adverse selection, which arises from asymmetric information, in an insurance market describes the fact that insurance buyers possess residual private information about their risk that insurers lack even after risk classification (Shi, Zhang, & Valdez, 2012). According to adverse selection theory, high-risk individuals are more likely to purchase higher levels of insurance coverage, as a result of which, these individuals will tend to have a higher probability of experiencing a loss. Riskier drivers will therefore tend to buy higher coverage and will also tend to submit more claims. This inference gives rise to a positive correlation between the amount of insurance coverage and the *ex post* occurrence of the insured risk.

Such a positive relationship is found between coverage and claims in several of the prior studies (Puelz & Snow, 1994; Shi et al., 2012; Li, Liu & Peng, 2013); however, there are also numerous examples of other studies where no evidence of adverse selection is discernible (Chiappori & Salanié, 2000; Dionne, Gouriéroux & Vanasse, 2001; Saito, 2006). The theory of 'propitious selection' (or 'advantageous selection') is one particular argument proposed in these studies as the means of explaining the absence of any relationship between coverage and claims. Similar to adverse selection, propitious selection is also a choice made by the insured party, although in this case, the choice is advantageous to the insurer.

Another major difference between the two theories is that those advocating propitious selection theories argue that there are factors relating to risk aversion that are important in determining insurance coverage purchase, factors that are not taken into consideration in adverse selection theory. These factors include, for example, the age and gender of the insured party and the age of the vehicle, each of which are considered to have an impact on the likelihood of incurring losses.

Unlike adverse selection theory, the argument in support of propitious selection theory posits that individuals who are highly risk-averse would be more likely to purchase greater insurance cover and take more physical precautions, thereby suggesting that such individuals will have a lower probability of being involved in RTAs (e.g., De Meza & Webb, 2001; Hemenway, 1990, 1992). Hemenway (1990) found weak evidence that automobile renters who wore their seat belt tend to buy collision damage waiver insurance offered by car rental companies. Consistent with the theory of propitious selection, Hemenway (1992) found that drivers who purchase vehicle liability coverage are less likely to engage in drink driving and are more likely to engage in risk reduction behaviors. Based upon this argument, we would expect to find a negative correlation between coverage and claims.

It is worthwhile to note that adverse selection theory and propitious selection theory make different predictions about the relationship between coverage and claims. The former predicts a positive relationship, while the latter predicts a negative relationship. The former argues that risky drivers would be more likely to purchase more insurance. Since risky drivers have a higher likelihood of being involved in RTAs, a positive relationship between coverage and claims can therefore be observed. Conversely, the latter argues that risk-averse drivers tend to purchase more insurance. Because these drivers are less likely to be involved in RTAs, a negative relation is thus observed.

Like adverse selection theory, moral hazard theory also predicts a positive relation between coverage and claims. In the case of adverse selection, drivers are assumed to have private information about their risk type and preference, conditional on the insurer's risk classification of the buyer of insurance. Those with private information that they are high risk or risk-loving would purchase more insurance than those with private information that they are low risk (Finkelstein & McGarry, 2006). However, moral hazard theory argues that once drivers purchase more insurance, they tend to be riskier and are more likely to be involved in RTAs. Spindler, Winter and Hagmayer (2013) indicated that moral hazard deals with 'hidden action', while adverse selection concerns 'hidden information.' In our paper, we attempt to distinguish between moral hazard and adverse selection by examining the correlation between prior and future claims (Abbring, Chiappori, Heckman & Pinquet, 2003; Cohen & Siegelman, 2010).

Using 1998–1999 data from the Auckland Car Crash Injury Study, Blows, Ivers, Connor, Ameratunga, and Norton (2003) found that uninsured drivers were more likely to suffer car crash injuries than insured drivers, a finding which provides some support for the theory of propitious selection, insofar as less risk-averse drivers may choose not to purchase insurance and yet have a greater likelihood of being involved in RTAs.

It should be noted that under both the adverse selection and propitious selection arguments, it is assumed that insured parties have an informational advantage over insurers. In the case of adverse selection theory, this advantage is revealed by riskier individuals purchasing higher levels of insurance coverage and having a greater likelihood of being involved in accidents; conversely, under the propitious selection argument, the advantage is revealed by risk-averse individuals purchasing higher levels of insurance, but with a lower probability of being involved in accidents.

The prior studies on traffic safety have tended to concentrate primarily on the prevention of road accidents, including a reduction in total injuries and fatalities, by addressing issues such as the use of seat-belts (Farmer & Wells, 2010), drinking and driving (Sloan, Chepke & Davis, 2013), speeding (Ardeshiri & Jeihani, 2014), the wearing of helmets (Bonander, Nilson & Andersson, 2014) and the use of child restraints (Romano & Kelley-Baker, 2015).

