Moral Hazard in the Australian Market for Comprehensive Automobile Insurance

David Rowell

This paper is a summary of research presented to the

School of Economics the University of Queensland for the degree of

Doctor of Philosophy

My academic supervisors were

Professor Luke Connelly and Dr Hong Son Nghiem

Disclaimer:

Acknowledgement:

I am grateful for the support provided by the Brian Gray Scholarship (jointly funded by the Australian Prudential Regulation Authority and the Reserve Bank of Australia).

1. Introduction

Each year, thousands of Australians experience some type of loss due to a road traffic crash (RTC). RTCs account for the second- largest source of preventable injury in Australia (Mathers *et al.* 1999). The consequences of an RTC can be devastating. Some people die and many more suffer injuries. These injuries can have deleterious consequences for both physical and psychological health, affecting the quality of life and economic well-being. Connelly and Supangan (2006) estimate that RTCs cost more than $17 billion per annum in Australia – an amount that is equal to approximately 2 percent of the Australian gross domestic product. The human and economic costs associated with RTCs have motivated the publication of a rich road safety literature. Factors including driver behaviour, automobile design, and road quality have been thoroughly researched and have been commonly been implicated with RTCs.

While the annual costs of RTCs to the Australian community are predictable, their impact on any one given individual, *ex ante* remains unknown. The community's response to this uncertainty has been the development of a market for automobile insurance. Automobile insurance enables risk averse individuals to defray those costs of an RTC, for which the driver is deemed legally liable, which include the cost of injury and property damage. The broader insurance literature has hypothesised a positive relationship between the act of purchasing insurance and the probability of claim, which has been termed moral hazard. However, within the road safety literature scant attention has been paid to the potential for automobile insurance to cause RTCs.

The motivation for this dissertation was to explore what causal relationship, if any, might exist between automobile insurance and the occurrence of an RTC. It has been recognised however, that a positive correlation between insurance and claim could be due to at least two processes observed in markets for insurance. The first, is a process of adverse selection, whereby high-risk individuals with private information about their risk-type purchase more insurance than they would otherwise. This behaviour can in part explain positive correlation between insurance and claim. The second, is a process known as moral hazard which has been defined as the "detrimental effect that insurance has on an individual's incentive to avoid losses." (Winter 1992, p. 61)

2. A Literature Review

2.1 A History of Moral Hazard

The term moral hazard, which was developed by the insurance-industry literature and subsequently used analysed within the economic literature, refers to the "impact of insurance on the incentives to reduce risk" (Winter 2000, p. 155). This concept has since been used to analyse a "wide variety of public policy scenarios, from unemployment insurance, corporate bailouts, to natural resource policy (Hale 2009). The phrase moral hazard has obvious and powerful rhetorical capabilities to moderate social attitudes towards the process of insurance. As the term moral hazard made the transition from the narrow confines of the insurance literature to the public domain, some social commentators have questioned the normative implications of the term. For example Tom Baker, a lawyer, has argued that:

> Today, moral hazard signifies the perverse consequence of the well-intentioned efforts to share the burdens of life, and it also helps deny that refusing to share those burdens is mean-spirited or self-interested. Indeed using the economics of moral hazard, it is but a short step to claim, in one economist-politician's memorable word, that "[s]ocial responsibility is a euphemism for individual irresponsibility. (pp. 239-240)

The non-economic literature has been strident in its criticism of economics and economists. For example, Baker states,

> (b)y "proving" that helping people has harmful consequences, the economics of moral hazard justify the abandonment of legal rules and social policies that try and help the less fortunate... (Baker, 1996, p. 240).

The overarching concern of this literature has been the capacity of the phrase "moral hazard" to influence social policy. Dembe and Boden have argued:

> Indeed, the concept of moral hazard is widely used and deeply entrenched in the practice of economics that little attention is paid to the underlying ethical and moralistic notions suggested by the use of that particular expression. (Dembe & Boden 2000, p. 258)

Benjamin Hale, a philosopher has claimed.

Figure 1 Teething Problems

> One thing that should be clear about the terminology of "moral hazard" is that the language invokes a normative notion. It suggests that there is a moral danger, a moral problem, associated with the over provision (or overprovision) of insurance. (Hale 2009, p. 2)

The Oxford English Dictionary defines an idiom "a group of words established by usage as having a meaning not deducible from those of the individual words." The English language contains many idioms. Clearly, not all 'teething problems' require a dentist (see Figure 1) and not all 'free trade' is free.

Pauly (1968, p. 531) has argued that "...the problem of 'moral hazard' in insurance has, in fact, little to do with morality but can be analyzed with orthodox economic tools." The purpose of the following historical review is to explore the underlying reasons for the divergence in what is understood by moral hazard.

Hale (2009) and Pearson (2002) have stated that the concept of moral hazard has developed along with insurance.

> Talk of moral hazards has been around since at least as long as the modern insurance industry, which some date as far back as 1662. (Hale 2009, p. 3)

However, the history of insurance does not support this claim. The earliest risk management techniques were used 7000 BC (Hart *et al.* 2007). Chinese merchants would disperse their cargo across several ships to spread their risk (Vaughan 1997). The oldest evidence of an insurance contract can be found in *The Hammurabi Code,* which was written, in Babylon in 1790 B.C. Law 48, states:

> If a man owe a debt, and the god Adad has flooded his field, or the harvest has been destroyed, or the corn has not grown through lack of water, then in that year he shall not pay corn to his creditor. He shall dip his tablet in water, and the interest of that year he shall not pay. (Edwards 1921, p. 20)

Greek and Roman merchants subsequently used 'bottomry' loans to transfer their risk to moneylenders, by borrowing money with a clause, which annulled their debt if their ship sunk (Hart *et al.* 2007). The earliest known European contract was underwritten in Genoa in 1343 (Ceccarelli, 2001) and the oldest preserved English contract was dated 1547 (Hart *et al.* 2007). In 1666, the Great Fire of London provided an impetus for fire insurance. Lloyd's of London was established in 1688, to enable slave merchants to insure for their losses at sea. Gamblers who congregated at the Lloyd's coffee house could accept liability for some proportion of shipping losses in exchange for a premium, by **writing** their name **under** the line; hence the origin of term the underwriter (Bernstein 1996). The industrial revolution saw the development of other lines of insurance including life insurance (Pearson 2002). Yet, the term moral hazard did not appear in the insurance-industry literature until 1865 (Baker 1996), suggesting that modern European insurance predated discussions of moral hazard by some 530 years. Clearly, the concept of moral hazard did not, as claimed by Hale (2009) and Pearson (2002) simply evolve with the development of insurance.

A surprisingly rich medieval literature had evolved to examine insurance. An intense theological debate was centred on whether insurance was a licit remedy for the commercial costs ensuing from acts of God. The Church considered random events to be outcomes of divine will and hence events that were not to be anticipated. Simony was condemned because it was viewed as the sale of Christ's charisma and usury was condemned because it was the sale of God's time. In 1234, Pope Gregory IX issued the *decretal Naviganti,* which stated that insurance was illicit (Ceccarelli 2001).

Theologians who supported the *Naviganti,* such as the Portuguese Carmelite João Sobrinho [1400- 1475], argued that the insurer was selling something that was not rightfully theirs to sell and that the safety of a venture proceeds only from God's will. Similarly, the fifteenth century French theologian Peter Tartaret, argued that the insurer should not profit from human presumption of safety since safety can only be granted by God (Ceccarelli, 2001).

Debate developed. As commercial insurance grew a counter argument evolved, which was sometimes promoted by theologians with familial ties to merchant traders. Thomas Aquinas [1225-1274] argued that since insurance did not affect ownership it was not usury. Domingo Soto argued that risk was an economic object, which insurance has enabled merchant and insurer to share. *Assecuratio* was licit because it allowed licit business to prosper (Ceccarelli 2001). By the sixteenth century, mercantilism had moderated the Church's opposition to insurance. However, theologians from both sides of the debate continued to accept that chance events, (e.g., a ship sinking), were the product of God's will. While ever this fundamentalist view of providence prevailed, the concept of moral hazard, which posits that individual behaviour can affect chance events, could not exist.

Febvre (1956) has argued that the growth of insurance changed our perception of nature. Future events were no longer solely attributed to God's will and individual behaviour was recognized as a co-determinant. The seventeenth century writings of the Flemish theologian, Leonardus Lessius, support Febvre's argument. Lessius argued that profits were derived from professional ability and not solely from God's will:

> In the case of the insurer having professional skill (i.e., has received relevant news by mail) or experience having with natural phenomena (i.e., he is aware that the weather will be good), who knows that the real value venture risks are less than what the current market place estimates and who does not reveal it to the other party, still the market price is to be considered just, because probable profits derive from his professional ability. (Lessius 1605) as quoted in Ceccarelli (2001, p. 627)

This transformation of thought was a necessary precursor to the eventual development of a theory of moral hazard.

Despite the emergence of insurance during the Middle ages, Dasaton (1988) has argued that the application of actuarial science by insurers was constrained because the science of probability remained in its infancy. Probability as an independent area of study did not formally commenced until the seventeenth century when the French mathematicians Pascal and Fermat analysed the fair division of stakes from an incomplete game of dice (Daston 1983). During the next 150 years three seminal probability texts followed, *On Reasoning in Games of Chance* [Huygens 1657], *The Art of Conjecturing* [Jakob Bernoulli 1713] and *Analytical Theory of Probability* [Laplace 1812]. Yet in the insurance industry, underwriting techniques remained rudimentary. For example, the mutual society, Amicable, whic was established in 1706, charged the same membership fee irrespective of age. Life insurers used the pseudo-science of physiognomy[1] to assess the health status of prospective policyholders. Royal Exchange sold life assurance policies, without medical review until 1838. Until 1850, London Assurance and Royal Exchange used only three risk classifications: (common, hazardous and doubly hazardous) to underwrite fire insurance (Pearson 2002).

Although, in principle, insurers from the eighteenth century differentiated between physical and moral risks, the lack of an actuarial science meant this distinction was abstract rather then concrete. Attributions such as character, probity, temperance, ethnicity and class were used to assess both physical and moral risks. English insurers, for example, identified Irish and Jewish populations as being morally suspect (Pearson 2002). The concept of moral hazard requires that insurers can differentiate between the risk of the insured and uninsured. Thus, the absence of the empirical tools to quantify risk was a practical constraint on the development of a concept of moral hazard.

It was not until 1865 that the term moral hazard first appeared in *The Practice of Fire Underwriting*, as:

> . . . the danger proceeding from motives to destroy property by fire, or permit its destruction. (Ducat 1865, pp. 164-165) in Baker (Baker 1996, p. 249)

The genesis of an idiom is an ill-defined process, which pairs concept with phrase. Baker (1996) contends that the creation of 'moral hazard' suited the times; from 'hazard' with its moral overtones of danger; and from 'moral', referring to the moral scientists who made chaste use of the odds. The contemporary definition of **hazard** is:

> [r]isk of loss or harm, peril, jeopardy. (OED 2008)

However, the word hazard ('*hassard*' or '*hasart*'), has French origins and first entered the English language in 1167 to describe a game of dice. It was not until 1618 that the word 'hazardous' as in 'perilous' first appeared in the English language (Harper 2001). Synergistically, Pascal and Fermat had developed probability theory to analyse a game of dice or *hassard*, Baker (1996) had observed Victorian England considered morally questionable.

