Estimating the Total Cost of Moral Hazard in a Market for Automobile Insurance: A Mixed-Process Estimator Approach

The 43 Seminar of the

European Group of Risk and Insurance Economists (EGRIE)

Nicosia, Cyprus

19-21 September 2016

David Rowell[1], Hong Son Nghiem[2] & Luke B. Connelly[1,3]

1. The University of Queensland, Centre for the Business and Economics of Health

2. Queensland University of Technology, Institute of Health and Biomedical Innovation

3. The University of Queensland, Faculty of Health and Behavioural Sciences

Abstract

Two types of moral hazard are of interest in the literature on risk and insurance: *ex ante* moral hazard is the effect that the possession of insurance may have on the loss due to an insurable event and *ex post* moral hazard is the effect that insurance may have on the loss once the insured event occurs. The expected loss of the insured event is the sum of these effects. In this paper, we present an empirical test for total moral hazard in a market for automobile insurance using a mixed-process estimator approach. To the best of our knowledge, ours is the first paper to do so. Our results suggest that total moral hazard increases the expected cost of automobile repairs by approximately 75 per cent. We estimate that the *ex ante* moral hazard and *ex post* moral hazard are responsible for approximately 34.6 per cent and 64.4 per cent of the total moral hazard effect, respectively.

Arrow (1985, p. 38) defines moral hazard as the effect that "hidden action" by the agent has on the probability distribution of a cost of a claim. Zweifel and Breyer (1997, p. 157) identified two distinct mechanisms by which insurance can affect the cost of a claim; *ex ante* moral hazard and *ex post* and moral hazard. *Ex-ante* moral hazard identifies the effect that insurance has on the probability of an accident and *ex post* moral hazard the affect that insurance has on the cost of the claim. The inter-relationship between *ex ante* and *ex post* moral hazard is expressed in the following identity for an expected loss.

Expected Loss:

$$\begin{aligned} EL &\equiv \rho L \\ &\equiv \rho f(Ins, ...) \times Lf(Ins, ...) \end{aligned} \tag{1}$$

where:

EL	= Expected loss
ρ	= Probability of a loss
L	= Loss size
Ins	= Insurance (=1 if the loss is insured; =0 otherwise)

Equation (1) is an identity, which states that the expected loss (EL) is given by the product of the probability of a loss (ρ) and the size of the loss (L). The first term of the identity captures the effect that insurance has on the probability of a loss; i.e., *ex ante* moral hazard. The second term of the identity captures the effect that insurance (Ins) has on size of loss; i.e., *ex post* moral hazard. Hence, insurance is causally linked to both the probability and size of the loss.

Equation (2), expresses the total moral hazard (TMH) as the difference between the expected loss with and without insurance

Total Moral Hazard:

$$\begin{aligned} TMH &\equiv E(L_1) - E(L_0) \\ &\equiv \rho_1 L_1 - \rho_0 L_0 \\ &\equiv (\rho_0 + \Delta\rho)(L_0 + \Delta L) - \rho_0 L_o \\ &\equiv \rho_0 L_0 + \rho_0 \Delta L + \Delta\rho L_0 + \Delta\rho\Delta L - \rho_0 L_0 \\ &\equiv \Delta L(\rho_0 + \Delta\rho) + L_0 \Delta\rho \\ &\equiv L_0 \Delta\rho + \rho_1 \Delta L \\ &\equiv Ex\,ante\,MH + Ex\,post\,MH \end{aligned} \tag{2}$$

Where

TMH	= Total moral hazard
Δ	= Changes in loss size or probability of a loss
Subscript	= Insurance status (1 = with insurance; 0 = without insurance)

Rearranging the terms in equation (2) it can be shown that *TMH* is equal to the sum of *ex ante* moral hazard (i.e., loss due to increase in the probability) and *ex post* moral hazard (i.e., loss due to increase in loss size given the increase probability has occurred). The total effect that insurance has on the expected loss is the sum of the *ex ante* and *ex post* moral hazard effects. The principal challenge, which besets empirical estimates of *ex ante* moral hazard, is that the decision to insure is assumed endogenous.

In this paper, we estimate the total effect that insurance has on the expected loss of a road traffic crash (RTC) in a market for automobile insurance, by estimating both *ex ante* and *ex post* moral hazard effects. To the best of our knowledge, ours is the first paper to do so. The work is arranged as follows. Section I reviews the literature on *ex ante* and *ex post* moral hazard in markets for automobile insurance. Section II describes our data and empirical approach. Section III presents our results. The final section concludes.

I. BACKGROUND

The literature, which has analyzed moral hazard in markets for automobile insurance, is composed of two discrete genres, *ex ante* moral hazard and *ex post* moral hazard.

Ex-ante Moral Hazard

Empirical tests *ex ante* moral hazard in markets for automobile insurance evolved out of a literature, which has tested for evidence for a coverage-risk correlation conditional upon a vector of controls that reflect the insurer's information set. Puelz and Snow (1994), the first to analyze individual automobile claims data, constructed an ordered logit model to reveal a correlation between risk-type and choice of deductible conditional upon the insurer's information set. They argued that the negative correlation coefficient constitutes empirical evidence of adverse selection; drivers with private information about their high-type purchase more insurance. It was subsequently argued that this result is also consistent with the hypothesis of *ex ante* moral hazard (Chiappori, 1999) and that the model was incorrectly specified (Dionne et al., 2001), which could lead to the identification of a conditional correlation when there is none. Dionne et al. (2001) demonstrate that when the nonlinearity of the risk classification variables was accounted for, the residual correlation disappeared.

Chiappori and Salanié (2000) were the first researchers to propose an explicit test for *ex ante* moral hazard in a market for automobile insurance, which recognized that the insurance decision was potentially endogenous. Using a French claims data set, they first proposed a test for asymmetric information using a bivariate probit model wherein the first probit equation predicts the level of insurance and the second probit equation predicts the occurrence of a claim. The null hypothesis of no asymmetric information as not rejected when tested using two null hypotheses: (i) zero covariance of residuals when estimating two probit estimators separately; and (ii) zero correlation coefficient of the two residuals when the model is estimated as a bivariate probit.

Chiappori and Salanié (2000) then proposed a test for *ex ante* moral hazard that exploited a “natural experiment” whereby adult children can inherit their parent’s bonus-malus coefficient if they state that their automobile is jointly owned. A dichotomous bonus-malus variable equal to one if the beginner driver inherits a bonus-malus coefficient of 0.5 is added to the coverage and claims equations. They argued that the sign of the coefficient on the bonus-malus variable from the claims equation may be used to identify the presence *ex ante* moral hazard. They report a negative coefficient, and reject the hypothesis of *ex ante* moral hazard.

Abbring et al. (2003) and Israel (2004) argued that while the conditional correlation approach offers a robust test for asymmetric information it cannot be used to distinguish moral hazard from adverse selection. Abbring et al. (2003) adapted the state-dependence approach by Heckman and Borjas (1980) to test for moral hazard. A proportional hazard model was used to compare: (i) the distribution of first and second claims across contracts over time, and (ii) the first and second claim times of each contract with two claims (or more) in a French claims data set. No evidence of moral hazard was reported. Israel (2004) has argued that Abbring et al. (2003) assumed that there are no other sources of state dependence. In particular, past accidents were explicitly assumed only to influence current behavior through their effects on the premium. Israel (2004) used a difference-in-difference approach to examine the occurrence of claims around the three-year insurance event and report a small but statistically significant *ex ante* moral hazard effect.

