C.I.R.P.É.E

Centre Interuniversitaire sur le Risque, les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **12-04**

A Review of Recent Theoretical and Empirical Analyses of Asymmetric Information in Road Safety and Automobile Insurance

Georges Dionne
Pierre-Carl Michaud
Jean Pinquet

Janvier/January 2012

Dionne: Corresponding author. Canada Research Chair in Risk Management, HEC Montréal, 3000, Cote-Ste-Catherine, Room 4454, Montreal, QC, Canada H3T 2A7. Tel.: 514 340-6596; Fax: 514 340-5019; and CIRPÉE georges.dionne@hec.ca
Michaud: UQAM, Canada
Pinquet: Université de Paris X, France

Abstract:
Road safety policies and automobile insurance contracts often use incentive mechanisms based on traffic violations and accidents to promote safe driving. Can these mechanisms improve road safety efficiently? Do they reduce asymmetric information between drivers and insurers and regulators? In other words, is there residual asymmetric information in observed distributions of accidents and infractions? We answer these questions in this chapter by reviewing recent theoretical and empirical results based on various data and methodologies. We present recent tests related to the identification of residual asymmetric information in road safety management and in automobile insurance contracting. We also propose a theoretical analysis of the foundations of point-record driver's licenses observed around the world.

Keywords: Road safety, point-record driver's license, asymmetric information, moral hazard, road safety management, automobile insurance contract, empirical test

JEL Classification: C12, C14, C23, C33, D81, G22, L90, R41

Introduction

Road safety policies and automobile insurance contracts often use incentive mechanisms based on traffic violations and accidents to promote safe driving. These mechanisms are monetary (fines, insurance premiums) and non-monetary (point-record driver's licenses). They are justified by the presence of asymmetric information in insurance contracting and road safety regulation. Drivers have more information on their driving environment and safety behavior than insurers and regulators do. Short-term and long-term relationships between principal (insurer or regulator) and agent (driver) are part of the incentive schemes that can be used to reduce the potential private and social costs associated with asymmetric information. Can these mechanisms improve road safety efficiently? Do they reduce asymmetric information between drivers and insurers and regulators? In other words, is there residual asymmetric information in observed distributions of accidents and infractions?

We answer these questions in this chapter by reviewing recent theoretical and empirical results based on various data and methodologies. The purpose of the review is to present recent tests related to the identification of residual asymmetric information in road safety management and in automobile insurance contracting. We also propose a theoretical analysis of the foundations of point-record driver's licenses observed around the world.

The remainder of the chapter is organized as follows. Section 1 provides recent statistics on road safety around the world. We then review the literature on the economics of incentives in road safety management. Section 3 presents the theoretical models of point-record driver's licenses corresponding to those observed in North America and in Europe. Section 4 proposes a review of recent models for testing the presence of residual asymmetric information in insurance markets. Section 5 reviews two tests for the presence of residual moral hazard in automobile insurance data in detail and Section 6 concludes the paper.

1. Road safety around the world

Since the 1970s, fatality rates due to road-traffic accidents have decreased steadily in developed countries, despite increased risk exposure (International Transport Forum, OECD, 2010). These countries experienced a steep decline in road accidents during the first decade of this century. This situation stands in sharp contrast to that in less developed regions, especially in low and middle income countries, where 90% of global road deaths occur. For example, the road fatality rate decreased by 48% in France during the period 2000-2009, whereas it increased by 12% in Malaysia during the same period. The corresponding variations are -19% in the United States and +12% in Argentina. As of 2009, the average number of fatalities per 100,000 persons was 18 in Argentina and 6.3 in Canada (11.1 in United States and 3.8 in United Kingdom). These numbers are consistently lower for OECD countries and higher for emerging countries for which we have official data. Many factors can explain these differences. Risk exposure is an important one, but even when we control for this factor, differences between developed and less developed regions remain. It seems that risk exposure in less developed regions has increased significantly in recent years. If we concentrate on road fatalities per billion kilometers driven in 2008, the risk of dying in a road accident is lowest in Sweden (5.1) and UK (5.2) while it is 8 in United States, 17.7 in Malaysia, and 20 in Korea.

The discrepancy between economic regions with respect to the social costs of road-traffic accidents can be explained by many factors other than risk exposure. For instance, the "Haddon matrix" (Haddon, 1968) provides a multi-factorial approach to road safety in which human, vehicle, and environment factors are crossed with three phases (before, during and after the accident). Again, all the factors clearly favor the most developed countries (WHO, 2004). The implied social cost of road accidents is very heavy, even in developed countries (Doyle, 2005). By 2020, road-traffic accidents should become the third cause of the disability-adjusted life years lost from disease or injury worldwide (Murray and Lopez, 1997). They ranked ninth in 1990.

2. The economics of incentives in road safety management

One major reason for the improvement of the situation in the OECD has been the development of incentives for safe driving. Insurers and regulators have introduced several contract mechanisms to reduce asymmetric information and improve road safety. Experience rating schemes used by the insurance industry have incentive properties (Boyer and Dionne, 1989; Dionne and Vanasse, 1989, 1992; Abbring et al, 2003). They are supplemented by point-record driver's licenses based on traffic violations. In many countries, each convicted traffic offense is filed with a specific number of demerit points. When the accumulated number of points exceeds a given threshold, the driver's license is suspended. Redemption clauses were added so that this penalty can be avoided in the long run. Bourgeon and Picard (2007) investigate the most important properties of point-record licenses in terms of road safety incentives and discuss how they can be combined with fines to design an optimal system that internalizes the social cost of road accidents. They do not take into account insurance pricing based on traffic violations and accidents. Dionne, Pinquet et al (2011a) extend their model by considering insurance pricing based on demerit points and compare the relative incentive efficiency of insurance pricing, fines and the point-record driver's license.

The North American continent preceded Europe in the design of such systems. Point record driver's licenses were introduced in 1947 in the USA. By comparison, they were introduced in Germany, Québec, France, and Spain in 1974, 1978, 1992 and 2005, respectively. Incentive mechanisms for road safety have been investigated in the economic literature for many years (Peltzman, 1975; Landes, 1982; Graham and Garber, 1984; Boyer and Dionne, 1987; Blomquist, 1988; Cummins and Weiss 1992, Cummins et al, 2001;, Devlin, 1992; Dionne and Laberge-Nadeau, 1999). Of the many mechanisms proposed, we consider fines, point-record driver's licenses, partial insurance and insurance experience rating. In the latter case, the individual driving history is usually summarized by past accidents or by point-records based on traffic offenses. We will not study the direct incentives effects of the fault system in detail.[1]

[1] See Shavell (2004) for an overview of this incentive mechanism, Fagart and Fluet (2009) for a recent theoretical contribution, and Cohen and Dehejia (2004) for a recent empirical analysis

In the presence of asymmetric information, insurers use partial insurance coverage or experience rating to improve resource allocation. Both schemes have proven to be efficient for handling moral hazard and adverse selection (see Holmstrom, 1979; Shavell, 1979; Pauly, 1974; Rothschild and Stiglitz, 1976, for partial insurance, and Chiappori et al, 1994; Dionne and Lasserre, 1985, for experience rating). A number of empirical tests have been proposed to measure the efficiency of such mechanisms for road safety (Sloan et al, 1995; Boyer and Dionne, 1989) or to measure the presence of residual asymmetric information problems in insurers' portfolios (Chiappori and Salanié, 2000; Dionne, Gouriéroux and Vanasse, 2001; Cohen and Siegelman, 2010; Dionne, Pinquet, Maurice and Vanasse, 2011a; Dionne, Michaud and Dahchour, 2011b). Abbring, Chiappori and Pinquet (2003) designed a new test based on the dynamics of insurance contracts to detect the presence of residual moral hazard. Their model makes it possible to separate the moral hazard effect on accidents from unobserved heterogeneity. They found no evidence of moral hazard in the French automobile insurance market. Their approach consists in analyzing the behavior of policyholders in a given insurance setting and contrasting this behavior with that predicted by theoretical models under moral hazard and adverse selection. Both information problems can be identified from a detailed analysis of the data. Accidents change the schedule of future premiums and incentives for road safety. Individuals who accumulate accidents are charged higher premiums and, under moral hazard, should improve their driving behavior and reduce their risk to regain their original premium. This result is obtained only when the pricing scheme is convex. A negative correlation should thus be observed between past accidents and future accidents under moral hazard, when optimal incentive schemes are present, or when the insurance pricing affects driving behavior. The empirical test is not so simple, however, because it involves the distinction between pure heterogeneity and state dependence. Dionne, Pinquet et al (2011a) have extended their methodology and applied their model to the point-record mechanism in Québec. They did not reject the presence of moral hazard in the data. We shall revisit this study below.