It appears, however, that little research, if any, has been undertaken with specific focus on the ways in which insurance purchase behavior may influence the probability of RTAs from the perspective of traffic safety; the present study therefore aims to fill this gap in the literature. The prior insurance literature also suggests that there are other factors, such as risk aversion, that may well offset the positive correlation between coverage and claims (Hemenway, 1990, 1992; Shi et al., 2012). In contrast to these studies, in the present study we argue that the absence of a relationship between coverage and claims is possibly due to the moderating effect arising from certain factors known to the insurer, such as the age of the vehicle.

Using data on vehicle damage insurance contracts in Taiwan, we find that drivers with higher levels of insurance coverage and a poor driving history are more likely to submit claims, thereby indicating the existence of adverse selection. However, we also find that this positive relationship between coverage and claims is weakened by the age of the vehicle, thereby suggesting a weaker coverage-claims correlation for the older vehicles.

The remainder of this paper is organized as follows. The background on Taiwanese vehicle insurance market is provided in section The taiwanese vehicle insurance market. The data and methodology adopted for the analyses undertaken in this study are described in section Data and methodology, followed in section Empirical results by the presentation and interpretation of our empirical results. Finally, the conclusions drawn from this study are presented in section Empirical results.

The taiwanese vehicle insurance market

In 2009, vehicle insurance is the major line of business for non-life insurance companies in Taiwan, accounting for 49.36% of the total premiums received by the whole non-life insurance industry (Taiwan Insurance Institute, 2015). Vehicle insurance includes three main types of insurance: vehicle damage, theft and liability insurance. In this paper we use data on vehicle damage insurance because it is voluntary in Taiwan. We do not use data on vehicle liability insurance because drivers are required to purchase compulsory vehicle liability insurance which only covers medical expenses associated with human injury and death. The maximum amount of compensation is approximately equivalent to US$ 73,000. Drivers who consider this amount inadequate would voluntarily purchase more coverage on vehicle liability. Drivers who wish to have insurance on damage inflicted to other drivers' cars need to buy voluntary vehicle liability insurance.

The aim of this paper is to examine the relation between insurance purchase and incidences of traffic accidents. It is therefore obvious that data on compulsory insurance cannot be used to examine this relation. Since all drivers in Taiwan need to purchase compulsory vehicle liability insurance, using voluntary vehicle liability is also inappropriate and the relation between insurance purchase and the frequency of traffic accidents would be less statistically significant.

It is worthwhile to note that the bonus-malus system, which is based on experience rating, has been effective in Taiwan since 1996. Under this system, the vehicle damage insurance premium rate of policyholders is determined by the insured- and insured vehicle-coefficients. The insured-coefficients include gender-age and past claims coefficients, while the insured vehicle-coefficients are related to the age and cubic capacity of the insured vehicle.

The bonus-malus system is symmetrical. To be more specific, the bonus element of this system encourages safe driving, while the malus element penalizes bad driving record. Drivers with no claims in the previous year have their premiums discounted by 20 per cent. Those without claims for two (three) consecutive years have their premiums discounted by 40 (60) per cent. Conversely, drivers with two (three, four, …) claims during the past three years will have their premiums increased by 20 (40, 60, …) percent.

Data and methodology

Data

The data on vehicle damage insurance contracts used in this study is obtained from the second largest non-life insurer in Taiwan. This insurer's market share was 11.17 per cent in terms of gross premiums written. Our study sample comprises of a total of 726 observations, with the contracts, covering the 2009 policy year, being written in the 2009 year and effective for the following 12-month period. Vehicle damage insurance provides protection for the insured vehicle if it is damaged in an accident.

A two-by-two contingency chart depicting the two-way frequency of the relationship between coverage and claims is presented in Table 1. As the table shows, 32.20 per cent (66/205) of all policyholders with high coverage levels are found to have submitted claims, whereas claims are found to have been submitted by only 18.62 per cent (97/521) of those policyholders with low coverage levels.

Table 1
Two-way frequency of variables.

		Claims		Totals
		0	1	
Coverage	0	424	97	521
	1	139	66	205
Totals		563	163	726

Note: Chi-squared statistic = 15.576 with statistical significance at the 1 per cent level ($p = 0.000$).

Of all drivers who had submitted claims, 40.49 percent (66/163) tended to purchase higher insurance coverage, and indeed, only 24.69 per cent (139/563) of drivers who had purchased higher insurance coverage did not submit claims. The Chi-squared (χ^2) statistic of 15.576 ($p = 0.000$) clearly indicates the existence of a positive relationship between coverage and claims, with this preliminary finding suggesting that insured parties with higher insurance coverage are more likely to submit claims.

Methodology

As indicated in Cohen and Siegelman (2010), an intuitive way of examining the relationship between coverage and claims is to run the probit regression as follows:

$$Claim_i = f(Coverage_i + CV_i) + e_i \tag{1}$$

where $Claim_i$ is a dummy variable which is equal to 1 if policyholder i submits one or more claims; otherwise 0; and $Coverage_i$ refers to the coverage choice of policyholder i; this variable is a dichotomous variable which is equal to 1 for high coverage and 0 for low coverage; CV_i is a set of control variables (to be defined below); e_i is a classic error term. We assume that the error term is independent of all explanatory variables and has the standard normal distribution.