Following Chiappori and Salanié (2000) further tests for asymmetric information using a bivariate probit specification were analyzed using Israeli (Cohen, 2005), Dutch (Zavadil, 2015), German (Spindler et al., 2014) and Japanese (Saito, 2006) claims data. The consensus was weak evidence of asymmetric information. Spindler et al. (2014) and Cohen (2005) reported coverage-risk correlations were limited to drivers of certain age groups. However, in a market for long-term care insurance Finkelstein and McGarry (2006) have demonstrated that zero evidence of a conditional correlation between insurance and accident/claim is not a necessary condition for the existence of the other dimensions of private information such as adverse selection and by implication *ex ante* moral hazard. They show that the inclusion of variables that may serve as proxies for private information that are not normally observable to the insurance firm could provide useful insights into behavior under insurance.

Recently, a small number of researchers have begun to look beyond claims data to specify a variety of tests for *ex ante* moral hazard. Dionne et al. (2013) used a three-year panel of survey data collected by the French market research firm SOFRES, to test for *ex ante* moral hazard using the conditional correlation approach. They also specified a bivariate probit model where the probabilities of accidents and insurance contract choice in the current period are the function of both outcomes in the previous period and other covariates. To specify a test for *ex ante* moral hazard they Dionne et al. (2013) argued:

> [w]e assume drivers differ in terms of risk type (or ability). In the model, agents first buy insurance without knowledge of their own risk. They learn about risk from their history of accidents. Accidents differ depending on whether the driver is at fault or not. Although the insurer observes the bonus-malus he does not learn as fast as the agent about his riskiness, partly because he/she observes claims only. Thus asymmetric learning develops, which may lead to pure adverse selection in contract choices within observable risk classes. (Dionne et al., 2013, p. 900)

A Granger causality test was used to test for *ex ante* moral hazard by examining the correlation between previous contract choice and an accident in the current period, conditional upon the insurer's information set. The authors report strong evidence of *ex ante* moral hazard for drivers with less than 15 years of driving experience. Importantly the ability to include an indicator of historical RTCs that were not observable by the insurer enabled Dionne et al. (2013) to specify their test for *ex ante* moral hazard.

Weisburd (2015) conducted tests for *ex ante* moral hazard, which analyzed automobile insurance coverage, which was distributed as an employee benefit to 67 per cent of motorists in their Israeli sample. Corporate automobile insurance provided an exogenous variation in insurance that was independent of adverse selection. Employer-determined coverage was estimated to reduce the average cost of an accident by $235. After controlling for driver and vehicle characteristics, it was reported that a $100 reduction in accident costs resulted in a 1.7 percentage point increase in the probability of an accident. Given an average accident rate of 16.3 percent, *ex ante* moral hazard was estimated to cause a 10 percent increase in RTCs.

Rowell et al. (Early view) present two tests for *ex ante* moral hazard in an Australian market for automobile insurance in the *Journal of Risk and Insurance*.[1] They specify (i) a recursive model and (ii) an instrumental variables model to address endogeneity with respect to policy selection in cross-sectional road traffic crash (RTC) survey data. The recursive model utilizes an identification strategy described by Dionne et al. (2013). The survey data set report previous RTCs that do not necessarily result in a claim. Rowell et al. (Early view) argued that previous RTCs were predetermined and therefore contemporaneously uncorrelated with the error term.

A second identification strategy used the insurance status of a second vehicle garaged at the household as an instrument for the insurance status of the primary vehicle. The rationale that underpins the selection of this IV was that shared household characteristics and joint vehicle operation result in a correlation between the IV and the endogenous variable but that the insurance of the secondary vehicle cannot elicit moral hazard while driving the primary car (Rowell et al., Early view).

These analyses provide corroborating evidence of *ex ante* moral hazard. Dionne et al. (2013) report evidence of *ex ante* moral hazard in a sub-set of inexperienced drivers and Weisburd (2015) and Rowell et al. (Early view) report *ex ante* moral hazard in their full samples. The reported marginal effects of insurance on the probability of an RTC were 10%, 12% and 17% (Rowell et al., Early view; Weisburd, 2015). Sensitivity analyses were conducted. Weisburd (2015) report no *ex ante* moral hazard with large (road) accidents vis-à-vis small (parking) accidents, arguing that as the principal cost of a parking accident was financial and not physical, *ex ante* moral hazard was more likely to be observed in small accidents. Rowell et al. (Early view) repeated their analyses in sub-samples of at-fault and non-at-fault RTCs. The anti-test produced no evidence of *ex ante* moral hazard in the sub-sample of not-at-fault RTCs, in which the true *ex ante* moral hazard may reasonably be assumed to be zero.

[1] Rowell, D., Nghiem, H.S. & Connelly, L.B., (2016) 'Two tests for *ex ante* moral hazard in a market for automobile insurance' is available online at the *Journal of Risk and Insurance* in early view.

Ex-post Moral Hazard

A parallel literature, which analyzed evidence of *ex post* moral hazard in markets for automobile insurance, also evolved. Zweifel and Breyer (1997) defined *ex post* moral hazard as the effect that insurance has on the cost of a claim. This catholic definition of *ex post* moral hazard, which we adopt, is also used by Arrow (1963) and Pauly (1968) in their discourse in the *American Economic Review* to describe a physician's ability to inflate expenditure on insured health care. However, within the automobile insurance literature the analysis of *ex post* moral hazard has instead generally been narrowly focused on fraud (Artís et al., 1999, 2002; Caron and Dionne, 1999; Subelj et al., 2011; Belhadji et al., 2000). These studies have analyzed closed claims, to estimate the prevalence of fraud or identify fraudulent claimants. Caron and Dionne (1999) have acknowledged, the identification of the fraudulent claims can be problematic. Belhadji et al. (2000) have relied on the subjective assessment claims assessor while Artís et al. (1999, 2002) have relied upon self-reported fraud following investigation by the insurer. Subelj et al. (2011) proposed an analysis of social networks to detect automobile insurance fraud, which uses "previously identified fraudsters" as the dependent variable.

Fraud is however is only one component of *ex post* moral hazard. Pauly's argument

> "... that the problem of "moral hazard" in insurance has, in fact, little to do with morality, but can be analyzed with orthodox economic tools (Pauly, 1968, p. 531),

implies that simple, legal, utility maximization by policyholders could result in significant *ex post* moral hazard. Empirical analysis, which failed to detect all fraudulent claims or utility maximization that did not satisfy the legal threshold of fraud, may thereby underestimate the true extent of *ex post* moral hazard. Despite strategies to reduce *ex post* moral hazard through inclusion of co-payments and supply side controls (e.g. minimum number of quotes and preferred repairers), the economic cost of *ex post* moral hazard could remain substantial.

In principle, one could test for evidence of *ex post* moral hazard by estimating equation (3) using a semilogarithmic regression, in a sample of RTCs.

$$lnC = f(Ins, \mathbf{S}) \tag{3}$$

Where:

lnC	= Log of cost of RTC repair
Ins	= Insurance (=1 if a comprehensive insurance policy is held; =0 otherwise)
S	= Severity of RTC

A statistically significant coefficient for the binary variable *Ins* could be accepted as evidence of *ex post* moral hazard. However, obtaining accurate information on **S** using claims data can be problematic. The true extent of the damage to the vehicle following an RTC is not fully observable to the insurer. Once a claim is submitted to the insurer, both policyholder and smash repairer faces an incentive to inflate the cost of RTC repairs. The smash repairer has an incentive to increase the price charged for a repair and the policyholder has an incentive to increase the quantity of repairs commissioned.