Insurance pricing may not be a sufficient tool for designing an optimal road safety policy because it may not create the appropriate incentives for reckless drivers (Sloan et al, 1995). Bourgeon and Picard (2007) show how point-record driver's license suspensions provide incentives for road safety among normal drivers (those who are affected by usual incentive schemes) in the presence of failures in the judicial system or in the insurance market to provide optimal incentives. Point-record driver's licenses also allow the government to incapacitate reckless drivers. Similar to insurance pricing, fines for traffic violations may be ineffective for reckless drivers when their amounts are bounded above, either because some drivers would not be able to pay them or for some equity reasons (see also Shavell, 1987). However, both fines and insurance pricing reinforce the efficiency of the point-record mechanism by providing normal drivers with more incentives. Dionne, Pinquet et al (2011a) compared the relative efficiency of fines, insurance pricing on demerit points and the point-record driver's license. They find that in Québec, during the 1985-1997 period, fines were on average the most efficient mechanism. However, their analysis was limited to a uniform incentive scheme based on a representative agent because they could not procure individual wealth information to introduce heterogeneity in the studied parameters.

Regulation of insurance premiums to make insurance rates affordable to drivers can affect road safety and insurance claim frequency by introducing distortions in the automobile insurance industry. One such regulation in the United States subsidizes insurance premiums of high-risk drivers and finances these premium reductions through surcharges to low-risk drivers. This type of regulation may increase market participation for high risk drivers and reduce that of low-risk drivers while increasing the rates of accidents and claims. Weiss, Tennyson and Regan (2010) obtain a positive correlation between rate regulation and insurance claim frequency, once they control for the endogenous choice of regulation. The authors also provide a detailed overview of rate regulation in the United States. In 2010, there was no bonus-malus system based on accident records in the United States although six states offer safer driver plans as part of their rate regulation.[2]

It is important to point out that with few exceptions (Chiappori et al, 1994), credit markets are not considered in this literature. In costless financial markets, savings can reduce the incentive effects of long-term contracting by granting wealthy bad risks facing

[2] On the French bonus-malus scheme, see Picard, 2000; Pinquet, 1999; Dionne, 2001. On the properties of the Tunisian bonus malus system see Dionne and Ghali, 2005. On the properties of other bonus malus systems around the world see Lemaire, 1995.

premium increases more flexibility. However, imperfect credit markets should temper the negative effects of savings on effort, given that transaction costs reduce the flexibility of savings over time. When savings are not monitored, the issue of contract renegotiation and savings on moral hazard is related to both the significance of the wealth effects and the imperfections in the credit markets that may reduce the negative effect of savings on safe behavior under full commitment.

Road safety and insurance markets are also subject to adverse selection and learning, which complicates the analysis of asymmetric information in insurance markets. Information asymmetries generally follow two distinct pathways. Adverse selection strongly predicts a positive correlation between policyholders' accident probability and the generosity of their insurance contracts. In the presence of moral hazard, the positive correlation is caused by the unobservability of effort to prevent accidents. Generous coverage reduces the expected cost of an accident and therefore reduces the incentives for safety. In the end, both pathways predict a positive correlation between accidents and coverage within a risk class. This suggests an empirical test for asymmetric information often referred to as the conditional correlation test. This test will be discussed in detail in Section 4. The evidence of the existence of residual asymmetric information in automobile insurance markets is not conclusive (Cohen and Siegelman, 2010). In Section 5, we will review a recent theoretical model that suggests statistical tests to separate moral hazard from asymmetric learning and adverse selection.

3. Incentive effects of point-record driver's licenses

3.1 Introduction

Point-record driver's licenses are non-monetary point-record mechanisms that provide incentives for safe driving. They do not involve direct monetary payments from the driver such as fines, which may incur significant indirect costs by reducing transportation flexibility, particularly for those who must drive to work. Each traffic violation is associated with a number of points that either accumulate (as in Québec) or deplete the individual capital (as in France). The number of points depends on the traffic violation severity. In our discussion, we consider accumulated demerit points. If the accumulated number of demerit points reaches or exceeds a given threshold, the driver's license is suspended. Before January 1990 this threshold was 12 in Québec and has been 15 since then.

To mitigate the social cost of license suspensions, point removal systems exist for most real-world point-record driver's licenses. In Québec, the demerit points related to a given driving offense are removed after two years. Hence, driver's license suspensions depend on the demerit points recorded during the last two years. Most point removal systems used by American states follow the same approach. The average number of demerit points per convicted offense is 2.4 in Québec. It takes about six traffic violations within two years to trigger a license suspension, an unlikely outcome when the annual traffic violation frequency is 16.9%. However, the heterogeneity of risks is high and a point-record driver's license is an incapacitating device for risky and reckless drivers through license suspensions. Another point removal system consists in cancelling all the demerit points after a given period of violation-free driving. This mechanism is used in most European countries. In France, for instance, traffic violations in the management of the point-record system are of both types. Until 2011, severe offences were removed globally after a three-year period of violation-free driving, whereas the duration was one year for less severe offenses. The increase in traffic control led to a substantial increase in the number of license suspensions, and the French authorities decided to strengthen point removal rules. Since 2011, the aforementioned durations are two years and six months respectively.

3.2 Basic model

Even if point-record driver's licenses always include a removal clause, we will first examine the polar case without removal. Bourgeon and Picard (2007) consider this case with an infinite horizon model that uses a binary effort variable. The cost of license suspension is the loss of driving utility during a period after which the driver is reinstated with a fresh zero-point record like that of a beginner. Dionne, Pinquet et al (2011a) extend this approach with a continuous effort level. The effort level that maximizes the driver's lifetime utility results from a maximization program that trades off the benefit of safe driving (reducing the accident probability) and the cost of unsafe driving (including a decrease in utility triggered by a traffic violation). The severity of violations (i.e. the related number of demerit points) is disregarded in these contributions to simplify the analysis. The two contributions show that the careful driving effort exerted by a rational policyholder increases with the number of accumulated demerit points under fairly general conditions. In other words, drivers that accumulate demerit points over time become safer to reduce the probability of losing their driver's license. This negative effect is possible only in the presence of moral hazard. There is no time effect on incentives, because the absence of point removal eliminates seniority effects of past traffic violations. These results are proved in Appendix A.1.

3.3 European model

We now consider more complex regimes with removal clauses. We first analyze the point-removal mechanism used in Europe (i.e. all the demerit points are removed after a given period of violation-free driving). Such systems are in place in France, Italy and Spain, etc. The time and event effects on incentives differ greatly from their counterpart in the North American setting presented in Section 3.4 (point removal applied to each traffic violation). The starting framework is that of the preceding section. The conclusions are the following:

- For every given number of demerit points, the incentives increase continuously with time. If optimal effort is greater than zero, it is strictly increasing.
- When all the demerit points are removed (i.e. after a violation-free driving record that reaches the threshold), the incentive level collapses to the minimum level.

Effort variation in the incentives after a convicted driving offense may be positive or negative. If the last traffic violation immediately follows another one, the variation is positive but decreases with the duration between the two last offenses. A drop in the effort level is expected if the duration is large enough. These results are proven in Appendix A.2. An informal proof of these results is the following. In a framework with point removal, the utility of a driver with an infinite horizon increases with time. In the European setting, the time counter is reset to zero after a traffic offence, and the utility after a traffic offense is constant. The incentive level is the difference between the current and the substitute utility (or the utility corresponding to a traffic offence); it therefore increases with time, as does the optimal effort level The event effect is less straightforward.

Incentive effects usually counteract revelation effects of unobserved heterogeneity. The risk level indicated by an event history increases with events and decreases with time, which motivates experience rating in non-life insurance from a fairness argument. Abbring et al (2003) investigate identification issues in the analysis of event histories in insurance, with the implicit assumption that incentive effects decrease with time and increase with events. In the present setting, however, the optimal effort level increases with time. Hence traffic violation risk decreases with time due to the incentives and to the revelation of unobserved heterogeneity. Figure 1 illustrates the French system.

Figure 1 here

3.4 North American model

We now analyze the incentive effects of the point-record driver's licenses in force in North America, where point removal is performed on each offence once a given seniority is reached. The regime is widely applied in the United States, United Kingdom, and Canada. The optimal behavior of the driver in this setting is more difficult to derive than with European rules, because all the seniorities of non-redeemed driving offenses must be included as state variables in the dynamic programming equations. Lifetime utility is expected to increase with time for a given number of demerit points accumulated. Optimal effort depends on the difference between the present utility and a substitute utility (i.e., that reached after an additional traffic violation). With the point removal system in force in Québec, the substitute utility increases with time, as does the present utility. Time should have more value for worse situations, which implies that the substitute utility should increase faster than the present utility. Optimal effort thus decreases with time. Appendix A.3 provides proof that optimal effort is continuous at the time of point removal. Optimal effort is then expected to increase with each traffic violation to compensate for the decreasing link between time and effort. To summarize, we expect the effort level to increase globally with the number of demerit points accumulated and to decrease with the seniority of non-redeemed traffic violations, if any. Figure 2 illustrates the case with two traffic offences.