Several related studies within the literature have identified a number of variables that are found to have some effect on the dependent variable; hence, these variables are included within the regressions as the control variables in the present study (Saito, 2006; Kim, Kim, Im & Hardin, 2009; Shi et al., 2012). These control variables include the age of the insured party, the gender of the insured party, the age of the insured vehicle, the 'bonus-malus' coefficient, the location of the insured vehicle, whether the vehicle is imported or locally produced, the vehicle capacity and whether the contract is a renewal contract.

Typically, young drivers are less experienced in driving and thus more likely to be involved in RTAs (Shi et al., 2012). Moreover, female drivers are also considered to be safer than their male counterparts. The 'bonus-malus' coefficient represents the policyholder's past claim history. It is expected that drivers with a worse claim history are more likely to file claims. Prior studies (e.g, Paefgen, Staake, & Fleisch, 2014) also argue that geographical zones are related to the probability of filing claims. It is obvious that driving in a busier area is more likely to have traffic accidents. Cohen (2005) found that the value of insured vehicle is positively related to the likelihood of having claims. Since domestically produced vehicles generally are cheaper than imported vehicles, we expect that domestically produced vehicles are less likely to have traffic accidents. Since the value of vehicles generally depreciates very quickly, drivers are less motivated to renew the insurance contracts on their vehicles. Those who do renew their insurance contracts generally are more risk-averse than those who do not. It is thus expected that policyholders that renew their contracts are less prone to submitting claims. All of the variables used in this study, along with their definitions, are presented in Table 2. It is worthwhile to note that these variables are observable to the insurer and used in pricing insurance policies.

As stated above, a positive relation between coverage and claims could indicate the existence of adverse selection or moral hazard. One approach for effectively distinguishing between adverse selection and moral hazard is to examine the correlation between prior and future claims (Abbring et al., 2003; Cohen & Siegelman, 2010). Under adverse selection, the prior claims records of insured parties reflect their level of riskiness, which would basically remain unchanged after either an initial insurance purchase or an increase in the level of insurance coverage; that is, insured parties with a history of numerous claims would continue to be prone to accidents in the future. Thus, we would expect to find a positive relationship between prior and future claims. In the case of moral hazard, since insured parties have less incentive to take precautions to prevent accidents from occurring, we would expect to find higher coverage leading to a lower level of caution, which, in turn, indicates a higher probablility of claims.

We also consider potential cross-effects on the dependent variable; one important dimension potentially moderating the relationship between coverage and claims is the age and the gender of the insured party. These two variables are also referred to as the risk-aversion variables, essentially because they are related to the level of risk aversion of the insured party. Individuals are generally considered to become more risk averse with an increase in age (Morin, & Suarez, 1983), and indeed, women are generally considered to be more risk averse than men (Borghans, Golsteyn, Heckman, & Meijers, 2009).

Both the theories of adverse selection and propitious selection take into consideration the risk aversion of the insured party. It is worthwhile to note that under adverse selection theory risk aversion is assumed constant across individuals, while it is not under propitious selection theory. A risk-averse person prefers a certain amount of wealth to a risky situation yielding the same expected wealth, so more risk-averse drivers would require a higher risk premium to induce them to accept the risk (Harrington & Niehaus, 2004). Thus, *ceteris paribus*, there will be a greater likelihood of a more risk-averse individual purchasing higher insurance coverage whilst also taking greater precautions to prevent any occurrence of loss.

The risk-aversion variables in this study, both of which are included in Equation (2), are the *Insured-Age* and *Gender*. We expect to find that the positive relationship between coverage and claims, as predicted by the theory of adverse selection, may be weakened by the risk-aversion variables, perhaps even becoming negative. We also anticipate that the positive relationship posited by adverse selection theory will again be weakened, or indeed, become a negative relationship, for older insured drivers and female drivers.

We further predict that the age of the insured vehicle may well prove to moderate the relationship between vehicle damage insurance coverage and claims; our prediction is primarily based upon the argument that old cars are generally driven by safe drivers (Shi et al., 2012). Since vehicles are insured at actual cash value (which is equal to replacement cost minus any depreciation), older vehicles tend to have smaller coverage than newer vehicles. Moreover, as the vehicle depreciates over time, one tends to purchase less coverage for it (Shi et al., 2012). We therefore further expect to find the age of the insured vehicle weakening the positive relationship between coverage and claims.

Equation (2), which includes the three moderator variables referred to above, is expressed as follows:

$$Claim_i = f(Coverage_i + Coverage_i \times Insured\_Age_i + Coverage_i \times Gender_i + Coverage_i \times Vehicle\_Age_i + CV_i) + e_i \tag{2}$$

As shown in Equation (2), the inclusion of the moderating variables may give rise to an inherent problem of collinearity; thus, in order to mitigate the possible effects arising from the collinearity between a product term and its components parts, the *Insured-Age* and *Vehicle-Age* variables are mean centered. However, the dummy variables, *Coverage* and *Gender*, are not mean centered for interpretational reasons (Jaccard, Turrisi, & Wan, 1990).