The contemporary meaning of the word **moral** is

[1]The Oxford English Dictionary defines physiognomy as the study of the features of the face, or of the form of the body generally, as being supposedly indicative of character; the art of judging character from such study.

Excellence of character or disposition... (OED 2008)

However, in the Middle Ages philosophical thought was often written in Latin and subsequently translated into English. While, Latin defines the word *mōrālis* as "... of or belonging to manners or morals, moral" (Lewis *et al.* 1879) its etymology comes from the word *mos* (Daston 1983), which is defined as

> ... manner, custom, way, usage, practice, wont, as determined not by the laws, but by men's will and pleasure, humor, self-will, caprice.(Lewis *et al.* 1879)

Thus the medieval use of the word 'moral' often retained its Latin sense, which did not always project the pejorative overtones associated with its English use. Baker asks,

> [w]hat combination of words could better signify the serious, scientific and highly proper –indeed moral– grounding of the insurance enterprise? (Baker, 1996, p. 248)

The idiom resonated and soon became entrenched within the insurance-industry literature. In 1867 the *Aetna Guide to Fire Insurance Handbook* developed a dual conception of moral hazard. The first was a moral hazard, due to character:

> Consider first the moral hazard....What is the general character borne by the applicant? Are his habits good? Is he an old resident, or a stranger and an itinerant? Is he effecting insurance hastily, or for the first time? Have threats been uttered against him? Is he peaceable or quarrelsome -popular or disliked? Is his business profitable or otherwise? Has he been trying to sell out? Is he pecuniarily embarrassed? Is the stock reasonably fresh and new, or old, shopworn and unsalable? When was the inventory last taken? Is the amount of insurance asked for, fully justified by the amount and value of the stock? Is a set of books systematically kept? (Aetna Insurance Co. (Aetna) 1867, p.21) as quoted in Baker (Baker 1996, p. 250)

The second conception was moral hazard due to temptation:

> Heavy insurance also increases the moral hazard, by developing motive for crime, where otherwise no temptation existed, and wrong was in no way contemplated. (Aetna Insurance Co. (Aetna) 1867 p.159) as quoted in Baker (Baker 1996, p. 251)

Baker (1996) claims that moral hazard, was due to either (i) a deliberate act of fraud or (ii) unintended act of carelessness; the former was immoral and the latter was not.

However, what Baker (1996) has sometimes described as moral hazard is upon closer inspection, is often, in fact, adverse selection. Reconsider Baker's quote from the *Aetna Guide*

> Consider first the moral hazard… (Baker 1996, p. 250)

It is only the penultimate sentence, which refers to the amount of insurance requested, that relates to moral hazard, in the modern sense. The dominant focus of this quote is the identification of pre-existing personal characteristics, which are correlated with the propensity to commit fraud. It is therefore adverse selection rather than moral hazard, which is the focus of this quote. The following sentence supports this conclusion.

> Character, or the individual predisposition for fraud or loss, is a dominant concern here. (Baker, 1996, p. 250)

When character is a **predisposition**, which precedes the purchase of insurance, clearly adverse selection rather than moral hazard is the focus of this analysis. Two further examples of moral hazard being used to describe adverse selection are provided below.

> If the moral hazard is not good, there are no considerations that would induce the company to accept the risk… (Tiffany 1882, p. 24) in Baker (Baker 1996, p. 253)
>
> Moral Hazard- The character of the applicant is usually of the first importance; and where this is not satisfactory, the applicant should be dismissed at once. (Aetna Insurance Co. (Aetna) 1867, p. 13) in Baker (Baker 1996, p. 253)

The following statement from Baker (1996) illustrates that he too recognised that there was some ambiguity in the way the phrase 'moral hazard' was used within the early insurance-industry literature.

> For nineteenth century insurers, moral hazard was a label that applied to people and situations. (Baker 1996, p. 240)

The parallel should be clear. When moral hazard applied to situations, it embodied moral hazard in the modern sense, i.e., a response to incentives. When moral hazard applied to people, it reflected an alternative meaning of moral hazard as, *de facto*, a process of adverse selection.

In 1905, E. U. Crosby who was employed as the General Agent for the North British and Mercantile Insurance Company published the following definition of moral hazard in *The Annals of the American Academy of Political and Social Science.*

> First, we have "direct moral hazard" where a property is fired by the owner for gain. Second, the "indirect moral hazard" where the owner may not be prospering or permanently located, and has little or no incentive for safekeeping hazard, keeping premises in repair and maintaining fire appliances, thus allowing the physical hazard to become abnormally high. (Crosby, 1905, pp. 225-226)

While this definition of moral hazard did refer to illegal acts such as arson, the primary focus of the analysis was on the incentives that produced the conduct rather than the identification of those high-risk individuals likely to file a claim. This definition of moral hazard is 'modern' in the sense that it does not imply that moral hazard is behaviour that is perpetrated by immoral persons. The following quote emphasises this point:

> The record of fire losses has clearly shown that moral hazard is frequently found among assured of means and of high social standing or with excellent mercantile ratings. (Crosby 1905, p 226)

At the beginning of the twentieth century, the insurance industry began to embrace a public role. Stone (2002), has argued that insurance is a social institution, which defines norms and values in political culture and ultimately shapes the way citizens think about issues of membership, community, responsibility and moral obligations. Community attitudes to risk aversion, fraud, propensity to claim and preventive effort can affect the viability of insurance. Insurers have an incentive to shape and define social norms, which promote individual and mutual responsibility and maximize commercial prosperity. Pearson (2002) has argued the insurance industry has promoted the idea that public resources should be committed to the amelioration of moral hazard.

> In an imperfect market, however, where costs and benefits were not precisely known, and in a society in which moral precepts dictated that property should be protected from fraud, theft, and arson whatever the cost, insurers may have felt that the allocation of both private and public resources to combating moral hazard should not be based on economic factors alone. (Pearson 2002, p. 8)

Throughout the first half of the twentieth, many examples of the pejorative use of the phrase 'moral hazard' can be found within the insurance-industry literature. In 1907, H.P.

Blunt wrote in the *Journal of the Insurance Institute* that:

> As regards the low-class alien population so much in evidence now-adays in our crowded centres, the rigid exclusion of these from their books is held by first-class offices to be a duty owing not only to their shareholders but also to the State, seeing that a policy in such hands is likely to be an incentive to crime. (Blunt 1907) in Pearson (2002, p. 35)

Moral hazard has variously described as "[m]en who steal or lie [or] magnify a slight injury, or be dilatory in resuming work when they are able to do so" (McNeill 1900) in Dembe and Boden (2000, p. 259) and "misrepresentation and negligence" (Campbell 1902) in Dembe and Boden (2000, p. 259). In 1935, the *Dictionary of Fire Insurance* stated that:

> [c]ertain features affect moral hazard abroad which are fortunately absent in Great Britain. For instance, Central America has long been recognised as a hotbed of serious moral hazard... A type known as Assyrians do not hesitate to adopt any means or make any statements so as to secure payment of policy moneys in full. (Remington & Hurren 1935, p. 328)

Other ethnocentric references to moral hazard include:

> Moral hazard was typically attributed to the (immoral) personal characteristics of individual, but some authorities claimed that it was more likely among certain ethnic and social groups like ... drug addicts and homosexuals. (Rupprecht 1940; Shepherd & Webster 1957) as quoted in Dembe and Boden (2000, p. 259)

In contrast the economic literature perceived moral hazard becuase of incentives rather criminality. Without explicitly using the idiom moral hazard in 1890, Alfred Marshall nevertheless recognised that insurance had the capacity to induce carelessness and fraud.

> Even as regards losses by fire and sea, insurance companies have to allow for possible carelessness and fraud; and must therefore, independently of all allowances for their own expenses and profits, charge premiums considerably higher than the true equivalent of the risks run by the buildings or the ships of those who manage their affairs well. (Marshall 1920, p. 231)

In 1895, John Haynes introduced the concept of moral hazard in *The Quarterly Journal of Economics* as:

> Lack of moral character gives rise to a class of risks known by insurance men as moral hazards. The most familiar example of this class of risks is the danger of incendiary fires. Dishonest failures, bad debts etc. would fall into this class, as well as all forms of danger from the criminal classes. (Haynes 1895, p. 412)

The following passage demonstrates that Haynes (1895) understood that moral hazard modified behaviour.

> Security is good, but security as well as [moral] hazard may have an unfavourable effect upon industry.
>
> (a) Intensity of effort is diminished...
>
> (b) Carelessness is encouraged by insurance...
>
> (c) The greatest disadvantage of technical insurance is the encouragement it gives to dishonesty.
>
> (Haynes 1895, p. 445)

Furthermore, Haynes (1895) did not view dishonesty is precondition for moral hazard to occur. He states:

> There would still remain the moral hazard of excessive estimates of loss where there was no dishonesty in the origin of the fire. (Haynes 1895, p. 445)

Not only did the twentieth century economic literature display a comparatively sanguine approach to moral hazard it also began to broaden the application of the concept beyond the limited confines of the private markets for insurance. In 1913, I. M. Rubinow wrote the following observation about [moral] hazard in his monograph *Social Insurance*:

> But the most damaging argument in the opinion of many is the charge that social insurance not only increases hazard, but vastly more stimulates the simulation of accidents or disease or unemployment; and that encourages the professional mendicant, demoralizes the entire working class by furnishing an easy reward for malingery. (Rubinow 1913, p 496)

In 1921, Frank Knight broadened the application of moral hazard to analyse the implications of incentives within different corporate structures.

> On the other hand, it is the inefficiency of organization, the failure to secure effective unity of interest, and the consequent large risk due to moral hazard

> when a partnership grows to considerable size, which in turn limit its extension to still larger magnitudes and bring about the substitution of the corporate form of organization. (Knight 1921, p 131)

Then in 1963, Kenneth Arrow piqued the interest of a new generation of economists in moral hazard when he outlined an efficiency argument for public intervention in the market for medical care based on moral hazard. He said,

> [t]he welfare case for insurance policies of all sorts is overwhelming. It follows that the government should undertake insurance in those cases where this market, for whatever reason [e.g., moral hazard], has failed to emerge. (Arrow 1963, p 961)

Today contemporary economic analysis uses moral hazard to analyse a diverse range of social issues, including worker's compensation (Butler & Worrall 1983) and disability benefits (Chelius & Kavanaugh 1988), share copping (Cheung 1969), stock market (Diamond 1967) and family behaviour (Becker 1981). There is scarcely an area of economic study where consideration of moral hazard and consequent incentives does not play a role (Coyle 2007). While, within the economic literature, this idiom has remained largely free of strong moral overtones a pejorative tone has persisted within the insurance-industry literature. For example, the insurance text *Risk Management* states:

> *Moral hazard*: refers to the increase in probability of loss associated which results from evil tendencies in the character of the insured person....
>
> *Morale hazard*: not to be confused with moral hazard, results from the insured person's careless attitude towards the occurrence of losses. The purchase of insurance may create a morale hazard, since the realization that the insurance company will bear the loss... (Vaughan 1997, p. 12)

This taxonomy, although not frequently used, may also be confusing. What insurers evidently have called morale hazard, 'carelessness due to incentives', is what economists such as Arrow (1963), Pauly (1968) and others have continued to call moral hazard. What the insurance-industry literature has termed moral hazard i.e. 'evil tendencies', has in this paper been called adverse selection, with the following caveat. The concept of asymmetric information is central to the economic concepts moral hazard and adverse selection. If the 'evil tendencies' were unobservable to the insurer then any unobserved self-selection by potential policyholders will result in a process of adverse selection. If the 'evil tendencies' were observable then the insurer will risk rate their policyholders accordingly and no process of adverse selection will occur.