Figure 2 here

4. Econometric measure of asymmetric information[3]

4.1 Introduction

The objective of this section is to present various tests for the presence of residual asymmetric information in automobile insurance markets. From the preceding sections, we know that the potential presence of asymmetric information between drivers and insurers regarding individual risks motivates experience rating, fines, and regulation of road safety. It is also well known from the insurance literature that risk classification is due, in part, to asymmetric information between the insurer and the insured (Crocker and Snow, 1986). Full efficiency in risk classification should separate individual risks and generate different actuarial insurance premiums that reflect these risks (Dionne and Rothschild, 2011). This means there should not be any residual asymmetric information between the insurer and the insured inside the risk classes. With actuarial premiums, full insurance should be the optimal contract, and there should be no correlation between insurance coverage and individual risk. However, in the real world of automobile insurance contracting, there may be numerous constraints that limit efficiency in risk classification. Incentive contracting and road safety regulation thus become important, and the empirical question is: how efficiently do these mechanisms improve road safety? In the following sections, we present the statistical methodology developed in recent years to detect residual asymmetric information in automobile insurance markets.

Cohen and Siegelman (2010) present a survey of empirical studies of adverse selection in insurance markets, including automobile insurance.[4] They argue that the coverage-risk correlation is particular to each market. Accordingly, the presence of a significant

[3] This section partly overlaps Section 8 in Dionne and Rothschild (2011) written for health insurance.

[4] See also Kim et al, 2009 and Saito, 2006.

coverage-risk correlation has different meanings in different markets, and even in different risk pools in a given market, depending on the type of insured service, the participants' characteristics, institutional factors, and regulation. This means that when testing for the presence of residual asymmetric information, one must control for these factors as well. What characteristics and factors explain the absence of coverage-risk correlation in automobile insurance markets? Some studies using the conditional correlation approach on cross-sectional data find evidence of asymmetric information (Dahlby, 1983, 1992; Puelz and Snow, 1994; Richaudeau, 1999; Cohen, 2005; Kim et al, 2009) while others did not (Chiappori and Salanie, 2000; Dionne et al, 2001, 2006). One major criticism of the conditional correlation approach with cross-sectional data is that it does not allow separation of adverse selection from moral hazard.

4.2 General tests for residual asymmetric information

Information problems are common in insurance markets. Usually, insured are better informed about their own characteristics or actions than their insurer. The two best-known information problems discussed in the economics literature are moral hazard and adverse selection (Arrow, 1963). Asymmetric learning is another information problem that can degenerate into adverse selection over time. Asymmetric learning is present when insured learn about their risk (by observing their accidents) more rapidly than the insurer (which observes only the claims). (See Cohen, 2005; Dionne, Michaud et al, 2011b.) Symmetric learning, in contrast, can create a full information situation between the parties after a given number of periods. Doing statistical tests on the link between insurance contracting and the presence of a given asymmetric information problem is therefore very complicated, because the same correlation between a contract characteristic and an observed risk can be attributed to more than one information problem. It may also be attributed to other characteristics that are not well controlled in the statistical test, such as risk aversion. The theoretical predictions must then be carefully established in a theoretical model (Dionne, Michaud, et al, 2011b) to distinguish the effect of each information problem.

Many theoretical contributions were published in the 1970s to account for stylized facts observed in insurance markets. The first models developed were one period or static.

Partial insurance, such as deductible and co-insurance contracts, can be justified by asymmetric information (Rothschild and Stiglitz, 1976; Shavell, 1979; Holmstrom, 1979). However, a deductible can be optimal for moral hazard, adverse selection, or proportional administrative costs. As mentioned, risk classification based on observable characteristics and multi-period relationships between principal and agent are other mechanisms associated with the presence of asymmetric information.

The first goal of empirical research on information problems in different markets is to determine whether residual asymmetric information remains in these markets when different incentive schemes are introduced. The empirical question in insurance can be summarized as follows: Is there any residual correlation between chosen insurance coverage and risk within risk classes? The second goal is to identify which information problem remains when the first test rejects the null hypothesis that there is no residual information problem. This step is important for both the insurer and the regulator because they must implement the appropriate instruments to improve resource allocation. A deductible efficiently reduces *ex ante* moral hazard, but not necessarily *ex post* moral hazard because often, the accident has already occurred when the action is taken. A high deductible can even have an adverse effect and encourage accident cost building (Dionne and Gagné, 2001). As is well known in the empirical literature, a positive correlation between insurance coverage and risk is a necessary condition for the presence of asymmetric residual information, but it does not shed light on the nature of the information problem. The third goal is to find ways to improve the contracts or the regulation and reduce the negative impact of asymmetric information on resource allocation. These resource allocation objectives must take into account other issues such as risk aversion, fairness, and accessibility of insurance. This last issue is particularly important in automobile insurance markets. A decrease in insurance coverage may reduce *ex ante* moral hazard because it exposes the insured person to risk, but it also significantly reduces accessibility to insurance protection for risky people who are not always responsible for their financial conditions.

In insurance markets, the distinction between moral hazard and adverse selection boils down to a question of causality (Chiappori, 2000). The theoretical literature on moral hazard states that unobserved actions of the insured result from the forms of contracts.

For example, a generous automobile insurance plan can reduce the incentives for prevention and increase the risk of having a road accident. With adverse selection, the nature of the risk already exists, but the nature of the contracts chosen is a function of the risks. There is therefore a reverse causality between the risk and the contract when we consider each problem separately, although the correlation between insurance coverage and the level of risk is positive in both cases.

Econometricians analyze two types of information when studying insurers' data (Boyer, Dionne, Vanasse, 1992; Puelz and Snow, 1994; Gouriéroux, 1999; Richaudeau, 1999; Dionne et al, 2006). The first type contains variables that are observable by both parties to the insurance contract. Risk classification variables are one example. Econometricians/insurers combine these variables to create risk classes when estimating accident distributions. Observed variables can be used to make estimates conditional on the risk classes or within the risk classes. The second type of information is related to what is not observed by the insurer or the econometrician during contract duration and at contract renegotiations, but can explain the insured's choice of contracts or actions. If we limit our interpretation to asymmetric information (either moral hazard or adverse selection), we can test the conditional residual presence of asymmetric information in an insurer's portfolio by testing for a correlation between the contract coverage and the realization of the risk variable during a contract period. Two parametric tests have been proposed in the literature (Chiappori and Salanié, 2000; Dionne et al, 2001; see Chiappori, 2000; and Chiappori and Salanié, 2003 for detailed analyses). One parametric test (Dionne at al. 2001) estimates the following relationship:

$$y_i = g\left(\alpha + \beta X_i + \gamma d_i + \delta E(d_i|X_i)\right) + \varepsilon_i, \qquad (1)$$

where y_i is the contract choice by individual i (level of deductible, for example), X_i is a vector of control variables such as the observable characteristics used in risk classification and control variables for risk aversion, β is a vector of parameters to be estimated, d_i is the realization of the random variable observed at the end of the contract period (accident or not, for example), $E(d_i|X_i)$ is the conditional expected value of the random variable obtained from the estimation of the accident distribution, and ε_i is the residual of the regression. A positive sign is usually anticipated for the coefficient of d_i (γ) when residual asymmetric information remains (higher coverage is related to more accidents or higher risk). The seminal theories of Rothschild and Stiglitz (1976) and Wilson (1977) strongly predict that such a correlation should be observed in the data in the presence of adverse selection, while Holmstrom (1979) and Shavell (1979) strongly predict that the correlation is due to moral hazard. Note that the dependent variable in the above regression can be the risk variable d_i while the coverage y_i is an independent variable. This symmetry is discussed in detail in Dionne et al (2006). The presence of the variable d_i is not necessarily exogenous in equation (1). It is often better to instrument this variable (See Dionne et al, 2009, 2010, and Rowell, 2011, for more details).

The presence of $E(d_i|X_i)$ is necessary to control for specification errors (missing variables) or for potential non-linearity not modeled in the equation. Without this control, the coefficient of d_i can be significant for reasons other than the presence of residual asymmetric information in the risk classes.

If the coefficient of d_i is not significant, one can reject the presence of residual asymmetric information in the risk classes when all other factors are well controlled. This does not mean that there is no asymmetric information in this market; rather, it means that the insurer's risk classification system eliminates asymmetric information efficiently, and that there is no *residual* asymmetric information within the risk classes. In other words, when risk classification is done properly, it is not necessary to choose the contract form within the risk classes or to modify the safety regulation to reduce asymmetric information.