It should be noted that our sample of drivers purchasing insurance coverage is not randomly determined; that is, each driver self-selects into purchasing a high or low level of insurance coverage, thereby giving rise to the commonly observed econometric phenomenon of 'self-selection bias', and indeed, the decision to purchase high or low insurance coverage may be systematically related to other unobserved factors. Furthermore, part or all of the variables potentially affecting the decision to purchase high or low levels of insurance coverage may simultaneously affect the probability of the submission of claims. This indicates that the decision to purchase high or low insurance coverage may be endogenous.

A consequence of the self-selection problem is that the estimated regression coefficients may be biased; thus, in order to avoid

Table 2
Variable descriptions.

Variable	Description
Dependent variables	
Claim	Policyholder filing one or more claims = 1; otherwise 0.
Explanatory variables	
Coverage	1 for high coverage (low deductible); 0 for low coverage (high deductible).
Insured_Age	Age of the insured party.
Coverage × *Insured_Age*	Product of *Coverage* and *Insured-Age*.
Gender	Gender of the insured party; Male = 1; Female = 0.
Coverage × *Gender*	Product of *Coverage* and *Gender*.
Vehicle_Age	Age of the insured vehicle.
Coverage × *Vehicle_Age*	Product of *Coverage* and *Vehicle-Age*.
Bonus-Malus	Bonus-Malus coefficient of the insured party.
RN	Insured vehicle located in Northern Taiwan = 1; otherwise 0
RC	Insured vehicle located in Central Taiwan = 1; otherwise 0
RS	Insured vehicle located in Southern Taiwan = 1; otherwise 0
RTH	Insured vehicle located in Taoyuan/Hsinchu = 1; otherwise 0
Domestic	Insured vehicle produced domestically = 1; otherwise 0
Cap1	Insured vehicle cubic capacity <1200 = 1; otherwise 0
Cap2	Insured vehicle cubic capacity 1201–1800 = 1; otherwise 0
Cap3	Insured vehicle cubic capacity 1801–2400 = 1; otherwise 0
Cap4	Insured vehicle cubic capacity 2401–3600 = 1; otherwise 0
Renew	Insurance policy renewal = 1; otherwise 0

the spurious effects of the decision to purchase high/low coverage on the likelihood of claims, effects that are potentially attributable to self-selection bias, we adopt the Heckman two-stage estimation approach for our analysis (Heckman, 1979). It should be noted that in our regressions on the *Coverage* and *Claim* variables, we use the same set of control variables as those used in the prior literature (e.g., Shi et al., 2012).

Heckman (1979) argued that bias in the estimated regression coefficients is attributable to an omitted variable, which is referred to as the inverse Mills ratio; thus, the first stage of our analysis involves running a probit regression of the treatment variable (the *Coverage* variable in our analysis) on the control variables included in Equation (1) in order to obtain the inverse Mills ratio. The second stage involves running a regression on the outcome variable (the *Claim* variable in our analysis); in this stage, we include the estimated inverse Mills ratio as an additional regressor to correct for the potential problem of selectivity. Details on the use of the Heckman two-step estimation approach can be found in Johnston and DiNardo (1997).

Empirical results

Univariate analysis

The summary statistics of all of the variables used in the present study are reported in Table 3. As the table shows, of our total sample of insured drivers, approximately 28 per cent were found to have purchased high levels of insurance coverage, whilst approximately 22 per cent were found to have submitted claims. A typical insured driver in our sample is a male, aged approximately 45 years, whilst the typical insured vehicle is approximately four years old.

The correlation matrix between each of the variable coefficients is presented in Table 4, from which a positive and highly significant relationship is clearly discernible between *Coverage* and *Claim* at the 1 per cent level; this indicates that drivers who have purchased higher levels of insurance coverage are more likely to submit claims, thereby indicating preliminary evidence of adverse selection. The interaction terms between *Coverage* and the moderator variables, *Insured-Age*, *Gender* and *Vehicle-Age*, are also found to be significant, at least at the 1 per cent level. Table 4 therefore provides an initial understanding of the likely interaction between *Coverage* and the moderator variables.

Table 3
Summary statistics.