Contemporary insurance texts have continued to advance definitions of moral hazard that embody self-selection. For example, moral hazard has been variously defined as:

> ... an imputed subjective characteristic of the insured that increases the probability of loss. (Mehr & Cammack 1976, p. 23)

and in an Australian text,

> ...moral hazards, such as dishonesty, carelessness, and lack of concern. (Hart *et al.* 2007, p. 1)

as opposed to those definitions found within economics literature (e.g., Mas-Colell (1995) and Varian (1992)), which focus on the role of incentives. Pindyck and Rubinfeld (1989), for example, simply state that the problem of moral hazard is that,

> ... behaviour may change after the insurance has been purchased. (p. 620)

This historical review offers at least three useful insights. First, two contrasting treatments of moral hazard by the economic and insurance-industry literatures sowed the seeds for an energetic public debate. The economic literature has applied a 'value-neutral' idiom, which has been used pejoratively within the confines of the insurance-industry literature, to an expanding range of topics of social interest. Deirdre McCloskey has argued that economists,

> ... want to persuade audiences, too, and therefore exercise wordcraft, in no dishonourable sense. (McCloskey 1994, p. 51)

The phrase 'moral hazard' has the potential to alienate readers. Economists therefore, need to be mindful of the potential this idiom has to obfuscate their message. Secondly, other social scientists interested in history of moral hazard and insurance, could consider the context in which the phrase moral hazard appears in the literature through the –moral hazard adverse selection– paradigm. Sometimes what at first appears to be a discussion of moral hazard on closer inspection may in fact more appropriately be viewed as a discussion of adverse selection.

Thirdly, to this day definitions of moral hazard found within the insurance-industry literature do not always correspond with the concept as it is used in economics. It may be beneficial therefore if conversation between the disciplines of economics and insurance standardized the meaning of moral hazard. A clear a distinction between moral hazard (insurance moderates behaviour) and adverse selection (behaviour moderates insurance status and choice of policy) would be a useful one. It is this last point, the differentiation of moral hazard from adverse selection, which is the central issue that is addressed in the forgoing empirical analysis.

2.2 Empirical Literature

The problem of moral hazard may arise when individuals purchase insurance under conditions such that their privately taken actions affect the probability distribution of the outcome (Holmstrom 1979). The economics literature contains a small number of studies that have attempted to estimate the effect of moral hazard in markets for RTC insurance. The differentiation of moral hazard from adverse selection is empirically challenging because both phenomena are associated with a positive correlation between the decision to purchase insurance and the probability of an accident. However, the directions of the causality are opposite. For example, in a market for RTC insurance moral hazard will induce those drivers who purchase insurance to have more accidents, while adverse selection will induce poor drivers to purchase more insurance, *ceteris paribus*. It has been stated:

> The disentanglement of adverse selection and moral hazard is probably the most significant and difficult challenge that empirical work on adverse selection [or moral hazard] in insurance markets faces. (Cohen & Siegelman 2010)

The modern debate on asymmetric information in auto insurance markets can be traced to the work of Puelz and Snow (1994), which used individual claims data to construct an ordered logit model, which showed a correlation, conditional upon the insurer's information set, between risk-type and choice of deductible. Puelz and Snow (1994) argued that this negative and statistically significant sign on its coefficient constitutes empirical evidence of adverse selection; however, two major criticisms were levelled at these results. First, this test does not distinguish adverse selection from moral hazard. Specifically, the result reported as adverse selection by Puelz and Snow (1994) is also consistent with the hypothesis that moral hazard (with or without adverse selection) exists in this market (Chiappori 1999). Second, several econometric issues were raised, the most important of which is that the model was incorrectly specified (Dionne, Gourieroux and Vanasse (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 are accounted for, the residual correlation, (which was interpreted as adverse selection), vanishes.

Chiappori and Salanié (2000) proposed an alternative test for asymmetric information using a bivariate probit model wherein the first probit predicts the level of insurance and the second probit predicts the occurrence of a claim. The null hypothesis of no asymmetric information was tested with two parametric tests of the following hypotheses: (i) $H_0 : cov(\varepsilon_i, \eta_i)$ when the two probit models are estimated separately and (ii) $H_0 : \rho = 0$ when the model is estimated as a bivariate probit. These authors used a French claims data set that contains 55 exogenous dummy variables, to control for the insurer's information set (Chiappori & Salanie 2000).

French law stipulates that the risk rating of policyholders is uniformly adjusted using the mandated *bonus-malus* coefficient, as follows. If a claim is submitted, the premium is increased by 25 percent and if no claim is submitted, the premium is decreased by 5 percent. The *bonus-malus* coefficient can range between a maximum of 3.5 and a minimum of 0.5. Chiappori and Salanié (2000) have argued that since the *bonus-malus* coefficient is observable to all insurers its omission would induce a bias, which over-estimates adverse selection, however its inclusion is also problematic since the *bonus-malus* is obviously endogenous. To circumvent this problem their analysis was restricted to samples of beginner drivers who have no claims history. Chiappori and Salanié (2000) report no evidence of asymmetric information in this sub-population of beginner drivers.

Chiappori and Salanié (2000) conclude with a specific test for moral hazard that exploits 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 probits. In the claims probit, Chiappori and Salanié (2000) argue that sign of the coefficient on the *bonus-malus* dummy differentiates three mutually exclusive options. They claim that (i) a negative sign implies that the parents' performances are positively correlated with the child's, (ii) nil correlation implies the parent's performances are uncorrelated with the child's and there is no mortal hazard and (iii) a positive sign implies that the parents' and child's performances are uncorrelated and there is some kind of moral hazard. Chiappori and Salanié (2000) report a negative coefficient, which they argue rejects the moral hazard hypothesis. In their conclusion Chiappori and Salanié (2000) suggest that 'exploiting dynamic data' may offer the best opportunities to test for moral hazard.

Following, the study by Chiappori and Salanié (2000) three distinct methodological responses can be identified. First, (Abbring *et al.* 2003) and Israel (2004) have eschewed conditional correlation to analyse instead longitudinal data. They have argued that while the conditional correlation approach that was used by Chiappori and Salanié (2000) offers a robust test for asymmetric information it fails to distinguish moral hazard from adverse selection.

Abbring *et al.* (2003) adapted a test for state-dependence (Heckman & Borjas 1980) to test for moral hazard. Abbring *et al.* (2003) argue that experience rating, as embodied in the *bonus-malus* scheme, implies a negative occurrence dependence of individual claims intensities under moral hazard: each claim increases the premium, which induces an increase preventative effort to avoid future claims. However, negative occurrence dependence is confounded by a positive correlation associated with the individual's risk type, i.e. policyholders who lodge a claim are poorer drivers and are hence more likely to lodge a future claim (Abbring *et al.* 2003). Abbring *et al.* (2003) analysed 79,684 contracts obtained from a French insurance company for the years 1987-89. In their sample, 4,831 policyholders lodged one claim and 287 lodged more than one claim. A proportional hazard model was used to compare (i) the distribution of first and second claims times across contracts and (ii) the first and second claims times of each contract with two claims (or more). No evidence of moral hazard was found (Abbring *et al.* 2003).

Israel (2004), however, 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 behaviour through their effects on the premium. To address this limitation Israel (2004) analysed a 10-year panel claims data set obtained from an insurance company that is domiciled in Illinois. The premia were risk-rated using drivers' claims histories for the previous three years. Pre 1997, the lodgement of a claim resulted in three risk classifications (i) a 10 percent increase in the premium if no claim had been lodged in the previous 3 years, (ii) a 20 percent increase if one claim had been lodged in the previous three years and (iii) a 50 percent increase in all other instances. Post 1997, the pricing structure for the lodgement of a claim was changed to (i) 10 percent (ii) 40 percent and (iii) 70 percent, increases respectively After three years, claims were removed from the policyholder's record. The hypothesis of moral hazard was tested by examining the occurrence of claims around the three-year insurance event. Israel (2004) finds a small but statistically significant moral hazard effect: as policyholders move from risk classification

(i) to (ii), the cost of the average 6-month premium ($250) is increased by $25, which is associated with a 0.1 percent decrease in the probability of further claim (Israel 2004).

A different approach, also employing panel data, was used by Dionne *et al.* (2004, 2006, 2007, 2010) to test for moral hazard. These authors argued that by limiting the analysis to beginner drivers, Chiappori and Salanié (2000) had omitted a measure of claims history that may conceal a conditional correlation, because this variable is both negatively correlated with contract choice and positively correlated with claims. Dionne *et al.* (2004, 2006, 2007, 2010) used a three-year panel data set to estimate a bivariate probit model that includes the *bonus-malus* coefficient. A Granger causality test was used to test for moral hazard by examining the conditional correlation between the insurance decision in period *t-1* and a claim in period *t* conditional upon the insurer's information set. Dionne (2004) states that switching from all-risk coverage to third-party only coverage reduces the annual probability of a claim by 5.9 percentage points. Subsequent versions of this research report that evidence of moral hazard was restricted to drivers with less than 15 years experience (Dionne *et al.* 2006, 2007, 2010).

An alternative approach was identified by Amy Finkelstein and Kathleen McGarry (2006) who conducted an empirical investigation of asymmetric information in a market for long-term care using individual-level survey data. They obtained Health and Retirement Study (HRS) data rather than claims data. They argue that an advantage of the data set is that it provides a rich description of the market for long-term nursing home insurance, and enables the econometrician to observe the insurer's information set (Finkelstein & McGarry 2006), along with other variables. Finkelstein and McGarry (2006) tested for asymmetric information using probit models for the utilization of long-term care and the purchase of long-term care insurance. They demonstrate that individuals possess private information, which is positively correlated with (i) actual admission to a nursing home and (ii) the purchase of insurance, conditioned upon the insurer's information set. They then use the bivariate model specified by Chiappori and Salanié (2000) to show no evidence of asymmetric information when these controls for the insurance company's prediction and application information are separately included in the model.

Finkelstein and McGarry (2006) hypothesize that individuals may possess other dimensions of private information, which are positively correlated with the preference for insurance and negatively correlated with the propensity to lodge a claim, thus confounding a test for asymmetric information. To test this hypothesis they utilize several variables including gender-appropriate preventive health care measures and seat-belt compliance as proxies for the individual's unobserved preference for insurance. Crucially, these data are not observable by the insurer. Finkelstein & McGarry (2006) then re-estimate both probit models and report that these data are positively correlated with nursing home insurance but negatively correlated nursing home admission. They demonstrate the existence of multiple dimensions of private information, which may potentially confound tests for asymmetric information using standard tests for conditional correlation. Their paper concludes with the following statement.