II. EMPIRICAL APPROACH

We estimate the combined effect of *ex ante* and *ex pos*t moral hazard, (i.e. total moral hazard) on the cost of RTC repairs using data from an Australian household survey.

Data

During a six-week period commencing from October 1999, EKAS Marketing Research Services conducted market research on behalf of IMRAS Consulting to analyze community attitudes to the Australian smash repair market. The resulting data, henceforth referred to as the IMRAS data set, were collected using computer assisted telephone interviews to contact 37,833 rural and metropolitan households in four Australian states (NSW, Victoria, Queensland and WA). Vehicle owners from 4,005 households (16.9 percent) completed the survey. The IMRAS data set was purchased for this analysis.

Although the IMRAS data set was not collected to test for evidence of asymmetric information *per se*, many of the variables that are necessary to conduct tests for moral hazard are available. Firstly, the IMRAS survey collected data on the insurance status of the respondent's automobile, as either (i) compulsory third-party (personal injury) only, (ii) (i) plus third-party property, (iii) (i) plus third-party property, fire and theft or (iv) (i) plus comprehensive property insurance. Only comprehensive insurance, which is discretionary, indemnifies the owner for the cost of smash repairs in an RTC for which he/she is at fault. Australia's comprehensive insurance policy is analogous to the French *assurance tous risques* as described by Chiappori and Salanié (2000). Importantly the data set also enabled the identification of 20 respondents who were uninsured when their RTC occurred and subsequently purchased comprehensive insurance.

Secondly, the IMRAS survey included a question about the historical involvement of the respondent in any RTC. A two-year recall period was used for this question both to ensure that sufficient data were collected on RTCs and smash repair experiences, and to minimize errors in respondent recall. In total, 994 of the respondents (24.8 percent) stated that they were involved in at least one RTC during the previous two-year period. This data field was utilized to create the dichotomous variable RTC, equal to one if any RTC was reported between 1997 and 1999 and zero if otherwise.

Thirdly, to conduct reliable tests for *ex ante* moral hazard using conditional correlation, it is necessary to identify a set of covariates that accurately reflects the insurers' information set. Two sources of information were reviewed. The first was the empirical literature, which identifies information commonly collected by predominantly French insurers to risk rate their policyholders. Secondly, data collected by the five largest underwriters in the IMRAS survey (NRMA Ltd., AAMI, GIO, RACV and Suncorp) were evaluated. These firms, which provided insurance coverage for 58.7 percent of the sample, each hosts a web page that enables the user to obtain a quote for comprehensive insurance. There is a considerable correspondence between the categories of data that are recognized as important in the (i) empirical and theoretical literature; (ii) data collected by insurance firms to generate premium quotations; and (iii) data included in the IMRAS data set. Controls for driver characteristics in the IMRAS data set included age, gender, young co-driver ($<$ 25 years of age), vehicle ownership type (i.e., private or corporate), location (metropolitan or rural), socioeconomic status (SES)) and years licensed, while controls for vehicle characteristics included the vehicle's value, age, make, body-type and engine size (4-, 6- and 8-cylinder).

While we believe these covariates provides a good approximation of the insurer's information set, it is possible that Australian insurers use data not reported in the IMRAS data set to risk-rate their policyholders, (for example, length of billing period (six-monthly or yearly)). Finkelstein and McGarry (2006) have stated that a conservative approach should be adopted when selecting covariates. In their analysis of the long-term care insurance market, they included variables that are not necessarily collected by all insurers, to ensure that their model adequately accounted for the insurer's information set. Analogously we identified two variables within the IMRAS data set: income and occupation type, which were not necessarily collected by all Australian insurers, but were identified as potential proxies for data otherwise collected by insurers to risk-rate their policyholders. As it has been established that tests for asymmetric information that use conditional correlation may produce spurious results if the explanatory covariates are inappropriately specified as a linear function of RTC, we have used combinations of dummy variables to reflect risk classification and, where data are continuous, flexible approximations (e.g. spline functions) have been substituted as recommended by Dionne et al. (2013).

Dionne et al. (2013) has stated that their ability to include an indicator of historical RTCs, which were not observable by the insurer, enabled them to specify a legitimate test for *ex ante* moral hazard. We therefore also use past RTCs as a proxy for unobserved risk-type to identify *ex ante* moral hazard in our system of equations. The IMRAS data set identifies historical RTCs that occurred between October 1994 and October 1997. Hence, we can create a dichotomous variable $RTC_{i,t-1}$ equal one if an RTC occurred between October 1994 and October 1997 and zero if otherwise.

Finally, the IMRAS data set also includes information on automobile repairs following an RTC. Respondents not only report the cost of repairs in Australian dollars but also data that reflects the latent severity of the RTC. For each RTC the IMRAS data set provide an itemized list of damaged parts up to a maximum of 35 distinct parts. Respondents also indicated the attendance of police, tow-truck or ambulance at the RTC, which attended 42 per cent of all RTCs. An average of 2.6 parts were repaired or replaced per RTC. Bumpers (front and rear) were the automobile component most frequently repaired. Descriptive data are reported in the appendices.

Econometric Approach

We specify two systems of equations to estimate the total effect that insurance has on expected RTC repairs using a mixed process estimator (MPE). The first (i.e. equations (4-6)) utilize a recursive model and the second (i.e. equations (7-9)) utilize an instrumental variable model. Estimation using a MPE allows endogenous variables and appropriate instruments to be included on right hand side of a system of equations.

Recursive Mixed Process Estimator:

$$RTC^*_{it} = \alpha_0 + \alpha_1 Ins_{it} + \alpha_2 \mathbf{X}_{it} + RT^*_{it} + PE^*_{it} + \varepsilon_{it} \tag{4}$$

$$Ins^*_{it} = \beta_0 + \beta_1 RTC_{i,t-1} + \beta_2 \mathbf{X}_{it} + RT^*_{it} + PE^*_{it} + \mu_{it} \tag{5}$$

$$lnC_{it} = \gamma_0 + \gamma_1 Ins_{it} + \gamma_2 \mathbf{S}_{it} + \epsilon_{it} \tag{6}$$

Where:

RTC = Road traffic crash (=1 if a road traffic crash occurs; =0 otherwise)

lnC = Log of cost of RTC repair

Ins = Insurance (=1 if comprehensively insured; =0 otherwise)

X = Insurer's information set (Vehicle and Driver)

S = Severity of RTC

i = the i^{th} individual

t = time

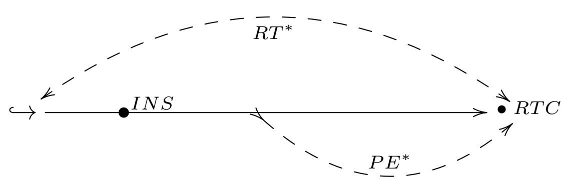
Equation (4)captures *ex ante* moral hazard. The dependent variable RTC_{it} is a dichotomous variable equal to one if an RTC was reported and zero if otherwise. The explanatory variable of interest is the dichotomous variable Ins_{it} equal to one if a comprehensive insurance policy is held; and zero if otherwise. The vector $\mathbf{X_{it}}$ is an exogenous set of policyholder and vehicular characteristics, which are observable to the insurer, and used to risk rate policyholders. RT^*_{it} + PE^*_{it} + ε_{it} is composite error term, which consists of unobserved risk type RT^*_{it}, unobserved preventive effort PE^*_{it} and a random noise term ε_{it}. The presence of RT^*_{it} and PE^*_{it} suggests that the error term is correlated with both Ins_{it} and RTC_{it} and hence INS_{it} is assumed endogenous. In this system, the null hypothesis of no *ex ante* moral hazard is given by $H_0 : \alpha_1 = 0$ in equation (4).

