

INCENTIVE MECHANISMS FOR SAFE DRIVING:
A COMPARATIVE ANALYSIS WITH DYNAMIC DATA

Jean PINQUET
Georges DIONNE
Charles VANASSE
Mathieu MAURICE

September 2009

Cahier n° 2009-27

DEPARTEMENT D'ECONOMIE

Route de Saclay

91128 PALAISEAU CEDEX

(33) 1 69333033

<http://www.enseignement.polytechnique.fr/economie/>
<mailto:chantal.poujouly@polytechnique.edu>

INCENTIVE MECHANISMS FOR SAFE DRIVING: A COMPARATIVE ANALYSIS WITH DYNAMIC DATA

Jean PINQUET¹
Georges DIONNE²
Charles VANASSE³
Mathieu MAURICE⁴

September 2009

Cahier n° 2009-27

Résumé: Les politiques de sécurité routière utilisent souvent des mécanismes incitatifs basés sur les infractions pour améliorer le comportement des conducteurs. Ces mécanismes sont par exemple des amendes, des primes d'assurance ou des permis à points. Nous analysons l'efficacité incitative de ces mécanismes. Nous obtenons leurs propriétés théoriques par rapport au nombre de points associés aux infractions et par rapport au temps contrat. Ces propriétés sont ensuite testées empiriquement avec des données issues du système public d'assurance au Québec. Nous concluons à la présence d'aléa moral dans les données, qui traduit le fait que les conducteurs qui accumulent les points deviennent plus prudents car ils sont plus sous risque de perdre leur permis. Par ailleurs, la prime indicée sur les points introduite en 1992 a réduit de 15% la fréquence d'infractions. Nous utilisons ce résultat pour calculer des équivalents monétaires pour les infractions et les retraits de permis.

Abstract: Road safety policies often use incentive mechanisms based on traffic violations to promote safe driving. Examples of mechanisms are fines, experience rating and point-record driving licenses. We analyse the effectiveness of these mechanisms in promoting safe driving. We derive their theoretical properties with respect to contract time and accumulated demerit points. These properties are tested empirically with data from the Quebec public insurance plan. We find evidence of moral hazard, which means that drivers who accumulate demerit points become more careful because they are at threat of losing their license. The insurance rating scheme introduced in 1992 reduced the frequency of traffic violations by 15%. We use this result to derive monetary equivalents for traffic violations and license suspensions.

Mots clés : Mécanismes incitatifs, permis à points, sécurité routière

Key Words : Point-record mechanisms, incentive effects, road safety

Classification JEL: D81, C23

¹ Université Paris 10 et Ecole Polytechnique (courriel: jean.pinquet@polytechnique.edu)

² HEC Montréal (courriel: dionne@hec.ca)

³ TD Asset Management

⁴ HEC Montréal

1 Introduction

Since the 1970s fatality rates due to road-traffic accidents have decreased steadily in developed countries, although risk exposure increased concomitantly (see OECD, 2005). For example, over the last ten years, the road fatality rate decreased by forty percent in France. However, the implied social cost of road accidents remains high (Doyle, 2005). A major reason for the improvement of the situation in the OECD has been the development of incentives for safe driving. Experience rating schemes used by the insurance industry have incentive properties (see Boyer and Dionne, 1989; 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. Point removal clauses are added so that this penalty can be avoided in the long run.¹ A point-record driver's license was implemented in Quebec in 1978, together with a no-fault insurance plan for bodily injuries which replaced a tort system.² The road fatality rate decreased by fifty percent during the fifteen years that followed.³

In Quebec, the *Société de l'Assurance Automobile du Québec* (referred to as SAAQ below) is a public monopoly that provides coverage for bodily injury. The SAAQ is also in charge of accident prevention and control, including the management of driver's licenses. Before 1992, the rating structure for bodily injury insurance was completely flat. Since December 1, 1992, the public authorities in Quebec have implemented an experience rating scheme based on accumulated demerit points. This mechanism was added to other incentives, i.e., fines, the point-record driver's license in force since 1978, and the private sector insurance rating for coverages other than bodily injury.

This paper analyzes the incentive properties of fines, point-record driver's licenses and experience rating based on traffic violations. Studies on incentive

mechanisms for road safety have appeared in the economic literature for many years (Peltzman, 1975; Landes, 1982; Boyer and Dionne, 1987). In the presence of asymmetric information, insurers use partial insurance or experience rating to improve resource allocation. Both schemes have proved to be efficient for handling moral hazard and adverse selection. Empirical tests have measured the effectiveness of such mechanisms for road safety (Sloan et al, 1995; Boyer and Dionne, 1989) and the presence of residual asymmetric information problems in insurers' portfolios (Chiappori and Salanié, 2000; Dionne, Gouriéroux and Vanasse, 2001). More recently, Abbring, Chiappori and Pinquet (2003) designed a new test based on the dynamics of insurance contracts to detect the presence of 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 car insurance market. The convex structure of the French "bonus-malus" system is used to show that the optimal effort level exerted by a rational policyholder increases after a claim at fault. In our study, insurance pricing is not the major incentive scheme but rather a measure used to complement fines and the point-record driver's license. Moreover, the pricing scheme of the Quebec public automobile insurance is an increasing step function of past demerit points.

Insurance pricing may not suffice as a 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 licenses provide incentives for road safety among normal drivers (those who respond to the usual incentive schemes) when the judicial system or the insurance market fail to provide optimal incentives. Point-record driver's licenses also allow the government to incapacitate reckless drivers because fines for traffic violations are bounded above in many jurisdictions. Bounded fines exist for different reasons: 1) many offenders are judgment proof and are unable to pay optimal fines; 2) many drivers are expected to escape from paying tickets

issued by the authorities when fines are very high; 3) society thinks it is unfair that rich and reckless drivers will pay high fines and continue to drive dangerously (Shavell, 1987a, 1987b). However, fines do reinforce the effectiveness of the point record mechanism by providing normal drivers with more incentives. In Bourgeon and Picard's model, which uses only two levels of prevention, the optimal fine must be set at the maximal level and must be neither progressive nor regressive. These authors also discuss the optimality of point removal mechanisms as a screening device. Public intervention can be justified when there is a significant difference between the private and the social cost of human lives (Viscusi, 1993). Finally, drivers may be unaware of their own accident or infraction probabilities or may misunderstand some features of the incentive environment.

We present the data base in Section 2 as well as our first empirical results related to the introduction of the new pricing policy implemented in 1992. The point-record mechanisms (driver's license suspensions and insurance pricing) are described in Section 3 and their incentive properties are investigated in an optimal behavior model in which time and effort are continuous. If incentives are caused by fines and by the point-record driver's license, we show that the optimal effort level increases globally with the number of demerit points accumulated and decreases with the seniority of non-redeemed traffic violations, if any. Traffic violation risk varies conversely, as it decreases with the effort level.

These results are compared with the data in Section 4. The observed dynamics on the drivers result from the incentive effects and from the revelation with time of hidden features in risk distributions, i.e. an unobserved heterogeneity effect. Let us compare these two effects at the time and event (accident or traffic violation) level of the data dynamics. Drivers with more traffic violations committed during a given period are riskier with respect to hidden features in risk distributions. Hence unobserved heterogeneity entails a risk reassessment after each event, and the event effect of risk revelation counters the corresponding incentive effect of the point-record driver's license. The revelation effect on

the risk level of a period without traffic violations is negative.⁴ The time effect of unobserved heterogeneity is also converse to that of the incentive effect, which raises an identification issue in the interpretation of the observed dynamics on the drivers. Abbring et al's (2003) test for moral hazard can be used if there are no time effects in the incentives. Because the time effects are important in Quebec's point-record system, we use another approach. We test for an increasing link between traffic violation risk and the number of demerit points. Rejecting this assumption amounts to finding evidence of moral hazard (*ab absurdo*, because of the increasing event effect created by unobserved heterogeneity). In Section 4.1, the incentives created by the threat of driver's license suspension are found to increase with accumulated demerit points. These findings confirm the theoretical analysis.

The insurance rating scheme introduced in 1992 reduced the frequency of traffic violations by 15%. In Section 4.2, we link the incentive levels of the three point-record mechanisms as derived from the theoretical analysis of Section 3 and the observed efficiency of these mechanisms. Monetary equivalents for traffic violations and license suspensions are derived from this analysis. Conclusions are drawn in Section 5 and technicalities are relegated to an appendix.

2 Presentation of the data base and preliminary empirical results

Our data base represents roughly one percent of the SAAQ portfolio. The panel covers the period from January 1, 1983 to December 31, 1996. A first sample of 40,000 license holders was selected at random at the beginning of 1983. About 300 young drivers were added each following year.⁵ The attrition rate per year is close to 1.5%, which is very low compared with the private sector. Due to the monopolistic status of the SAAQ, leaving the company means that drivers

lose their license, which explains the low attrition rate. The endogenous attrition is not very high. It was estimated from a bivariate probit model on traffic offenses and departures from the sample. A score test for the nullity of the correlation coefficient between the two equations⁶ was performed with the regression components set used in Section 4. The null hypothesis was not rejected at a five percent significance level. Hence the attrition risk adds no significant information to the assessment of traffic violation risk.

The personal characteristics of each driver are available on the driver's license for the current period. These characteristics are used as regression components in the empirical study. Several types of events are recorded in the database; they are listed below with related variables in addition to the date. 1) Accidents that have led to a police report. Only those with bodily injury are compensated by the SAAQ. 2) Convicted violations of the Highway Safety Code, together with the number of demerit points used in the point-record mechanisms. The number of demerit points is based on the severity of the traffic violation. Their distribution is given in Section 3.4. 3) Driver's license suspensions, which are spells rather than events, and 4) Premium payments, which since the 1992 reform are related to accumulated demerit points. These payments are made every two years on the policyholder's birthday.