> There are many examples of information not priced by insurance companies that the econometrician may observe in survey data, such as wealth which is not priced for annuities, occupation which is not priced for auto insurance, and preventive health measures which are often not priced for health insurance. These types of disparities between the information used by insurance companies and that available to the econometrician suggest that this test may find widespread applicability. (Finkelstein & McGarry 2006, p. 952)

The Finkelstein and McGarry (2006) paper has important implications for the empirical estimation of *ex ante* moral hazard, adverse selection, or other multiple dimensions of private information that may exist. Although they did not attempt to estimate the moral hazard effect *per se*, a consequence of their result is that if an econometrician wants to estimate a dimension of private information (e.g., moral hazard), then the collection of data, which control for other dimensions of private information is crucial to its identification.

3. A Proposed Methodology

The capacity of cross sectional data to identify moral hazard may have been discarded with undue haste. The discipline of economics has recognized that variable omission has the potential to compromise statistical inference when the error term is correlated with the explanatory variable for almost 80 years. Econometricians have developed several techniques to address this issue. Working's (1927) canonical empirical analysis of the demand function for pig-iron argued that the omission of variables which capture supply side responses to a change in price can lead to spurious results. Wright (1929) describes an early application of an instrumental variable to analyse the effect of levying a duty on consumption of a commodity. Given the existence of these well-established techniques to address endogeneity in cross-sectional data sets, it remains unclear why the identification of moral

Table 1: Anatomy of an endogenous variable

Generic cross sectional data	Insurance data
Correlation	Asymmetric information
Causation	Moral hazard
Omitted variable bias and selection bias	Private information
	-Adverse selection
	-Advantageous selection

hazard has proved so intractable to empirical estimation.

Conceptually, testing for moral hazard in the setting of insurance can be understood within the paradigm of endogeneity in cross sectional data, more broadly. Insurance however, has developed its own taxonomy, which can be used to understand this phenomenon. The following schematic outlined in Table 1 has matched the terminology that is used to conceptualize endogeneity in generic cross sectional data with the terminology that is used in an insurance setting.

Consider a standard test for asymmetric formation using conditional correlation on claims data.

$$Claim = \alpha_0 + \alpha_1 INS + \alpha_2 \mathbf{X} + RT^* + \varepsilon_i \tag{1}$$

Where

Claim = 1 if a claim lodged and zero if otherwise

INS = 1 if insured and zero if otherwise

X = Vector of variables reflecting the insurer's information set

RT^* = Unobserved risk type

$\varepsilon_i, \eta_i, \mu_i, \varpi_i$= Random error terms

where *Claim* is the dependent variable which is equal to one if a claim is lodged and *INS* is a dichotomous explanatory variable of interest equal to one if a motorist has an insurance policy and ***X*** is a vector of variables which reflect the insurer's information set. The error term has been disaggregated to include one dimension of private information RT^* which denotes the motorist's unobserved risk type and ε_i, which is a random disturbance term. The variable *INS* is endogenous if it is correlated with the unobserved variable RT^*. A positive correlation between insurance and claim could indicate the presence of moral hazard, adverse selection or both.

The endogeneity that characterizes equation (1) can be dealt with by using either (i) a proxy variable or (ii) an instrumental variable (IV). The proxy variable approach requires that a variable such as *RT_Prx*, be included as a proxy for latent variable risk type RT^* in equation (2) as follows:

$$Claim = \beta_0 + \beta_1 INS + \beta_2 \mathbf{X} + \beta_3 RT\_Prx + \eta_i \tag{2}$$

The inclusion of an appropriate proxy could enable the econometrician to test moral hazard with the null hypothesis $H_0 : \beta_1 = 0$. However, claims data sets are problematic because all data, that are available to the econometrician, are provided by the insurer. Therefore, claims data cannot contain a variable that can satisfactorily function as a proxy for RT^*.

Similarly, an IV approach would also require that the candidate IV for insurance *INS_IV* be unobservable to the insurer. The rationale is as follows. To function as an effective instrument *INS_IV* should be correlated with the endogenous variable *INS* but be uncorrelated with the error term ε_i. If the variable *INS_IV* is to provide an unbiased estimate of the coefficient α_1 in equation (1) it must be correlated with latent variable risk type RT^*. Thus a specification such as equation (3) will capture some proportion of the variation in RT^* that was unexplained by equation (1).

$$Claim = \gamma_0 + \gamma_1 INS\_IV + \gamma_2 \mathbf{X} + \mu_i \tag{3}$$

Now consider a specification such as equation (4), which includes all the explanatory variables: *INS*, *INS_IV* and *X*.

$$Claim = \delta_0 + \delta_1 INS + \delta_2 INS\_IV + \delta_3 \mathbf{X} + \varpi_i \tag{4}$$

Equation (4) will more accurately predict the dependent variable *Claim*, than equation (1), because equation (4) includes all the explanatory variables included in equation (1) and the variable *INS_IV* which is correlated with the otherwise latent variable, risk type RT^*. Thus if the variable *INS_IV* improves the prediction of *Claim* and is observable to the insurer, it will be included as a sub-set of the insurers information set ***X*** in the original estimation of equation (1). It follows, therefore, that no matter how rich a claims data set, it will not include data which could serve as either a proxy for RT^* or an instrument for *INS*. Thus, claims data do not include variables that enable one to empirically test for moral hazard. Claims data are the wrong data.

The empirical literature on moral hazard has been subject to considerable debate regarding the ability of various approaches to disentangle adverse selection from *ex ante* moral hazard. The analysis of the HRS for evidence for asymmetric information by Finkelstein & McGarry (2006) has some important implications for the identification of moral hazard. Finkelstein and McGarry (2006) have first demonstrated that standard test for asymmetric information (i.e., a conditional correlation between insurance and accident/claim) is not a necessary condition for the existence of the other dimensions of private information. Secondly, they demonstrated that the inclusion of variables that may proxy for private information that is not normally observable to the insurance firm, could provide useful insights into behaviour under insurance. Therefore, in this paper, survey data rather than claims data will be used for the reasons that were advanced by Finkelstein and McGarry (2006).

4. Data

4.1 Survey Data Set

Over a six-week period commencing in October 1999, EKAS Marketing Research Services conducted market research on behalf of IMRAS Consulting to analyse community attitudes to the Australian smash repair market. The resulting data, henceforth referred to as the IMRAS data set, used 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.

Although these data were not collected to analyse insurance, many of the variables that are necessary to analyse asymmetric information, moral hazard and adverse selection are available. A two-year recall period was selected to ensure that sufficient data were collated on RTCs and smash repair experiences. These data were commercially available, and purchased for this study. Critically, the IMRAS data set contains data on the variables that are of principal interest. Firstly, the survey identified the incidence of RTCs that occurred during the preceding 2 years. 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.

Secondly, the IMRAS survey collected data on the insurance status of the respondent's automobile, which was identified as (i) compulsory third-party, (ii) third-party property, (iii) third-party property plus fire and theft or (iv) comprehensive insurance. Only comprehensive insurance indemnifies the owner for the cost of smash repairs in a crash for which he/she is at fault. Importantly, if the respondent had an RTC, the data set can be used to identify if the respondent was insured by (i) the same firm, (ii) a different firm or (iii) was uninsured, at the time of the RTC. Thus, drivers who change their insurance following an RTC are identifiable.

Thirdly, to conduct a reliable test for asymmetric information using conditional correlation, it is necessary to define 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 (as covariates) the data commonly collected by predominantly French insurers on their policyholders. Secondly, data collected by Australian insurance industry was reviewed. The five most frequent insurance carriers for survey respondents were the NRMA Ltd., AAMI, GIO, RACV and Suncorp. These firms, which provided cover 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 considerable congruence 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. Table 1 presents the descriptive statistics from the IMRAS data set and indicates the richness of the data set with respect to the characteristics of the respondents, their vehicles and insurance policies, as well as the standard categorical variables on RTCs and insurance.

Demographic characteristics in the data set include driver age, gender, age of co-driver (=1 $<$ 25 years) and vehicle ownership (=1 if private). Measures of location include dichotomous variable metropolitan (=1 if lives in city) and postcode. A measure of socioeconomic status (SES) was obtained for each postcode from the Socioeconomic Indices for Areas (SEIFAs) developed by the Australian Bureau of Statistics (ABS) (ABS 2006). The SEIFA index is used by the ABS to rank regions according to their levels of social and economic well-being. The ABS reports SEIFA indices by collection district (CD), which are the geographical regions used to gather census data. Each postcode is comprised of a number of collection districts. A weighted SEIFA index of Advantage-Disadvantage ($P.C._{AD\_index}$) was constructed for each postcode (PC) as follows.

$$PC_{AD-index} = \frac{\sum \left(\frac{SEIFA_{Pop.} * SEIFA_{AD-index}}{Pop.} \right)}{CD/PC} \tag{5}$$

The term in parentheses is a weighted SEIFA Advantage-Disadvantage index for each postcode that controls for the estimated resident population. A categorical variable, which measures the latent SES, was constructed using quartiles of the constructed index.

Table 2 reports descriptive data.

Table 2: Descriptive statistics (n=4005)

Variables	Obs.	Freq.	Freq.(%)	RTC (%)	Insured (%)
Comprehensively insured	4005	3163	79.0	25.0	n.a.
RTC history					
RTC 1997-99 †	4005	994	24.8	n.a.	79.5
Driver Characteristics					
Aged 18 to 24 years †	3971	356	9.0	32.9	52.8
Aged 25 to 34 years †	3971	826	20.8	25.5	76.8
Aged 35 to 44 years †	3971	1031	26.0	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	49.0	23.7	76.0
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	22.0	76.9
SES poor †	4005	1092	27.3	24.4	75.6
SES rich †	4005	1110	27.7	27.0	81.0
SES richest †	4005	997	24.9	25.2	82.0
Licensed 0 to 5 years †	3945	338	8.6	33.4	51.5
Licensed 6 to 10 years †	3945	440	11.2	27.0	71.1
Licensed 11 to 15 years †	3945	434	11.0	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