An equivalent model was proposed by Chiappori and Salanié (2000). Here, two equations are estimated simultaneously, one for contract choice and the other for accident distribution. An example is the bivariate probit model:

$$y_i = f(X_i, \beta) + \varepsilon_i \quad (2)$$

$$d_i = g(X_i, \beta) + \eta_i. \quad (3)$$

The test consists in verifying whether there is dependence between the residuals of the two equations. An absence of conditional correlation is interpreted as an absence of residual asymmetric information in the data. The authors present an additional non-parametric test that is independent of the functional forms of the above models. It is based on a Chi-square test of independence. However their test seems to be limited to discrete variables, contrary to the two parametric tests presented above. (See Su and Spindler, 2010, for a longer discussion).

Fang et al (2008) do not reject asymmetric information in the medical insurance market, but do not find evidence of adverse selection. Their results are consistent with multidimensional private information along with advantageous selection (de Meza and Webb, 2001). They obtain a *negative* correlation between risk and insurance coverage. Risk aversion is not a source of advantageous selection in their data. The significant sources are income, education, longevity expectations, financial planning horizons, and most importantly, cognitive ability.

To separate moral hazard from adverse selection, econometricians need a supplementary step. An additional market relationship can be estimated to look for adverse selection (conditional on the fact that the null hypothesis of no asymmetric information was rejected), as Dionne et al (2009) did for auctions. In insurance markets, dynamic data are often available. Time adds an additional degree of freedom to test for asymmetric information (Dionne and Vanasse, 1992; D'Arcy and Doherty, 1990; Dionne and Doherty, 1994; Hendel and Lizzeri, 2003). This information can be used in many insurance markets where past experience information is available and when it is possible to use it. For ethical reasons, this information is not utilized on an individual basis in health insurance and for bodily injury insurance in many countries. Experience rating works at two levels in insurance. Past accidents implicitly reflect unobservable characteristics of the insured (adverse selection) and introduce additional incentives for prevention (moral hazard). Experience rating can therefore directly mitigate problems of adverse selection and moral hazard, which often hinder risk allocation in the insurance market.

Experience rating not only provides additional information on risk, but may also play an important role in the dynamic relationship between policyholders' insurance claims and contract choice. The theoretical literature on repeated insurance contracting over time clearly indicates that these features may help overcome problems of moral hazard when risks known to the policyholder (endogenous) are unobservable by the insurer (Winter, 2000) or when exogenous characteristics are unobservable (Dionne et al, 2000). Contract choice is influenced by the evolution of the premium, which is closely linked to the insured's risk or past experience. Because increased insurance coverage tends to lower the expected cost of accidents for the insured, incentives for safe behavior are weakened for all risks. Under experience rating, the subsequent rise in accidents increases the marginal costs of future accidents when experience rating is taken into account. Experience rating may therefore offset the disincentive effect created by single-period insurance coverage.

The above empirical tests are conducted in a static framework, which fails to recognize the dynamics that experience rating introduces in contractual relationships. Chiappori and Salanié (2000) discuss in detail how the omission of the experience-rating variable, even in tests with one-period data, must plausibly explain the failure to detect asymmetric information. Dionne, Michaud et al (2004) tested this conjecture by adding a bonus-malus variable in equations similar to those presented above. They affirmed that the bonus-malus coefficient was indeed negatively related to the level of insurance coverage (through fluctuations in the premium) and positively correlated to claims (potentially through unobserved heterogeneity). The coefficients thus appear to hide the link between claims and contract choice, which is exactly what Chiappori and Salanié (2000) argue. This is apparent in traditional cross-sectional tests, as well as in extrapolations using longitudinal data models that simply pool repeated observations or permit the correlation of unobserved independent factors with each contract observed over time. The additional time factor thus improves the power of the test to detect asymmetric information.

Abbring et al (2003) apply a multi-period incentive mechanism by focusing on the dynamics of claims, but not on the dynamics of contract choice (because of data limitations). Proposing specific assumptions about the wealth effects of accidents to policyholders who differ only in their claim records (thus their experience rating), their model predicts that subjects with the worst claims records should try harder to increase safety, and thereby, ceteris paribus, file fewer claims in the future. However, their data do not support the presence of moral hazard. Dionne, Pinquet et al (2011a) extend their model and do not reject the presence of moral hazard, using a different data set. The potential presence of adverse selection in their data was not a real problem because all drivers must be insured for bodily injuries (see also Abbring et al, 2008, and Rowell, 2011, for other tests of moral hazard).

Dionne, Michaud, and Dahchour (2011b) show that failure to detect residual asymmetric information, and more specifically, moral hazard and adverse selection in insurance data, is due to the failure of previous econometric approaches to model the dynamic relationship between contract choice and claims adequately and simultaneously when looking at experience rating. Intuitively, because there are at least two potential information problems in the data, an additional relationship to the correlation between risk and insurance coverage is necessary to test for the causality between risk and insurance coverage. Using a unique longitudinal survey of policyholders from France, they propose a methodology to disentangle the historical pathways in claims and premiums. They show how causality tests can be used to differentiate moral hazard from asymmetric learning (and eventually adverse selection). They do not reject moral hazard for a given group of policyholders, and do not reject asymmetric learning for younger drivers.

5. Testing for moral hazard in the automobile insurance market

This section presents in more detail two contributions that found residual moral hazard in automobile insurance, described above. The emphasis is on the econometric methodology.

5.1 Moral hazard in Québec automobile insurance data

A point-record driver's license was implemented in Québec in 1978, together with a no-fault insurance regime for bodily injuries, which replaced a tort system. We now concentrate on this regime in analyzing moral hazard in function demerit points. Because no-fault environments are common in the North American continent, traffic violations are events likely to be used in experience rating schemes. Increases in premiums are often triggered by claims at fault in the motor insurance sector. Here we concentrate on the use of demerit points.

In Québec, the public insurer in charge of the compensation of bodily injuries utilizes an experience rating scheme based on demerit points. The same public enterprise is also in charge of the point-record license system. Dionne, Pinquet et al (2011a) showed that the new insurance pricing reduced the number of traffic violations by 15%. They also verified that there is still residual *ex ante* moral hazard in the management of road safety. We now concentrate the discussion on the methodology they developed for obtaining this result.

As mentioned, the methodology extends the model of Abbring et al (2003). Over time, observed demerit points of a driver illuminates two effects: an unobserved heterogeneity effect and an incentive effect. Drivers with more demerit points accumulated during a period are riskier with respect to hidden features in risk distributions. Hence, unobserved heterogeneity is a form of risk reassessment over time in the sense that those who accumulate demerit points represent higher risks over time. This effect is in the opposite direction of the incentive effect. Moreover, as shown in Section 3.4, the time effect of unobserved heterogeneity is also converse to that of the incentive effect.

The model, proposed by Dionne, Pinquet et al (2011a), tests for an increasing link between traffic violation and the number of accumulated demerit points over time. Rejecting the positive link is evidence of moral hazard. They estimate the following hazard function (Cox, 1972):

$$\lambda_i(t) = \exp\left(x_i(t)\beta\right) + g\left(adp_i(t)\right) \times h\left(c_i(t)\right) \tag{4}$$

where $\lambda_i(t)$ is the hazard function for driver i at time t, $x_i(t)$ is a vector of control variables, β represents the corresponding coefficients, and $adp_i(t)$ is the number of demerit points accumulated over the two previous years at time t.

In the absence of moral hazard, g should be increasing. They found that g is decreasing when drivers have accumulated more than seven demerit points. This means that beyond seven demerit points, they become safer drivers if they do not lose their driver's license. This is evidence of the presence of moral hazard in the data. It also means that these drivers were negligent when the accumulated record was below seven demerit points.

5.2 Moral hazard in French automobile insurance data

To separate learning leading to adverse selection (asymmetric learning) from moral hazard, Dionne, Michaud et al (2011b) consider the case where information on contracts and accidents is available for multiple years in the form of panel data. They exploit dynamics in accidents and insurance coverage controlling for dynamic selection due to unobserved heterogeneity. They construct two additional tests based on changes in insurance coverage. Coupled with the negative occurrence test of Abbring et al (2003) and Dionne, Pinquet et al (2011a), these tests allow them to separate moral hazard from asymmetric learning (which should become adverse selection in the long run).

They analyze the identification of asymmetric learning and moral hazard within the context of a tractable structural dynamic insurance model. From the solution of their theoretical model, they simulate a panel of drivers behaving under different information regimes or data generating processes (with or without both moral hazard and asymmetric learning). They validate their empirical tests on simulated data generated from these different information regimes. They then apply these tests to longitudinal data on accidents, contract choice and experience rating for the period 1995-1997 in France. They find no evidence of information problems among experienced drivers (more than 15 years of experience). For drivers with less than 15 years of experience, they find strong evidence of moral hazard but little evidence of asymmetric learning. They obtain evidence of asymmetric learning, despite the small sample size, when focusing on drivers with less than 5 years of experience. To obtain these results, they estimated the following model.