Equation (5) adapts an identification strategy described by Dionne et al. (2013) to control for the endogenously determined decision to insure. The dependent variable is Ins_{it}. The explanatory variable $RTC_{i,t-1}$, a dichotomous variable equal to one if the respondent reported a past RTC and zero if otherwise. We argue that $RTC_{i,t-1}$ (i) proxies unobserved risk type, and hence is able to capture the adverse selection effect and (ii) it is predetermined in the sense that it proxies the lag of RTC, and hence is exogenous by definition.

Equation (6) captures *TMH,* which is the effect that Ins_{it} has on the log cost of an RTC repair lnC_{it} conditional upon *ex ante* moral hazard. The dependent variable lnC_{it} is the log of the cost of a smash repair following an RTC. Costs were logged because (i) the Shapiro and Wilk (1965) test for normality indicated that the residuals were not normally distributed [W = 0.656, p-value $<$ 0.01] and (ii) the reported coefficients are interpretable as elasticities. The explanatory variable of interest is Ins_{it} and **S** is a vector set of exogenous covariates, which reflect the underlying severity of the RTC (e.g., road service attendances). The null hypothesis of zero TMH is $H_0 : \gamma_1 = 0$ in equation (6).

Instrumental Variable Mixed Process Estimator:

The temporal relationship between *hidden action* and *hidden information* in a market for [RTC] insurance, as described by Arrow (1985), is illustrated using a path diagram (see Figure 1). Solid lines denote causal links and arrows indicate causal direction and the dashed arcs identify confounding processes (Pearl, 2009). In Figure 1, unobserved risk type affects the insurance decision, which in turn can affect unobserved preventive effort, and both can affect the probability distribution of losses due to the insured event.



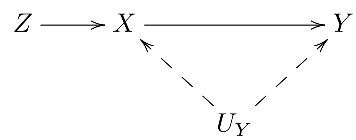
Source: Adapted from Pearl (2009, p.90)

Figure 1: A path diagram for a market of insurance

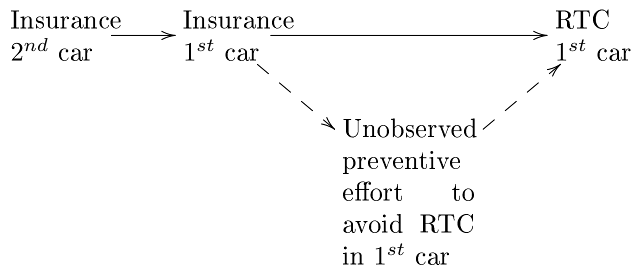
Figure 2, are path diagrams that illustrate the rationale for our IV choice. Figure 2a illustrates the requirements that a legitimate instrumental variable should satisfy (Pearl, 2009). To be valid, an IV should satisfy two criteria; (i) it should be correlated with the endogenous variable and (ii) it should be uncorrelated with the error term (Wooldridge, 2000). In our data the insurance status of a secondary vehicle in a sub-sample of two vehicle households with be correlated with the insurance status of the primary vehicle, but will have no independent effect on the level of preventive effort invested to avoid an RTC in the primary vehicle (See Figure 2b).

In Figure 2a, Y is the dependent variable and X an explanatory variable, which is endogenous because the effect that X has on Y is confounded by the unobserved effect U_Y has on both X and Y. More formally:

> [t]he traditional definition qualifies a variable Z as instrumental (relative to the pair (X, Y)) if (i) Z is not independent of X and (ii) Z is independent of all variables (including the error term) that have an influence on Y that is not mediated by X



(a) Generic IV (Pearl 2009 p. 248)



(b) IV for insurance: (Adapted from Pearl 2009 p. 248)

Figure 2: Path diagram for an instrumental variable

Pearl (2009, p.247).

Figure 2b applies the relationships illustrated in 2a to the setting of a market for comprehensive automobile insurance. The dependent variable Y corresponds to the probability of an RTC in the primary vehicle (RTC 1^{st} car). The explanatory variable X corresponds to the probability that the primary vehicle was comprehensively insured (Insurance 1^{st} car). The economic theory states that *ex ante* moral hazard occurs when insurance firm is unable to observe or contract for the agent's preventive effort (Hölmstrom, 1979; Marshall, 1976; Mirrlees, 1999; Pauly, 1974; Shavell, 1979; Winter, 1992). Thus, U_Y corresponds to the unobserved preventative effort used to avoid an RTC in the primary vehicle. For further discussion on the selection of IV, see Rowell et al. (Early view)

Equations 7 to 9 are a MPE, which uses the insurance status a second vehicle garaged at the household is used as an instrument for the insurance status of the primary vehicle.

$$RTC^*_{it} = \alpha_0 + \alpha_1 Ins\_Car1_{it} + \alpha_2 \mathbf{X}_{it} + RT^*_{it} + PE^*_{it} + \varepsilon_{it} \tag{7}$$
$$Ins\_Car1_{it} = \beta_0 + \beta 1 Ins\_Car2_{it} + \beta_2 \mathbf{X}_{it} + RT^*_{it} + PE^*_{it} + \mu_{it} \tag{8}$$
$$lnC_{it} = \gamma_0 + \gamma_1 Ins\_Car1_{it} + \gamma_2 \mathbf{S}_{it} + \epsilon_{it} \tag{9}$$

Where:

Ins_Car1 = Insurance for the primary car (=1 if comprehensively insured; =0 otherwise)

Ins_Car2 = Insurance for the secondary car (=1 if comprehensively insured; =0 otherwise)

In Equation (7), RTC^*_{it} is the dependent variable. The explanatory variable of interest, Ins_Car1, is a dichotomous variable equal to one if the primary vehicle is insured and zero if otherwise. As described in the Recursive MPE, $\mathbf{X_{it}}$ captures the insurer's information set and RT^*_{it} and PE^*_{it} denote unobserved risk type and preventive effort, respectively. In Equation (8), the dependent variable Ins_Car1_{it} is insurance status of the primary vehicle and Ins_Car2_{it} is our instrument, a dichotomous variable equal to one if the secondary vehicle is comprehensively insured and zero if otherwise. Equation (9) captures *TMH,* i.e. the effect that Ins_Car1_{it} has on lnC_{it}, the log cost of an RTC repair conditional upon *ex ante* moral hazard. The null hypothesis of nil TMH $H_0 : \gamma_1 = 0$ in equation (9).