Variables	Obs.	Freq.	Freq.(%)	RTC (%)	Insured (%)
Comprehensively insured	4005	3163	79.0	25.0	n.a.
Income < $20,000 p.a.	3250	332	10.2	20.2	70.8
Income $20,000 to $39,999p.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	35.0	86.9
Income > $150,000 p.a.	3250	117	3.6	25.6	88.0
Income Refused to divulge	3250	724	22.3	18.9	77.9
Profession lower White	4005	1166	29.1	30.0	85.6
Profession upper Blue	4005	761	19.0	26.0	81.2
Profession lower Blue	4005	612	15.3	22.2	72.9
Profession home duties	4005	172	4.3	19.2	66.3
Profession student	4005	403	10.1	18.6	76.7
Profession retired	4005	164	4.1	35.4	51.8
Profession unemployed	4005	571	14.3	18.9	85.3
Refused to divulge profession	4005	71	1.8	22.5	64.8
Occupation refused to divulge	4005	85	2.1	23.5	70.6
4-cylinder vehicle †	4005	2570	64.2	27.2	79.5
6-cylinder vehicle †	4005	1273	31.8	21.0	78.4
8-cylinder vehicle †	4005	162	4.0	17.9	75.3
Make Ford †	4005	795	19.9	21.1	77.5
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	21.0	80.8
Body-type Sedan †	4005	3385	84.5	25.5	78.8
Body-type Commercial †	4005	250	6.2	21.6	68.0
Body-type 4 WD †	4005	295	7.4	20.0	87.5
Body-type Sports car †	4005	81	2.0	24.7	87.7

Variables	Obs.	Freq.	Freq.(%)	RTC (%)	Insured (%)
Comprehensively insured	4005	3163	79.0	25.0	n.a.
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	80.0
Car age > 12 years †	4005	1069	26.7	22.0	51.3
Value < $2000 †	3505	423	12.1	24.1	39.5
Value $2001 to $5999 †	3505	795	22.7	25.0	63.9
Value $6001 to $10000 †	3505	700	20.0	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

Note:

1. Variables marked with a † are commonly observed to be collected by insurers in Australia.

2. All sets of dummy variables are mutually exclusive, except for Body-type.

Additional variables that are not usually collected by Australian insurance firms, such as income and occupation type, are also available in the IMRAS data set. The literature has emphasized the importance of including a measure of claims history. A dichotomous variable $RTC_{1994-97}$ (=1 if RTC occurred 1994-97) was created and used in preference to the no-claim bonus variable because it is applicable both to insured and uninsured drivers. The use of this variable also obviates any concerns about differences in the insurance rules that insurers may apply to awarding no-claim bonuses and so on.

4.2 Australian Market Data

Unlike the United States of America where voluminous amounts of data are published at the level of individual insurers, and the UK where "freedom with publicity" was for many decades the reason for relatively limited regulatory involvement, typically very little data has been published in Australia at class of business level except in aggregate (such as that from APRA, the insurance council and ISA). (Laganiere *et al.* 2008, p. 3)

Australia is a wealthy nation. In 2003-04, the mean and median net household wealth was reported to be $494,346 and $311,550, respectively (ABS 2007b). With a population of 22 million people, Australia claims sovereignty to a continent with an area of 7.7 million square kilometres. The transport system is comprised of 810,000 kilometres of roads (Austroads 2005). In 1999, Australia's fleet contained 12.3 million vehicles. The average age of the fleet was 10 years and it was comprised of 9.7 million passenger vehicles (Productivity Commission 2005).

Driving an automobile entails a risk. Currently 1500 motorists die annually. While RTC fatalities have been declining since the 1980s (see Figure 2) the financial costs of RTCs remain significant.

Figure 2 Australian road fatalities
Source: Australian Bureau of Statistics (2007a, p. 538)

In 1996, the BTE (2000) estimated the national costs of RTC to be $15 billion. The total cost of vehicle repair was estimated to be $3.89 billion (27.4 percent) and average cost per repaired vehicle was $3,100. If towed the mean cost was $7,069 and if not towed the mean cost was $2,070 (Bureau of Transport Economics 2000). In 1996 it was estimated that one in every 7.8 registered vehicles (12.8 percent) was involved in an RTC annually (Bureau of Transport Economics 2000). In 1998, the Community Attitude to Road Safety (CARS) surveyed 1359 people (response rate 69 percent). They report that 18 percent of the population aged 15 years and over had been involved in an RTC during the three years from 1995 to 1998 (Mitchell-Taverner 1998). This equates to an annualized probability of 6 percent, which given a stated average of 1.95 persons per RTC this implies that each registered vehicle has an 11.7 percent probability of being involved in an RTC. Australian motorist are risk averse. The Insurance Council of Australia has stated that 87 percent to 90 percent of all vehicles are comprehensively insured

In Australia, approximately 40 domestic automobile insurance providers underwrite 10 million policies annually. In 2002, the five largest direct insurers accounted for 78 percent of earned premium (Productivity Commission (2005). A Herfindahl-Hirschman Index of 3,433 suggests a highly concentrated market. However, data supplied by the Productivity Commission implies a more competitive market structure. Figure 3 reports that in 2001-02 aggregate premium revenue exceeds claims expenses, for domestic and commercial vehicle insurance, by a 6.5 percent. Once the impact of administrative overheads are considered it appears that insurance firms in Australia earned near 'zero' economic profits in 2001-02.

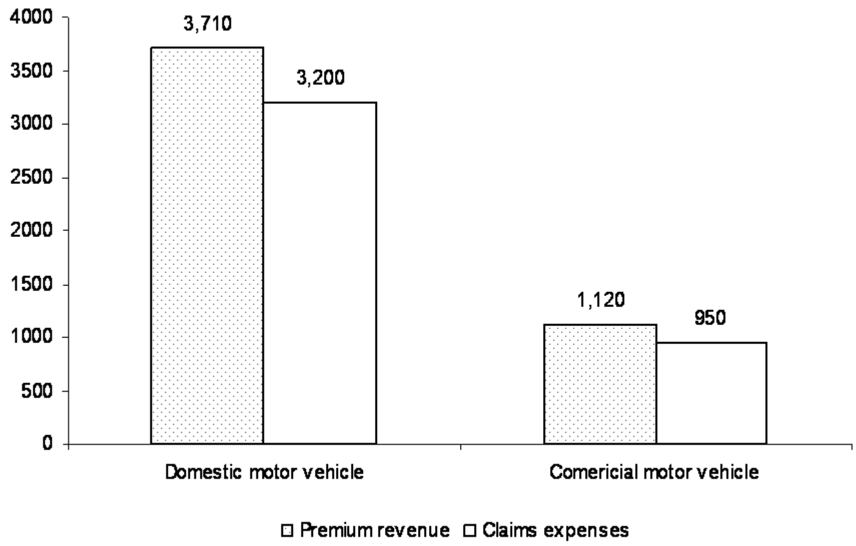

Figure 3: Premium revenue and claims expense, 2001-02 Domestic and commercial vehicle insurance, $ million
Source: The Productivity Commission (2005, p. 14).

5. Econometric Approach

> Reports that say that something hasn't happened are always interesting to me, because as we know, there are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns – the ones we don't know we don't know. And if one looks throughout the history of

our country and other free countries, it is the latter category that tend to be the difficult ones.

(Rumsfeld 2002)

A structural model of asymmetric information in a market for automobile insurance that simultaneously explains the choice of insurance (*INS**) and the probability of a crash (*RTC**), conditional upon a set of observable variables, is constructed as follows:

$$RTC_i^* = \alpha_0 + \alpha_1 INS_i + \alpha_2 \mathbf{X}_{1i} + \varepsilon_i \quad (6)$$

$$INS_i^* = \beta_0 + \beta_1 RTC_i + \beta_2 \mathbf{X}_{2i} + \mu_i \quad (7)$$

where RTC_i and INS_i are observed dummy variables, equal to one if the individual i has an *RTC* or purchases comprehensive insurance, respectively and $\mathbf{X}$ is a vector of variables which reflects the insurre's information set. Two methods to estimate the structural model above: (i) transform it to a recursive simultaneous system using a pre-determined variable, via *RTC* in the previous period, as a proxy for risk type, and (ii) use the insurance status of the second car in two-car households as an instrumental variable. Theoretically, the extent to which (i) or (ii) constitutes a better specification of the model depends on the extent to which one believes that the insurer's information set is likely to be exhaustive with respect to the classification of risk types.

5.1 Recursive simultaneous system

Past RTCs are correlated with the incidence of RTCs in the current period. The IMRAS data set identifies previous RTCs, i.e. crashes that occurred from 1994 to 1997 ($RTC_{1994-97}$). It can be argued that the variable $RTC_{1994-97}$ has two important properties that are necessary for its use in a recursive model. First, it proxies unobserved risk type, and hence is able to capture the adverse selection effect. Second, it is predetermined in the sense that it proxies the lag of RTC, and hence is exogenous by definition.

Under this specification, the above simultaneous system is modified to a recursive system as follows:

$$RTC_{it}^* = \alpha_0 + \alpha_1 INS_{it} + \alpha_2 \mathbf{X}_{1it} + \varepsilon_{it} \quad (8)$$

$$INS_{it}^{*} = \beta_0 + \beta_1 RTC_{i,1994-97} + \beta_2 \mathbf{X}_{2it} + \mu_{it} \quad (9)$$

In this system, the null hypothesis of no moral hazard is given by $H_0 : \alpha_1 = 0$ in equation (8). The coefficient β_1 in equation (9) captures the adverse selection effect (i.e., high-risk drivers are more likely to buy comprehensive insurance).

Greene (1998; 2000) showed that recursive simultaneous system above can be estimated efficiently and consistently using bivariate probit approach, ignoring the simultaneity and endogeneity issues since the likelihood function remains the same when these issues are taken into account. Greene (2000) also demonstrated detailed steps to calculate the marginal effect calculations in the above system of equations include both direct effects (i.e., effects on the probability that RTC_{it}=1) and indirect effects (i.e., effects on the probability that INS_{it}=1, which is in turn, transmitted to the probability that RTC_{it}=1).

The extent to which the parameter β_1 in equation (9) captures adverse selection is determined by the degree to which $RTC_{i,1994-97}$ can capture unobserved risk type, RT_i^*. Conceptually, the variable $RTC_{i,1994-97}$ is comprised of RTCs of two types (i) those which were reported the insurance firms via claims and (ii) those that were unreported. If the variable $RTC_{i,1994-97}$ includes a comprehensive array of minor RCTs, which are otherwise unobserved by insurance firms, then unobserved risk type RT^* will, in part, be captured. Alternatively, if $RTC_{i,1994-97}$ is fully observable to the insurers then no new information identifying risk-type is provided and hence β_1 does not reflect adverse selection effect.

The proportion of RTCs occurring three to five years ago that are observable to insurers generally is unknown although, anecdotally, insurers commonly ask applicants for policies to report their claims over the past three years. However, in the current period, 65.2 percent of insured drivers who report a RTC also lodge a claim. If reporting behaviour were constant over time, this would imply that the variable $RTC_{1994-97}$ does provide some additional information on risk-type. Alternatively, one could argue that recall bias ensures that only major RTCs are recalled and no new information on unobserved risk type is provided. In this case, an alternative approach is to estimate moral hazard while controlling for adverse selection would use an instrumental variable approach as follows.

5.2 An instrumental variable (IV) model

In the structural model above equation (8) captures one type of asymmetric information *ex ante* moral hazard and equation (9), which is comprised of the same set of variables, captures the effect of second type asymmetric information, adverse selection. Equation (8), which is the model of interest, is a probit model where RTC_{it} is a function of INS_{it}, a variable, which is binary and endogenous. A model with an endogenous variable of this type can be estimated as a bivariate probit with full information maximum likelihood (FIML) (Wooldridge 2002). Note that the bivariate probit model specified above has no exclusions; this is consistent with other econometric models that have been specified in this literature (Chiappori & Salanie 2000; Cohen 2005; Dionne *et al.* 2004, 2006, 2007). To ensure that this model is just identified, though, one instrumental variable is required.