Between 1985 and 1996, the average yearly frequencies of accidents with bodily injuries, accidents of all types (not including joint reports filed with private insurers) and traffic violations are equal to 1.4%, 6.7% and 16.9% respectively. Figure 1 represents the relative frequencies derived from a one year centered moving average.⁷

Insert Figure 1 about here

There is an overall decline in the frequency of accidents, whereas the frequency of traffic violations remains more stationary. This may seem surprising, but it is explained by the evolution of the traffic control environment. For instance, the number of traffic control devices such as radars increased during the

1980s and 1990s. An increase in the rate of traffic offenses recorded by devices or police officers among those committed explains this relationship. Figure 1 shows evidence of several periods where the frequency of traffic violations increased along with opposite variations in the frequencies of accidents. A step-up in traffic control during these periods may well explain such observations.

A traffic violation committed by a driver must be selected twice in order to be filed with demerit points. It must first be recorded by a control device or a police officer. We already mentioned that the related selection rate had previously increased. Second, the offender must be convicted of the recorded traffic violation. The filing of a traffic violation is somewhat discretionary. Since the 1992 reform for instance, drivers are being forced to pay more in premiums given demerit points, and we might expect policemen to be more hesitant to hand them out, and to give warnings instead.⁸ The conviction rate is less likely to vary with time than the recording rate.

In Figure 1, a downturn is also observed in the frequency of traffic violations just before the date (December 1, 1992) of the reform that introduced the experience rating structure based on demerit points. The reform was announced in August 1, 1992, which may explain why the downturn starts slightly before December 1, 1992. The experience-rated premium was first applied at the contract anniversary following December 1, 1992 and used the complete two year history of demerit points. Hence the new incentives for safe driving took effect for most of the drivers at August 1, 1992 (i.e. drivers with their next birthday after December 1, 1992).⁹ On average, the annual frequency of traffic violations was equal to 17.6% before the reform and 15.4% afterwards, which corresponds to a 12.5% decrease. The 1992 reform can be interpreted as a laboratory experiment to test whether experience rating based on demerit points reduces traffic violations. Nonetheless, the lower rate of traffic violations following the 1992 Quebec reform may be due to the changes in other factors that influence driver behavior. Identifying the influence of these factors necessitates a control group

that is not affected by the policy change. Unfortunately, we do not have access to such a control group because the insurer is a monopoly and bodily injury insurance is compulsory.¹⁰ In Section 4.2, we shall link the average decrease in the frequency of traffic violations before and after the 1992 reform to the overall effectiveness of incentive schemes.

Monetary and non-monetary incentives for safe driving are based on traffic violations as well as the optimal behavior models presented in Section 3. However the actual social cost of road traffic is caused by accidents. To reconcile these two approaches, two results are worth mentioning. First, demerit points are good predictors of accidents. This is well documented in the literature and is confirmed on our data in Section 4.1. Second, the global stationarity of convicted traffic violation frequency observed in Figure 1 concurs with a probable decrease in the frequency of committed traffic violations (see the aforementioned developments on selection rates). Lowering traffic violation risk through point-record mechanisms should also lower accident risk and the related social cost.

Finally, Figure 1 shows that accidents with bodily injuries evolve in much the same way as those recorded in the SAAQ file. We include accidents of all types in the empirical analysis in order to obtain more stable results.¹¹

3 Incentive effects of point-record mechanisms

3.1 Point-record mechanisms in Quebec

In this section, we describe Quebec's point-record mechanisms which are based on traffic violations, both monetary (insurance premiums) and non-monetary (point-record driver's license). Comparisons are made with respect to the mechanisms used by other countries.¹² We investigate the incentive properties of point-record mechanisms in Sections 3.2 to 3.4.

In many countries, driver's license suspensions are based on demerit points.

In Quebec, demerit points are assigned to convictions for traffic offenses, and their number depends on the traffic violation severity. When the accumulated number of demerit points reaches or exceeds a given threshold, the driver's license is suspended. Before January 1990 this threshold was set at twelve in Quebec and has since been increased to fifteen.

In order to mitigate the social cost of license suspensions, point removal systems exist for most real-world point-record driver's licenses. In Quebec, 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 of point removal systems used by American states follow the same approach. The average number of demerit points per convicted offense is 2.4 in Quebec. 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 also an incapacitating device for risky and reckless drivers through the license suspensions. Another point removal system consists in cancelling all the demerit points after a given period of violation-free driving. This mechanism was recently implemented in Spain, with a two year seniority. The French driver's license uses the two point removal systems.

The experience rating structure introduced by the SAAQ in December 1, 1992 links each premium paid every two years to the demerit points accumulated over the previous two years. The rating structure is given in Section 3.4. Once the premium is paid, the driver is reinstated with a fresh zero point record. Thus the length of the record relevant to the derivation of optimal behavior never exceeds two years.

3.2 Basic model for a point-record driver's license without point removal

Bourgeon and Picard (2007) analyze the incentive effects of point-record driver's licenses. Their model uses a binary effort variable. We extend their approach with a continuous effort level. Hence the effectiveness of effort may also be a continuous function of contract time, a desirable property for empirical validation. We show that under fairly general conditions, a rational policyholder's safe driving effort will increase with the number of demerit points accumulated.

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 section. 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.¹³ The duration D may be fixed or random in the model. In Quebec, a license suspension is of random length because drivers must pass a new exam after a given period before recovering their license.¹⁴ 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 $e(t)$.¹⁵ 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 section, we suppose 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 on the expected utility leads to

$$u_n = \frac{d_u}{r} - \frac{c_\lambda(u_n - u_{n+1})}{r}, \quad (0 \leq n < N), \quad (1)$$

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

$$c_\lambda(\Delta u) \stackrel{def}{=} \min_{e \geq 0} e + [\lambda(e) \times \Delta u]. \quad (2)$$

Technical details can be found in Appendix A.1. From equation (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. Values of this type are derived in Section 4.2. The function c_λ minimizes the sum $e + [\lambda(e) \times \Delta u]$, 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 (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 that

$$u_N = \beta u_0, \quad \beta = E[\exp(-rD)]. \quad (3)$$

The utilities are then derived from the recurrence equations (1) and (3). Optimal effort depends on the variation of lifetime utility Δu because it minimizes the function defined in equation (2). The variable Δu (the argument of c_λ in equation (2)) which determines optimal effort will be referred to later as the 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 shown in Appendix A.1 that e_n increases with n for any given value of N . The related

frequency of 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).¹⁶

Fines represent another monetary incentive scheme applied in Quebec throughout the period investigated in this study. Let us denote \overline{fa} as the average fine for a traffic violation conviction. Given that fines and premiums are low in comparison to average wealth, risk aversion is not significant. If fines are combined with the preceding point-record driver's license, the incentive level is equal to $\overline{fa} + u_n - u_{n+1}$. This means that the average fine is added to the utility loss in the variable which determines optimal effort. Besides, the optimal effort still increases with n for a given value of the average fine.

3.3 Point-record driver's licenses with point removal

In Quebec, each traffic violation is redeemed at the end of a two-year period. Integrating this feature in the optimal behavior model is difficult, 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 Quebec, 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 should thus decrease with time. Appendix A.2 provides proof that optimal effort is continuous at the time of a point removal. Optimal effort is then expected to increase with each traffic violation in order 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.

Insert Figure 2 about here

Figure 2 illustrates the changes in the optimal traffic violation risk during a period that includes two violations and their point removal. The continuous time effect of incentives balances the downward jump in risk after each violation. The incentive effects at the time and event level are the opposite of those created by the revelation of unobserved heterogeneity. The incentive properties of the point removal system in force in Quebec are similar to those of a mechanism without point removal as regards the number of demerit points accumulated. In Section 4, we will use the model of Section 3.2 as a proxy for the incentive environment in Quebec.

3.4 Incentive effects of premiums indexed on demerit points: The example of Quebec

Table 1 presents the rating structure enforced for each driver's license on the first contract anniversary following December 1, 1992. The premium paid every two years after this date depends on the number of demerit points accumulated in the last two years. It does not represent the total premium for bodily injury insurance but rather the additional premium related to demerit points. This average premium is equal to \$54.60, and complements a yearly driver's license fee for insurance coverage equal to \$107.

Insert Table 1 about here

In this section, the incentive properties of this rating structure are analyzed separately from the point-record driver's license. An important input is the distribution of demerit points for a given driving offense, which we left out in Section 3.2. Denoting f_j as the proportion of traffic violations with j demerit points, we have the following values

$$f_1 = 4.71\%; f_2 = 52.32\%; f_3 = 38.34\%; f_4 = 2.83\%; f_5 = 1.80\%. \quad (4)$$

Note that f_5 actually refers to offenses with five points and more. Table 1 shows that the premium is a step function of the accumulated demerit points. Because of the local non-convexity of the premium, the incentives may not always increase with the number of demerit points accumulated. Consider, for instance, a policyholder just before her contract anniversary. The incentive level will be stronger with two accumulated demerit points than with four. With four points, it is indeed less than likely that the next traffic offense will trigger an increase in premium. The corresponding probability is $2.83 + 1.80 = 4.63\%$, if we assume that the distribution of the f_j is independent of the accumulated demerit points. The incentives for safe driving are stronger at a two point level because the probability of climbing a step in the rating structure after a traffic offense is close to one. The aforementioned result is in contrast to that obtained by Abbring et al (2003) for the French "bonus-malus" scheme and its exponential structure. Hence, step pricing schemes may have poor incentive properties. They are employed because they are simple.