Variables	Mean	S.E.	Min.	Max.
Claim	0.2245	0.4176	0.00	1.00
Coverage	0.2824	0.4505	0.00	1.00
Insured_Age	44.8774	9.5597	22.00	78.00
Coverage × *Insured_Age*	0.3487	4.9411	−20.88	33.12
Gender	0.7135	0.4524	0.00	1.00
Coverage × *Gender*	0.2080	0.4062	0.00	1.00
Vehicle_Age	3.7273	2.3253	0.00	13.00
Coverage × *Vehicle_Age*	−0.1544	1.2082	−3.73	8.27
Bonus-Malus	−0.3394	0.2454	−0.60	0.60
RN	0.2452	0.4305	0.00	1.00
RC	0.2025	0.4021	0.00	1.00
RS	0.2645	0.4414	0.00	1.00
RTH	0.1281	0.3344	0.00	1.00
Domestic	0.7961	0.4031	0.00	1.00
Cap1	0.0110	0.1045	0.00	1.00
Cap2	0.4449	0.4973	0.00	1.00
Cap3	0.4284	0.4952	0.00	1.00
Cap4	0.0826	0.2755	0.00	1.00
Renew	0.7190	0.4498	0.00	1.00

Multivariate analysis

The estimation results using binomial probit regressions for Models (1) and (2), respectively relating to Equations (1) and (2), are presented in Table 5. The Chi-squared (χ^2) statistic in Model (1) is found to be 99.654 (p-value = 0.000) whilst that in Model (2) is 111.016 (p-value = 0.000). Statistical significance is found for both models at the 1 per cent level, thereby indicating that the fitted models are better than a null model without explanatory variables. The McFadden pseudo R^2 values are 0.129 in Model (1) and 0.144 in Model (2).

The regression results from Equation (1) reveal that the coverage variable is positive and statistically significant at the 5 per cent level, thereby indicating a greater likelihood of an insured party with higher coverage submitting claims. This finding provides preliminary evidence of the presence of asymmetric information; however, the positive relationship between *Coverage* and *Claim* may also suggest the existence of either adverse selection or moral hazard.

In the present study, the *Bonus-Malus* variable, which represents the prior claims history of the policyholder, is found to be positive and highly significant, thereby indicating that an insured party with a poor claims history has a higher probability of being involved in RTAs (Shi et al., 2012; af Wåhlberg, 2012). This evidence, together with our prior finding of a positive relationship between *Coverage* and *Claim*, would seem to indicate the existence of adverse selection (Cohen & Siegelman, 2010).

In order to examine the moderating effects of the *Insured-Age*, *Gender* and *Vehicle-Age* variables on the relationship between *Coverage* and *Claim*, interaction terms are subsequently added into our regressions. The adverse effects potentially arising from the problem of multi-collinearity are alleviated by mean centering the *Insured-Age* and *Vehicle-Age* variables, although the dummy variables, such as *Coverage* and *Gender* are not mean centered.

The results obtained from Equation (2) reveal that *Coverage* remains positive, albeit insignificant (p-value = 0.113), which suggests that the signficant results reported earlier will probably have been weakened by the moderator variables, *Insured-Age*, *Gender* and *Vehicle-Age*. We also find that the coeffieient on the interaction term between *Coverage* and *Vehicle-Age* is negative and highly significant at the 1 per cent level, thereby suggesting that the impact of *Coverage* on *Claim* is weaker for insured drivers with older vehicles.

Consistent with Bair, Huang and Wang (2012), we find that the probability of the occurrence of RTAs is potentially affected by the location of the insured vehicle, with three out of the four variables on the insured vehicle location being found to be significnt at the 1 per cent level. We also find evidence to show that the cubic capacity of the insured vehicle has an impact on the likelihood of occurrences of RTAs.

The estimation results based upon the Heckman two-stage approach are presented in Table 6. As noted earlier, to obtain the inverse Mills ratio, we first of all run a probit regression on the *Coverage* variable along with several explanatory variables, from which we find that the χ^2 statistic is 63.459 and highly significant, thereby indicating good model fitness. The McFadden pseudo R^2 is 0.073.

The inverse Mills ratio is subsequently added into Equation (2) as a regressor. In the *Claim* regression, the χ^2 statistic is found to be 111.440 (p-value <0.001), with the null hypothesis (that all of the slope coefficients would be equal to zero) being duly rejected. However, the inverse Mills ratio is not found to be significant, with a p-value of 0.5165, which indicates that selection bias does not appear to be a problem of any significance in our analysis. The McFadden Pseudo R^2 is 0.144.

Table 4
Correlation matrix.

Variables		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)
Claim	(a)	–								
Coverage	(b)	0.146***	–							
Insured_Age	(c)	−0.054	0.081**	–						
Coverage × *Insured_Age*	(d)	−0.009	0.113***	0.519***	–					
Gender	(e)	−0.017	0.032	0.154***	0.062*	–				
Coverage × *Gender*	(f)	0.107***	0.817***	0.100***	0.157***	0.325***	–			
Vehicle_Age	(g)	−0.090**	−0.148***	0.053	−0.020	−0.025	−0.156***	–		
Coverage × *Vehicle-Age*	(h)	−0.188***	−0.204***	−0.020	−0.030	−0.068*	−0.235***	0.528***	–	
Bonus-Malus	(i)	0.174***	0.162***	0.002	0.080**	0.045	0.183***	−0.453***	−0.233***	–
RN	(j)	0.031	−0.059	0.002	−0.010	−0.028	−0.048	0.092**	0.037	−0.036
RC	(k)	−0.148***	−0.027	0.000	0.024	0.008	−0.030	−0.025	0.063*	−0.035
RS	(l)	0.029	0.082**	0.059	0.058	0.083**	0.108***	−0.063*	−0.089**	0.096***
RTH	(m)	0.258***	0.107***	−0.092**	−0.081**	−0.113***	0.027	−0.001	−0.056	0.030
Domestic	(n)	−0.096***	−0.047	−0.009	0.034	0.042	−0.010	−0.087**	0.003	−0.084**
Cap1	(o)	0.006	0.080**	0.041	0.080**	−0.021	0.043	−0.005	0.017	0.006
Cap2	(p)	−0.030	−0.075**	−0.029	0.059	0.114***	−0.001	−0.058	−0.006	−0.009
Cap3	(q)	0.008	0.001	0.031	−0.065*	−0.079**	−0.053	−0.003	−0.041	−0.005
Cap4	(r)	0.066*	0.012	−0.057	−0.034	−0.064*	0.006	0.104***	0.063*	0.020
Renew	(s)	−0.068*	−0.084**	0.052	0.035	−0.037	−0.072*	0.407***	0.303***	−0.355***