Both systems of equations (4-6) and (7-9) were estimated using Stata's conditional mixed process (CMP) estimator. In both systems, the decision to insure is assumed endogenous. Equations (4-6) use a recursive system to obtain an consistent estimate of *ex ante* moral hazard in equation (4), while equations (7-9) use an instrumental variables approach to obtain an consistent estimate *ex ante* moral hazard in equation (7). TMH, the effect that insurance has on the expected cost of repairs conditional upon *ex ante* moral hazard, is captured by γ_1 in equations (6) and (9).

III. RESULTS

The results from (i) MPE with recursive structure and (ii) MPE with instrument, are presented in Table 1.

The statistically significant coefficients for both of the specifications reported in Table 1 are generally similar. In both models, vehicle owners are risk-averse; vehicle value is positively correlated with comprehensive insurance decisions and vehicle age is negatively correlated with comprehensive insurance. While driver age is not correlated with an RTC or insurance, years-licensed –a closely related construct– is strongly correlated with the dependent variables in each model. Vehicle *value* and vehicle *age* are correlated with an RTC. Some differences between the models do exist, though. For example, SES is correlated with an RTC in the recursive MPE

Table 1: Estimation of Total Moral Hazard: Results

	Recursive (n=2697)			IV (n=1011)		
Explanatory variables	RTC (0/1) [eq. 4]	Ins (0/1) [eq. 5]	C(log $) [eq. 6]	RTC (0/1) [eq. 7]	Ins (0/1) [eq. 8]	C (log $) [eq. 9]
Insured (0/1)	-	-	0.563***	-	-	0.821***
RTC Severity						
Ambulance (0/1)	-	-	-0.03	-	-	-0.244
Tow Truck (0/1)	-	-	0.543**	-	-	0.566*
Police (0/1)	-	-	0.303**	-	-	0.712***
Towed away (0/1)	-	-	0.63**	-	-	0.537
No. of damaged parts	-	-	0.242***	-	-	0.24***
Insured (0/1)	1.48***	-	-	1.567***	-	-
Second car insured (0/1)	-	-	-	-	0.694***	-
RTC History						
RTC 1994-97	-	0.155*	-	0.414***	-0.374**	-
Driver characteristics observable to insurer						
Aged 25 to 34 years	-0.125	0.105	-	-0.092	0.991***	-
Aged 35 to 44 years	0.031	0.021	-	-0.271	0.681	-
Aged 45 to 54 years	-0.011	0.142	-	-0.557	0.682	-
Aged over 55 years	-0.013	0.259	-	-0.558	0.786	-
Male	0.025	-0.326***	-	-0.01	-0.112	-
Metropolitan	0.157*	0.045	-	0.025	0.168	-
Other driver $<$ 25 years	0.256**	0.07	-	-0.004	0.272	-
Private registration	-0.085	0.265	-	-0.401**	-0.186	-
SES 2nd quartile	-0.086	0.201**	-	0.075	0.087	-
SES 3rd quartile	-0.103	0.224**	-	0.143	-0.034	-

	Recursive (n=2697)			IV (n=1011)		
Explanatory variables	RTC (0/1)	Ins (0/1)	C(log $)	RTC (0/1)	Ins (0/1)	C (log $)
	[eq. 4]	[eq. 5]	[eq. 6]	[eq. 7]	[eq. 8]	[eq. 9]
SES 4th quartile	-0.194*	0.32***	-	0.077	0.422**	-
Licensed 6 to 10 years	-0.507***	0.39**	-	-1.083***	0.663**	-
Licensed 11 to 15 years	-0.587***	0.717***	-	-1.197***	0.471	-
Licensed 16 to 20 years	-0.756***	0.849***	-	-1.014***	1.044**	-
Licensed 21 to 25 years	-0.746***	0.822***	-	-0.931**	0.963**	-
Licensed > 25 years	-0.781***	0.747***	-	-0.684	0.936**	-
Vehicle's Characteristics						
6-cylinder vehicle	-0.007	0.027	-	0.025	0.168	-
8-cylinder vehicle	-0.175	-0.022	-	0.116	-0.193	-
Make GM Holden	-0.025	-0.107	-	-0.142	-0.255	-
Make Toyota	0.162	0.113	-	0.077	0.106	-
Make Mitsubishi	0.051	0.01	-	-0.088	0.035	-
Make Asian	0.181	-0.102	-	0.264	-0.277	-
Make European	0.006	0.022	-	-0.473	-0.085	-
Body-type Commercial	0.029	-0.11	-	-0.022	-0.38**	-
Body-type 4 WD	-0.219	0.001	-	-0.448*	-0.088	-
Body-type Sports car	-0.069	0.341	-	0.109	0.522	-
Car age 3 to 7 years	0.03	-0.035	-	0.142	-0.598**	-
Car age 7 to 12 years	0.244**	-0.501***	-	0.28	-0.988***	-
Car age > 12 years	0.362**	-0.995***	-	0.419*	-1.587***	-
Value $2,001-$6,000	-0.264**	0.515***	-	-0.445*	0.951***	-
Value $6,001-$10,000	-0.468***	0.81***	-	-0.424*	1.264***	-
Value $10,001-$16,000	-0.535***	1.03***	-	-0.752***	1.71***	-
Value $16,001-$25,000	-0.439**	1.015***	-	-0.286	1.266***	-
Value > $25,000	-0.44**	1.032***	-	-0.376	1.963***	-

Explanatory variables	Recursive (n=2697) RTC (0/1) [eq. 4]	Ins (0/1) [eq. 5]	C(log $) [eq. 6]	IV (n=1011) RTC (0/1) [eq. 7]	Ins (0/1) [eq. 8]	C (log $) [eq. 9]
Additional driver/vehicle characteristics not observable to insurer						
Income $20,000-$39,999	0.307**	0.001	-	0.032	-0.074	-
Income $40,000-$59,999	0.189	0.341	-	0.046	0.316	-
Income $60,000-$79,999	0.257	-0.035	-	-0.04	0.11	-
Income $80,000-$99,999	-0.029	-0.501***	-	-0.546	0.935**	-
Income $100,000-$149,999	0.325	-0.995***	-	-0.018	-0.179	-
Income > $150,000	0.257	0.515***	-	-0.316	0.398	-
Income not divulged	0.113	0.81***	-	-0.104	0.649**	-
Profession lower white	-0.114	1.03***	-	-0.264	0.131	-
Profession upper blue	-0.152	1.015***	-	-0.127	-0.318	-
Profession lower blue	-0.063	1.032***	-	-0.45	-0.018	-
Profession home duties	-0.332**	0.001	-	-0.736***	0.042	-
Profession student	-0.175	0.178	-	-0.332	-0.205	-
Profession retired	-0.281*	0.109	-	-0.585**	0.33	-
Profession unemployed	0.099	0.581***	-	0.2	-0.739**	-
Profession not divulged	-0.07	0.165	-	0***	-1.536***	-
Km per year driven	0.005**	0.422*	-	0.003	0.008	-
Km per year driven squared	-	0.159	-	-	-0.00015	-
Constant	-1.442***	-1.442***	5.684***	-0.442	-1.166**	5.531***

Note:

(i) Interaction terms and coefficients are not reported but available upon request.[2]

(ii) The level of statistical significance denoted by *** @ 1%, ** @ 5% and * @ 10%

[2]Finkelstein and McGarry (2006) have argued that, to allow for possible nonlinearities among variables, controls including interaction terms should be included. Interaction terms, other claims (vehicle stolen, vehicle broken into, vehicle burnt, a car-part was stolen) and RTCs and dummy variables for year of RTC (1999, 1998 or 1997) were included. See Rowell *et al.* (in press) for details.

but not in the IV MPE, possibly due to the smaller sample size for the latter specification. In the Recursive MPE RTC_{t-1}, was a proxy for unobserved risk type, and was, as expected, positively correlated with the decision to insure (equation (6)). In the IV MPE RTC_{t-1} was included as a proxy for the *bonus malus* (or no claim bonus, as it is known in Australia). The coefficient was, as expected, positively correlated with RTCs in equation (7) but was negatively correlated with insurance in equation (8). A number of explanations may exist. Advantageous selection is one. Another could be previous RTCs have increased premiums, reducing the demand for insurance. Neither MPE includes controls for driver learning as described by Dionne et al. (2013).