Let us design an optimal behavior model based on this rating structure. Once the premium is paid, the driver is reinstated with a fresh zero point record. Hence the optimal control model can be designed with the next contract anniversary as the horizon. Let π_n be the premium paid for n demerit points accumulated during a period of $T = 2$ years between two contract anniversaries. As we disregard the point-record driver's license and its possible deprivation, the driving utility is no longer a parameter. However, we retain fines in the in-

centives for safe driving. We denote $v_n(t)$ as the expected disutility of premiums and fines paid until the next contract anniversary, where t is the seniority of the last anniversary and n is the number of demerit points accumulated since that date. We have the terminal conditions

$$v_n(T) = \pi_n, \forall n = 0, \dots, N. \quad (5)$$

If incentives are related to fines and insurance premiums, the incentive level is the sum of the average fine \overline{fa} and of the expected variation of $v_n(t)$ after a traffic offense (see Appendix A.3). Hence optimal effort is determined by $\overline{fa} + \Delta v_n(t)$, with $\Delta v_n(t) = \left(\sum_{j / f_j > 0} f_j v_{\min(n+j, N)}(t) \right) - v_n(t)$. Derivations show that the average of $\Delta v_n(t)$ with respect to n does not vary much with time. The terminal values of Δv_n are derived from equations (4), (5) and Table 1. For $n = 0, 2, 4$, we obtain

$$\Delta v_0(T) = \$2.32; \Delta v_2(T) = \$47.65; \Delta v_4(T) = \$3.43. \quad (6)$$

The incentives with two points accumulated are much stronger than with four points, which confirms the analysis following equation (4). The overall average of $\Delta v_n(t)$ with respect to t and n is close to \$12, a 9% increase for an \$130 average fine. In Section 4.2, this increase will be compared with the variation in traffic violation frequency before and after the reform.

4 Empirical results on the incentive effects of point-record mechanisms

4.1 Point-record driver's license

In this section, we analyze the data before the 1992 reform that introduced the experience rating scheme based on demerit points. Thus the point-record driver's license interacts only with fines. Regressions are performed from January

1985 (we need a two year history to derive the accumulated demerit points) to December 1992, the date of the reform enforcement. We try to confirm the theoretical findings of Sections 3.2 and 3.3 (i.e., the effort level increases globally with the number of demerit points accumulated and decreases with the seniority of non-redeemed traffic violations, if any), and to find evidence of moral hazard in the data.

We estimate the hazard functions of convicted traffic offenses and accidents with a proportional hazards model (Cox, 1972). We retained the following specification

$$\lambda_i^j(t) = \exp((x_i(t)\beta_j) + g_j(adp_i(t))) \times h_j^{S_i(t)}(c_i(t)). \quad (7)$$

In equation (7), $\lambda_i^j(t)$ is the hazard function of type j ($j = 1$: traffic violation or $j = 2$: accident) for driver i at calendar time t . Regression components are denoted by the line-vector $x_i(t)$. We retained the gender, driver's license class, place of residence, age of the driver, calendar effects related to years and months, and the number of past license suspensions.¹⁷ The number of demerit points accumulated in the last two years is denoted as $adp_i(t)$. Let us comment the expected shape of g_1 , which links traffic violation risk to accumulated demerit points. The revelation effect of unobserved heterogeneity is increasing in adp , and is converse to the incentive effect. In the absence of moral hazard, g_1 should then be increasing. We will test the null hypothesis that g_1 is weakly increasing against the alternative that it is not (i.e. g_1 strictly decreases for at least some values). Rejecting the null amounts to finding evidence of moral hazard.¹⁸

Effort is expected to decrease with time only if the number of demerit points accumulated is greater than zero. Hence we specified a stratified proportional hazards model.¹⁹ The baseline hazard functions $h_j^{S_i(t)}$ depend on the risk type j and on the stratum $S_i(t)$. There are two strata, depending on whether the variable $adp_i(t)$ is equal to zero or not. Lastly, contract time $c_i(t)$ is integrated into the baseline hazard function h_j . The function c_i is set at zero at the beginning of the whole period. Then it is reset to zero at each event that trig-

gers a variation of the accumulated demerit points (i.e., traffic violation or point removal). This event-driven operation should eliminate interactions between calendar and contract-time effects for the stratum associated with $adp > 0$.

Insert Table 2 about here

Table 2 shows that \hat{g}_1 is increasing until adp reaches the seven point threshold, at which point the increasing property is no longer fulfilled, and \hat{g}_1 reaches a maximum at seven points. It is worth mentioning that the SAAQ warns the policyholders when their accumulated demerit points increase beyond a seven point threshold.²⁰ When the accumulated demerit points are less than or equal to seven, the unobserved heterogeneity effects outweigh the incentive effects because \hat{g}_1 is increasing. Controlling for other covariates, a driver with seven accumulated demerit points is $\exp(0.974) = 2.65$ times riskier than a driver with a violation-free record. Hence unobserved heterogeneity effects strongly outweigh incentives to drive carefully between zero and seven points. This result is reversed beyond seven points. For instance, a driver with seven accumulated demerit points that gets a supplementary two point offense has a traffic violation risk that is 19% lower. As the unobserved heterogeneity effect is increasing with adp , the estimated risk reduction due to the incentive effect is greater than 19%. A 20% value for the risk reduction after a traffic offense will be retained in the next section, in order to derive monetary equivalents for traffic violations and license suspensions. Table 2 presents tests where the null assumption is $g_1(7) \leq g_1(adp)$, for $adp > 7$. The null is rejected for these values of adp at significance levels greater than 1.3%, except for $adp = 11$. We also performed a global test on the weakly increasing shape of g_1 . The test statistic is a squared distance between the unconstrained estimation of g_1 given in Table 2, and the set of weakly increasing functions of adp (see Appendix A.4 for more details). A graphical illustration is given in Figure 3. We reject the increasing shape assumption for every usual significance level. As an absence of incentive effects

entails an increasing shape for g_1 (unobserved heterogeneity effect), rejecting the null shows evidence of moral hazard.

Insert Figure 3 about here

Table 2 shows less evidence of moral hazard on accident risk than on traffic violation risk, as the estimation of g_2 increases until $adp = 10$. A possible interpretation is that we cannot separate at fault from no-fault accidents. In the literature, the incentive effect is usually higher with at fault accidents. Besides, drivers nearing the license suspension threshold might also apply opportunistic strategies regarding traffic violations (e.g. paying more attention to radars) without otherwise modifying their attitude towards road traffic risk.

The estimated hazard functions from the traffic violation equation globally decrease with time, including the one related to the stratum $adp > 0$.²¹ This means that the continuous time effect created by the revelation of unobserved heterogeneity outweighs the incentive effects (see Figure 2 and the discussion that follows). This is not surprising, given that the average increase of g_1 in Table 2 (a vast majority of drivers have less than seven demerit points) must be balanced by a negative time effect in order to obtain stationary risk levels on average.

4.2 Incentive effects of the 1992 reform and monetary equivalents for traffic violations and license suspensions

In this section, we assess the efficiency of the 1992 reform. We compare its incentive level to those of fines and of the point-record driver's license. We relate the efficiency of the reform to its relative weight in the three incentive mechanisms, and we infer results on the shape of the link between the cost and efficiency of effort. Lastly, we derive monetary equivalents for traffic violations

and license suspensions, from the estimated efficiency of incentives created by a supplementary traffic violation.

In Section 2, we mentioned a 12.5% decrease in the average frequency of traffic violations before and after the reform which introduced the experience rating structure based on demerit points. This result changes slightly if we control with the regression components used in Table 2. A regression estimated from 1985 to 1996 with the covariates of Section 4.1 and a dummy related to the period following December 1, 1992 associates the reform with a 15% decrease.²² The results of Section 3.4 (for instance equation (6)) suggest that the number of demerit points accumulated since the last anniversary should influence the effectiveness of the 1992 reform. However, we did not obtain significant results in this direction. The drivers' limited knowledge of the environment could explain this poor result, a point developed later.

We perform an overall comparison of the three incentive schemes. We use the model without point removal of Section 3.2 to analyze the incentives for drivers in Quebec. Before the 1992 reform, fines were supplemented by a point-record driver's license. Optimal effort after n non-redeemed traffic violations depends on the incentive level, i.e. the argument of the function c_λ defined in equation (2). This level is equal to $\overline{fa} + u_n - u_{n+1}$ (see Section 3.2). The average fine \overline{fa} is equal to \$130, and the 1992 reform entails an average increase in the incentive level equal to \$12 from Section 3.4. At this point, it seems interesting to relate optimal frequency risk and the incentive level. This relation can be assessed from the elasticity between the former and the latter variable. When the incentives are effective, it can be shown that this elasticity is less than -1 if and only if $\log(\lambda)$ is a concave function of effort (elasticity and concavity are considered locally: see Appendix A.5 for a proof). A global elasticity equal to -1 is linked to an exponential decay of λ . With $\lambda(e) = \lambda(0) \times \exp(-\alpha e)$, the optimal risk level as a function of $\overline{fa} + \Delta u$ (the incentive level) is equal to $1/(\alpha \times (\overline{fa} + \Delta u))$ if the incentives are effective.

We now apply this result to the 1992 reform. As the reform entailed a significant reduction in traffic violation risk regardless of the number of demerit points accumulated, we can assume that incentives are effective for a representative driver.²³ This leads us to analyze the elasticity between traffic violation risk and the incentive level. Suppose that we leave out the modifications of incentive levels due to the aggregation of incentive mechanisms. Then we can relate:

- On the one hand, a 15% reduction in the frequency of traffic violations since the 1992 reform.
- On the other hand, a relative increase in the incentive level ranging between 9% and 10%. Indeed, the 1992 reform entails a \$12 average increase in the incentive level. This increase supplements the other components of the incentives, i.e., the \$130 average fine and the utility variation for the point-record driver's license. In Table 2, the point-record driver's license offers significant incentives beyond a seven point threshold, a result corresponding to only a minority of drivers (1.4%). The contribution of the point-record driver's license to the incentives is low as compared with that of fines.