Variables		(j)	(k)	(l)	(m)	(n)	(o)	(p)	(q)	(r)
Claim	(a)									
Coverage	(b)									
Insured_Age	(c)									
Coverage × *Insured_Age*	(d)									
Gender	(e)									
Coverage × *Gender*	(f)									
Vehicle_Age	(g)									
Coverage × *Vehicle_Age*	(h)									
Bonus-Malus	(i)									
RN	(j)	–								
RC	(k)	−0.287***	–							
RS	(l)	−0.342***	−0.302***	–						
RTH	(m)	−0.218***	−0.193***	−0.230***	–					
Domestic	(n)	−0.069*	0.102***	0.063*	−0.174***	–				
Cap1	(o)	−0.029	−0.020	−0.003	−0.001	0.021	–			
Cap2	(p)	−0.066*	0.004	0.004	−0.078**	0.157***	−0.094**	–		
Cap3	(q)	0.018	0.035	0.017	0.010	0.099***	−0.091**	−0.775***	–	
Cap4	(r)	0.073**	−0.027	−0.033	0.080**	−0.270***	−0.032	−0.269***	−0.260***	–
Renew	(s)	0.078**	0.025	−0.021	−0.026	0.117***	0.037	−0.026	0.052	−0.035

Note: *** indicates statistical significance at the 1 per cent level; ** indicates statistical significance at the 5 per cent level; and * indicates statistical significance at the 10 per cent level.

Table 5
Probit regressions of vehicle damage insurance claims.

Variables	Model (1)		Model (2)	
	Coeff.	S.E.	Coeff.	S.E.
Constant	−1.4943***	0.4788	−1.4664***	0.5139
Coverage	0.3061**	0.1245	0.3758	0.2372
Insured_Age	−0.0073	0.0060	−0.0112	0.0073
Coverage × *Insured_Age*	–	–	0.0106	0.0130
Gender	0.0456	0.1256	0.1028	0.1512
Coverage × *Gender*	–	–	−0.2547	0.2775
Vehicle_Age	−0.0249	0.0290	0.0233	0.0323
Coverage × *Vehicle_Age*	–	–	−0.1917***	0.0601
Bonus-Malus	0.7376***	0.2482	0.8147***	0.2528
RN	0.6773***	0.2024	0.6441***	0.2036
RC	0.0694	0.2240	0.0462	0.2265
RS	0.6115***	0.1997	0.5664***	0.2013
RTH	1.3034***	0.2210	1.2691***	0.2222
Domestic	−0.2237	0.1490	−0.1925	0.1504
Cap1	0.9254	0.6307	0.8674	0.6248
Cap2	0.8771**	0.3722	0.7610**	0.3732
Cap3	0.8514**	0.3699	0.7207*	0.3713
Cap4	0.9204**	0.3898	0.8366**	0.3909
Renew	0.0215	0.1385	0.1061	0.1436

Note: *** indicates statistical significance at the 1 per cent level; ** indicates statistical significance at the 5 per cent level; and * indicates statistical significance at the 10 per cent level.

Table 6
Heckman two-stage regressions of vehicle damage insurance claims.

Variables	Claim	
	Coeff.	S.E.
Constant	−2.2975*	1.3827
Coverage	2.1533	2.7495
Insured_Age	−0.0170	0.0116
Coverage × *Insured_Age*	0.0078	0.0137
Gender	0.0064	0.1611
Coverage × *Gender*	−0.2672	0.2790
Vehicle_Age	0.0509	0.0533
Coverage × *Vehicle_Age*	−0.1682**	0.0705
Bonus-Malus	0.4929	0.5557
RN	0.5581**	0.2419
RC	−0.0852	0.3025
RS	0.3334	0.4100
RTH	0.8998	0.6089
Domestic	−0.2348	0.1640
Cap1	0.8548	0.6266
Cap2	1.4177	1.0800
Cap3	1.3201	0.9971
Cap4	1.3701	0.9118
Renew	0.1167	0.1446
IMR	−1.0502	1.6189

Note: *** indicates statistical significance at the 1 per cent level; ** indicates statistical significance at the 5 per cent level; and * indicates statistical significance at the 10 per cent level.