Recall that an IV should be correlated with the endogenous variable (insurance) but uncorrelated with the error term. To be a credible IV, the candidate variable must not be observable to the firm, otherwise one would expect the firms to use the observable information to rate the premium, although it is known that exceptions exist (Finkelstein & McGarry 2006). An analysis of claims data, supplied by an insurance firm, would preclude the identification of an effective IV. The household survey contains data that are typically unobserved by insurers. This presents opportunities that are exploited here.

The insurance status of supplementary vehicles was also collected in the IMRAS survey. The insurance status of additional vehicles garaged within the household is utilised as an IV: insurance status of the second vehicle garaged in a two-vehicle household is used to instrument for the insurance status of the principal vehicle. The justification for this choice of instrument is outlined as follows.

Firstly, driving ability may be familially correlated and therefore, so may within-household decisions to purchase comprehensive insurance. Chiappori and Salanié (2000) provided empirical evidence of a familial relationship in respect of driving abilities. Secondly, in a household where driving abilities are not correlated but use of the vehicles is shared, a correlation between choices of insurance is likely to develop. Typically, vehicle owners share access to their vehicle with their spouse, adult child or other household members. Therefore, the decision to purchase insurance is partly determined by the ability of the co-drivers. Shared driving experiences are likely to ensure that asymmetric information with regard to driving abilities, within the household, is minimal.

The analysis was thus restricted to those 1,776 households in the IMRAS data set that had two vehicles. By restricting the sample to households with two vehicles, the insurance status of the 'other' vehicles could be expressed as indicator variable equal to one if comprehensively insured and equal to zero if otherwise for a single second vehicle. For the insurance status of the second automobile to function as an effective instrument for the insurance status of the first automobile, within in the household, the two variables should be correlated (Wooldridge 2000, p. 463). In the sample of 1,776 households with two vehicles, the pair-wise correlation, with the p-value in parentheses, for the two binary variables insurance status of the first automobile and insurance status of the second automobile was 0.285 (< 0.01). Thus, the IV candidate is connected with the target variable, which economic theory assumes is endogenously determined. The second condition that an IV should satisfy is that it should be uncorrelated with the error term (Wooldridge 2000, p. 463). This second condition cannot be proved empirically; however, the following intuitive argument is offered. Any correlation that exists between the insurance decision on the second automobile and the incidence of RTC in the first automobile may reflect adverse selection, but cannot reflect moral hazard, since no amount of insurance purchased for the second automobile will induce the driver of the first automobile to exercise less care when driving it. Thus, this IV is uncorrelated with the error term, by assumption. As such, this IV passes both conditions that are required for the defensible application of the IV approach.

Therefore, the following bivariate probit model will be estimated with the insurance status of the second garaged automobile within the household instrumented for the insurance status of the first automobile, within a sample of two-car households, as follows:

$$RTC_i = \alpha_0 + \alpha_1 INS_{i,1st\_car} + \alpha_2 \mathbf{X}_i + \varepsilon_i \tag{10}$$

$$INS_{i,1st\_car} = \beta_0 + \beta_1 INS_{i,2nd\_car} + \beta_2 \mathbf{X}_i + \eta_i \tag{11}$$

The null hypothesis of no moral hazard is given by $H_0 : \alpha_1 = 0$ in equation (10). A test for residual asymmetric information is given by $H_0 : \rho = 0$

5.3 Insurer's Information Set

To test for asymmetric information using conditional correlation, a vector of variables that reflect the insurer's information set must be included. Controls for driver characteristics included, age, gender, young co-driver, ownership, location, SES and years of licensure, while controls for vehicle characteristics included, vehicle value, age, make, body-type and engine size. Chiappori and Salanié (2000) have argued 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. See, for example Puelz &

Snow (1994). To circumvent this problem, 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.* (2004, 2006, 2007). The variables marked with a † in Table 2 represent the insurer's minimal information set.

While, this set of covariates provides a good approximation of the insurer's information set, it possible that Australian insurers collect and use data that is unavailable within the IMRAS data set to risk-rate their policyholders. For example, insurers are observed to collect data identifying whether or not the vehicle was garaged and the billing period i.e., yearly or six-monthly. These data may be used by insurers to identify risk types more accurately. 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 two identified 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, were used to risk-rate policyholders.

An examination of the IMRAS data set identified two potential sources of bias. The first concern was that recall bias might confound the results. The participants were asked if any of the following indemnifiable incidents (car stolen, car broken into, car burnt, car-part stolen, or RTC) occurred from October 1997 to November 1999. A dichotomous variable $RTC_{1997-99}$ was constructed if the respondent indicated that an RTC had occurred. The theory of recall bias posits that RTCs that occurred in 1997 might be less frequently recalled than RTCs that occurred in 1999. However, the survey also asked (i) which incident-type occurred most recently and (ii) in which year did this incident-type occur. This information identifies four mutually exclusive RTC sub-types, which are used to control for potential recall bias. The first three sub-types are comprised of those RTCs that occurred most recently in 1999, 1998, and 1997. A fourth, RTC sub-type was comprised of those residual RTCs, which were preceded by another incident-type (e.g. car stolen). Figure 4 provides a breakdown of RTC by time. The shaded area denotes the period of the survey. Thus, a set of dichotomous variables identifying the four RTC sub-types three of which identify how recently the RTC occurred where included to explanatory variables to control for any possible recall bias.

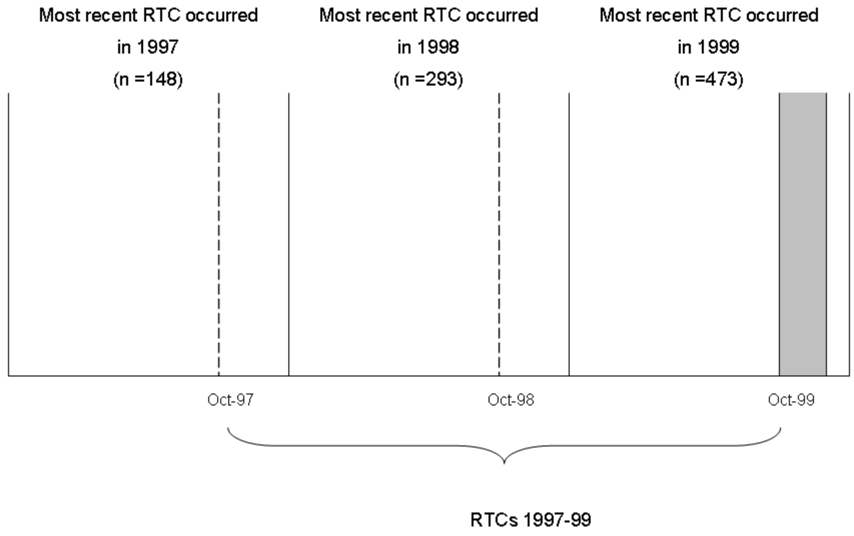

Figure 4: Reported RTCs from October 1997 to November 1999

A second concern was that the occurrence of other indemnifiable incidents (e.g. theft of car) might confound the analysis in unpredictable ways. For example, automobile theft may either increase the demand for insurance if the individual makes an upwards revision to his or her risk status or reduce the supply of insurance (e.g., if insurers drop claimants at the end of contract). Finkelstein and McGarry (2006) have argued that to allow for possible nonlinearities among variables, controls including interaction terms should be included. In this spirit, dichotomous variables identifying if the automobile was stolen, broken into, burnt, or car-part stolen with interaction terms identifying if an RTC occurred concurrently, were also included.

Finally, note that while the recursive model utilizes the full data set (n=4005), the bivariate probit model with a single IV is restricted to a sample of two-car households (n=1776), to potentiate the effectiveness of the IV. To control for the possibility that two-car households might be systematically different from other households a dichotomous variable equal to one if a two-car household was included in the recursive model. This enabled results from both models to be compared.

6. Results

To begin, a test for asymmetric information in the Australian market of automobile insurance using a bivariate probit model of the form specified by Chiappori and Salanié (2000) is estimated.

$$RTC_i = \alpha_0 + \alpha_1 \mathbf{X}_i + \varepsilon_i \quad (12)$$

$$INS_i = \beta_0 + \beta_1 \mathbf{X}_i + \nu_i \quad (13)$$

RTC and INS are dichotomous variables and is a vector of variables, denoted † in Table 2, which reflect the insurer's information set. The null hypothesis of no asymmetric information is given by $H_0 : \rho = 0$. The bivariate probit model reports ρ= 0.02 with a *p*-value = 0.74. Thus, one cannot reject the null hypothesis. This result concurs with that of Chiappori and Salanié (2000), and Dionne *et al.* (2004, 2006, 2007, 2010) who also report no evidence of asymmetric information.

Table 3 below, reports the results for two tests of *ex ante* moral hazard. The first four columns contain the results derived from the recursive model and the next four columns contain the results derived from the bivariate probit model with one IV. Arbitrarily the reference individual was aged 18 to 24 years, with less than 5 years licensure and lived in a postcode with the lowest SES. He/she drove a 4-cylinder, Ford sedan that was less than 3 years old and valued at less than $2000.

Table 3: Recursive model and bivariate probit model with one IV Preliminary Results

	Recursive (n=4005)				Biprobit with IV (n=1776)			
	RTC 1997-99		First car insured		RTC 1997-99		First car insured	
Variables	Coeff.	*p*- value	Coeff.	*p*- value	Coeff.	*p*- value	Coeff.	*p*- value
First car insured	1.46	<0.01	n.a.	n.a.	1.14	0.02	n.a.	n.a.
Second car insured	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	0.65	<0.01
RTC History								
RTC 1994-97 †	n.a.	n.a.	0.15	0.08	0.47	<0.01	-0.13	0.39
Constant	-1.35	<0.01	-0.30	0.29	-0.74	0.13	-1.29	0.01

Note: The coefficients for the covariates are not reported

The focus is the relationship between insurance and a RTC. The coefficients for the variable First car insured are 1.455 (*p*-value $<$0.01) and 1.141 (*p*-value $<$0.01) in the recursive and bivariate probit models, respectively. These results suggest that conditional upon the insurer's information set, the purchase of insurance is correlated with an increased probability of an RTC. Thus, the null hypothesis of nil *ex ante* moral hazard in the market for comprehensive vehicle insurance can be rejected in both models.

6.1 Tests for a Weak Instrument

Two tests for weak instruments are conducted. First, the change in the Pseudo-R^2 between equations (14) and (15), which estimate the endogenous variable $INS_{i,1st\_car}$, regressed on vector of exogenous variables $\mathbf{X}$, with and without the selected IV, $INS_{i,2nd\_car}$ and a likelihood ratio (LR) test on the null hypothesis of no correlation between the IV and the endogenous regressor is reported.

$$INS_{i,1st\_car} = \beta_0 + \beta_1\mathbf{X} \tag{14}$$

$$INS_{i,1st\_car} = \beta_0 + \beta_1\mathbf{X} + \beta_2 INS_{i,2nd\_car} \tag{15}$$

The statistical analysis showed that the Pseudo-R^2 increased from from 0.3367 to 0.3688 and the LR test, $\chi 2(1)$ = 39.36, which was statistically significant at the one percent level suggests that the selected IV is not weak.