The coefficients for our tests for TMH were 0.563 and 0.821 in the Recursive MPE (equation (6)) and IV MPE (equation (9)), respectively. To correctly interpret the coefficient for dummy variable in a semi-log function Halvorsen and Palmquist (1980) have stated that the relative effect (g) of a coefficient for dummy variable (c) on a dependent variable (Y) is given by $g = \exp(c) - 1$. Hence, exponentiation of the coefficients implies that TMH increased the expected cost of RTC repairs between 75% and 127%, depending on which model is chosen. Our preference is for the Recursive MPE approach as itanalyzes a random sample of 2,697 households, as opposed to 1,011 households with two vehicles.

A Decomposition of Total Moral Hazard:

We decompose our estimate of TMH as follows. First we derive an estimate of *ex post* moral hazard by estimating equation (10) in a sample restricted to 683 RTC repairs.

$$lnC_{it} = \alpha_0 + \alpha_1 Ins_{it} + \alpha_2 \mathbf{S}_{it} + \alpha_3 \mathbf{V} + \epsilon_{it} \tag{10}$$

In equation (10), the dependent variable lnC_{it}, is the log cost of RTC repairs and Ins_{it} the explanatory variable of interest. **S** is vector of exogenous variables, which capture the severity of the RTC (i.e attendance of ambulance, two truck, police and vehicle towed away). **V** is vector of vehicle characteristics, which we posit may affect the cost of an RTC repair and is comprised of a comprehensive set of dichotomous variables indicating vehicle type, plus *value* and *age*.

Table 2: *Ex post* moral hazard: Results

lnC_Repair (n=683)	Coefficient	p-value
Insurance	0.29	0.04
RTC severity		
Ambulance	0.05	0.86
Tow Truck	0.73	0.01
Police	0.46	0.02
Towed away	0.77	0.01
Vehicle characteristics		
Make Audi	-1.31	0.18
Make BMW	-0.52	0.44
Make Chrysler/Jeep	-0.95	0.23
Make Daewoo	-1.22	0.02
Make Daihatsu	-0.43	0.32
Make Holden/GMH	-0.11	0.53
Make Honda	-0.63	0.03
Make Hyundai	-0.08	0.81
Make Jaguar	0.51	0.71
Make Kia	0.56	0.68
Make Land Rover	-1.31	0.1
Make Mazda	0.16	0.56
Make Mercedes	-1.72	0.03
Make Mitsubishi	-0.14	0.49
Make Nissan	-0.21	0.32
Make Peugeot	-0.8	0.4
Make Proton	-0.52	0.59
Make Renault	2.22	0.1
Make Rover	0.38	0.78
Make Saab	-0.88	0.38
Make Seat	3.25	0.02
Make Subaru	0.47	0.16
Make Suzuki	-0.12	0.85
Make Toyota	-0.11	0.5
Make VW	-0.1	0.88
Make Volvo	0.65	0.24
Value	2.36E-05	<0.01
Value squared	-3.68E-11	0.04
Age of vehicle	-0.01	0.22
Constant	6.38	<0.01
R-squared	0.26	

In our specification of equation (10) it is assumed that unobserved risk-type, (RT^*) has no *ex post* effect on the cost of a repair. Hence the null hypothesis of no *ex post* moral hazard is $H_0 : \alpha_1 = 0$. When equation (10) was estimated the coefficient $\alpha_1$was 0.29 and statistically significant (see Table 2). The transformed coefficient (Halvorsen and Palmquist, 1980), implied that *ex post* moral hazard increased the cost of repair by 34.3 per cent or 1.34.

The estimates for *ex ante* moral hazard are given by coefficients for α_1 from equations (4) and (7) and are reported in Table 1 as 1.48 and 1.57, respectively. These coefficients are directly interpretable. A movement of from 0 to 1 in α_1 produces a $100 * \alpha_1$ percentage point change in the probability that a RTC occurs (i.e. *ex ante* moral hazard). The IMRAS data reports that the annual probability of an RTC is 0.12. Hence, a coefficient of 1.48 reported by the Recursive MPE implies that *ex ante* moral hazard increased the probability of an RTC by approximately 1.18 (i.e. 1.48 * 0.121); and for the IV MPE, an estimate of 1.19 (i.e. 1.57 * 0.121).

The identity for *TMH* outlined in equation (1) can be represented as follows

$$
\begin{aligned}
EL &\equiv \rho L \qquad (11)\\
\Delta EL &\equiv \Delta(\rho L)\\
TMH &\equiv \Delta\rho \times \Delta L
\end{aligned}
$$

Substituting 1.18 for $\Delta\rho$ and 1.34 for ΔL, implies that *TMH* is 1.58. That is insurance increased the expected loss of an RTC by 58 per cent This result is consistent with the findings obtained from our Recursive MPE approach (75%). Furthermore, we estimate that the relative contribution of of *ex ante* moral hazard and *ex post* moral hazard to TMH are 34.6 per cent and 64.4 per cent , respectively.[3]

[3]Ex ante % $= \frac{0.18}{0.18+0.34} \times 100$ & Ex post % $= \frac{0.34}{0.18+0.34} \times 100$

IV. DISCUSSION

It has been claimed that "[t]he topic [of moral hazard] divides naturally into *ex ante* moral hazard and *ex post* moral hazard" (Winter, 2000, p. 207). Within the insurance economics literature, the empirical analysis of *ex ante* and *ex post* moral hazard in markets for automobile insurance has invariably been segregated and analyzed separately. The incentives, which influence policyholder behavior, are clearly distinct. *Ex ante* moral hazard occurs when insurance reduces unobserved preventive effort resulting in a correlation between insurance and the insured event, conditional upon the insurer's information set. *Ex post* moral hazard occurs when conditional upon a claim, the policyholder and or service provider has an incentive to increase the cost of an RTC repair.

No doubt, the separation of empirical analyses of *ex ante* and *ex post* moral hazard can partially be attributed to a desire to address the endogeneity, which has vexed efforts to quantify *ex ante* moral hazard. It has been claimed that

> [t]he disentanglement of adverse selection and [*ex ante*] moral hazard is probably the most significant and difficult challenge that empirical work on adverse selection in insurance markets faces.(Cohen and Siegelman, 2010, p. 71),

and indeed, several highly respected economists (Chiappori and Salanié, 2000; Abbring, 2003; Dionne et al., 2013) have been motivated to publish empirical tests for *ex ante* moral hazard in markets for automobile insurance.