The second step regression results on the *Claim* variable reveal that *Coverage* is positive, but insignificant. The interaction term between *Coverage* and *Vehicle-Age* is found to be negative and significant at the 1 per cent level, thereby indicating the moderating effect of *Vehicle-Age* on the relationship between the *Coverage* and *Claim* variables. These main results are generally consistent with those reported in Table 5; thus, as previously stated, the potential problem of self-selection bias is not found to have any significant impact on our data. Furthermore, we suggest that the two-stage Heckman approach may, in fact, be quite inefficient, as compared to the maximum likelihood equivalent (Johnston & DiNardo, 1997); thus, for reasons of brevity, we provide no further discussion on the results reported in Table 6.

For robustness checks, we estimated the conditional relationship between coverage and claims using a bivariate regression. Our conclusion remains unchanged. Our estimate for ρ in the regression is 0.17531, with a standard error of 0.08029, and an associated p-value of 0.0290, which indicates that the association between *Coverage* and *Claim* is significantly positive at the 5 per cent level. The null hypothesis of independence is therefore rejected.

We also run a bivariate regression with interaction terms. Consistent with the insignificant results presented above, we find that our estimate for ρ is 0.19860, with a standard error of 0.16416, and an associated p-value of 1.210, which indicates that the association between *Coverage* and *Claim* is insignificant at conventional levels.

Conclusions

To the best of our knowledge, the present study represents the first attempt to simultaneously examine: (i) whether insurance purchase behavior has effects on the probability of occurrences of road traffic accidents; and (ii) whether this relationship will tend to be moderated by the age and gender of the insured party and the age of the vehicle. Using a dataset obtained from a major property and casualty insurance company in Taiwan, we set out to examine two important issues relating to the relationship between vehicle insurance purchase and occurrences of RTAs.

The first issue is concerned with whether drivers who purchase higher levels of insurance coverage are more likely to be involved in RTAs; and indeed, our analysis reveals a positive relationship between the level of insurance coverage and the probability of the submission of claims. We also find that insured drivers with a poor driving history are more likely to be involved in RTAs. Taken together, our results suggest the existence of adverse selection.

The second issue relates to whether this coverage-claims relationship will be strengthened or weakened by the age and gender of the insured party and the age of the vehicle; we find that when considering these moderators, the original positive relationship becomes insignificant. Of particular importance is the finding that the positive relationship is significantly weakened by the age of the vehicle.

The general implication of our results is that drivers with a poor history of driving who are also found to purchase higher levels of insurance coverage for their new vehicles are likely to be involved in RTAs and are also likely to submit more insurance claims. We consider that our study has important implications for those with responsibility for the prevention of road traffic accidents, including traffic regulatory authorities and practitioners, and suggest that the relevant authorities should pay particular attention to this particular category of drivers in order to reduce the overall incidences of road traffic accidents.

Collaboration between traffic authorities and insurers is strongly encouraged; for instance, more education on road safety is required for such accident-prone drivers; indeed, insurers may choose to provide discounts on the premiums of drivers who attend such road safety lessons. Finally, we suggest that future research may go on to further investigate whether or not the positive coverage-claims relationship is potentially moderated by other factors.

References

Abbring, J., Chiappori, P.-A., Heckman, J., & Pinquet, J. (2003). Adverse selection and moral hazard in Insurance: Can dynamic data help to distinguish? *Journal of the European Economic Association, 1*(2–3), 512–521.

af Wåhlberg, A. E. (2012). Changes in driver celeration behaviour over time: Do drivers learn from collision? *Transportation Research Part F, 15*, 471–479.

Akerlof, G. A. (1970). The market for 'Lemons': Qualitative uncertainty and the market mechanism. *Quarterly Journal of Economics, 84*, 488–500.

Ardeshiri, A., & Jeihani, M. (2014). A speed limit compliance model for dynamic speed display signs. *Journal of Safety Research, 51*, 33–40.

Bair, S., Huang, R. J., & Wang, K. C. (2012). Can vehicle maintenance records predict automobile accidents. *Journal of Risk and Insurance, 79*(2), 567–584.

Blows, S., Ivers, R. Q., Connor, J., Ameratunga, S., & Norton, R. (2003). Car insurance and the risk of car crash injury. *Accident Analysis and Prevention, 35*, 987–990.

Bonander, C., Nilson, F., & Andersson, R. (2014). The effect of the swedish bicycle helmet law for Children: An interrupted time-series study. *Journal of Safety Research, 51*, 15–22.

Borghans, L., Golsteyn, B. H. H., Heckman, J. J., & Meijers, H. (2009). Gender differences in risk aversion and ambiguity aversion. *Journal of European Economic Association, 7*(2–3), 649–658.