Secondly, following Cameron and Trivedi (2009), the system of equations (10) and (11) is re estimated using two-stage least squares (2SLS) with robust standard errors. [2]In the just-identified system with one endogenous regressor, a test the significance of the instrument $INS_{i,2nd\_car}$ in the first-stage regression (equation (11)) is conducted. The null hypothesis that the instrument is weak is given by $H_0 : \beta_1 = 0$, and we reject the null hypothesis if the *F* statistic is greater than 10 (Staiger and Stock 1997). Thus, the computed an *F* statistic of 36.5 enables the null hypotheses of a weak instrument to be rejected.

6.2 A Falsification Test

It is assumed that while at-fault RTCs were a function of driver effort, not-at-fault RTCs occurred randomly and were not a function of driver effort. Shavell (1979) has argued that *ex ante* moral hazard occurs when insurance causes a reduction in unobserved preventive

[2]Angrist (2006) has argued that 2SLS can be used to estimate binary probability models with dummy endogenous variables because linear 2SLS estimates have a robust causal interpretation that is insensitive to possible nonlinearity introduced by the dummy dependent variables.

effort. A corollary to this proposition is that *ex ant*e moral hazard should be more evident in at-fault, *vis-à-vis* not-at-fault RTCs.

Thus, the following sensitivity analysis was constructed. In the event of an RTC, the IMRAS survey asked participants "Were you able to prove that someone else was at-fault?". The sample was comprised of 994 RTCs; of which the other driver was not-at-fault on 471 occasions and the other driver was at-fault on 432 occasions. The sample was thereby stratified: in the first sub-sample, a binary variable was constructed to be equal to one if the participant indicated the other driver was not-at-fault and equal to zero if no RTC occurred. The not-at-fault RTCs were then removed from the sample. Similarly, in the second sub-sample, a binary variable was constructed equal to one if the participant indicated that the other driver was at-fault and zero if no RTC occurred. The at-fault RTCs were then removed from the sample. See Figure 5.

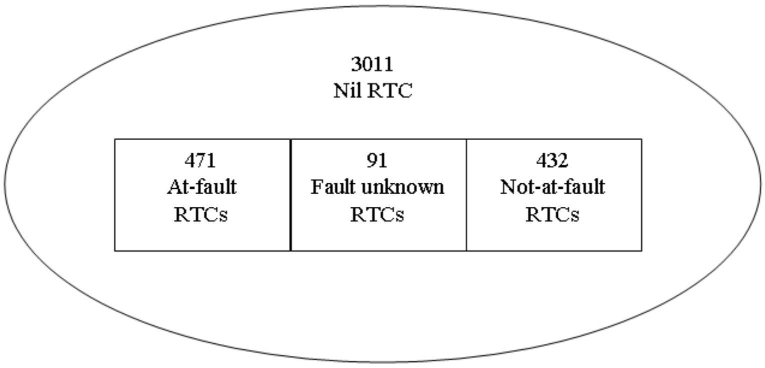

Figure 5: RTCs by fault, Full sample

Figure 6 repeats the stratification of RTCs for the sub-sample of two-vehicle households.

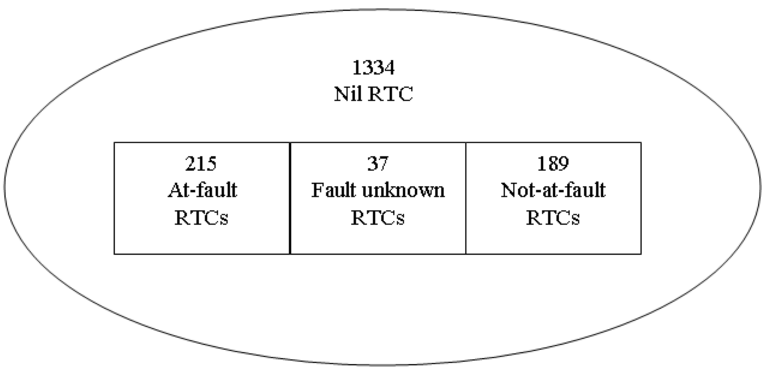

Figure 6: RTCs by fault, Two-vehicle household

Each sub-sample was analysed using the bivariate probit mode. The coefficients with *p*-values are reported for the explanatory variable of interest *First car insured* in Table 4.

Table 4: Bivariate probit model with one IV in samples of at-fault and not-at-fault RTCs Premliminary Results

	Recursive model				Biprobit model with IV			
	At-fault RTC (n=2396)		Not-at-fault RTC (n=2372)		At-fault RTC (n=1085)		Not-at-fault RTC (n=1067)	
	Coeff.	*p*-value	Coeff.	*p*-value	Coeff.	*p*-value	Coeff.	*p*-value
First car insured †	1.40	<0.01	0.86	0.26	1.64	<0.01	0.63	0.50
Constant	-1.65	<0.01	-1.56	<0.01	-0.90	0.09	-1.48	0.06

Note: The coefficients for the covariates are not reported

The results in Table 4 are consistent with economic theory. While evidence of *ex ante* moral hazard persists in the sample of at-fault RTCs 1.642 (p-value <0.01) evidence of *ex ante* moral hazard in the sample of not-at-fault RTCs 0.627 (p-value =0.5) is no longer present, thus providing strong corroboration of the evidence of *ex ante* moral hazard that was presented in Table 3.

7. Discussion

Moral hazard has been of theoretical interest to economists for five decades. In markets for vehicle insurance, testing for moral hazard has been vexed because the decision to insure has been assumed determined endogenously. The literature has, to date, identified several approaches to address this concern. Chiappori & Salanie (2000) restricted their analysis to a sample of beginner drivers with no history of a claim and use a natural experiment to test for moral hazard. Other analysts have used longitudinal data to test for moral hazard (Abbring *et al.* 2003; Israel 2004), or have used panel data to report a Granger causality test for moral hazard (Dionne *et al.* 2004, 2006, 2007). The current state of the literature suggests a preference for empirical analyses to be conducted with panel data (Chiappori & Salanie 2000; Dionne *et al.* 2007).

In this paper, two conventional econometric strategies were used to control for endogeneity in cross-sectional survey data. The first approach used a recursive model whereby the endogenous variable insurance is estimated using the predetermined variable $RTC_{i1994-97}$, a proxy for risk type and a vector of observable covariates. This approach assumes that $RTC_{i1994-97}$ is predetermined and therefore uncorrelated with the error term. However, RTC history is a variable that is widely used by insurers to risk rate their policyholders. Therefore, the extent to which the variable will be uncorrelated with the error term depends on the degree to which reported RTCs (including minor RTCs otherwise unreported to insurance firms) can fully capture risk type.

Cognisant of the possibility that our first approach may not fully account for unobserved risk type a second model which used the insurance status of the second car garaged at the household as an IV for the insurance status of the first car in a sub-sample of two car household was estimated. The intuitive 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 second car cannot elicit moral hazard while driving the first car.

The coefficients for insurance in the recursive model 1.428 (p-value < 0.01) and bivariate probit model 1.202 (p-value < 0.01) both imply that moral hazard is evident in the market for vehicle insurance. An advantage of using the recursive model is that the entire data set is analysed, however to address the inherit endogeneity within the structural model is assumed to be predetermined. On the other hand, the bivariate probit model, which uses an IV to address the endogeneity, analyses a restricted sample of two car households. Results derived using this approach could be confounded if the existence of moral hazard in insured drivers living in two-car household were systematically different from drivers who do not live in a two-car household. However, the inclusion of a binary variable two-car household in the recursive model was not found to be statistically significant with respect either to an RTC or to insurance, which allays this concern.

The ability to differentiate between at-fault from not-at-fault RTCs in this data set enabled these results to be subjected to further testing. The correlation between insurance and RTC implies *ex ante* moral hazard. If the correlation were spurious, it should presumably exist in both sub-samples. However, this was found not to be the case. While no evidence of *ex ante* moral hazard was found in the sub-sample of not-at-fault RTCs, evidence of *ex ante* moral hazard persists in the sub-sample of at-fault RTCs. These findings are consistent with our theoretical expectations for evidence of moral hazard in each sub-sample. This falsification test provides some reassurance of the veracity of the results.

The decision to follow the methodological lead of Finkelstein and McGarry (2006) and analyse survey data as opposed to claims data was one important part of this approach: both econometric models rely on the use of data that are unobserved by insurers and hence unavailable in claims data sets. The recursive model assumes that the RTC history is a measure of the unobserved risk-type, as it includes some RTCs not necessarily reported to insurer. The bivariate probit model utilises the insurance status of the second car, a variable that is not normally observable to the insurer, as an instrument. It is difficult therefore to see how either method could be applied to claims data since all pertinent and observable data would be subsumed into the vector of covariates representing the insurer's information set.

The cross-section-ness of insurance data was not the issue. Rather cross sectional analysis has been utilizing the wrong data. Claims data can be enticing. However, as sirens perched on the rocks, these data can result in frustration for the intrepid econometrician in search of moral hazard. At first glance, the econometrician will see a clean and accurate data set, with a large number of observations and rich arrays of variables. The appeal is understandable. While these data may be capable, of identifying asymmetric information (notwithstanding the findings of Finkelstein and McGarry (2006)) they do not contain the necessary information to discriminate between the various dimensions of private information.

Econometricians who wish to test for moral hazard should consider the advantages of survey data in general and household data in particular rather than embark on the more traditional analysis with claims data. While any variable, which is unobservable to the insurer, could serve as a viable instrument for any range of esoteric reasons, household surveys may offer a particularly fertile source of information. The basis for this claim is founded in the economics of asymmetric information. Pertinent information about the policyholder, which is ordinarily unobservable to the insurer, may nevertheless be observable to other household members. Household behaviour, which reflects a common propensity to insure but not to claim, is a potentially valuable source of information for econometricians seeking to test for moral hazard.

Bibliography

Abbring, Jaap H., Chiappori, Pierre Andre, & Pinquet, Jean. 2003. Moral Hazard and Dynamic Insurance Data. *Journal of the European Economic Association*, **1**(4), 767–820.

ABS. 2006. *Census of Population and Housing: Socio-Economic Indexes for Areas (SEIFA), Australia - Data Cube only, 2001*.

Angrist, J. 2006. Instrumental variables methods in experimental criminological research: what, why and how. *Journal of Experimental Criminology*, **2**(1), Angrist, J. (2006). ”Instrumental variables methods in experimental criminological research: what, why and how.” Journal of Experimental Criminology 2(1): 23–44.

Arrow, K. J. 1963. Uncertainty and the Welfare Economics of Medical Care. *The American Economic Review*, **LIII**(5), 941–973.

Australian Bureau of Statistics. 2007a. *2007 Year Book of Australia Number 89 ABS cat. no. 1301.0*. Tech. rept. Australian Bureau of Statistics (ABS), Canberra.