They consider a joint parametric model for the probabilities of accidents and contract choice. Each equation corresponds to a dynamic binary choice model with pre-determined regressors and an error component structure. The error component structure is important given the likelihood of serial correlation in contract and accident outcomes. They use the solution proposed by Wooldridge (2005) to take the potential left censoring effect into account.

More specifically, the process for accidents is specified as:

$$\begin{gathered} n_{it} = I(x_{it}\beta_n + \phi_{nd}d_{it-1} + \phi_{nn}n_{it-1} + \phi_{nb}b_{it} + \varepsilon_{n,it} > 0) \\ i = 1,\ldots,N, t = 1,\ldots,T \end{gathered} \tag{5}$$

where $\varepsilon_{n,it}$ has an error component structure $\varepsilon_{n,it} = \alpha_{ni} + v_{n,it}$, n_{it} is a binary variable for the number of accidents of individual *i* at time *t*, d_{it-1} is his contract choice in period $t-1$, n_{it-1} is his number of accidents in period $t-1$, and b_{it} is his bonus-malus score at period *t*. The presence of moral hazard would be confirmed by a positive sign for ϕ_{nd} (more insurance coverage-more accidents) and a negative sign for ϕ_{nb} (a higher malus creates more incentives for safe driving as for the test presented in the previous section with accumulated demerit points.) Here a high malus means an accumulation of accidents over the previous periods. They specify a similar equation for contract choice:

$$\begin{gathered} d_{it} = I(x_{it}\beta_d + \phi_{dd}d_{it-1} + \phi_{dn}n_{it-1} + \phi_{db}b_{it} + \varepsilon_{d,it} > 0) \\ i = 1,\ldots,N, t = 1,\ldots,T \end{gathered} \tag{6}$$

where again $\varepsilon_{d,it} = \alpha_{di} + v_{d,it}$. The asymmetric learning test is a test of whether an accident in the last period, conditional on the bonus-malus, leads to an increase in coverage this period. The driver thus learns he is riskier than anticipated and increases his insurance coverage. It is a test of whether $\phi_{dn} > 0$ or not.

6. Conclusion

We have studied several recent modeling frameworks in detail to illustrate the possible trade-offs between efficient insurance provision and regulation and incentives for road safety. We have also analyzed statistical models used to detect residual informational asymmetries in different data sets. The results indicate a presence of residual moral hazard and asymmetric learning (for young drivers). This means there is still room for improving resource allocation in the developed countries studied even if road-traffic accidents have significantly decreased in recent years.

Many extensions to the existing models have to be considered. Very few studies have considered the fleets of vehicles and particularly the trucking industry, which is responsible for much of the risk exposure on the roads. Here the relevant trade-off between insurance coverage and incentives must take into account the profitability of road safety in a hierarchical framework comprising insurers, regulators, fleet owners and drivers. Symmetric learning where the econometrician and the insurer learns the risk type of the driver by observing only the claims while the driver observes all accidents is another difficulty not well treated in the literature. Finally, "La soif du bonus" (Lemaire, 1985), associated with the incentive to claim or not in order to keep a good bonus-malus record, may create another bias in the estimation of residual asymmetric information in the data.

The statistics clearly show that the social cost of road safety is much higher in less developed regions. We have suggested that the development of incentives for safe driving might be an explanation for the difference between developed and less developed regions. However, other explanations are also linked to the lower wealth of these regions and consequently the lower quality of roads and vehicles, and to the absence of high quality driver's education.

References

Abbring, J., Chiappori, P.A., Pinquet, J., 2003. Moral Hazard and Dynamic Insurance Data. Journal of the European Economic Association 1(4), 767-820.

Abbring, J.H., Chiappori, P.A., Zavadil, T., 2008. Better Safe than Sorry? *Ex ante* and *Ex post* Moral Hazard in Dynamic Insurance Data. Mimeo, VU University of Amsterdam.

Blomquist G.C., 1988. The Regulation of Motor Vehicle and Traffic Safety. Kluwer Academic Publishers, Boston.

Boyer, M., Dionne, G., 1987. The Economics of Road Safety. Transportation Research, 21B (5), 413-431.

Boyer, M., Dionne, G., 1989. An Empirical Analysis of Moral Hazard and Experience Rating. The Review of Economics and Statistics 71, 128-134.

Boyer, M., Dionne, G., Vanasse, C., 1992. Econometric Models of Accident Distributions. In: Dionne, G. (Ed.), Contributions of Insurance Economics, Kluwer Academic Publishers, Boston, pp. 169-213.

Bourgeon, J.M., Picard, P., 2007. Point-Record Driver's license and Road Safety: An Economic Approach. Journal of Public Economics 91, 235-258.

Chiappori, P.A., 2000. Econometric Models of Insurance Under Asymmetric Information. In: Dionne, G. (Ed.), Handbook of Insurance, Kluwer Academic Publishers, Boston, pp. 365-393.

Chiappori, P.A., Jullien, B., Salanié, B., Salanié, F., 2006. Asymmetric Information in Insurance: General Testable Implications. RAND Journal of Economics 37, 783-798.

Chiappori, P.A., Macho, I., Rey, P., Salanié, B., 1994. Repeated Moral Hazard: The Role of Memory, Commitment, and the Access to Credit Markets. European Economic Review 38, 1527-1553.

Chiappori, P.A., Salanié, B., 2003. Testing Contract Theory: A Survey of Some Recent Work. In: Dewatripont, M., Hansen, L.P., Turnovsky, S. (Eds.), Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress, Vol. 1, Cambridge University Press, Cambridge, pp. 115-149.

Chiappori, P.A., Salanié, B., 2000. Testing for Asymmetric Information in Insurance Markets. Journal of Political Economy 108, 56-78.

Cohen, A., 2005. Asymmetric Information and Learning: Evidence from the Automobile Insurance Market. The Review of Economics and Statistics 87(2), 197-207.

Cohen, A., Dehejia, R., 2004. The Effects of Automobile Insurance and Accident Liability Laws on Traffic Fatalities. Journal of Law and Economics 47(2), 357-393.

Cohen, A., Siegelman, P., 2010. Testing for Adverse Selection in Insurance Markets. Journal of Risk and Insurance, 77(1), 39-84.

Cox, D.R., 1972. Regression Models and Life Tables. Journal of the Royal Statistical Society, Series B, 34, 187-220.

Crocker, K., Snow, A., 1986. The Efficiency Effects of Categorical Discrimination in the Insurance Industry. Journal of Political Economy 94, 321-344.

Cummins, J.D., Phillips, R., Weiss, M.A., 2001. The Incentive Effects of No-Fault Automobile Insurance. Journal of Law and Economics, 44(2), 427-464.

Dahlby, B., 1992. Testing for Asymmetric Information in Canadian Automobile Insurance. In: Dionne, G. (Ed.), Contributions to Insurance Economics, Kluwer Academic Publishers, Boston, pp. 423-443.

Dahlby, B., 1983. Adverse Selection and Statistical Discrimination: An Analysis of Canadian Automobile Insurance. Journal of Public Economics 20, 121-130.

D'Arcy, S.P., Doherty, N.A., 1990. Adverse Selection, Private Information, and Lowballing in Insurance Markets. Journal of Business 63, 145-164.

De Meza, D., Webb, D.C., 2001. Advantageous Selection in Insurance Markets. RAND Journal of Economics 32(2), 249-262.

Devlin, R.A., 1992. Liability Versus No-fault Automobile Insurance Regimes: An Analysis of the Experience in Québec. In: G. Dionne (Ed.), Contributions to Insurance Economics, Kluwer Academic Publishers, Boston, 499-520.

Dionne, G., 2001. Insurance Regulation in Other Industrial Countries. In: Cummins, J.D. (Ed.), Deregulating Property-Liability Insurance, AEI-Brookings Joint Center For Regulatory Studies, Washington, pp. 341-396.

Dionne, G., Doherty, N.A., 1994. Adverse Selection, Commitment and Renegotiation: Extension to and Evidence from Insurance Markets. Journal of Political Economy 102, 209-235.

Dionne, G., Doherty, N.A., Fombaron, N., 2000. Adverse Selection in Insurance Markets. In: Dionne, G. (Ed.), Handbook of Insurance, Kluwer Academic Publishers, Boston, pp. 185-243.

Dionne, G., Gagné, R., 2001. Deductible Contracts Against Fraudulent Claims: Evidence from Automobile Insurance. Review of Economics and Statistics 83(2), 290-301.

Dionne, G., Ghali, O., 2005. The (1992) Bonus-malus System in Tunisia: An Empirical Evaluation. Journal of Risk and Insurance 72(4), 609-633.