In this paper, however, we have argued that *ex ante* moral hazard only captures a portion of the total moral hazard effect. Economists seeking to quantify the effect that insurance has on the expected cost of RTC repairs, should consider both *ex ante* and *ex post* moral hazard, as both variants could result in an efficiency loss in this market. We presented an empirical test for total moral hazard, which encompassed *ex ante* and *ex post* moral hazard, in a market for automobile insurance using a mixed-process estimator approach. Our results indicate that total moral hazard increases the expected cost of automobile repairs by between 75 per cent and 127 per cent, depending upon the chosen estimator. Due to the larger sample size we have greater confidence in the former result. We estimated that the relative effects of *ex ante* moral hazard (1.18) and *ex post* moral hazard (1.34) in the Recursive MPE were responsible for approximately 34.6 per cent and 64.4 per cent of the total moral hazard effect, respectively.

The principal strength of our analysis lies with our data. The IMRAS survey data set included a rich set of covariates that provided a good approximation of the insurer's information set. In addition, the survey reported data not typically collected by insurers. We therefore exploited past RTCs (Dionne et al., 2013; Rowell et al., Early view) and the insurance status of a secondary vehicle (Rowell et al., Early view) to include unbiased estimates of *ex ante* moral hazard in our MPE. Furthermore, access to a diverse array of post-RTC data including, attendance of emergency services and extent of RTC repair, enabled *ex post* moral hazard to be incorporated in in our estimate to TMH.

The key finding from our preliminary analysis is that the total moral hazard effect in markets for automobile insurance appears to be considerably larger than stand-alone estimates of *ex ante* moral hazard (Weisburd, 2015; Rowell et al., Early view; Dionne et al., 2013) or estimates of *ex post* moral hazard, limited to fraudulent behavior, have previously suggested. When possible, we would urge other economists to consider conjointly analyzing the impact of *ex ante* and *ex post* moral hazard in markets for insurance.

Bibliography

Abbring, Jaap. H. & Heckman, James. J. & Chiappori Pierre. Andre. & Pinquet Jean. (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.

Abbring, Jaap H., Pierre-André Chiappori and Jean Pinquet (2003), 'Moral hazard and dynamic insurance data', *Journal of the European Economic Association* **1**(4), 767–820.

Arrow, K. J. (1963), 'Uncertainty and the welfare economics of medical care', *The American Economic Review* **LIII**(5), 941–973.

Arrow, K. J. (1985), The economics of agency, *in* J. W.Pratt and R.Zeckhauser, eds, 'Principals and agents: The Structure of Business', Harvard Business School Press, Boston, chapter 2, pp. 37–51.

Artís, M., M. Ayuso and M. Guillen (2002), 'Detection of automobile insurance fraud with discrete choice models and misclassified claims', *The Journal of Risk and Insurance* **69**(3), 325–340.

Artís, Manuel, Mercedes Ayuso and Montserrat. Guillén (1999), 'Modelling different types of automobile insurance fraud behaviour in the spanish market', *Insurance: Mathematics and Economics* **24**, 67–81.

Belhadji, El Bachir, Georges Dionne and Faouzi Tarkhani (2000), 'A model for the detection of insurance fraud', *The Geneva Papers on Risk and Insurance* **25**(4), 517–538.

Caron, Louis and Georges Dionne (1999), Insurance fraud estimation: More evidence from qubec automobile insurance industry, *in* G.Dionne and C.Laberge-Nadeau, eds, 'Automobile insurance: Road Safety, New Drivers, Risks, Insurance Fraud and Regulation (pp. 175-182)', Kluwer Academic Publishers, Boston.

Chiappori, Pierre-André (1999), Asymmetric information in automobile insurance: An overview, *in* G.Dionne and C.Laberge-Nadeau, eds, 'Automobile insurance: Road Safety, New Drivers, Risks, Insurance Fraud and Regulation', Kluwer Academic Publishers, Boston, chapter 1, pp. 1–11.

Chiappori, Pierre-André and Bernard Salanié (2000), 'Testing for asymmetric information in insurance markets', *The Journal of Political Economy* **108**(1), 56–78.

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

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

Dionne, Georges, Christian Gourieroux and Charles Vanasse (2001), 'Testing for evidence of adverse selection in the automobile insurance market: A comment', *The Journal of Political Economy* **109**(2), 444–453.

Dionne, Georges, Pierre-Carl Michaud and Maki Dahchour (2013), 'Separating moral hazard from adverse selection and learning in automobile insurance: Longitudinal evidence from france', *Journal of the European Economic Association* **11**(4), 897–917.

Finkelstein, Amy and Kathleen McGarry (2006), 'Multiple dimensions of private information: Evidence from the long-term care insurance market', *The American Economic Review* **96**(4), 938–958.

Halvorsen, R. and R. Palmquist (1980), 'The interpretation of dummy variables in semilogarithmic equations', *The American Economic Review* **70**(3), 474–475.

Heckman, James J. and George J. Borjas (1980), 'Does unemployment cause future unemployment? definitions, questions and answers from a continuous time model of heterogeneity and state dependence', *Economica* **47**(187), 247–283.

Hölmstrom, B. (1979), 'Moral hazard and observability', *The Bell Journal of Economics* **10**(1), 74–91.

Israel, Mark (2004), *Do we drive more safely when accidents are more expensive? Identifying Moral Hazard from Experience Rating Schemes (Mimeo)*, Kellogg School of Management, Northwestern University.

Marshall, John M. (1976), 'Moral hazard', *The American Economic Review* **66**(5), 880–890.

Mirrlees, J. A. (1999), 'The theory of moral hazard and unobservable behaviour: Part i', *The Review of Economic Studies* **66**(1), 3–21.

Pauly, Mark V. (1968), 'The economics of moral hazard: Comment', *The American Economic Review* **58**(3), 531–537.

Pauly, Mark V. (1974), 'Overinsurance and public provision of insurance: The roles of moral hazard and adverse selection', *The Quarterly Journal of Economics* **88**(1), 44–62.

Pearl, Judea (2009), *Causality Model, Reasoning and Inference*, Cambridge University Press, New York.

Puelz, Robert and Arthur Snow (1994), 'Evidence on adverse selection: Equilibrium signaling and cross-subsidization in the insurance market', *The Journal of Political Economy* **102**(2), 236–257.

Rowell, D., H. S. Nghiem and L.B. Connelly (Early view), 'Two tests for ex ante moral hazard in a market for automobile insurance', *Journal of Risk and Insurance* (DOI:10.1111/jori.12161).

Saito, Kuniyoshi (2006), 'Testing for asymmetric information in the automobile insurance market under rate regulation', *Journal of Risk and Insurance* **73**(2), 335–356.

Shapiro, S. and M. Wilk (1965), ‘An analysis of variance test for normality (complete samples)’, *Biometrika* **52**(3-4), 59–611.

Shavell, Steven (1979), ‘On moral hazard and insurance’, *The Quarterly Journal of Economics* **93**(4), 541–562.

Spindler, Martin, Joachim Winter and Steffen Hagmayer (2014), ‘Asymmetric information in the market for automobile insurance: Evidence from germany’, *The Journal of Risk and Insurance* **81**(4), 781–801.

Subelj, Lovro, Stefan Furlan and Marko Bajec (2011), ‘An expert system for detecting automobile insurance fraud using social network analysis’, *Expert Systems with Applications* **38**(1), 1039–1052.