Australian Bureau of Statistics. 2007b. *Household Wealth and Wealth Distribution 2005-06 Australia ABS cat. no. 6554.0*. Tech. rept. Australian Bureau of Statistics (ABS), Canberra.

Austroads. 2005. *Road Facts 2005 An overview of the Australian and New Zealand road systems*. Tech. rept. Austroads Incorporated, Sydney.

Baker, Tom. 1996. On the genealogy of moral hazard. *Texas Law Review*, **75**(2), 237.

Becker, Gary S. 1981. *A treatise on the family*. Cambridge, Massachusetts: Harvard University Press.

Bernstein, David E. 1996. Procedural Tort Reform: Lessons from Other Nations. *Regulation*, **19**(1), 71–81.

Blunt, H.P. 1907. The Moral Hazard in Fire Insurance. *Journal of the Insurance Institute of Great Britain and Ireland*, **10**, 97–126.

Bureau of Transport Economics. 2000. *Road Crash Costs in Australia BTE Report 102*. Tech. rept. Bureau of Transport Economics (BTE), Canberra.

Butler, RJ, & Worrall, JD. 1983. Workers' compensation: Benefit and injury claims rates in the seventies. *The Review of Economics and Statistics*, **65**(4), 580–589.

Cameron, A., & Trivedi, P. 2009. *Microeconometrics using stata*. Stata Press.

Campbell, A.C. 1902. *Insurance and Crime*. New York: G.P. Putnam's Sons.

Ceccarelli, Giovanni. 2001. Risky Business: Theological and Canonical Thought on Insurance from the Thirteenth to the Seventeenth Century. *Journal of Medieval and Early Modern Studies*, **31**(3), 607–658.

Chelius, JR, & Kavanaugh, K. 1988. Workers' compensation and the level of occupational injuries. *Journal of Risk and Insurance*, 315–323.

Cheung, SNS. 1969. *The Theory of Share Tenancy: With Special Application to Asian Agriculture and the First Phase of Taiwan Land Reform*. The University of Chicago Press.

Chiappori, Pierre-Andre. 1999. Asymmetric Information in Automobile Insurance: An Overview. *Pages 1–11 of:* Dionne, Georges., & Laberge-Nadeau, Claire. (eds), *Automobile insurance: Road safety, new drivers, risks, insurance fraud and regulation*. Boston: Kluwer Academic Publishers.

Chiappori, Pierre-Andre, & Salanie, Bernard. 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, & Siegelman, Peter. 2010. Testing for Adverse Selection in Insurance Markets. *Journal of Risk and Insurance*, **77**(1), 39–84.

Connelly, L. B., & Supangan, R. 2006. The economic costs of road traffic crashes: Australia, states and territories. *Accident Analysis and Prevention*, **38**(6), 1087–1093.

Coyle, Diane. 2007. *The soulful science : what economists really do and why it matters*. Princeton, N.J.: Princeton University Press.

Crosby, E. U. 1905. Fire Prevention. *The Annals of the American Academy of Political and Social Science*, **26**(2), 224.

Daston, Lorraine. 1983. Mathematical Probability and the Reasonable Man of the Eighteenth Century. *Annals of the New York Academy of Sciences*, **412**(History and Philosophy of Science: Selected Papers), 57–72.

Daston, Lorraine. 1988. *Classical probability in the Enlightenment*. Princeton, N.J.: Princeton University Press.

Dembe, Allard. E., & Boden, Leslie. I. 2000. Moral Hazard: A Question of Morality? *New Solutions: A Journal of Environmental and Occupational Health Policy*, **10**(3), 257–279.

Diamond, PA. 1967. The role of a stock market in a general equilibrium model with technological uncertainty. *The American Economic Review*, **57**(4), 759–776.

Dionne, Georges, Gourieroux, Christian, & Vanasse, Charles. 2001. Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment. *Journal of Political Economy*, **109**(2), 444–453.

Dionne, Georges, Michaud, Pierre-Carl, & Dahchour, Maki. 2004. Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France. **Available at SSRN: http://ssrn.com/abstract=583063 or doi:10.2139/ssrn.583063 (Accessed June 2010)**.

Dionne, Georges, Michaud, Pierre-Carl, & Dahchour, Maki. 2006. Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France. **Available at http://www.aria.org/meetings/2006papers/Dionneseparating06-06-01.pdf (Accessed June 2010)**.

Dionne, Georges, Michaud, Pierre-Carl, & Dahchour, Maki. 2007. Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France. **Available at http://www-gremaq.univ-tlse1.fr/seminaires/michaud.pdf (Accessed June 2010)**.

Dionne, Georges, Michaud, Pierre-Carl, & Dahchour, Maki. 2010. Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France CIRRELT 2010-40. **Available at**

https://www.cirrelt.ca/DocumentsTravail/CIRRELT-2010-40.pdf (Accessed November 2010).

Ducat, Arthur. C. 1865. *The Practice of Fire Underwriting*. 4th edn.

Edwards, Chilperic. 1921. *The Hammurabi code : and the Sinaitic legislation*. 3rd edn. London: Watts.

Febvre, Lucien. 1956. Pour l'histoire d'un sentiment: le besoin de securite. *Annales: Economies, sociatas, civilisations*, **11**, 244–47.

Finkelstein, Amy, & McGarry, Kathleen. 2006. Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market. *The American Economic Review*, **96**(4), 938–958.

Greene, W. H. 1998. Gender economics courses in liberal arts colleges: Further results. *Journal of Economic Education*, 291–300.

Greene, W. H. 2000. *Econometric analysis*. 4th edn. Upper Saddle River, N.J.: Prentice Hall. William H. Greene. ill. ; 25 cm. Accompanying CD-ROM catalogued separately.

Hale, Benjamin. 2009. What's so moral about the moral hazard. *Public Affairs Quarterly*, **23**(1), 1–25.

Harper, Douglas. 2001. *Online Etymological Dictionary*.

Hart, D.G., Buchanan, R.A., & Howe, B.A. 2007. *Actuarial Practice of General Insurance*. 7th edn. Sydney: Institute of Actuaries of Australia.

Haynes, John. 1895. Risk as an Economic Factor. *The Quarterly Journal of Economics*, **9**(4), 409–449.

Heckman, James J., & Borjas, George J. 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.

Holmstrom, 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**.

Knight, Frank H. 1921. *Risk, uncertainty and profit.* Chicago: University of Chicago Press.

Laganiere, Benoit., Mackic, Nevena., & Minty, David. 2008. The role of industry statistics in Australian insurance portfolio management. *In: Institute of Actuaries of Australia 16th General Insurance Seminar.*

Lessius, Leonardus. 1605. *De Iustitia et Iure.*

Lewis, Charlton. Thomas., Short, Charles., Freund, Wilhelm., & Andrews, Ethan. Allen. 1879. *A Latin dictionary : founded on Andrews' edition of Freund's Latin Dictionary, reviewed and in great part rewritten by Charlton T. Lewis and Charles Short.* Oxford: Clarendon.

Marshall, Alfred. 1920. *Principles of economics : an introductory volume.* 8th edn. London: Macmillan.

Mas-Colell, Andreu, Whinston, Michael D., & Green, Jerry R. 1995. *Microeconomic theory.* New York: Oxford University Press.

Mathers, C., Vos, T., & Stevenson, C. 1999. *The burden of disease and injury in Australia, cat. no. PHE 17.* Tech. rept. AIHW.

McCloskey, Deirdre. 1994. *Knowledge and persuasion in economics.* Cambridge: Cambridge University Press.

McNeill, G.E. 1900. *A Study of Accidents and Accident Insurance.* Boston: Insurance Topics Co.

Mehr, R. I., & Cammack, E. 1976. *Principles of Insurance.* 6 edn. Illinois: Homewood.

Mitchell-Taverner, P. 1998. *Community Attitudes to Road Safety, Community Attitudes Survey Wave 11.* Tech. rept. Federal Office of Road Safety.

OED. 2008. *Oxford English Dictionary.* Oxford: Oxford University Press.

Pauly, Mark V. 1968. The Economics of Moral Hazard: Comment. *The American Economic Review,* **58**(3), 531–537.

Pearson, Robin. 2002. Moral Hazard and the Assessment of Insurance Risk in Eighteenth- and Early-Nineteenth-Century Britain. *The Business History Review,* **76**(1), 1–35.

Pindyck, Robert S., & Rubinfeld, Daniel L. 1989. *Microeconomics*. New York: Macmillan ; London : Collier Macmillan.

Productivity Commission. 2005. *Smash Repair and Insurance*. Tech. rept. 34. The Productivity Comission.

Puelz, Robert., & Snow, Arthur. 1994. Evidence on Adverse Selection: Equilibrium Signaling and Cross-Subsidization in the Insurance Market. *The Journal of Political Economy*, **102**(2), 236–257.

Remington, B.C., & Hurren, H. G. 1935. *Moral hazard abroad*.

Rubinow, I. M. 1913. *Social Insurance*. New York: Henry Holt & Co.

Rumsfeld, Donald H. 2002 (Feburary). *US Department of Defense News Breifing*. Accessed at (http://www.defense.gov/transcripts/transcript.aspx?transcriptid=2636).

Rupprecht, C.F. 1940. The Modern Fire Underwriter. *The Spectator*.

Shavell, Steven. 1979. On Moral Hazard and Insurance. *The Quarterly Journal of Economics*, **93**(4), 541–562.

Shepherd, P., & Webster, A.C. 1957. *Selection of Risks*. New York: Society of Actuaries.

Staiger, D., & J., Stock. 1997. Instrumental variables regression with weak instruments. *Econometrica*, **65**(3), 557–586.

Stone, Deborah. 2002. Beyond Moral hazard; Insurance as Moral Opportunity. *Pages 52–79 of:* Baker, Tom., & Simon, Jonathan. (eds), *Embracing Risk The Changing Culture of Insurance and Responsibility*. Chicago: The University of Chicago Press.

Tiffany, H.S. 1882. *Tiffany's instruction book for fire insurance agents 46*. 10 edn.

Varian, Hal R. 1992. *Microeconomic analysis*. 3rd edn. New York: Norton.

Vaughan, E. J. 1997. *Risk Management*. New York: Wiley.

Winter, Ralph. A. 1992. Moral Hazard and Insurance Contracts. *Pages 61–96 of:* Dionne, Georges. (ed), *Contributions to Insurance Economics*. Boston: Kluwer Academic Publishers.

Winter, Ralph. A. 2000. Optimal Insurance under Moral Hazard. *Pages 155–183 of:* Dionne, Georges. (ed), *Handbook of Insurance*. Boston: Kluwer.

Wooldridge, Jeffrey M. 2000. *Introductory econometrics : a modern approach*. South-Western College.

Wooldridge, Jeffrey M. 2002. *Econometric analysis of cross section and panel data*. Cambridge, Mass.: MIT Press.

Working, E.J. 1927. What Do Statistical 'Demand Curves' Show? *The Quarterly Journal of Economics*, **41**, 212–35.

Wright, Sewall. 1929. Fisher's Theory of Dominance. *The American Naturalist*, **63**(686), 274–279.