Dionne, G., Gouriéroux, C., Vanasse, C., 2001. Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment. Journal of Political Economy 109, 444-453.

Dionne, G., Laberge-Nadeau, C. (Eds), 1999. Automobile Insurance: Road Safety, New Drivers, Risks, Insurance Fraud and Regulation, Kluwer Academic Publishers, Boston.

Dionne, G., La Haye, M., Bergerès, A.S., 2010. Does Asymmetric Information Affect the Premium in Mergers and Acquisitions? Working paper 10-03, Canada Research Chair in Risk Management, HEC Montréal.

Dionne, G., Lasserre, P., 1985. Adverse Selection, Repeated Insurance Contracts and Announcement Strategy. Review of Economic Studies 70, 719-723.

Dionne, G., Vanasse, C., 1989. A Generalization of Automobile Insurance Rating Models: The Negative Binomial Distribution with a Regression Component. ASTIN Bulletin 19, 199-212.

Dionne, G., Maurice, M., Pinquet, J., Vanasse, C., 2005. The Role of Memory in Long-Term Contracting with Moral Hazard: Empirical Evidence in Automobile Insurance. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=764705.

Dionne, G., Michaud P.C., Dahchour, M, 2011(b). Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France. Working paper 10-05, Canada Research Chair in Risk Management, HEC Montreal. Forthcoming in Journal of the European Economic Association.

Dionne, G., Michaud P.C., Dahchour, M, 2004. Separating Moral Hazard from Adverse Selection and Learning in Automobile Insurance: Longitudinal Evidence from France. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=583063.

Dionne, G., Pinquet, J., Maurice, M., Vanasse, C., 2011(a). Incentive Mechanisms for Safe Driving: A Comparative Analysis with Dynamic Data. Review of Economics and Statistics 93(1), 218-227.

Dionne, G., Rothschild, C.G. (2011). Risk Classification and Health Insurance. Working paper 11-05, Canada Research Chair in Risk Management, HEC Montreal.

Dionne, G., St-Amour, P., Vencatachellum, D., 2009. Asymmetric Information and Adverse Selection in Mauritian Slave Auctions. Review of Economic Studies 76, 1269-1295.

Doyle, J.J., 2005. Health Insurance, Treatment and Outcomes: Using Auto Accidents as Health Shocks. The Review of Economics and Statistics 87(2), 256-270.

Fagart, M.C., Fluet, C., 2009. Liability Insurance under the Negligence Rule. RAND Journal of Economics 40(3), 486-508.

Fang, H., Keane, M.P., Silverman, D., 2008. Sources of Advantageous Selection: Evidence from the Medigap Insurance Market. Journal of Political Economy 116, 303-350.

Graham J.D., Garber, S., 1984. Evaluating the Effects of Automobile Safety Regulation. Journal of Economic Policy and Management 3(2), 206-224.

Haddon, W. Jr., 1968. The Changing Approach to the Epidemiology, Prevention, and Amelioration of Trauma: The Transition of Approaches Etiologically Rather than Descriptively Based. American Journal of Public Health 58, 1431-1438.

Hendel, I., Lizzeri, A., 2003. The Role of Commitment in Dynamic Contracts: Evidence from Life Insurance. Quarterly Journal of Economics 118, 299-327.

Holmström, B., 1979. Moral Hazard and Observability. The Bell Journal of Economics 10(1), 74-91.

Israel, M., 2007. Do We Drive More Safely When Accidents are More Expensive? Identifying Moral Hazard from Experience Rating Schemes. Working paper, Wharton School, University of Pennsylvania.

Kim, H., Kim, D., Im, S., 2009. Evidence of Asymmetric Information in the Automobile Insurance Market: Dichotomous Versus Multinomial Measurement of Insurance Coverage. Journal of Risk and Insurance, 76(2), 343-366.

Landes, E.M., 1982. Insurance, Liability, and Accidents: A Theoretical and Empirical Investigation of the Effect of No-fault Accidents. Journal of Law and Economics 25, 49-65.

Lemaire, J., 1995. Bonus-malus Systems in Automobile Insurance. Kluwer Academic Publishers, Boston, 283 p.

Lemaire, J., 1985. Automobile Insurance: Actuarial Models. Kluwer Academic Publishers, Boston, 248 p.

Michaud P.C., Tatsiramos, K., 2011. Fertility and Female Employment Dynamics in Europe: The Effect of Using Alternative Econometric Modeling Assumptions. Journal of Applied Econometrics 26, 4, 641-668.

Murray, C.J.L., Lopez, A.D., 1997. Alternative Projections of Mortality and Disability by Cause 1990-2020: Global Burden of Disease Study. The Lancet 349, 1498-1504.

OECD International Transfort Forum (2010): A Record Period for Road Safety. URL: http://cemt.org/IRTAD/.

Pauly, M.V., 1974. Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection. Quarterly Journal of Economics 88(1), 44-62.

Peltzman, S., 1975. The Effects of Automobile Regulation. Journal of Political Economy 83, 677-725.

Picard, P., 2000. Les nouveaux enjeux de la régulation des marchés d'assurance. Working Paper no 2000-53, THEMA, Université Paris X-Nanterre.

Pinquet, J., 1999. Une analyse des systemes bonus-malus en assurance automobile. Assurances 67(2), 241-249.

Pinquet, J., Guillén, M., Bolancé, C., 2003. Time-varying credibility for frequency risk models: Estimation and tests for autoregressive specifications on the random effects. Insurance: Mathematics and Economics 33, 273-282

Pinquet, J., Guillén, M., Bolancé, C., 2001. Allowance for the Age of Claims in Bonus-Malus Systems. ASTIN Bulletin 2(31), 337-348.

Pinquet, J., 2000. Experience Rating through Heterogeneous Models. In: Dionne, G. (Ed.), Handbook of Insurance, Kluwer Academic Publishers, Boston.

Priest, G.L., 1987. The Current Insurance Crisis and Modern Tort Law. Yale Law Journal 96(7), 1521-1541.

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

Richaudeau, D., 1999. Automobile Insurance Contracts and Risk of Accident: An Empirical Test Using French Individual Data. Geneva Papers on Risk and Insurance Theory 24, 97-114.

Rothschild, M., Stiglitz, S., 1976. Equilibrium in Insurance Markets: An Essay on the Economics of Imperfect Information. Quarterly Journal of Economics 90(4), 629-649.

Rowell, D., 2011. Moral Hazard: Empirical Evidence in the Australian Market for Automobile Insurance. PhD thesis, University of Queensland.

Saito, K., 2006. Testing for Asymmetric Information in the Automobile Insurance Market under Rate Regulation. Journal of Risk and Insurance 73(2), 335-356.

Shavell, S., 2004. Foundations of Economic Analysis of Law. The Belknap Press of Harvard University Press, 737 p.

Shavell, S., 1979. On Moral Hazard and Insurance. Quaterly Journal of Economics 93(4), 541-562.

Sloan, F.A., Reilly, B.A., Schenzler, C., 1995. Effects of Tort Liability and Insurance on Heavy Drinking and Driving. Journal of Law and Economics 38, 49-77.

Su, L., Spindler, M., 2010. Nonparametric Test for Asymmetric Information. Working paper, Singapore Management University.

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

Winter, R., 2000. Optimal Insurance under Moral Hazard. In: Dionne, G. (Ed.), Handbook of Insurance, Kluwer Academic Publishers, Boston, pp. 155-183.

WHO (World Health Organization), 2004. World Report on Road Traffic Injury Prevention, WHO Press, Geneva.

Weiss, M., Tennyson, S., Regan, L., 2010. The Effects of Regulated Premium Subsidies on Insurance Costs: An Empirical Analysis of Automobile Insurance. Journal of Risk and Insurance 77, 597-624.

Wooldridge, J.M., 2005. A Simple Solution to the Initial Conditions Problem in Dynamic, Nonlinear Panel Data Models with Unobserved Heterogeneity. Journal of Applied Econometrics 20, 39-54.

Appendixes

A.1: Point-record driver's license without point removal

We assume that the driver's license is revoked when the driver reaches a total of N demerit points. For the sake of simplicity, each convicted traffic violation is linked to one supplementary demerit point in this appendix. A driver with a suspended driver's license is reinstated after a period D with a fresh zero-point record like that of a beginner. The duration D may be fixed or random in the model. In Québec, a license suspension is of random length because drivers must take the exam again after a given period before their license is reinstated. A rational driver maximizes his expected lifetime utility expressed in $ and derived from:

- An instantaneous driving utility, d_u.
- A time-dependent disutility of effort, denoted as $e(t)$. This effort level is linked to an instantaneous traffic violation frequency risk, denoted as $\lambda(e(t))$. The hazard function $\lambda(e)$ corresponds to a probability $p(e)$ in static or discrete time incentive models, and is assumed to be a positive, decreasing and strictly convex function of the effort level.