Weisburd, Sarit (2015), ‘Identifying moral hazard in car insurance contracts’, *The Review of Economics and Statistics* **97**(2), 301–313.

Winter, Ralph A. (1992), Moral hazard and insurance contracts, *in* G.Dionne, ed., ‘Contributions to insurance economics’, Kluwer Academic Publishers, Boston, chapter 3, pp. 61–96.

Winter, Ralph A. (2000), Optimal insurance under moral hazard, *in* G.Dionne, ed., ‘Handbook of insurance (pp. 155-183)’, Kluwer, Boston, chapter 6.

Wooldridge, Jeffrey M. (2000), *Introductory econometrics: A Modern Approach*, South-Western College Publishing, Mason, Ohio.

Zavadil, Tibor (2015), ‘Do the better insured cause more damage? testing for asymmetric information in car insurance’, *The Journal of Risk and Insurance* **82**(4), 865–889.

Zweifel, Peter and Friedrich Breyer (1997), *Health Economics*, Oxford University Press, New York.

Appendix 1: Descriptive Statistics for RTCs

Description	Frequency	(%)	Insured (%)
RTCs 1997-99	994	100	79.5
Services attending RTC			
Ambulance	45	4.5	80
Tow truck	142	14.3	83.8
Police	126	12.7	77.8
Vehicle towed away	104	10.5	84.6
Damage to vehicle			
Front bonnet	318	32	82.1
Front bumpers	284	28.6	83.1
Front lights	231	23.2	82.7
Front windscreen	84	8.5	84.5
Left side doors	161	16.2	78.9
Front wheel	61	6.1	80.3
Rear wheel	33	3.3	90.9
Left side front panel/mudguard	174	17.5	79.3
Left side back panel/mudguard	111	11.2	81.1
Right side doors	170	17.1	77.6
Right side front panel/mudguard	148	14.9	85.1
Right side back panel/mudguard	120	12.1	85.8
Rear boot/hatch back	148	14.9	87.2
Rear bumper	225	22.6	88
Rear windscreen	43	4.3	88.4
Rear lights	113	11.4	81.4
Interior front dashboard	32	3.2	87.5
Interior other	22	2.2	95.5
Roof of car	99	10	89.9
Engine	49	4.9	85.7
Other	20	2	90
Aerial	4	0.4	100
Framework	9	0.9	77.8
Bodywork	16	1.6	81.3
The whole rear section	1	0.1	100
The bull-bar	6	0.6	100
Side windows	36	3.6	88.9
It was burnt out/ written off	22	2.2	86.4
Mirrors	4	0.4	75
Accessories	22	2.2	81.8
Door locks	47	4.7	78.7
Seats	5	0.5	100
Muffler/ exhaust	3	0.3	66.7
Type of damage unknown	6	0.6	83.3
Fuel tank	3	0.3	100

Appendix 2: Descriptive Statistics from IMRAS Data Set

Variables	Obs.	Freq.	Freq.(%)	RTC(%)	Insured (%)
Comprehensively insured	4005	3163	79	25	n.a.
RTC 1997-99	4005	994	24.8	n.a.	79.5
Driver Demographics					
Aged 18 to 24 years	3971	356	9	32.9	52.8
Aged 25 to 34 years	3971	826	20.8	25.5	76.8
Aged 35 to 44 years	3971	1031	26	25.5	81.6
Aged 45 to 54 years	3971	848	21.4	25.1	83.1
Aged over 55 years	3971	910	22.9	20.3	84.6
Male	4005	1964	49	23.7	76
Nominated Driver $<$ 25 years	4005	475	11.9	32.4	74.7
Private registration	3985	3770	94.6	24.7	78.6
Metropolitan/rural	4005	2520	62.9	27.6	80.6
SES poorest	4005	806	20.1	22	76.9
SES poor	4005	1092	27.3	24.4	75.6
SES rich	4005	1110	27.7	27	81
SES richest	4005	997	24.9	25.2	82
Licensed 0 to 5 years	3945	338	8.6	33.4	51.5
Licensed 6 to 10 years	3945	440	11.2	27	71.1
Licensed 11 to 15 years	3945	434	11	26.3	80.4
Licensed 16 to 20 years	3945	609	15.4	25.5	83.4
Licensed 21 to 25 years	3945	485	12.3	24.5	82.3
Licensed $>$ 25 years	3945	1639	41.5	22.1	83.8
Income $<$ \$20,000 p.a.	3250	332	10.2	20.2	70.8
Income \$20,000 to \$39,999 p.a.	3250	612	18.8	21.1	72.5
Income \$40,000 to \$59,999 p.a.	3250	592	18.2	24.3	80.9
Income \$60,000 to \$79,999 p.a.	3250	397	12.2	27.5	81.1
Income \$80,000 to \$99,999 p.a.	3250	262	8.1	25.6	89.3
Income \$100,000 to \$149,999 p.a.	3250	214	6.6	35	86.9
Income $>$ \$150,000 p.a.	3250	117	3.6	25.6	88
Income Refused to divulge	3250	724	22.3	18.9	77.9
Profession upper white	4005	1166	29.1	30	85.6
Profession lower white	4005	761	19	26	81.2
Profession upper blue	4005	612	15.3	22.2	72.9
Profession lower blue	4005	172	4.3	19.2	66.3
Profession home duties	4005	403	10.1	18.6	76.7
Profession student	4005	164	4.1	35.4	51.8
Profession retired	4005	571	14.3	18.9	85.3
Profession unemployed	4005	71	1.8	22.5	64.8
Refused to divulge profession	4005	85	2.1	23.5	70.6
Vehicle characteristics					
4-cylinder vehicle	4005	2570	64.2	27.2	79.5
6-cylinder vehicle	4005	1273	31.8	21	78.4
8-cylinder vehicle	4005	162	4	17.9	75.3
Make Ford	4005	795	19.9	21.1	77.5

Variables	Obs.	Freq.	Freq.(%)	RTC(%)	Insured (%)
Make Holden	4005	750	18.7	22.5	75.9
Make Toyota	4005	788	19.7	28.6	81.3
Make Mitsubishi	4005	437	10.9	24.3	80.3
Make Asian	4005	1006	25.1	27.1	79.6
Make European	4005	229	5.7	21	80.8
Body-type Sedan	4005	3385	84.5	25.5	78.8
Body-type Commercial	4005	250	6.2	21.6	68
Body-type 4 WD	4005	295	7.4	20	87.5
Body-type Sports car	4005	81	2	24.7	87.7
Car age 0 to 3 years	4005	992	24.8	25.4	93.3
Car age 3 to 7 years	4005	994	24.8	26.3	93.5
Car age 7 to 12 years	4005	950	23.7	25.9	80
Car age $>$ 12 years	4005	1069	26.7	22	51.3
Value $<$ \$2000	3505	423	12.1	24.1	39.5
Value \$2001 to \$5999	3505	795	22.7	25	63.9
Value \$6001 to \$10000	3505	700	20	26.7	84.1
Value \$10001 to \$16000	3505	673	19.2	26.3	91.8
Value \$16001 to \$25000	3505	575	16.4	24.5	93.6
Value $>$ \$25000	3505	339	9.7	23.6	94.7
RTC history					
RTC 1994-97	3984	603	15.1	30.5	81.6