This suggests that the elasticity between the optimal frequency risk and the incentive level is less than -1 in this case. This result is linked to a locally concave shape of $\log(\lambda)$ for the representative driver. However external effects could also explain the reduction in the frequency of traffic violations. We cannot eliminate these effects because there is no control group that is not affected by the reform. Besides, the elasticity would be modified if the distribution of demerit points for a given driving offense (see equation (4)) was wrongly perceived by the drivers. The \$12 contribution of the reform to the incentive level is low, because of the high frequency of drivers without demerit points since the last birthday (87%), and the low incentive level of the reform for these drivers (see equation (6)). This

level depends largely on the probability of moving up a step in the premium schedule after an additional traffic violation, which must be associated with four demerit points or more. If the perceived frequency of corresponding traffic violations was greater than the actual one (i.e., $f_4 + f_5 = 2.83 + 1.80 = 4.63\%$), the variation of the incentive level induced by the reform would increase. In that case the elasticity would be closer to zero.

Lastly, let us assess monetary equivalents for a traffic violation and a license suspension. The monetary equivalent of a traffic violation is the component of the incentive level related to the point-record driver's license. It is equal to the loss of lifetime utility after a traffic violation, which depends on the number of traffic violations accumulated. A value can be derived from the effectiveness of effort estimated in Table 2 and from the aforementioned link between efficiency of effort and the incentive level. An additional traffic violation beyond seven accumulated demerit points entails a reduction of traffic violation frequency of close to 20%. Although these drivers cannot be seen as representative, we will apply the elasticity derived from the preceding developments. If the 9% increase in the incentive level induced by the 1992 reform entails a 15% reduction in the frequency of traffic violations, a 20% decrease in traffic violation frequency is associated with a 12% increase in the incentive level. The implied loss of lifetime utility depends on the traffic violation frequency risk λ but mostly on the discount rate r (see Appendix A.6). With $\lambda = 0.15$, the monetary equivalent of an additional traffic violation for these drivers belongs to the interval [$\$120$, $\$195$] if $r = 3\%$, and to [$\$41.1$, $\$55.7$] if $r = 6\%$. Besides, the growth rate of this monetary equivalent with respect to the number n of non-redeemed traffic violations falls between r/λ_n and r/λ_{n+1} , where λ_n is the optimal traffic violation frequency related to n . Monetary costs for license suspensions are then obtained by adding the costs of traffic violations until the crossing of the demerit point threshold. Starting from a zero-point record and assuming that six traffic violations are needed to trigger a license suspension, the monetary

cost of a license suspension is bounded by \$700 and \$1178 if $r = 3\%$, and if $\lambda_0 = 0.17$; $\lambda_1 = 0.17$; $\lambda_2 = 0.16$; $\lambda_3 = 0.15$; $\lambda_4 = 0.12$; $\lambda_5 = 0.09$.²⁴ A misperception of the environment could modify the monetary equivalents of traffic violations and license suspensions. In the discussion on the elasticity between the optimal frequency risk and the incentive level, we argued that an overestimation of the frequency of severe traffic violations would increase the perceived incentive effect of the 1992 reform. The \$12 average effect of the reform on the incentive level should then be upgraded, and the monetary equivalents of traffic violations and license suspensions should also be upgraded.

5 Conclusion

In this article, we analyze the properties of policies designed to promote safe driving. Three important incentive mechanisms for road safety are used in Quebec. When the Quebec government introduced a no-fault insurance regime in 1978, it implemented a point-record driving license to maintain incentives for road safety. Because prevention activities are not observable by the public insurer, the new mechanism uses drivers' accumulated demerit points to approximate past safety activities and to increase incentives to reduce infractions. Under moral hazard, drivers who accumulate demerit points become more careful because they are at threat of losing their license. This result was confirmed from the data. Otherwise, we would have accepted an increasing link assumption between accumulated demerit points and traffic violation risk.

Fines are on average the most efficient device, but they are associated to a uniform incentive level. We designed our incentive models with a representative driver, but there is of course heterogeneity in the individual parameters, such as the threshold beyond which the incentives are effective. We did not have wealth variables at hand, and an interesting empirical issue would have been to cross such variables with a reform dummy in risk assessment.

The experience-rated premium based on accumulated demerit points is a monetary point-record mechanism. The empirical results exhibit a rather uniform effectiveness after its enforcement in 1992, i.e. a 15% decrease in the frequency of traffic violations. Its incentive effects do not strictly increase with the accumulated demerit points, however, because of the steps in the rating structure. The actual incentive effect of the reform looks more like that induced by an increase in the average fine. The SAAQ modified its rating policy in 2008, with a premium increase from the first demerit point.

Regulations are not self-enforcing. To successfully achieve a reform, the regulator must implement enforcement activities and bear the costs thereof. In our application to road safety, the Police is the most important enforcer of the different incentive schemes. Police officers have a limited capacity to modify the penalties for a given infraction but may affect the detection probabilities based on the aggregate convictions cycles, the driver's accumulated demerit points or other tasks. When analyzing our results, we must consider the effect of these potential behaviors. For example, is it possible that when accumulated demerit points are greater in a given file, police officers have an incentive to issue them more or less often because of risk of the driver's losing his license.

We did not have precise information to control for these factors. It is well documented that the Police collaborates very well with the Quebec public insurer and Government. There is consequently a low probability that aggregate safety scores would have affected the Police behavior without the agreement of the Government, which is assumed to be fully committed to its road safety policies. Regarding the effect of accumulated demerit points, we do not believe there is a general policy in the Police force on the behavior to adopt based on drivers' past experience although some individuals may modify their behavior in one direction or the other. This effect is implicitly considered as a purely random effect in our analysis and should not have affected our results on the point record incentives.

In this study we have not examined the long-term evolution of accidents in detail, because we did not have access to the control variables of interest. In recent years, many road safety initiatives have had an impact on accidents but did not necessarily have any effect on violations. These initiatives include measures such as occasional campaigns to prevent fatal accidents, increased police patrols to reduce speeding and designated driver campaigns to prevent drinkers from getting behind the wheel. The decline in deaths and serious injuries can also be explained by vehicular improvements and the wearing of seat belts. All such measures are complementary to those studied in this article.

Notes

¹These clauses and their incentive properties are detailed in Section 3.

²The North American continent preceded Europe in the design of such systems. Point-record driver's licenses were introduced in 1947 in the USA. Germany, France, and Spain implemented these mechanisms in 1974, 1992 and 2005, respectively.

³We do not have information on factors different from the incentive schemes analyzed in this article that might have explained this decrease. This study is performed in a no-fault environment. For a recent comparison of strict liability and a negligence rule for risk-incentives trade-off, see Fagart and Fluet (2009).

⁴Increases in premium after an event and "no claims discount" clauses observed in non-life insurance can be explained by the revelation effect of unobserved heterogeneity.

⁵Selecting at random one percent of the new license holders every year would evidently have been a preferable sampling procedure. One thousand new license holders would then have been selected every year, as the entry rate in the SAAQ portfolio is close to 2.5%.

⁶Binary variables related to traffic offenses and attrition were created on a monthly basis, and explained with the covariates used in Section 4.1. The score test statistic is equal to 0.34. Hence we do not reject the nullity of the correlation coefficient at the usual significance levels.

⁷We begin in 1985 in order to match the regressions that follow, because a two-year history is needed to obtain the accumulated demerit points. Data are first averaged over one year, to account for strong seasonal effects. A centered moving average derived on five fortnights is then performed twice in order to reduce the volatility of the series.

⁸We thank a referee for suggesting this interpretation.

⁹The SAAQ had the information on the drivers' traffic violation history since 1978. Drivers with a contract anniversary falling between the announcement of the reform and its enforcement are not incited by the experience-rated premium before this anniversary. Incentives exist otherwise (for these drivers after their birthday, and for all the other drivers since the announcement of the reform). A referee suggested using this natural experiment in order to disentangle the incentive effects of the reform from calendar effects. We did not obtain significant results. Four months is a short period, and on average only one driver out of twelve was not incited by the rating scheme during the period.

¹⁰See Manning et al (1987) for the use of a control group in the assessment of a cost sharing modification in the health insurance market.

¹¹Important variables in the regressions such as the number of accumulated demerit points have low frequencies for the highest values. It is hard to make an accurate estimation if the frequency of events is low, as is the case for accidents with bodily injury.

¹²We focus on license suspensions in the comparisons. However financial penalties are also associated with demerit points in some American states.

¹³This reinstatement can be seen as a removal of demerit points. In the present paper, we consider a point removal mechanism to be a cancellation of demerit points applied before the suspension of the driver's license.

¹⁴The failure rate at the first exam after the license suspension is 25%.

¹⁵Safe driving effort can also reduce the expected disutility of accidents. If $e \rightarrow e - \delta(e)$ is the implied decrease in the disutility flow, replacing e by $e - \delta(e)$ in the model includes the influence of safe driving effort on accident disutility.

¹⁶We thank a referee for suggesting this connection. A "strike" is a conviction for a serious felony, the equivalent of a demerit point. Three strikes entail a prison sentence of 25 years to life, which in our context would be the license suspension. Comparing the time to rearrest for criminals with one or two strikes in California, the authors conclude that those with two strikes are less risky than those convicted twice, with one strike only. They thus observe an increase in the deterrence effect at the second strike.

¹⁷Comprehensive regressions based on two-year periods can be found in Dionne, Maurice, Pinquet, and Vanasse (2001).

¹⁸We thank a referee for suggesting this formulation of the test.

¹⁹Stratification in a proportional hazards model means that Cox likelihoods (of a multinomial logit type) are derived for each stratum and then multiplied together. In other words, an individual with an observed event is assumed to have competed only with other individuals

in the same stratum and at risk at the same date. However, the same coefficients for the covariates are used across all strata.

²⁰Drivers are not informed when offenses are redeemed.