Chekijian, S., Paul, M., Kohl, V. P., Walker, D. M., Tomassoni, A. J., Cone, D. C., & Vaca, F. E. (2014). The global burden of road Injury: Its relevance to the emergency physician. *Emergency Medicine International*. http://dx.doi.org/10.1155/2014/139219.

Chiappori, P. A., & Salanié, B. (2000). Testing for asymmetric information in insurance markets. *Journal of Political Economy, 108*(1), 56–78.

Cohen, A. (2005). Asymmetric information and Learning: Evidence from the vehicle insurance market'. *Review of Economics and Statistics, 87*(2), 197–207.

Cohen, A., & Siegelman, P. (2010). Testing for adverse selection in insurance markets. *Journal of Risk and Insurance, 77*(1), 39–84.

De Meza, D., & Webb, D. C. (2001). Advantageous selection in insurance markets. *RAND Journnal of Economics, 32*(2), 249–262.

Dionne, G., Gouriéroux, C., & Vanasse, C. (2001). Testing for evidence of adverse selection in the automobile insurance market: A comment. *Journal of Political Economy, 109*, 444–473.

Farmer, C. M., & Wells, J. K. (2010). Effect of enhanced seat belt reminders on driver fatality risk. *Journal of Safety Research, 41*, 53–57.

Finkelstein, A., & McGarry, K. (2006). Multiple dimensions of private Information: Evidence from the long-term care insurance market. *American Economic Review, 96*(4), 938–958.

Harrington, S. E., & Niehaus, G. R. (2004). *Risk management and insurance* (2nd ed.). Singapore: McGraw Hill.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*, 153–161.

Hemenway, D. (1990). Propitious selection. *Quarterly Journal of Economics, 105*(4), 1063–1069.

Hemenway, D. (1992). Propitious selection in insurance. *Journal of Risk and Uncertainty, 5*, 247–251.

Jaccard, J., Turrisi, R., & Wan, C. K. (1990). *Interaction effects in multiple regression* (1st ed.). London: Sage Publications.

Johnston, J., & DiNardo, J. (1997). *Econometric methods* (4th ed.). Singapore: McGraw-Hill.

Karagyozova, T., & Siegelman, P. (2012). Can propitious selection stabilize insurance markets? *Journal of Insurance Issues, 35*(2), 121–158.

Kim, H., Kim, D., Im, S., & Hardin, J. W. (2009). Evidence of asymmetric information in the automobile insurance Market: Dichotomous versus multinomial measurement of insurance coverage. *Journal of Risk and Insurance, 76*(2), 343–366.

Li, C., Liu, C., & Peng, S. (2013). Bundled automobile insurance coverage and accidents. *Accident Analysis and Prevention, 50*, 64–72.

Morin, R. A., & Suarez, A. F. (1983). Risk aversion revisited. *The Journal of Finance, 38*(4), 1201–1216.

Nguyen-Hoang, P., & Yeung, R. (2014). Dollars for Lives: The Effect of Highway Capital Investments on Traffic Fatalities. *Journal of Safety Research, 51*, 109–115.

NHTSA. (2010). *Highlights of 2009 motor vehicle crashes*. US Department of Transportation: National Highway Traffic Safety Administration.

Paefgen, J., Staake, T., & Fleisch, E. (2014). Multivariate exposure modeling of accident risk: Insights from pay-as-you-drive insurance data. *Transportation Research Part A, 61*, 27–40.

Puelz, R., & Snow, A. (1994). Evidence on adverse Selection: Equilibrium signaling and cross-subsidization in the insurance market. *Journal of Political Economy, 102*(2), 236–257.

Romano, E., & Kelley-Baker, T. (2015). Child passengers injured in motor vehicle crashes. *Journal of Safety Research, 52*, 1–8.

Rothschild, M., & Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics, 90*(402), 629–649.
Saito, K. (2006). Testing for asymmetric information in the automobile insurance market under rate regulation. *Journal of Risk and Insurance, 73*(2), 335–356.
Shi, P., Zhang, W., & Valdez, E. A. (2012). Testing adverse selection with two-dimensional information: Evidence from the Singapore auto insurance industry. *Journal of Risk and Insurance, 79*(4), 1077–1114.
Sloan, F. A., Chepke, L. M., & Davis, D. V. (2013). Race, gender and risk perceptions of the legal consequences of drinking and driving. *Journal of Safety Research, 45*, 117–125.
Spindler, M., Winter, J., & Hagmayer, S. (2013). Asymmetric information in the market for automobile Insurance: Evidence from Germany. *Journal of Risk and Insurance, 81*(4), 781–801.
Taiwan Insurance Institute. (2015). *Statistics of non-life insurance premium income*. Retrieved October 8, 2015, from http://www.tii.org.tw/fcontent/information/information03_01.asp?P2b_sn=23.
WHO. (2012). *Global Status Report on road safety 2011*. World Health Organization. Retrieved from http://www.who.int.
WHO. (2014). *Global Status Report on road safety 2013*. World Health Organization. Retrieved from http://www.who.int/violence_injury_prevention/road_safety_status/2013/en/.