In this appendix, we posit that there is no point removal mechanism. In that case, the lifetime expected utility (we assume an infinite horizon) depends only on the number n of accumulated demerit points, and is denoted as u_n. The Bellman equation (A.1) on the expected utility leads to:

$$u_n = \frac{d_u}{r} - \frac{c_\lambda(u_n - u_{n+1})}{r}, \ (0 \le n < N), \tag{A.1}$$

where r is a subjective discount rate, and where c_λ is defined as follows:

$$c_\lambda(\Delta u) \underset{def}{=} \min_{e \ge 0} e + \left[\lambda(e) \times \Delta u\right]. \tag{A.2}$$

From equation (A.2), incentives are effective if we have $\Delta u > -1/\lambda'(0)$. In this equation, Δu is the lifetime utility loss between the current state and the one reached after an additional traffic offense. Once quantified, Δu is the monetary equivalent of this traffic violation. The function c_λ minimizes the sum $e + \left[\lambda(e) \times \Delta u\right]$, which is the disutility flow of both effort (short-term component) and the expected lifetime utility loss (long-term component). All the u_n are lower than $u_{\max} = d_u / r$, the private lifetime driving utility without the point-record driver's license. Equation (A.1) means that $c_\lambda(u_n - u_{n+1})/r$ is the minimal private utility cost of the point-record mechanism for a driver with n demerit points. The cycle of lifetime utilities is closed with a link between u_0 and u_N, the lifetime expected utility just after the suspension of the driver's license. For instance, if the private disutility of driver's license suspension is only the loss of driving utility during a period D, we have

$$u_N = \beta u_0,\ \beta = E[\exp(-rD)]. \tag{A.3}$$

The optimal utilities are then derived from the Bellman equation (A.1), and equation (A.3). Optimal effort depends on the variation of lifetime utility Δu because it minimizes the function defined in equation (A.2). The variable Δu (the argument of c_λ), which determines optimal effort, will be referred as the optimal incentive level. In this setting, optimal effort depends on the number n of accumulated demerit points but not on time, and we denote it as e_n. It is easily shown that e_n increases with n for any given value of N. The related frequency of traffic violations $\lambda_n = \lambda(e_n)$ thus decreases with n. The intuition behind this result is the following. A given reduction in traffic violation risk is more efficient as the threat of the license suspension gets closer. Hence the efficiency of effort increases with the number of demerit points accumulated, and we obtain an increasing link between this number and the optimal effort level. A parallel can be drawn with the "three strikes and you're out" rule enforced in California to deter crime. The deterrence effect increases from one to two strikes (Helland and Tabarrok, 2007).

A.2: Point-record driver's license with global point removal

In this appendix, we derive the incentive properties of the point-removal clause of the European type (global cancellation of demerit points once a given violation-free driving period is reached). The expected utility is denoted as $u_n(t)$, where t $\left(0 \le t \le T\right)$ is the seniority of the last convicted traffic violation. The expected utility for $n=0$ does not depend on time, and is denoted as u_0. All the demerit points are redeemed if $t=T$, and we have $u_n(T)=u_0$ for $n=1,\ldots,N-1$. With the assumptions and notations of the preceding section, the Bellman equation is (A.4).

$$\begin{aligned} &u_n^{'}(t)=\left(ru_n(t)-d_u\right)+c_\lambda\left(u_n(t)-u_{n+1}(0)\right)\left(0 \le n < N\right) \\ &\Leftrightarrow u_n^{'}=f_{u_{n+1}(0)}\left(u_n\right),\ f_v\left(u\right)=\left(ru-d_u\right)+c_\lambda\left(u-v\right). \end{aligned} \tag{A.4}$$

Hence $u_n(t)$ is defined implicitly by equation (A.5).

$$\int_{u_n(t)}^{u_n(T)=u_0} \frac{du}{f_{u_{n+1}(0)}(u)} = T-t,\ \forall t \in [0,T]. \tag{A.5}$$

The functions f_v are strictly increasing and the integral of $1/f_v$ diverges in the neighborhood of the value, which nullifies f_v because we have:

$$r < f_v^{'}(u) \le r + \lambda(0)\ \forall u, v.$$

Hence $u_{n+1}(0)$ is defined from $u_n(0)$ from equation (A.4) with $t=0$, which implies that:

$$f_{u_{n+1}(0)}(u_n(0)) > 0.$$

This condition holds if $f_v(u_n(0)) > 0$ for positive values of v, which amounts to:

$$u_n(0) > \underline{u},\ with\ f_0(\underline{u}) = (r\underline{u} - d_u) + c_\lambda(\underline{u}) = 0.$$

Then:

$$f_{u_{n+1}(0)}\left(u\right)>0\ \forall\, u\in\left[u_n\left(0\right),\ u_n\left(T\right)=u_0\right]$$

and u_n is strictly increasing on $[0,T]$ from (A.4).

A result with an obvious economic interpretation is that the sequence $\left(u_n\left(0\right)\right)_{n=0,\ldots,N}$ is decreasing. Indeed, we have:

$$f_{u_{n+1}(0)}\left(u_n\left(0\right)>0\Leftrightarrow u_n\left(0\right)-u_{n+1}\left(0\right)>c_\lambda^{-1}(d_u-ru_n\left(0\right)\right)>0.$$

We now prove the concavity of the sequence $(u_n(0))_{n=0,\ldots,N}$, which means that the sequence $\left(u_n\left(0\right)-u_{n+1}\left(0\right)\right)_{n=0,\ldots,N}$ is increasing. The recurrence equation (A.6):

$$u_{n+1}(0)=w(u_n(0)),\ with\int_v^{u_0}\frac{du}{(ru-d_u)+c_\lambda(u-w(v))}=T \qquad \text{(A.6)}$$

follows from (A.5) with $t=0$. Consider v_0,v_1 with $\underline{u}<v_0<v_1<d_u/r$.

We have:

$$\int_{v_1}^{u_0}\frac{du}{(ru-d_u)+c_\lambda(u-w(v_1))}=T;\ \int_{v_0}^{u_0}\frac{du}{(ru-d_u)+c_\lambda(u-w(v_1)+v_1-v_0)}>T.$$

The first equality results from (c). The second integrand is greater than the first one if the comparison starts from the lower bound of the integral, and the wider range of the second integration reinforces the inequality. Because these integrals increase with w, we have:

$$v_0<v_1\Rightarrow w(v_1)+v_0-v_1>w(v_0).$$

Hence the function $v\to v-w(v)$ is decreasing. We obtain the desired result with $v_0=u_{n+1}(0)$ and $v_1=u_n(0)$.

We can then prove the results on the optimal effort level, which we denote as $e_n(t)$. We have (A.7):

$$e_n(t) = e_{opt}(u_n(t) - u_{n+1}(0)), \tag{A.7}$$

where e_{opt} is the optimal effort level expressed as a function of the incentive level. Given that the functions u_n are strictly increasing on $[0,T]$, equation (A.7) implies the optimal effort level increases with time for a given number of demerit points. From the definition of e_{opt}, the optimal effort is strictly increasing if it is greater than zero.

If the duration between demerit points n and $n+1$ is equal to t, the transition between the optimal effort levels is:

$$e_n(t) = e_{opt}(u_n(t) - u_{n+1}(0)) \rightarrow e_{n+1}(0) = e_{opt}(u_{n+1}(0) - u_{n+2}(0)).$$

If the last traffic violation immediately follows another one, the variation is positive because of the concavity of the sequence $(u_n(0))_{n=0,\ldots,N}$. Because u_n is increasing, the variation of optimal effort after a convicted traffic violation decreases with the duration between the two last offenses.

Figure 1 in the text gives an example where:

$$u_n(t) - u_{n+1}(0) > u_{n+1}(0) - u_{n+2}(0),$$

which implies a drop in the effort level and an increase in risk if the driver is convicted with a supplementary demerit point. This result is in contrast with what is expected with the redeeming mechanisms of the North American type.

Lastly, the optimal effort after a redemption of all the demerit points is equal to $e_0 = e_{opt}(u_0 - u_1(0))$. This value is lower or equal to any optimal effort level. Indeed, we have:

$$u_0 - u_1(0) \leq u_n(0) - u_{n+1}(0) \leq u_n(t) - u_{n+1}(0)$$

from the properties proved on the utility functions.