²¹Results are available from the authors upon request.

²²We retained the covariates used in Table 2 except for dummies related to calendar effects and to the number of past license suspension spells. The estimated parameter for the reform dummy is equal to -0.163, and the related standard deviation is equal to 0.008. Hence the reform effect is conclusive at the usual tests significance levels.

²³From Section 3.2, a sufficient condition to have this result is that the average fine is higher than the threshold $-1/\lambda'(0)$ beyond which the incentives are effective.

²⁴These values correspond to an effort level that increases with the number of demerit points accumulated.

References

- [1] Abbring, Jaap, Pierre-André Chiappori, and Jean Pinquet, “Moral Hazard and Dynamic Insurance Data,” *Journal of the European Economic Association* 1 (2003), 767-820.
- [2] Boyer, Marcel and Georges Dionne, “The Economics of Road Safety,” *Transportation Research*, 21B (5) (1987), 413-431.
- [3] Boyer, Marcel and Georges Dionne, “An Empirical Analysis of Moral Hazard and Experience Rating,” *The Review of Economics and Statistics* 71 (1989), 128-134.
- [4] Bourgeon, Jean-Marc and Pierre Picard, “Point-Record Driving License and Road Safety: An Economic Approach,” *Journal of Public Economics* 91 (2007), 235-258.
- [5] Chiappori, Pierre-André and Bernard Salanié, “Testing for Asymmetric Information in Insurance Markets,” *Journal of Political Economy* 108 (2000), 56-78.

- hal-00414479, version 1 - 9 Sep 2009
- [6] Cox, David R., "Regression Models and Life Tables," *Journal of the Royal Statistical Society, Series B*, 34 (1972), 187-220.
 - [7] Dionne, Georges, Christian Gouriéroux, and Charles Vanasse, "Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment," *Journal of Political Economy* 109 (2001), 444-453.
 - [8] Dionne, Georges, Mathieu Maurice, Jean Pinquet, and Charles Vanasse, "The Role of Memory in Long-Term Contracting with Moral Hazard: Empirical Evidence in Automobile Insurance," Working paper 01-05, HEC Montréal (2001). URL: <http://neumann.hec.ca/gestiondesrisques/01-05.pdf>
 - [9] Doyle, Joseph J., "Health Insurance, Treatment and Outcomes: Using Auto Accidents as Health Shocks," *The Review of Economics and Statistics* 87, 2 (2005), 256-270.
 - [10] Fagart, Marie-Cécile, and Claude Fluét, "Liability Insurance under the Negligence Rule," *Rand Journal of Economics* 40 (2009), 486-509.
 - [11] Helland, Eric and Alexander Tabarrok, "Does Three Strikes Deter? A Nonparametric Estimation," *The Journal of Human Resources* 52, 2 (2007), 309-330.
 - [12] Landes, Elisabeth M. "Insurance, Liability, and Accidents: A Theoretical and Empirical Investigation of the Effect of No-fault Accidents," *Journal of Law and Economics* 25 (1982), 49-65.
 - [13] Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, Arleen Leibowitz, and Susan Marquis, "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment," *American Economic Review* 77 (1987), 251-277.

- hal-00414479, version 1 - 9 Sep 2009
- [14] OECD International Road Traffic and Accident Data Base (2005). URL: <http://cemt.org/IRTAD/>.
 - [15] Peltzman, Sam, "The Effects of Automobile Regulation," *Journal of Political Economy* 83 (1975), 677-725.
 - [16] Shavell, Steven, "The Optimal Use of Nonmonetary Sanctions as a Deterrent," *American Economic Review* 77 (1987a), 584- 592.
 - [17] Shavell, Steven, "Economic Analysis of Accident Law," (1987b) Harvard University Press, Cambridge.
 - [18] Sloan, Frank A., Bridget A. Reilly, and Christoph Schenzler, "Effects of Tort Liability and Insurance on Heavy Drinking and Driving," *Journal of Law and Economics*, 38 (1995), 49-77.
 - [19] Viscusi, Kip W., "The Value of Risks to Life and Health," *Journal of Economic Literature*, 31 (1993), 1912-1946.

TABLE 1.—SAAQ INSURANCE PREMIUMS FOR BODILY INJURY AS A FUNCTION OF ACCUMULATED DEMERIT POINTS SINCE THE LAST CONTRACT ANNIVERSARY

Accumulated demerit points (last two years)	Premium for the next two years (Canadian \$)	Frequency (%)
0,1,2,3	50	93.7
4,5,6,7	100	4.9
8,9,10,11	174	1.1
12,13,14	286	0.2
15 and more	398	0.1

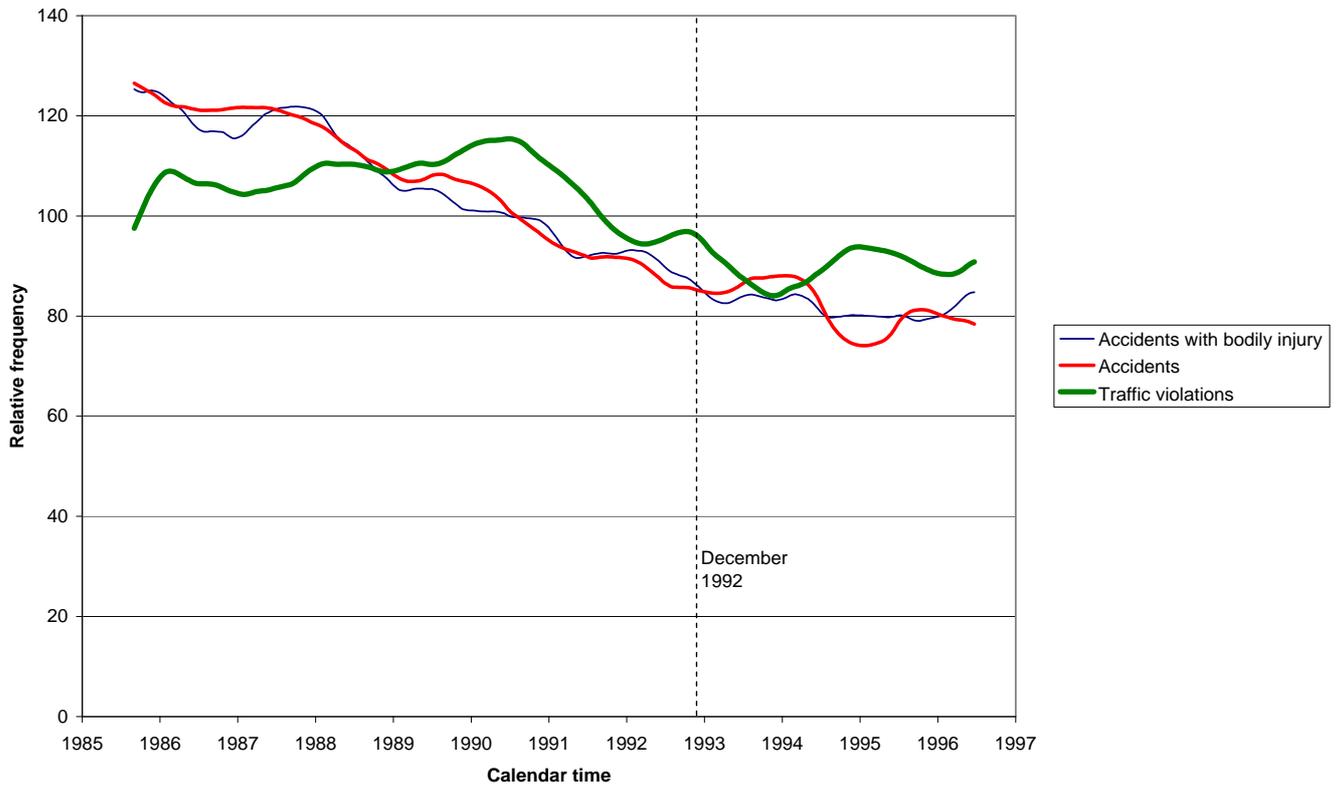
Notes: The premium is a step function of accumulated demerit points. The incentives for safe driving depend on the accumulated demerit points. For example, we document in the text that the incentive level is stronger with two accumulated demerit points than with four because the probability of climbing a step in the rating structure after a traffic violation is higher.

TABLE 2.—ESTIMATION OF THE HAZARD FUNCTION FOR TRAFFIC VIOLATION AND ACCIDENT FREQUENCY RISKS

Variable	Level	Frequency (%)	Traffic violation risk (g_1)	P-value (H0: $g_1(7) \leq g_1(adp)$, $adp > 7$)	Accident risk (g_2)
<i>adp</i> :	0 point (*)	76.60	0		0
Number of accumulated demerit points (last two years)	1 point	0.39	0.311		0.243
	2 points	9.36	0.500		0.328
	3 points	6.23	0.522		0.449
	4 points	1.92	0.794		0.496
	5 points	2.09	0.854		0.595
	6 points	1.25	0.873		0.647
	7 points	0.72	0.974		0.601
	8 points	0.55	0.876	0.0130	0.736
	9 points	0.43	0.764	< 0.0001	0.797
	10 points	0.32	0.786	0.0002	0.842
	11 points	0.06	0.906	0.2328	0.522
	12 points	0.04	0.378	< 0.0001	0.810
	13-14 points	0.04	0.571	0.0004	0.222

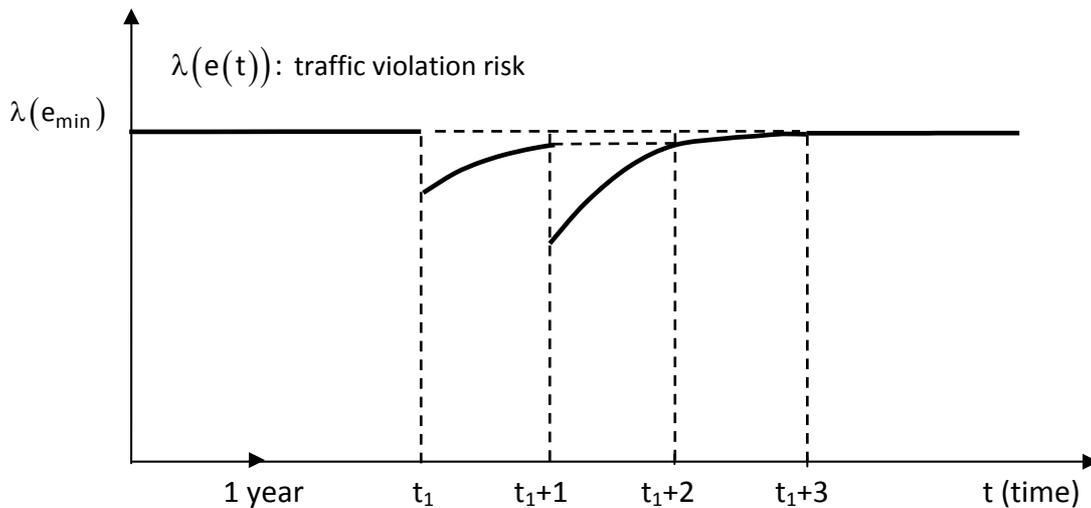
Notes: (*) is the reference level. Controlling for other variables, a driver with seven accumulated demerit points who gets a supplementary two point offense has a traffic violation risk reduced by $1 - \exp(0.764 - 0.974) = 19\%$. As the unobserved heterogeneity effect is increasing with *adp*, the estimated risk reduction due to the incentive effect is greater than 19%. The "p-values" refer to tests for the null assumption: $g_1(7) \leq g_1(adp)$ (with $adp > 7$) vs. the alternative, with one-sided rejection regions. For instance, the null assumption related to $adp = 8$ is rejected at the level α if α is greater than the "p-value" 1.3%, and accepted otherwise. The results suggest that g_1 is not increasing beyond $adp = 7$, which shows evidence of moral hazard. A global test for the increasing shape of g_1 is presented in Figure 3. Additional regression variables are: gender, driver's license class (9 levels), place of residence (16 levels), age of the driver (5 slopes), number of past suspension spells (3 levels), as well as calendar effects related to years (8 levels), and months (12 levels). Number of observations: 3,492,868 duration-event indicator pairs, derived from 41,290 driver's licenses. The durations are bounded above by one month.

FIGURE 1.—RELATIVE FREQUENCIES (IN PERCENTAGE) FOR TRAFFIC VIOLATIONS AND ACCIDENTS



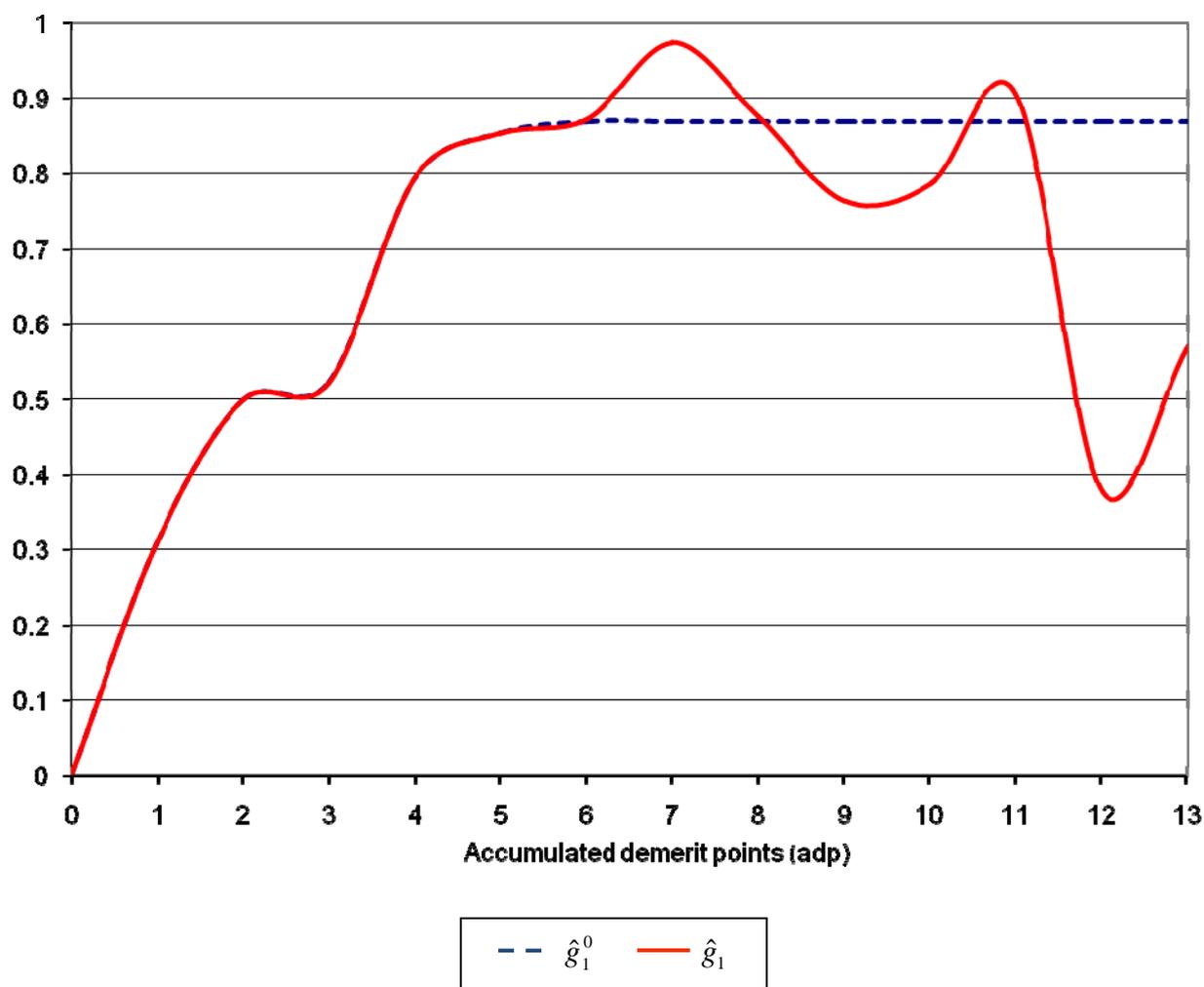
Notes: Figure 1 represents relative frequencies of traffic violations and accidents derived from a one-year centered moving average. The reform introducing the insurance experience rating structure based on demerit points took effect on December 1, 1992.

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



Notes: The function $t \rightarrow \lambda(e(t))$ 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.

FIGURE 3.—GLOBAL TEST FOR THE INCREASING SHAPE OF THE TRAFFIC VIOLATION RISK
AS A FUNCTION OF ACCUMULATED DEMERIT POINTS



Notes: This figure presents a graphical test where the null is the weakly increasing shape of g_1 as a function of adp , the number of accumulated demerit points. The test statistic is a squared distance between the unconstrained estimation of g_1 (thick line) given in Table 2, and the set of weakly increasing functions of adp . The metric is that of the Wald asymptotic test, and the weakly increasing function which is closest to g_1 is represented by the dotted line. The statistic is equal to 43.18 and its asymptotic distribution is less than that of a $\chi^2(13)$ in the sense of first order stochastic dominance. This leads us to reject the increasing shape assumption for every significance level greater than 42×10^{-6} . As an absence of incentive effects entails an increasing shape for g_1 (unobserved heterogeneity effect), rejecting the null shows evidence of moral hazard.

A Appendix

A.1 Incentive effects of point-record driver's licenses: Model without point removal

The Bellman equation on the expected utility is

$$u_n = \max_{e \geq 0} (d_u - e)dt + (\exp(-rdt) \times [((1 - \lambda(e)dt) \times u_n) + (\lambda(e)dt \times u_{n+1}) + o(dt)]).$$

We then obtain

$$0 = \max_{e \geq 0} (d_u - e) - (r + \lambda(e))u_n + \lambda(e)u_{n+1},$$

and equation (1).

We give the main properties of the function

$$c_\lambda(\Delta u) \stackrel{def}{=} \min_{e \geq 0} e + [\lambda(e) \times \Delta u] = \min_{e \geq 0} h(\Delta u, e),$$

with λ a positive, decreasing and strictly convex hazard function. The related optimal effort level is equal to

$$e_{opt}(\Delta u) = \arg \min_{e \geq 0} h(\Delta u, e) \Rightarrow$$

$$e_{opt}(\Delta u) = 0 \text{ if } \Delta u \leq \frac{-1}{\lambda'(0)}; \quad e_{opt}(\Delta u) = \left(\lambda'\right)^{-1} \left(\frac{-1}{\Delta u}\right) \text{ if } \Delta u \geq \frac{-1}{\lambda'(0)}. \quad (8)$$

Hence the function c_λ is defined on the real line as the optimal effort. From the last equation, we obtain

$$\Delta u \leq \frac{-1}{\lambda'(0)} \Rightarrow c_\lambda(\Delta u) = \lambda(0) \times \Delta u, \quad (9)$$

and c_λ is linear in the neighborhood of 0, which corresponds to no effort. The function c_λ is strictly increasing because λ is strictly positive. If $\Delta u \geq 0$, we have that:

$$c_\lambda(\Delta u) = h(\Delta u, e_{opt}(\Delta u)) \geq e_{opt}(\Delta u) \Rightarrow \lim_{\Delta u \rightarrow +\infty} c_\lambda(\Delta u) \geq \lim_{\Delta u \rightarrow +\infty} e_{opt}(\Delta u) = +\infty.$$

Hence c_λ is an increasing homeomorphism on the real line.