A.3: Continuity of effort in a point-record driver's license with a point removal of the North American type

The Bellman equation with a point-driver's license of the North-American type can be written as follows

$$d_u - ru(S) + \left(\frac{d}{dt}\left[u(S_t)\right]\right)_{t=0^+} = c_\lambda\left(u(S) - E\left[u\left(TR(S)\right)\right]\right). \tag{A.8}$$

The state variables S are the seniorities of each non-redeemed traffic offense (if any) and the related demerit points. The related lifetime utility is $u(S)$. The state S_t is reached from S with an eventless history (no traffic offense, point removal or contract birthday) of duration t. The expectation $E\left[u\left(TR(S)\right)\right]$ is the lifetime utility averaged with transition probabilities on the state(s) reached from S after a traffic offense.

Let us now consider a redeeming system of the North American type. The state variables are the seniorities of each non-redeemed traffic offense, if any. Let us denote these variables as:

$$S = (t_1, \dots, t_n),\ 0 \le t_1 < \dots < t_n < T.$$

The corresponding lifetime utility and optimal effort are denoted as $u(S)$ and $e(S)$. Then the states reached without traffic offense before the next redemption are:

$$S_t = (t_1 + t, \dots, t_n + t),\ 0 \le t < T - t_n.$$

We denote the state reached from S after an additional traffic offense (if $n < N$) as:

$$(0, t_1, \dots, t_n) = TR(S).$$

The Bellman equation on lifetime utility can be written as follows:

$$d_u - ru(S) + \left(\frac{d}{dt}\left[u(S_t)\right]\right)_{t=0^+} = c_\lambda\left(u(S) - u\left(TR(S)\right)\right).$$

Hence we have that:

$$e(S) = e_{opt}\left(u(S) - u\left(TR(S)\right)\right).$$

Results are similar to those derived for redeeming systems of the European type do not seem simple to obtain from the Bellman equation. For instance the monotonicity of the map $t \to e(S_t)$ seems questionable. The following result suggests that this map should be globally decreasing.

We now prove the continuity of optimal effort after a redemption in a system of the North American type. Because the lifetime utility is continuous after a redemption, we have the following result:

$$n \geq 1 : \lim_{t \to (T-t_n)^-} u(S_t) = u(S^R),\ S^R = (t_1 + T - t_n, \ldots, t_{n-1} + T - t_n).$$

The state S^R is reached from S if there is no traffic offense before the first redemption. Then it is easily seen that:

$$\lim_{t \to (T-t_n)^-} u\left[TR(S_t)\right] = u\left[TR(S^R)\right] = u(0, t_1 + T - t_n, \ldots, t_{n-1} + T - t_n).$$

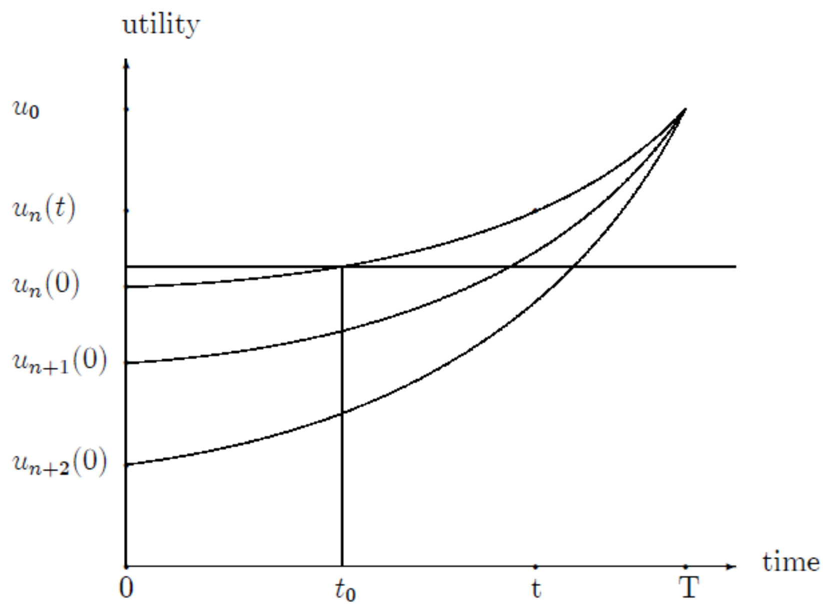
This means that the left continuity at $T - t_n$ of the map $t \to u(S_t)$ also holds for the map $t \to u\left[TR(S_t)\right]$, which is associated with the states reached after an additional traffic offense. The reason is that redemption of past offenses occurs regardless of the future individual history.

From the three last equations, we obtain:

$$\lim_{t \to (T-t_n)^-} e(S_t) = e(S^R)$$

and the continuity property of the optimal effort level. Because we expect a global increasing link between optimal effort and the accumulated demerit points, the time-effect should globally be decreasing to fulfill this continuity property.

FIGURE 1.—UTILITY FUNCTIONS AND VARIATION OF THE OPTIMAL EFFORT LEVEL.

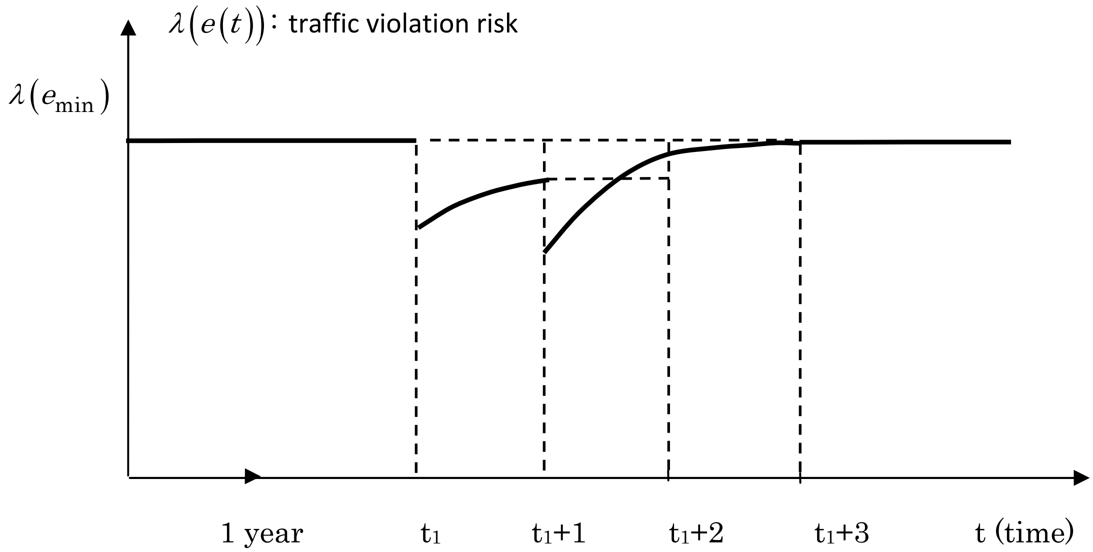


Optimal effort decreases after a traffic violation in the state (n, t)
if and only if

$$u_n(t) - u_{n+1}(0) > u_{n+1}(0) - u_{n+2}(0) \Longleftrightarrow t > t_0.$$

EXPLANATION: (t) is the lifetime utility in the state (n,t) $(n$ is the number of accumulated demerit points and t is the seniority of the last traffic violation, with $t < T$). All traffic violations are removed if the seniority of the last one reaches the threshold T. The state reached from (n,t) with a traffic offence is $(n+1,0)$, and the corresponding utility is $u_{n+1}(0)$. The incentive level at the state (n,t) is equal to $u_n(t) - u_{n+1}(0)$, which is the difference between the current utility and that reached after a traffic violation. The state reached from $(n+1,0)$ with a traffic offence is $(n+2,0)$, and the incentive level at the state $(n+1,0)$ is equal to $u_{n+1}(0) - u_{n+2}(0)$.

FIGURE 2.—TIME EVOLUTION OF TRAFFIC VIOLATION RISK RELATED TO OPTIMAL EFFORT. EXAMPLE WITH TWO TRAFFIC OFFENSES



EXPLANATION: In this figure, the function $t \to \lambda(e(t))$ is represented by the thick line. In this example, the first traffic violation occurs at t_1, the second at t_1+1. The two demerit points are removed at t_1+2 and t_1+3. The optimal effort level is denoted as $e(t)$. It is determined by all the seniorities of non-redeemed traffic violations, if any. The minimum effort level is denoted as e_{min} and may be greater than zero, depending on the individual characteristics of the driver. After each traffic violation, there is an upward jump in the optimal effort level and an implied drop in traffic violation risk if incentives are effective. The continuous time effect of incentives counters the event-driven effect (i.e. $e(t)$ decreases with t and $\lambda(e(t))$ increases). Before t_1, the effort level is minimal. It increases after the first traffic violation. The drop in traffic violation risk is the opposite of the unobserved heterogeneity effect. The effort is maximum after the second offense, and then we have $e(t_1+2) = e(t_1+1)$ at the time of the first point removal (one non-redeemed traffic violation with a one-year seniority in both cases). We also have $e(t_1+3) = e_{\min}$ at the time of the second removal.