The function c_λ is concave. From the envelope theorem, we have

$$h'_{\Delta u}(\Delta u, e) = \lambda(e) \Rightarrow c'_\lambda(\Delta u) = h'_{\Delta u}(\Delta u, e_{opt}(\Delta u)) = \lambda(e_{opt}(\Delta u)). \quad (10)$$

Hence c_λ is concave from the assumptions on λ and from the properties of e_{opt} .

We give a proof of the increasing property of the optimal effort level as a function of accumulated demerit points. From equation (1), we obtain

$$u_n - u_{n+1} = c_\lambda^{-1}(r(u_{\max} - u_n)), \quad u_{\max} = \frac{d_u}{r} \quad (0 \leq n < N). \quad (11)$$

The sequence $(u_n)_{0 \leq n \leq N}$ is decreasing because we have $u_{\max} \geq u_n$. Plugging this result into equation (11) implies that the sequence $(u_n - u_{n+1})_{0 \leq n < N}$ is increasing. The optimal effort level is denoted as e_n , and expressed as

$$e_n = \arg \min_{e \geq 0} e + [\lambda(e) \times (u_n - u_{n+1})] = e_{opt}(u_n - u_{n+1}),$$

for $0 \leq n < N$, where e_{opt} is defined from (8). As e_{opt} is an increasing function, the optimal effort is an increasing function of the number of demerit points for any given value of the license suspension threshold.

Let us specify the condition under which incentives are effective. From (11) and (9), we obtain

$$e_n > 0 \Leftrightarrow u_n - u_{n+1} = \frac{d_u - ru_n}{\lambda(0)} > \frac{-1}{\lambda'(0)} = \underline{\Delta u}. \quad (12)$$

If fines are included in the incentives, u_{n+1} is replaced by $u_{n+1} - \overline{fa}$ in equation (1), which leads to the recurrence equation (see Figure 4)

$$\begin{aligned} d_u - ru_n &= c_\lambda(u_n - u_{n+1} + \overline{fa}) \\ \Leftrightarrow u_{n+1} &= u_n + \overline{fa} - c_\lambda^{-1}(d_u - ru_n) = g(u_n). \end{aligned} \quad (13)$$

The fixed point of g is the lifetime driving utility if fines were the only incentive scheme, i.e.

$$\tilde{u}_{\max} = \frac{d_u - c_\lambda(\overline{fa})}{r}.$$

We evidently assume that $d_u > c_\lambda(\overline{fa})$, i.e. $\tilde{u}_{\max} > 0$. If the two incentives are mixed, we have $u_n \leq \tilde{u}_{\max}$ and we deduce from (13) the properties of utilities and of optimal effort levels as functions of n that we obtained in the first place. In addition, we have

$$e_n > 0, \forall n, \Leftrightarrow \overline{fa} + u_n - u_{n+1} > \underline{\Delta u}, \forall n.$$

This condition is fulfilled if

$$\overline{fa} > \underline{\Delta u} = -1/\lambda'(0),$$

in which case the incentives are effective at every level.

A.2 Incentive effects of point-record driver's licenses: Model with point removal

The Bellman equation on a holistic incentive model can be written as follows

$$d_u - ru(S) + \left(\frac{d}{dt} [u(S_t)] \right)_{t=0^+} = c_\lambda(\overline{fa} + u(S) - E[u(TR(S))]). \quad (14)$$

The state variables S are the seniorities of each non-redeemed traffic offense (if any), the related demerit points and the seniority of the last contract anniversary if the premium is included in the incentives. 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 anniversary) of duration t . The parameters d_u and \overline{fa} are the driving utility flow and the average fine, and $E[u(TR(S))]$ is the lifetime utility averaged with transition probabilities on the state(s) reached from S after a traffic offense. Continuity equations on utility at the time of a point removal or of a contract anniversary (in the latter case, the increase in lifetime utility is equal to the disutility of the premium) and the equation linking the utility of a beginner and the utility just after a license suspension define the solution together with equation (14).

Let us prove the continuity of optimal effort after a point removal in a system where each traffic violation is redeemed beyond a given seniority threshold, equal to T . We suppose that each traffic violation is associated with one demerit point, and that incentives are related to fines and to the point-record driver's license. The state variables are then the seniorities of each non-redeemed traffic offense, if any. Let us denote these variables as

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

The corresponding optimal effort is denoted as $e(S)$. The states reached without traffic offense before the next point removal are then

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

We denote the state reached from S after an additional traffic offense (if $n < N$) as $(0, t_1, \dots, t_n) = TR(S)$. As the lifetime utility is continuous after a point removal, we have the following result:

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

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

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

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

From the three last equations, we obtain

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

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

A.3 Incentive effects of the experience rating system

Let us derive the Bellman equation on the expected disutility function given in (15), including an average fine of fa_j for a j demerit point traffic violation. The optimal disutility function is obtained from the program

$$v_n(t) = \min_{e \geq 0} edt + (\exp(-rdt) \times (1 - \lambda(e)dt) \times v_n(t + dt)) \\ + \left(\exp(-rdt) \times \left[\sum_{j / f_j > 0} f_j \lambda(e)dt \times [v_{\min(n+j,N)}(t + dt) + fa_j] \right] \right) + o(dt),$$

which leads to

$$0 = v'_n(t) + c_\lambda \left(\overline{fa} + \left(\sum_{j / f_j > 0} f_j v_{\min(n+j,N)}(t) \right) - v_n(t) \right) - rv_n(t),$$

with $\overline{fa} = \sum_{j / f_j > 0} f_j \times fa_j$ the average fine. Then we obtain the Bellman equation

$$v'_n(t) = rv_n(t) - c_\lambda (\overline{fa} + \Delta v_n(t)), \quad (0 \leq n \leq N). \quad (15)$$

The argument of c_λ , $\overline{fa} + \Delta v_n(t)$, is the incentive level.

A.4 Test for the increasing shape of a vector of parameters, and for evidence of moral hazard

Let θ be the vector of values reached by the function g_1 , where g_1 is defined in equation (7). We have $\theta_{adp} = g_1(adp)$, $1 \leq adp \leq 13$, with adp the number of demerit points accumulated in the last two years. In what follows, θ can be

replaced by g_1 , but we use here the usual notations for asymptotic tests. The level $adp = 13$ also includes the 14 point total. As $adp = 0$ is the default level in the regression, the increasing shape of g_1 is related to the null assumption

$$g_1(0) = 0 \leq \theta_1 \leq \dots \leq \theta_{13}. \quad (16)$$

Rejecting this assumption amounts to finding evidence of moral hazard. The null assumption is obviously expressed as a set of inequalities of the type $f_i(\theta) \geq 0$ ($i = 1, \dots, 13$), with

$$f(\theta) = \underset{i}{vec}(f_i(\theta)) = J \times \theta, \quad \underset{13,13}{J} = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ -1 & 1 & \ddots & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & 0 & \ddots & 1 & 0 \\ 0 & \dots & 0 & -1 & 1 \end{pmatrix}.$$

A usual asymptotic test for the nullity of $f(\theta)$ derived from $\hat{\theta}$, the maximum likelihood estimation (m.l.e.) of θ – estimated here together with nuisance parameters – is the Wald test. The related statistic writes as follows

$$W = \|f(\hat{\theta})\|_M^2 = d_M^2(f(\hat{\theta}), 0), \quad M = \left(\hat{V} \left[f(\hat{\theta}) \right] \right)^{-1} = \left[J_f(\hat{\theta}) \times \hat{V}(\hat{\theta}) \times {}^t J_f(\hat{\theta}) \right]^{-1},$$

where $J_f(\hat{\theta})$ is the Jacobian of f ($J_f(\hat{\theta}) = J$ as f is linear), and where $\hat{V}(\hat{\theta})$ is derived from the asymptotic efficiency of the m.l.e. Under the null, the Wald statistic follows a $\chi^2(13)$ distribution asymptotically. The null (i.e. $f(\theta) = 0 \Leftrightarrow \theta = 0$ as f is one-to-one; as J is invertible, we have $W = \|\hat{\theta}\|_{[\hat{V}(\hat{\theta})]^{-1}}^2$) is rejected from our data at every usual level, as $W = 1933.12$.

The increasing shape of g_1 is the null to be tested for with inequalities, and is expressed as follows

$$H_0 : f_i(\theta) \geq 0 \quad \forall i = 1, \dots, 13 \Leftrightarrow f(\theta) \in O^+,$$

where O^+ is the positive orthant of \mathbb{R}^{13} . In order to test for H_0 against the alternative, a natural statistic derived from the Wald test uses the distance

between the unconstrained estimation of positivity constraints and the positive orthant, i.e.

$$W^+ = d_M^2(f(\hat{\theta}), O^+) = \min_{z \in O^+} \|f(\hat{\theta}) - z\|_M^2 = \|f(\hat{\theta}) - z_0\|_M^2, \quad z_0 = P_{O^+}(f(\hat{\theta})).$$

The metric M was already used for the Wald statistic. As 0 belongs to the positive orthant, we have $W^+ \leq W$. The derivation of the projection $P_{O^+}(f(\hat{\theta}))$ of $f(\hat{\theta})$ on the positive orthant O^+ can be seen numerically as a quadratic programming problem. As this projection belongs to the frontier of the positive orthant, the issue is to find the set of coordinates that are strictly positive.

The asymptotic properties of the statistic W^+ under the null are described by Wolak (1991). The statistic is not asymptotically distribution free under the null, as is the case in the equality setting. The limit distribution is a mixture of $\chi^2(i)$ distributions ($0 \leq i \leq 13$), where $\chi^2(0)$ is the unit mass (Dirac) at zero. If $f(\theta)$ belongs to the interior of O^+ , the limit distribution of W^+ is obviously equal to $\chi^2(0)$, due to the pointwise convergence of $f(\hat{\theta})$ towards $f(\theta)$. Otherwise the weights of the mixture depend on the position of $f(\theta)$ at the frontier of the positive orthant. As a $\chi^2(i+1)$ variable is obtained as the sum of a $\chi^2(i)$ variable and of a nonnegative variable, the $\chi^2(i)$ distributions increase with i in the sense of first order stochastic dominance. Then all the limit distributions of W^+ will be first order-dominated by a $\chi^2(13)$ distribution, owing to their definition as mixtures. This result is also simply obtained from the inequality $W^+ \leq W$.

On our data, the projection $P_{O^+}(f(\hat{\theta}))$ is found equal to the projection of $f(\hat{\theta}) = J \times \hat{\theta}$ on the subspace of \mathbb{R}^{13} defined by $\{z \in \mathbb{R}^{13} / z_7 = \dots = z_{13} = 0\}$. We have: $P_{O^+}(f(\hat{\theta})) = f(\hat{\theta}^0) \Rightarrow \hat{\theta}^0 = J^{-1} \times [P_{O^+}(J \times \hat{\theta})]$, with

$$\begin{aligned} {}^t\hat{\theta} &= \left(\begin{array}{cccccccccc} 0.311 & 0.500 & 0.522 & 0.794 & 0.854 & 0.873 & 0.974 & \dots & 0.571 \end{array} \right); \\ {}^t\hat{\theta}^0 &= \left(\begin{array}{cccccccccc} 0.312 & 0.500 & 0.523 & 0.795 & 0.855 & 0.870 & 0.870 & \dots & 0.870 \end{array} \right); \end{aligned}$$

The missing components of $\hat{\theta}$ can be found in Table 2, and we have $\hat{\theta}_6^0 = \dots = \hat{\theta}_{13}^0$. Hence $\hat{\theta}^0$ is the vector of parameters with a weakly increasing shape that is the

closest to our estimation with respect to the metric ${}^tJMJ = [\widehat{V}(\widehat{\theta})]^{-1}$. Both vectors are represented in Figure 3 (with θ replaced by g_1). From our data, the statistic W^+ is equal to

$$W^+ = \|f(\widehat{\theta}) - f(\widehat{\theta}^0)\|_M^2 = \|\widehat{\theta} - \widehat{\theta}^0\|_{[\widehat{V}(\widehat{\theta})]^{-1}}^2 = 43.18,$$

The statistic W^+ is a squared distance between the unconstrained estimation of g_1 and the set of weakly increasing functions of adp . If F is the distribution function of a $\chi^2(13)$ variable, we have that $F(43.18) = 0.999958$. If the null assumption: $f(\theta) \in O^+$ is tested for with rejection regions of the type $[W^+ > a]$, it will be rejected at any significance level greater than $1 - 0.999958 = 42 \times 10^{-6}$. To see this, it is enough to apply the first order stochastic dominance property given before. Then the increasing shape of g_1 is rejected at every usual significance level, which shows evidence of moral hazard.

A.5 Elasticity between the optimal frequency risk and the incentive level

We relate the shape of λ and the elasticity between optimal frequency risk and the incentive level. Let us perform a local expansion around a value Δu^0 of the incentive level, in a situation where the incentives are effective (i.e. $\Delta u^0 > \underline{\Delta u} = -1/\lambda'(0)$). If we write

$$e^0 = e_{opt}(\Delta u^0), \quad e^0 + de = e_{opt}(\Delta u^0 + d\Delta u),$$

the equations

$$1 + \lambda'(e^0)\Delta u^0 = 0; \quad 1 + \left[\lambda'(e^0 + de) (\Delta u^0 + d\Delta u) \right] = 0$$

lead to

$$de = \frac{-\lambda'(e^0)}{\lambda''(e^0)\Delta u^0} d\Delta u + o(d\Delta u),$$

and to

$$\frac{d\lambda}{\lambda(e^0)} = \frac{\lambda'(e^0)}{\lambda(e^0)} de = \frac{[-\lambda'(e^0)]^2}{\lambda(e^0) \times \lambda''(e^0)} \times \frac{d\Delta u}{\Delta u^0}.$$

Hence the aforementioned elasticity is equal to $(\lambda')^2 / \lambda\lambda''$. Now we have that

$$(\log \lambda)'' = \frac{\lambda''}{\lambda} - \left(\frac{\lambda'}{\lambda}\right)^2 = \frac{\lambda''}{\lambda} \left(1 + \frac{(\lambda')^2}{\lambda\lambda''}\right).$$

Then the conclusions given in Section 4.2 are easily obtained.

A.6 Monetary equivalents of traffic violations

Let us suppose that the increase in the argument of c_λ is close to 12% after a traffic violation. This is the value retained in Section 4.2 for a driver with seven demerit points accumulated, which corresponds to $n = 3$ traffic violations on average. As the argument of c_λ in the model without point removal and with fines is equal to $\bar{f}a + u_n - u_{n+1}$ (see equation 13) we have that

$$\bar{f}a + u_{n+1} - u_{n+2} = 1.12 \times (\bar{f}a + u_n - u_{n+1}). \quad (17)$$

We shall compare the utility losses $u_{n+1} - u_{n+2}$ and $u_n - u_{n+1}$ from the recurrence equation on lifetime utility, and obtain a monetary equivalent of an additional traffic violation from a derivation of the utility loss $u_n - u_{n+1}$. We have

$$u_{n+1} = g(u_n), \quad g(u) = \bar{f}a + u - c_\lambda^{-1}(d_u - ru)$$

(see Figure 4). From the equality $c'_\lambda(\Delta u) = \lambda(e_{opt}(\Delta u))$ (see equation (10)), we obtain

$$g'(u_n) = 1 + \frac{r}{\lambda_n}, \quad \lambda_n = \lambda(e_{opt}(\bar{f}a + u_n - u_{n+1})).$$

The parameter λ_n is the frequency risk corresponding to the optimal effort exerted with n traffic violations accumulated. As λ_n decrease with n , we have that

$$1 + \frac{r}{\lambda_n} \leq \frac{u_{n+1} - u_{n+2}}{u_n - u_{n+1}} = \frac{g(u_n) - g(u_{n+1})}{u_n - u_{n+1}} \leq 1 + \frac{r}{\lambda_{n+1}}. \quad (18)$$

From equations (17) and (18), we obtain

$$\left(\frac{r}{\lambda'_n} - 0.12\right) \times (u_n - u_{n+1}) = 0.12 \times \overline{fa} = 15.6 \$, \lambda_{n+1} \leq \lambda'_n \leq \lambda_n.$$

The monetary equivalent of an additional traffic violation is then bounded as follows:

$$\Leftrightarrow \frac{15.6 \$}{\frac{r}{\lambda_{n+1}} - 0.12} \leq u_n - u_{n+1} \leq \frac{15.6 \$}{\frac{r}{\lambda_n} - 0.12}.$$

Section 4.2 provides numerical examples with $\lambda_n = 0.15$, $\lambda_{n+1} = 0.12$. The monetary cost of a license suspension follows from a sum of the items related to traffic violations and from the inequalities given in (18).

References

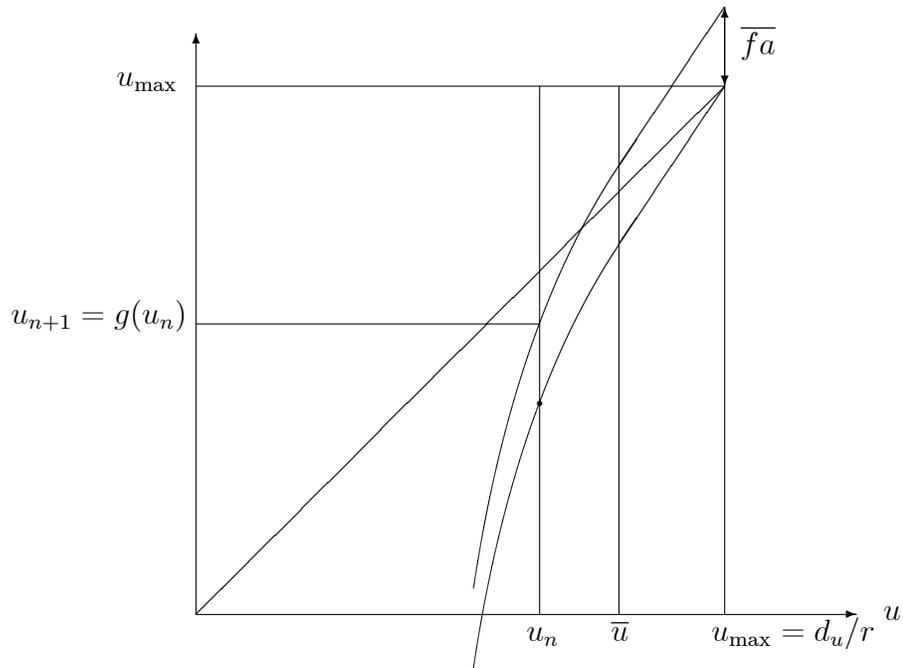
- [1] Wolak, Frank A., "The Local Nature of Hypothesis Tests Involving Inequality Constraints in Nonlinear Models," *Econometrica* 59 (1991), 981-995.

Figure 4

Recurrence equation on the lifetime utility function

Point-record driver's license without fines: $u_{n+1}^0 = f(u_n^0)$

Point-record driver's license with fines: $u_{n+1} = g(u_n)$



$$f(u) = u - c_\lambda^{-1}(r(u_{\max} - u)), \quad g(u) = f(u) + \bar{f}a.$$

Effective incentives condition with and without fines

$$e_n > 0 \Leftrightarrow u_n < \bar{u} = u_{\max} \left(1 + \frac{\lambda(0)}{\lambda'(0) \times d_u} \right).$$